Is the episodic buffer of Baddeley's model independent of the central executive? A new measure of the episodic buffer

J. Óscar Vila-Chaves, Francisco Gutiérrez-Martínez, and J. Antonio García-Madruga

Department of Developmental and Educational Psychology. Universidad Nacional de Educación a Distancia (UNED).

Título: ¿El Retén Episódico de Baddeley es independiente del Ejecutivo Central? Una nueva medida del Retén Episódico.
Resumen: El Retén Episódico (RE) está tomando un creciente papel central en las explicaciones sobre el funcionamiento de la memoria operativa. De hecho, los últimos estudios de Baddeley y sus colaboradores sitúan al RE en el corazón del sistema de memoria. Recientemente la discusión también atañe a si este componente de la memoria operativa presenta una naturaleza independiente respecto a los recursos del ejecutivo central. Algunos estudios muestran como en tareas automatizadas, la construcción y mantenimiento de elementos almacenados en el RE no requieren de recursos desde el ejecutivo central.
El presente trabajo pretende analizar esta cuestión para lo que se toman diferentes variables y se ha diseñado una nueva prueba para medir el RE. En esta prueba de doble tarea, la tarea de procesamiento consiste en la lectura de textos sencillos con unos espacios en blanco que deben completarse con las palabras clave. Los resultados muestran cómo a pesar del aumento del procesamiento debido al incremento de la longitud de los textos, no se produce un aumento en la carga del ejecutivo central, ni en la creación de los agrupamientos de información, ni en su mantenimiento. Por ello, pensamos que el RE bajo ciertas circunstancias opera de forma independiente al ejecutivo central.
Palabras clave: Memoria Operativa. Retén Episódico. Ejecutivo Central. Lectura. Memoria a Largo Plazo.

Abstract: The Episodic Buffer (EB) is taking a growing central role in explanations regarding the functioning of working memory. In fact, in the most recent studies by Baddeley and his collaborators, the EB has situated itself at the core of this memory system. Recently, the discussion also concerns whether this component of working memory seems to demonstrate an independent nature with respect to central executive resourcing. Some studies show that in automatic tasks the creation and maintenance of elements stored in the episodic buffer do not require resources from the central executive.
The current work attempts to evaluate this assumption for what different variables are taken and a new test has been developed to measure the EB. In this double task test, the processing task consists of reading short simple texts that contain missing words. The results show that further processing due to increasing the length of the texts does not correspond to higher load demands made on the central executive, nor in the creation of chunks or their maintenance. Thus, we think that the EB is under certain circumstances independent of the central executive.
Keywords: Working Memory. Episodic Buffer. Central Executive. Reading. Long-term memory.

Introduction

Working Memory (WM) has become an essential component in explaining cognitive functioning. It is understood primarily as an active memory system in which information required for performing cognitive tasks is temporarily held and processed. Despite the importance it holds for the principal models of cognition, there is still no widespread agreement on their nature and conceptualization (Miyake & Shah, 1999). In any case, the most important explanatory models of WM do agree on stressing two characteristic features: 1) the participation of a domain general Central Executive that is mainly in charge of focusing and shifting attention, activating and updating representations, and inhibiting automatic processing and irrelevant information (Baddeley, 2007; Cowan, 2005; Engle, 2002; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000); and 2) a limited set of resources available to the system in undertaking its two basic functions: the temporary storage of information relevant to the current task at hand and, simultaneously, processing this and any other concurrent information required by the task (Baddeley & Logie, 1999; Just & Carpenter, 1992).

Work in recent years has revived the debate regarding the way in which WM is related to content stored within Long Term Memory (LTM). It is unclear how WM uses prior knowledge related to an ongoing task in order to improve its performance. Nor is it clear whether it does so by optimizing its processing function, its storage function, or both. Nevertheless, experts do agree that working memory nourishes itself from information held in LTM, and particularly relevant is the grammatical and semantic knowledge involved (Baddeley, 2010). In this regard, for example, we know that memory for familiar words is substantially better than it is for unknown words, demonstrating the influence knowledge stored within LTM has. Similarly, it is easier to remember words that form a complete sentence than it is to remember a list of unrelated words (as found in the work of Brener, 1940).

In the same way, various studies show how expert performance is based on the use of more efficient recall cues that enable a reduction in processing time and foster the ability to handle more information. The improvement is such processing results in fewer demands on cognitive resources, which in turn increases the "residual" capacity of available resources for storage (Daneman & Carpenter, 1980; Ericsson & Kintsch, 1995; Kane, Hambrick, Tuholski, Wil-

* Correspondence address [Dirección para correspondencia]:
J. Óscar Vila Chaves, Department of Developmental and Educational Psychology, School of Psychology. UNED. C/Juan del Rosal, 10. Madrid (Spain). E-mail: jovila@psi.uned.es
(Article received: 02-10-2018, revised: 24-06-2019, accepted: 16-11-2019)

helm, Payne & Engle, 2004). In a similar way, the use of prior knowledge also translates into an overall increase in storage capacity, and this may be related to the use of strategies that are directly geared to improved storage. For instance, since the 20th Century we have known how strategies such as information "grouping" (or "chunking") allows a person to "extend" their capacity to remember things over the short term (see, for instance, Miller, 1956; Tulving & Patkau, 1962; Ericsson & Kintsch, 1995).

In short, while the involvement and incidence of prior knowledge in WM is unquestionable, what remains unclear is the way in which information within LTM relates and integrates with information currently active that is managed, manipulated and maintained within WM. Today, what is of particular interest is related to the processes "linking" or "binding". Such "binding" may serve as a necessary mediator in the construction of "chunking", as happens naturally during the process of understanding a text, and appears may be one of the main operational features of WM.

The multi-component model proposed by Baddeley & Hitch

Among the most important of theories regarding WM, and one that stands out for its particular influence, is the multi-component theory proposed by Baddeley and Hitch (1974; Baddeley, 1986, 2000). This will serve as a main reference for the current work. This theory presents a WM system that includes four main components: two slave subsystems for temporal storage (the phonological loop and the visual-spatial sketchpad), one system of attentional control (the central executive), and one temporal storage unit in charge of coordinating present information active in WM with that of LTM (the Episodic Buffer) (see Figure 1.).

The phonological loop and the visual-spatial sketchpad are subsystems specialized in the transitory processing and maintaining, respectively, of verbal and visual-spatial information. Both of these subsystems are composed of passive storage and a processing unit in order to refresh the information.

Figure 1
The multi-component model of working memory (Baddeley, 2000).
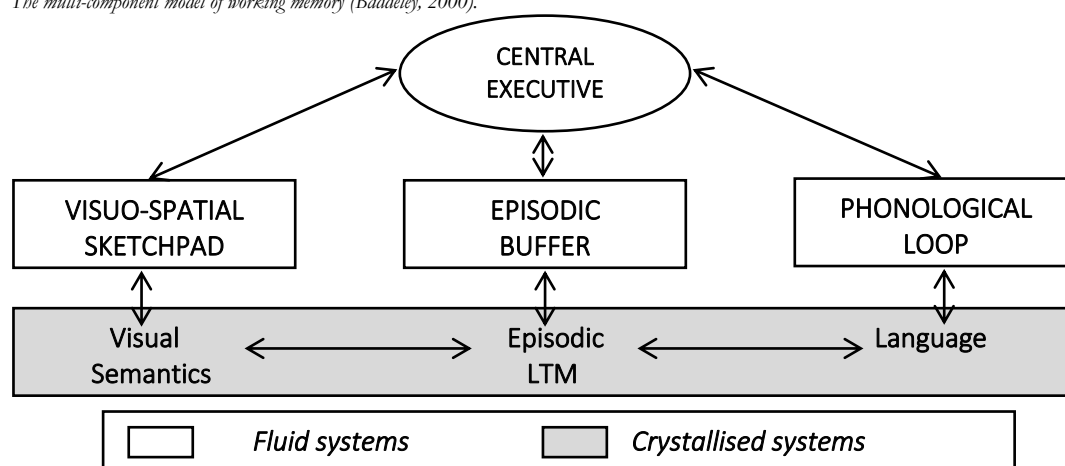
On the other hand, the central executive (CE) coordinates and manages other WM components responsible for controlling attentional resources and monitoring the processing of information. It is thus ultimately responsible for the management and regulation of all such activity within the cognitive system. Initially, the theory conceived of CE as that responsible for processing and storage itself. However, later, these functions became more associated with belonging to different components. In this way, the CE became considered as that in charge of resource processing and control, while storage became considered as an additional process dependent on another component of the system, one known as the Episodic Buffer (EB).

The EB was added by Baddeley in 2000 and is characterized as a temporary storage component that uses a multi-dimensional code, is accessible to consciousness, and has limited capacity (Baddeley, Hitch & Allen, 2009; Allen, Bad-

deley & Hitch, 2014). It serves as the "mental space" in which information from the slave subsystems (the phonological loop and visual-temporal sketchpad) is integrated with linguistic and semantic information stored within LTM relevant to whatever task is at hand. It is this function that permits the combination of all information and its integration into multi-dimensional clusters (i.e., "chunks"), which Baddeley has called "episodes." To do so requires an episodic storage space that, according to recent studies (Baddeley et al., 2009; Baddeley, Allen, & Hitch, 2011), seems to function more passively than that initially proposed (Baddeley, 2000).

Specifically, what Baddeley has recently proposed is that under certain conditions (low level of complexity, common and/or related materials) the processing required to integrate information is basically automated and does not require additional support from the CE. In this way, the maintenance of information deposited as "chunks" in the EB is in no way

affected by processing demands (Baddeley et al., 2009). Nor should any interference occur between the cognitive load required to processes information and that it required for its storage in the EB. In this work, Baddeley *et al.* present four experiments in which, from different conditions and concurrent tasks, they compare the execution of the subjects in the memorization of three types of word sets. From the same set of words, they are constructed: 1) "*constrained lists*", are lists of words without syntactic structure, 2) "*constrained sentences*", phrases with a certain syntactic meaning (see Table 1). In addition, in some experiment 3) phrases from news are also used, which therefore have a correct syntactic structure and do not use the same words as the other two sets "*open sentences*". The objective of using the same word pool is that proactive interference will minimize the contribution of episodic and semantic long-term memory. Thus, participants are forced to rely on the existing temporary grouping in working memory to support his memory. That is, they must be based on the syntactic structure of the phrases that are constructed with those words. In one case the structure exists and in other cases it is almost non-existent. In this way, the semantic components are limited or equalized to avoid their facilitating effect on the integration of information (binding). As concurrent task, different tasks, such as counting backward tasks, N-back with different conditions, test articulatory suppression, and visuospatial tasks are used in the different experiments.

In general terms, the work shows how despite the use of different concurrent tasks to load the EC, the results are quite stable showing that the best execution occurs when the words are presented in a more recognizable syntactic struc-

ture, without finding differences between an oral or visual presentation of the items. Likewise, the highest proportion of semantic errors occur to a greater extent when the syntactic structure is adequate indicating that the reading is carried out in a more superficial way guided by the syntax. The other relevant issue is that the execution in the concurrent task is not affected, so the lack of involvement of the CE in the primary task is confirmed.

Table 1
Examples of sequences used in Baddeley y cols. (2009).

	Example of items
Constrained List	A. JOHN CAR INSTANTLY LARGE BORROWED WHITE B. SOLDIER TALL WAITER TEACHER FOLLOWS AND OLD SAD
Constrained sentence	A. LUCY the OLD PILOT RAPIDLY BORROWED the SMALL RED BOOK B. TALL SOLDIER FOLLOWS WAITER AND OLD SAD TEACHER

This newly formulated model (Baddeley et al., 2011) now places the Episodic Buffer in the core of WM (see Figure 2), involving the CE in the functions of integration and maintenance of information within the EB, but only when the processes that are required for the integration of information (i.e., the processes of "binding") have not been automated. When these processes are automated, they no longer require the CE's participation (Allen & Baddeley, 2009; Baddeley et al., 2009; Allen, Baddeley & Hitch, 2006).

Figure 2
A revised model of working memory in fluid systems (Baddeley, Allen & Hitch, 2011). The new proposal only includes those systems of a fluid character, but the authors have not modified the way in which these systems are related to the crystalized systems in LTM.
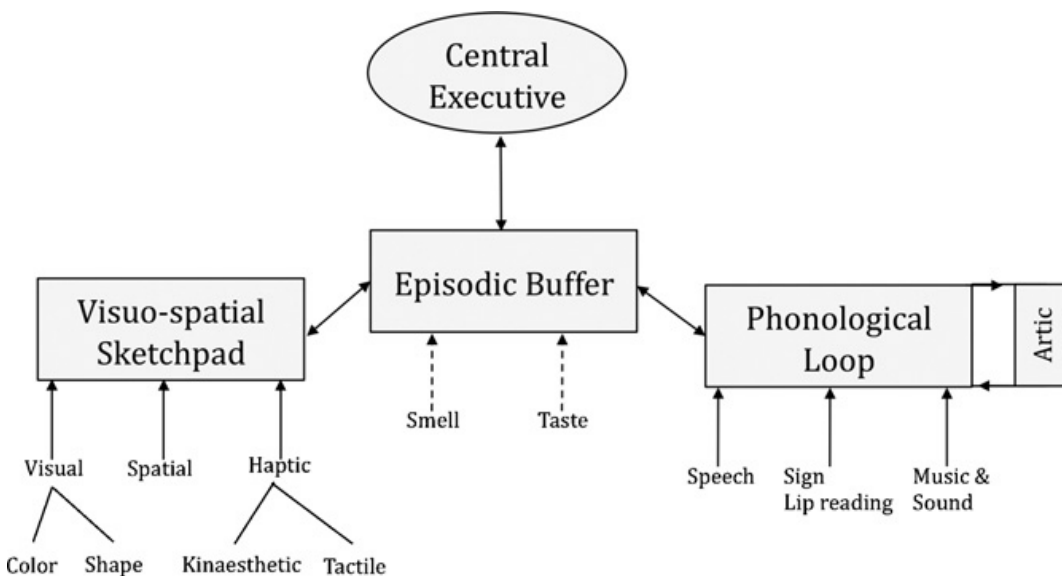
It is in this context that our main interest lies: to examine whether, as Baddeley et al. (2009, 2011) have laid out, the Episodic Buffer is partially independent of the central executive and, if so, it can operate in certain conditions without requiring its assistance. This brings us to the debate on whether or not the two-basic function of WM (processing and storage) are independent or not. Baddeley's new proposal contrasts somewhat with other theoretical models that argue that both functions, processing and storage, are interrelated. These models assume that in performing any single task, the CE manages a single pool of resources, and that the management of available resources implies a trade-off of between the two functions (Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007; see also, e.g., Daneman & Carpenter, 1980; Case, 1985; Turner & Engle, 1989; Just & Carpenter, 1992; Anderson, Reder, & Lebière, 1996).

In particular, our work aims to continue in the line of the studies of Baddeley et al. (2009) (see Table 1) discussed above, where the storage capacity of the Episodic Buffer was measured on memory tasks for distinct types of word lists that would limit its semantic properties and thus prevent the facilitation effect of the semantic integration of information (i.e., prevent "semantic binding"). These studies have shown how increasing processing load in *constrained sentences* can presumably occur without any interference in storage, since word maintenance is significantly higher compared to its incidence on the restricted list.

Specifically, our work aims to contrast in a different way the results presented by Baddeley et al. (2009) where they take as measure of the EB the storage capacity of the subjects in recall tasks on different types of word lists while performing a concurrent task that requires CE resources. As we described above, the idea is to try to show how the processing load in the secondary (concurrent) task can be increased without significantly interfering with the storage and maintenance of the information in the primary task.

In this paper, we explore a new recall task in which participants must comprehensively read a few simple narratives having phrases designed such that understanding requires few attentional resources. In this way, the CE is minimally involved, just as there will exist minimal interference in the creation of "chunks" in the Episodic Buffer (EB). In this way, the CE is minimally involved, because the demand for attention resources by the EB will be minimal, since the texts require the construction of very simple mental models. Simple texts facilitate the organization of information (binding) and probably also the necessary semantic grouping for storage (chunking). In this way, our proposal has similar goals as that of the studies of Baddeley et al. (2009), but in the opposite direction and with certain differences. Instead of minimizing the effect of the participation of the CE by avoiding the semantic features of sentences and relying and prioritizing the syntactic ones, we propose a task that utilizes a very simple semantic context in order to minimize the participation of the CE. Thus, we intend to measure the passive capacity of the Episodic Buffer through the use of easy-to-

understand sentences coordinated in a context of global understanding. Moreover, unlike demands of the classic task, in this way CE resources can be to a greater extent dedicated to the "re-activation" of words previously used in the Analogy span task. This can also be understood in terms of the priming effect, as it "pre-activates" the words in the analogies task that later are "re-activated" in simplified form within the context of the reading. We think, that in any case, the re-activation could be carried out with the support of the situational representation generated during comprehension, as often occurs in everyday reading tasks. As detailed below, this provides a more "ecological" approach of WM functioning in general, and of the EB itself in particular, given all resources available to the participant will be measured in a more natural way.

This contextualization allows that the "binding" and "chunking" processes to be carried out on informative elements with semantic coherence in the context of a text. We think it is important that the measurements be carried out based on cognitive tasks similar to those that participants perform in daily life, in particular the measurement of EB due to its characteristics of integration of the MLP information. Therefore, unlike Baddeley et al. (2009), we will not use word lists to be "organized and/or grouped", but semantic contexts (texts) that allow the creation of more significant mental models for participants and facilitate the integration of information (*binding processes*) and the creation of groups (*chunking processes*).

Measuring the Episodic Buffer

Our proposal aims to analyze the performance of the EB and its dependence or not on general attentional processes. To do this, we should try to specifically measure the EB, that is, measure the capacity of the EB to store the chunks that are being developed and maintained.

But in addition, we think that this measure should be done in a more natural cognitive context than that involved in the traditional dual-task paradigm −or "concurrent processing"− directed at compromising CE resources. To do this, we use a habitual task of cognitive processing, such as reading comprehension, which requires coordinating the active information in WM with the related information that is stored in the LTM in order to integrate all information and create "semantic fragments" that must maintained simultaneously; that is, a task that necessarily implies a "natural" use of EB. Likewise, we cannot forget the close and well-known relationship between MO measures and reading comprehension (see, for example, Daneman & Merikle, 1996; Hannon & Daneman, 2001, 2004; Cain, Oakhill, & Bryant, 2004; Vukovic & Siegel, 2006; Carretti, Borella, Cornoldi y De Beni, 2009; García-Madruga et al., 2013; García-Madruga, Gómez-Veiga, & Vila, 2016).

In the work we present, a verbal test is used in which the participants must read several texts, taking into account that according to Baddeley (2000, 2007, 2012), the EB must be

the mental space in which they must be carried out the processes involved in reading comprehension.

Reading comprehension requires the participation of the processes that carry out the integration of information (binding processes), and the processes necessary to build semantic clusters (chunks) that allow the understanding of the text. It is likely that these "representational chunks" facilitate the recall of the concrete items (words) that have been activated to build them.

Also, to ensure the participation of EB, the memory that subjects must do in this task happens after at least 20 seconds of the last activated keyword. The texts contain an introductory part of the reading to force the delay time thus facilitating the incorporation of more long-term resources (Alloway & Ledwon, 2014; Conway & Engle, 1994; Cowan, Wood, Nugent & Treisman, 1997; Cowan 2001). After the introduction phrases, the texts present a series of incomplete sentences that must be filled in with the previously activated keywords.

The test we propose has been designed based on the Analogies Reasoning Span Test (Gutiérrez-Martínez, García-Madruga, Carriedo, Vila & Luzón, 2005; García-Madruga, Gutiérrez-Martínez, Carriedo, Luzón & Vila, 2007). This is a classical dual-task measure in which participants read aloud a series of simple verbal analogies, find and read out loud the solution, and remember the word-solution of each analogy at the end. In the new test, one second measurement is added, so that the procedure includes two phases which will permit us to take two successive measures of WM: one index based on the cited analogies test show that basically measures CE participation, and another that presumably measures the Episodic Buffer capacity.

To summarize, in addition to the Analogies span task, the new procedure requires participants to perform a second recall task that consists of comprehensively reading short texts that contain some missing words. To reach the correct understanding of the texts, participants must then complete the gaps in the texts with words inferred and stored previously in the analogies task.

Thus, the new procedure consists of two consecutive stages: stage one made up of the analogies span test, which we will call ANALOGY; and a second stage with the new task of short texts, which we will call CONTEX:

1) ANALOGY. Test participants begin by solving a growing series of analogies (2, 3, 4 and 5) the result of which should be stored in memory and recalled at the end of each series in the correct order of appearance (they must recall all stored words in each series; 2, 3, 4 or 5).

2) CONTEX. Then participants are presented with text containing a brief introduction and a number of phrases corresponding to the number of analogies each series contains. Each sentence ends with a blank space that must be filled in with the keywords previously stored during the analogy task, thus offering a second chance to remember them. In this way, the memory for keywords is recorded at two times: one immediately after solving the analogies, and a second delayed memory that must be integrated into the texts (see Figure 3).

Figure 3
Basic procedural structure of the new test (a level 2 example).
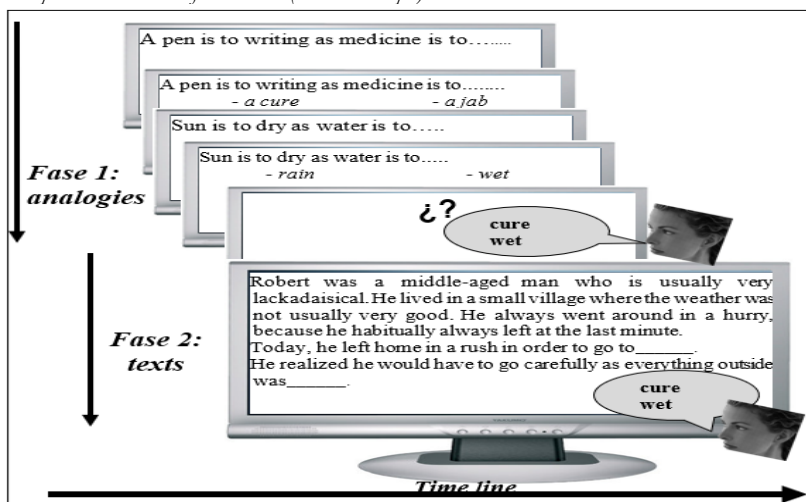
Note: This example is a translation of the Spanish task so it can several some of the controlled conditions.

As we have described above, we assume that the first criterion that must fulfill the new task is that it registers the processes specific to the Episodic Buffer with minimum participation by the CE. Accordingly, we developed this task by following an approach that somehow reverses that of Baddeley et al. (2009). That is, instead of avoiding semantic processes, we used a task that necessarily requires this kind of process in order to achieve comprehension. The key to this is the simplicity of texts used in this new measure and that results from previous studies have not encountered significant differences in neither reading speed nor level of comprehension in adolescent participants (Gutiérrez-Martínez et

al., 2005; García-Madruga et al., 2007). Therefore, it seems normal readers may perform the semantic integration in semi-automatic ways, constructing the representational model of the text in a simple manner, both within and between sentences. Thus, our task makes the binding processes necessary to complete it even more automatic, by using syntactically simple and similar sentences that have less demanding content. However, it is important to note, as pilot studies have shown, that the semantic content of sentences would not allow participants to guess the keywords without prior pre-activation into ANALOGY task, given that those texts were designed with blanks that might be completed with different words to the keywords (keywords were remembered with a success rate of less than 5%).

In sum, for an undemanding task for expert readers (i.e., university-level), we believe that its correct completion would require minimal activation of the CE's attentional resources. This would permit us to attribute any differences found in the new task to the independent or specific capabilities of the Episodic Buffer. Thus, Baddeley´s new proposal which affirms the relative independence of EB and its eminently passive character could be tested.

CONTEX, an ecological proposal to measure Episodic Buffer Span

Additionally, our idea also revolves around the possibility of making a more natural or ecological, or, let us say, contextualized index of the Episodic Buffer. We understand that this new contextualization task occurs via two different pathways. First, the words that must be retrieved have been previously activated and stored in the task of analogies, which can be a facilitating element. Also, in the second part of the task, the representational model that must be constructed during the comprehension of the text, requires the activation of previous knowledge that supposes a favorable semantic context to "re-activate" the previously stored keywords. We believe that the activations and chunking caused by comprehensive reading can favor the recovery of stored words.

We have tried to reproduce the context for the task in which WM resources can be utilized naturally, just as when used in daily tasks in which processing and storage functions operate together. In this way, we seek to distance ourselves from procedures that employ concurrent tasks that are not naturally related, and thus serve different purposes. For example, in the classic Reading Span Test (the RST of Danneman & Carpenter, 1980). The subject must read a series of sentences and store the last word of each sentence and retrieve them at the end in their correct order. Both tasks are independent, and the final goal is to remember a sequence of specific words (without an ulterior motive), an uncommon occurrence in daily activities. Indeed, the researcher might choose a different word to remember instead of the last, and the result would remain the same, given that the tasks are not connected. The task can even be solved correctly with-

out completing understanding the sentences; a superficial reading is enough. Nevertheless, comprehensive reading done in a natural context involves storing those important elements that favor comprehension, not the items located in a particular place.

In our proposal, the new test has been designed in order that the information stored in the first task is itself the ultimate goal of the second task: full understanding of the new text. With this, we procure that the characteristic functions of WM, processing and storage, occur naturally in the context of a routine cognitive task. In this case, memory per se is not the final goal of the task, yet it serves a higher purpose, of natural and habitual character in line with completely understanding a text.

In order to analyze whether the processes involved in CONTEX related to the episodic buffer depend or not on the CE resources, the reading time of the sentence was measured to confirm whether it remains stable despite the increase in processing as the task progresses. If in CONTEX the reading time remains stable at all levels, we would have one more indicator of the low intervention of the EC. Other studies have confirmed that in double-task tests that use reading tasks, when storage demand increases, subjects use more time during processing as they must share their resources.

The design of a task that we think measures EB, requires contrasts with other WM tasks and with measures of high-level cognitive processes such as intelligence, reasoning and reading comprehension.

Approach and objectives of the study

The current study has as main objective to analyze if the EB can make the necessary integrations to store and maintain the resulting "fragments" without significant participation of the EC. To do this, we will measure whether the activations required to create the "representational chunks" facilitate the recovery of words that make up those chunks. A daily task will be used to take advantage of the facilitating effect of the context and in which there is a low participation of the CE in the process of organizing information by using very simple texts that require a semi-automated understanding. For this purpose, it is necessary to check whether the context of "reading comprehension" with simple texts does not really increase the processing load, and hence provokes at least a minimum of participation by the CE. If so, and consistent with Baddeley, we can say that the basic functions of WM (processing and storage) are practically independent in these types of tasks in which the processing component is highly automated. It is in the context that we present a new task (CONTEX) that measures WM capacity from those processes specifically related to the Episodic Buffer, and that also meets another one of the objectives outlined above: the assessment of WM in a more natural and ecological way.

The results in this task will be contrasted with another two measures of WM used in various studies, and which

mainly report on the participation of processes more closely related with the CE (the Reading Span Test—RST, and the Reasoning Span Test based on Analogies—ANALOGY). Thus, we think that the new task measures the most specific processes of the EB, while the other measures of MO measure processes of the CE of a more general nature.

In addition, this paper also includes a reading comprehension test in order to analyze the specific task of reading texts containing the new measure. Finally, to provide even stronger validity of the WM measures, we have included two higher-level cognitive variables, intelligence and reasoning, given that numerous studies have illustrated the close relationship WM has with these variables (see, for example, Cornoldi, 2006; Oberauer, Schulz, Wilhelm & SüB, 2005 on fluid intelligence; and Kane & Engle, 2002, García-Madruga et al., 2007 on reasoning). Although it is not the main interest of the work, we think it is important to confirm that the tasks of WM maintain the usual relationships with the measurements of the high-level cognitive processes.

Thus, our hypotheses can be divided into two groups. The first two related to the primary objective of the work, analyze the relationship between the Episode Buffer and the Central Executive. And the last two hypotheses will focus on the validity of the tests used, analyzing whether the usual reactions found in the literature between intelligence, reasoning and reading comprehension are maintained, taking special interest in the behavior of the new task. Thus, 4 main hypotheses are specified:

1. The EB does not require the intervention of the CE in certain tasks. The texts included in CONTEX do not involve a significant increase in processing load, confirming the low participation of the CE. Evidence for this will include the fact that the average reading times will be similar across all levels.
2. Differences will exist between the different WM tests. (a) ANALOGY will be more difficult than the RST, as it involves greater processing load on CE resources. (b) Despite having higher demand for processing, CONTEX will be easier than ANALOGY, as it is a second subsequent recall task and occurs in context of reading. Thus, in CONTEX the recovery of words that were not previously remembered in ANALOGY is expected. (c) Although ANALOGY and CONTEX are part of the same test, we expect a smaller score difference between CONTEX and RST as the tasks with less involvement of the CE resources.

3. Correlations between the diverse WM measures will be positive and significant. However, there will exist differences between the score on CONTEX and the other two WM measures, as these assess different components of WM.
4. WM measures will correlate significantly with the higher-level cognitive variables (intelligence, reasoning, and reading comprehension), thus illustrating their predictive power.

Method

Participants

The sample consisted of 60 college freshman psychology students with an average age of 24.1 years ($SD = 8.5$ months). The participants were rewarded for their participation with a bonus towards their qualifying practicum in developmental psychology. A total of 6 participants were removed from the initial sample as for various reasons they did not complete all of the tasks.

Measures

Working memory

- Reading Span Test (RST)

The RST test, designed by Daneman and Carpenter (1980), applied in its Spanish version (Elosúa et al., 1996), consists in reading a series of 12-word long sentences. The subject must remember the last word of each sentence composing the series and in the right order of appearance. The task includes various levels that increase progressively. The initial series requires remembering 2 words from two sentences in correct order, up to a maximum of 5 words from 5 sentences (you can see an example of level 2 in Table 2). Each level is composed of 3 series. Points are awarded for correct answers when recalling all of the words in correct order, and for incomplete answers when remembering the words but in a different order. Omitting or substituting any one key word is considered an error. Minimum score is 2 and the maximum is 5.8. (Further details about this test, as well as the criteria for awarding points, can be found in Gutierrez-Martínez et al., 2005.) All of the WM tests were scored by following this same procedure.

Table 2
Example of a level 2 RST.

Prueba de Amplitud Lectora (PAL)

The participant should read aloud, completely and without hesitation, the words that appear consecutively on the screen. For example:

1	*The commissioner informed the president that terrorists were planning on killing him.*
2	*The historical monuments are numerous and well presented in the guide.*
3	*?*

The subject must remember, before being asked, the last word in each sentence and the order of their presentation. For instance: "him" then "guide".

Note: ☐ *indicates the order in which screens are presented.*

- Test of Contextualized Span (PA-CONTEX)

The newly designed procedure we have named PA-CONTEX involves solving and remembering the word-solutions of a few simple analogies (see Figure 3). After the first recalling of these key words as described above, the participant must then read a few texts containing blanks the participant is to fill out with the words previously stored. To facilitate identifying the results at each stage, we have named the points obtained in the first memory task ANALOGY, and those obtained in the second task CONTEX.

To prepare for this set of tasks, a few pilot studies were run in order to confirm that the texts themselves did not lend "*guessing*" or intuiting the key words. Furthermore, in order to create very simple texts, low difficulty words across the frequency spectrum were always used. Similarly, the average time participants took to read each text was homogenized (around 15 seconds), as well as the approximate time each participant took to fill out each word blank (5 seconds). These 20-second delays with respect to the first memory (in the first phase) may involve more long-term resources than otherwise the phonological loop or the visual-spatial sketchpad, without refreshing, can handle (see Alloway & Ledwon, 2014; Conway & Engle, 1994; Cowan et al., 1997; Cowan 2001). Thus, as Baddeley has proposed, the maintenance of this information must occur in the Episodic Buffer with the help of LTM. The criterion followed to parameterize the extent of the sentences was to equalize the length of sentences found in texts with those used in the Spanish version of the RST (an average of 12 words). The full content of the task is detailed in the Appendix.

Intelligence

As an index of general intelligence (*g* Factor), the Raven's Standard Progressive Matrices (SPM) was used, as taken by participants in university. It is recognized as providing a scale for estimating "g factor" or fluid intelligence.

The task re-quires participants to reason about the relationships that make up an incomplete set of abstract forms (in a 3 x 3 matrix) in order to select the item that correctly complete the set. The test includes two booklets that contain individual sets of matrices of increasing difficulty. The first booklet contains 12 matrices and was used to train participants. The second booklet contains 36 elements, and participants scores were obtained from the number of correctly solved exercises from this booklet (Raven et al., 1996).

Reasoning

The test of reasoning used was that taken from the studies of Gutiérrez-Martínez et al., (2005), by selecting conditional deductive reasoning problems of an either factual (if A then B) or counterfactual (if A had been, then would have been B). Participants were presented with the four classical conditional inference types (Modus Ponens, Affirmation of the Consequent, Denial of the Antecedent, and Modus Tollens) The test consists of two examples of each kind of statement, resulting in a total of 16 statements randomly ordered for each participant: 2 statements for each of 4 inference types (8) and two examples of each (8 x 2 = 16). Final results were obtained from directly scoring the results of each correct answer.

Table 3 shows an example of the content of each of the statements.

Table 3.
Example of statements in the Reasoning task

Conditional Factual Statement
(Example of Modus Ponens)
If Hugo goes to Granada, then Pilar goes to Galicia.
Hugo goes to Granada
What might you then conclude? _____
Conditional Counterfactual Statement
(Example of Modus Tollens)
If Paco had gone to Burgos, Juan would have gone to León.
Juan did not go to León.
What might you then conclude?_____ _____

Reading comprehension

To measure reading comprehension was employed a Spanish reading comprehension test. The subtask of textual reading comprehension from Ramos and Cuetos' standardized test of reading processes was used (the PROLEC-SE) (Cuetos, Rodríguez & Ruano, 2001). Analyses were performed on the points directly obtained from the 10 inferential questions involved in the subtask, as responses regarding literal memories produced no errors.

Procedure

A within-subject design was used in which each participant completed all of the tasks. In the first session paper-and-pencil based tests of intelligence, reasoning and reading comprehension, in this order were conducted in groups. The second session was completed individually, and participants completed the tests of WM (RST and PA-CONTEX: ANALOGY + CONTEX). The WM texts were administered on a computer using the *E-Prime* software (Schneider, Eschman, & Zuccolotto, 2002). The order of administering the two WM tests was alternated.

Data analysis

The reliability of the tests has been estimated in terms of "internal consistency" by means of Cronbach's alpha coefficient.

The validity of the construct of the new task will be analyzed based on the differences in time to read within each level and between levels, and the total scores obtained in

WM tests. In order to test for significance in performance differences, we applied the parametric Student t-test for related samples.

The expected linear relationship between the tests was estimated by calculating the Pearson correlation coefficient between the total scores. This correlation analysis was also performed with respect to the criterion measures (General Intelligence, Reasoning and Reading Comprehension) in order to verify the predictive capacity of the tests studied.

In order to contrast the relationship between WM and the criteria variables, we carried out a confirmatory factorial analysis. Previously, the degree of association between variables was analyzed by "Bartlett's test of sphericity" and a measure of sampling adequacy by the "Kaiser-Meyer-Olkin" (KMO).

Finally, to assess the predictive power of WM measures in relation to our criteria variables, three stepwise multiple regression analyses were performed.

Results

Hypothesis 1 and 2

Reading times of sentences making up WM tasks were recorded in order to test the possible independence between processing (secondary task) and storage (primary task). This was done to confirm whether reading time is affected by increasing the load on storage. Table 4 illustrates the mean reading times for levels 2 and 4 (only a small number of participants reached level 5, thus it was not considered). By using the difference between these levels as criteria, RST and ANALOGY times show a significant increase (RST: $t = -12.26$; ANAOLOGY: $t = -6.67$; $df = 53$ and $p < .00$ in both cases). This data confirms that there is a greater load when the task level is increased (from 2 to 4). However, as expected, the new task (CONTEX) shows no significant reading times differences between levels ($t = -1.90$, $df = 53$, $p > 0.5$), and thus we can say that processing load is similar across both task levels.

Table 4
Mean differences in sentence reading (Student')

$N=54$		Mean	SD	Error Tip. Mean	Difference in measures ($df = 53$)
Par 1	Reading RST N2	4328.45	541.20	73.65	
	Reading RST N4	5011.04	738.31	100.47	-12.26**
Par 2	Reading ANALOGY N2	3172.86	802.49	109.21	
	Reading ANALOGY N4	3667.04	878.96	119.61	-6.67**
Par 3	Reading Sentences N2_CONTEX	5140.16	1348.16	183.46	
	Reading Sentences N4_CONTEX	5480.76	834.78	113.60	-1.90

**$p < .01$. *$p < .05$

On the other hand, while the sentences used in RST and CONTEX have the same length, the time participants spent reading on the RST ($M = 4774.19$ms., $SD = 668.21$ms) is significantly less than spent in CONTEX ($M = 5344.91$ms., $SD = 940.38$ ms), which suggests that reading was done more superficially ($t = -5.38$, $df = 53$, $p < .01$). Other studies have also found that RST allows a superficial reading of sentences freeing up resources in processing to facilitate storage. (Gutiérrez-Martínez et al., 2005; García-Madruga et al., 2007).

We should note that we identified outlier data more than three standard deviations above the mean, as per convention. Seven outliers were identified and each was replaced by the upper limit (the mean plus three standard deviations).

In order to test whether the WM measures indicated the utilization of different system processes, we analyzed the descriptive statistics, the mean (M) and standard deviation (SD) of which can be seen for all measures in Table 5. The data shows that scores were highest for the RST and lowest for ANALOGY, making the latter the most complex task (hypothesis 2a). These results agree with those found in other studies (García-Madruga et al., 2007; Gutiérrez-Martínez et al., 2005). The new measure, CONTEX, occurs in intermediate position and has a score close to that of RST. Reliability measures of CONTEX also reflect a high degree of internal validity for this measure (Cronbach's alpha = .78).

Table 5
Descriptive statistics of measures.

$N = 54$	CONTEX	ANALOGY	RST	RAVEN	REASONING	PROLEC
MEAN	3.38	2.75	3.48	24.33	10	7.74
Std. Dev.	.78	.50	.69	4.77	3.51	1.56

***$p < .001$, ** $p < .01$, *$p < .05$; one tailed

As expected, significant differences between the scores on all measures was found, as reflected in the comparison of means using Student's t-test: between RST and ANALOGY ($t = 9.98$, $df = 53$; $p < .01$), between RST and CONTEX ($t = -2.51$, $df = 53$; $p = .01$), and between two measures of our new procedure, ANALOGY and CONTEX ($t = 7.44$, $df = $

53; $p < .01$). However, we think that the most relevant result to confirm our hypotheses is that the participants recalled in CONTEX words they had not been able to recall in ANALOGY (hypothesis 2b). These results reflect a certain independence between the two measures and confirm that the results of the new procedures designed here in both phase 1 and 2 are indeed different in some way. In fact, CONTEX gets scores closer to RST than to ANALOGY, see Table 5 (hypothesis 2c).

Hypothesis 3 and 4

Inter-correlations between WM measures are highly significant, indicating that despite differences in their scoring, they are founded upon an important common base (see Table 6). We emphasize the level of significant obtained by CONTEX (< .001), especially, with respect to other WM measures.

Correlations between criteria variables reflect the relationships that may exist between them, as we did expect. In particular the test of intelligence correlated very significantly with that of reasoning and textual comprehension, the highest of which was between reasoning and intelligence and, on the other hand, the lowest of which was between reasoning and comprehension.

Also, as expected, the correlations of WM measures with the higher-level cognitive variables were also positive and significant in every case (see Table 6). Among them, data again confirm CONTEX as having the highest correlation across all cases. The intelligence test (RAVEN) is more highly correlated with CONTEX than it is with any other measure, and the only one that is highly significant. Our reasoning measure (REASONING) is more highly correlated with CONTEX, even if very similar to ANALOGY, and is least correlated with RST (in line with other studies, see García-Madruga et al., 2007 and Gutiérrez-Martínez et al., 2005). As the for the reading comprehension test (PROLEC), the data also support our predictions: RST obtains the lowest correlations and CONTEX obtains the highly significant correlational indices.

Table 6
Pearson's Correlations between measures.

$N = 54$	1	2	3	4	5
1. CONTEX					
2. ANALOGY	.63***				
3. RST	.49***	.46***			
4. RAVEN	.39**	.23*	.29*		
5. REASONING	.36**	.35**	.30*	.42**	
6. PROLEC	.36**	.30*	.26*	.38**	.25*

***$p < .001$; **$p < .01$; *$p < .05$; one tailed

In order to deepen the understanding of the relationship between WM and the criteria variables, we also carried out a confirmatory factorial analysis using principal components to evaluate the fit of a two-factor model. However, the sample size means that these results must be taken as merely exploratory. We think that in this model the WM tasks should cohere together, on the one hand; and other the other, our criteria variables, especially intelligence and reasoning, as these are tasks that assess more general characteristics of a higher order. Studies confirm that these are related but independent variables.

Prior to carrying out the factorial analysis, the degree of association between variables was analyzed by Bartlett's test of sphericity, indicating that the Chi Squared index forms the best fit to the data ($X^2 = 77.307$, $df = 15$, $p = .00$). Likewise, a measure of sampling adequacy, the KMO (Kaiser, Meyer and Olkin) reflects a score of .778, over and above the recommended score for factorial analysis, which is .75.

The two-factor generated by the model explain more than 63% of the variance (see Table 7). Using Varimax rotation shows Factor 1 to be basically made up of WM measures and explains 34% of the variance, with CONTEX being the greatest contributor to this factor. Factor 2 explains 29% of the variance, with RAVEN being the measure that makes up the largest contribution to this component. It

is worth noting that CONTEX still participates in Factor 2 with a weight of .34, while the other WM measures are below .15, again indicating that CONTEX has more predictive power relative to the others.

Table 7
Factorial Analysis with extraction through an analysis of Principal Components. Method of rotation: Kaiser's Varimax rotation.

		Kaiser's Varimax rotation *	
	Extraction	Component 1	Component 2
CONTEX	.722	.780	.338
ANALOGY	.753	.855	
RST	.592	.747	
RAVEN	.735		.851
REASONING	.494		.642
PROLEC	.487		.666

Total variance explained (63.05%): Component 1 (33.88%), Component 2 (29.17%)
*Variables with weights less than .30 were omitted.

Finally, to assess the predictive power of WM measures in relation to our criteria variables, three step-wise multiple regression analyses were performed. The main variables were WM measures, and intelligence, reasoning, and textual comprehension were dependent variables. Previously, we calculated an index of multi-collinearity (VIF), which resulted in less than two for all cases (and below the index of 10 cited by Kleinbaum), and illustrated a condition index below 15

(and below the limit of 20 calculated by Belsley to indicate a range of weak co-linearity).

Regression analysis generated a model that explained 16% of the variance ($F(1, 52) = 9.579, p < .01$; Error = 4.43) for the intelligence measure, with CONTEX the only independent variable introduced ($B = 2.36$; $\beta = .394, p < .01$). For reasoning, the model generated explained 13% of the variance ($F(1, 52) = 7.80, p < .01$; Error = 206), again with CONTEX the only participatory independent variable ($B = .099$; $\beta = .361$; $p < .01$). The same goes for textual comprehension, explaining 13% of the variance ($F (1, 52) = 7.478$; $p < .01$; Error = 1.469), with CONTEX the only variable ($B = .691$; $\beta = .355, p < .01$).

Discussion and conclusions

The main objective of this study was to explore the nature of the Episode Buffer as proposed by Baddeley (2000) in order to analyze its operation and its relationship to the Central Executive (CE). Specifically, we wanted to analyze whether the Episodic Buffer serves as an independent passive memory store, from which CE processes oriented towards information integration, coordination and grouping that take place in this space do not interfere. The data presented in the current study, as a whole, seem to support these conclusions.

We first look at our first hypothesis (H1) concerning the low processing load when reading texts in the new task. Reading times for sentences that make up the texts in CONTEX did not significantly increase in difficulty as a function of increasing storage level. For the RST and ANALOGY, on the other hand, a significant increase in difficulty with increasing level did occur. These results seem to indicate that increasing the difficulty in storage-memory interferes with reading sentences (as in the RST) or with the processing of analogies (as in ANALOGY). Yet thanks to the near-automatic semantic processing supported by processes of a more long-term nature, this same kind of interference does not exist for CONTEX. In fact, the results seem to show the opposite effect, concurrent reading processing helps the re-activation of mnesic traces by involving the MLP processes without interfering with the task by not consuming CE resources. In this way, it seems to confirm Baddeley's approach specifying that storage and processing are undertaken independently (H1) under some circumstances.

In this regard, another interesting result is the difference found between reading sentences in RST versus CONTEX, despite the fact that sentences in both are of the same length level of difficulty. That subjects take significantly less time confirms our idea that RST allows for a superficial reading, whereas CONTEX requires a more deeply semantic reading, even though both texts are very simple and assume an equally low processing load. Understanding the texts in CONTEX require specific processes of information integration (in the short and long-term), such as the grouping necessary

for storage and the maintenance of this information. In other words, functions specific to the Episodic Buffer.

Now let's consider the second (H2) and third hypotheses (H3). High inter-correlations were found between the three tests of WM (H3), which confirms the construct validity of the measures. However, the most interesting part consists in analyzing these correlations along with the differences found between the scoring on these tasks and the low levels of co-linearity between them. Together, these results confirm that although the measures recorded have certain shared processes in common, they also have certain differences and specific characteristics. This is particularly important in the case of ANALOGY and CONTEX, given that both measures are undertaken with the same two-step procedure. In this case, the results support the hypothesis that the capacity of WM storage is operating with the participation of different processes, even if we cannot rule out the fact these might be different storage units or systems.

The higher scores of CONTEX compared to ANALOGY supports that idea that this new measure records an execution that is based on relevant information taken from long-term memory that has been pre-activated in some way, whether more efficiently or by taking further action. We should remember that CONTEX involves a larger number of elements to process (even though this processing does not increase system load), and the latency between processing and the memory task is greater, something that should foster a higher dissipation of mnemic traces.

We cannot ignore that the ANALOGY test was performed before that of CONTEX, which facilitates the maintenance and over-activation of remembered items in the task. However, the fact that there were words remembered in the CONTEX test that were not in the ANALOGY test leads us to believe that the task as designed enabled the creation of contextual cues that allow for the recovery of disabled elements (or elements insufficiently activated) in the first phase of the procedure. It is important to remember, in contrast, that the text themselves do not evoke these key words. To our knowledge, this reactivation of unremembered (forgotten) items in the analogies task is key, as it seems to show that the experimental conditions of the new task (which approximate real life more closely) allow participants to use all of their mental resources as they might in their daily life. This is to say, contextualizing the task fosters or allows for the participation of relevant information previously experienced. Additionally, the capacity for information storage and retrieval is displayed regardless of the processes involved in reading (i.e., information integration and grouping). Thus, this provides new data regarding the capacity of the Episodic Buffer and its limits when considering 4 elements and the best participants.

As we have argued, the present study is consistent with other work that reports no interaction between superficial processing and the deeper processes of integration, storage and information recall (Baddeley et al., 2009; Baddeley et al., 2011; Alloway et al., 2004). They thus have proposed that

the integration of words occurs via automatic processes that both implicate Long Term Memory and are independent of the Central Executive (Allen & Baddeley, 2009; Baddeley et al., 2009). Additionally, that maintenance of these stored items in WM requires no important additional attentional support in order to recall (Baddeley et al., 2011; Delvenne, Cleeremans, & Laloyaux, 2010; Fougnie & Marois, 2009; Gajewski & Brockmole, 2006; Johnson et al., 2008).

Ultimately, the improvement in CONTEX shows that the use of knowledge stored in LTM promotes the activation of WM and allows for the use of all available resources to reach a maximum capacity of storing grouped elements, which appears to be around 4. Thus, it is parsimonious to believe that the Episodic Buffer is a passive mental space independently utilized to store these items, and that the recall of the chunks depends on the contextual cues that are used to this effect. In certain aspects, this converges with the models described by Engle and Cowan (Engle, 2001, 2002; Cowan, 1999, 2005; Unsworth & Engle, 2007) that propose a single storage unit that has different levels of activation. However, one cannot rule out the possibility that we are recording to different passive stores, one more short-term (that recorded in RST and ANALOGY) and another with more long-term characteristics (consistent with the Episodic Buffer, as reflected in the results of CONTEX), in line with the classical proposal of expert performance described by Ericsson and Kintch (1995).

Finally, this study also provides new data towards the debate surrounding the relationships the are established between the diversity of WM tasks and the various higher-order cognitive skills such as reasoning, reading comprehension, and fluid intelligence. The data again show performance on WM tasks constitute a good predictor of the level of cognitive ability in our participants. However, what is remarkable is that CONTEX is the task that shows the greatest predictive capacity, as it is highly correlated with these three variables and highly significant and is the only measure that is significant in the regression analysis.

In a sense, our results also support the hypothesis that CONTEX involved the Episodic Buffer. By requiring the participation of LTM, it is in this mental space (in the Episodic Buffer) where the higher cognitive processes are carried out: comprehension processes, solving deductive reasoning problems and the matrices making up fluid intelligence problems. This might explain the why the results obtained on CONTEX given the three criterion variables were so similar, both in their correlations and in the percentage of variance explained in each variable. Besides confirming the construct validity of the task itself, these results also confirm the criterion validity of CONTEX as a test itself.

However, it may be necessary that future studies design other tasks in which general processing and the specific task processes such as integration and chunking are matched, in order to continue evaluating the storage capacity of the Episodic Buffer. It is also necessary to evaluate the difference between carrying out the CONTEX test after that of ANALOGY (as was done in the current study) with that of carrying it out without having previously asked for the recalled word solution to the analogies test. Finally, further studies should seek to increase the sample size and broaden its scope to include school-age children, given that the predictive capacity of CONTEX regarding the three criterion variables are positively related to variables such as academic performance and all that this involves from a practical point of view.

References

Allen, R.J., & Baddeley, A.D. (2009). Working memory and sentence recall. In A. Thorn, & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 63-85). Hove, East Sussex, UK: Psychology Press.

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., Baddeley, A.D., & Hitch, G.J. (2014). Evidence for two attentional components in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(6), 1499–1509.

Alloway, T.P., & Ledwon, F. (2014). Working memory and sentence recall in children. *International Journal of Educational Research, 65,* 1-8.

Alloway, T.P., Gathercole, S.E., Willis, C., & Adams, A.M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal Experimental Child Psychology, 87*(2), 85-106.

Anderson, J.R., Reder, L.M., & Lebière, C (1996). Working memory: activation limitations on retrieval. *Cognitive Psychology, 30,* 221-256.

Baddeley, A.D. (1986). *Working memory.* Oxford, UK: Oxford University Press.

Baddeley, A.D (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423.

Baddeley, A.D. (2007). *Working memory, thought, and action.* Oxford, UK: Oxford University Press.

Baddeley, A.D. (2010). Working memory. *Current Biology, 20,* 136-140.

Baddeley, A.D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63,* 1–29.

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G.A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic.

Baddeley, A.D., & Logie, R.H. (1999). Working memory: The multicomponent model. In A. Miyake y P. Shan (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control.* Cambridge: Cambridge University Press.

Baddeley, A.D., Hitch, G.J., & Allen, R.J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language, 61,* 438–456.

Baddeley, A.D., Allen, R.J., & Hitch, G.J. (2011). Binding in visual working memory: The role of the episodic buffer. *Neuropsychologia, 49,* 1393–1400.

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 570-585.

Brener, R. (1940). An experimental investigation of memory span. *Journal of Experimental Psychology, 26,* 467-482.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*(1), 31–42.

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific

reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences*, *19*(2), 246–251.

Case, R. (1985). *Intellectual development: Birth to adulthood*. New York: Academic Press.

Conway, A.R.A., & Engle, R.W. (1994). Working memory and retrieval: A resource dependent inhibition model. *Journal of Experimental psychology: General, 123*, 354-373.

Cornoldi, C. (2006). The contribution of cognitive psychology to the study of human intelligence. *European Journal of Cognitive Psychology*, *18*(1), 1–17.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge, UK: Cambridge University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-185.

Cowan, N. (2005). *Working memory capacity*. New York: Psychology.

Cowan, N., Wood, N.L., Nugent, L.D., & Treisman, M. (1997). There are two word length effects in verbal short-term memory: Opposed effects of duration and complexity. *Psychological Science, 8,* 290-295.

Cuetos, F., Rodríguez, B., & Ruano, E. (2001). *PROLEC-SE. Procesos Lectores*. Madrid: TEA.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450–466.

Daneman, M., & Merikle, P.M. (1996). Working memory and comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422–433.

Delvenne, J.F., Cleeremans, A., & Laloyaux, C. (2010). Feature bindings are maintained in visual short-term memory without sustained focused attention. *Experimental Psychology, 57*(2), 108-116.

Elosua, M.R., Gutiérrez, F., Garcia-Madruga, J.A., Luque, J.L. & Gárate, M. (1996). Adaptación española del "Rea- ding Span Test" de Daneman y Carpenter. *Psicothema, 2,* 383-395.

Engle, R.W. (2001). What is working memory capacity? En H.L. Roediger, J.S. Naime, I. Neath, & A.M. Supremant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowde* (pp. 297–314). Washington, DC: American Psychological Association.

Engle, R.W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23.

Ericsson, K.A., & Kintsch, W. (1995). Long-term working memory. *Psychological review*, *102*(2), 211.

Fougnie, D., & Marois, R. (2009). Attentive tracking disrupts feature binding in visual working memory. *Visual Cognition, 17*, 48-66.

Gajewski, D.A., & Brockmole, J.R. (2006). Feature bindings endure without attention: Evidence from an explicit recall task. *Psychonomic Bulletin & Review, 13*, 581–587.

García-Madruga, J.A., Elosúa, M.R., Gil, L., Gómez-Veiga, I., Vila, J.O., Orjales, I., Contreras, A., Rodríguez, R., Melero, M.A., & Duque, G. (2013). Reading Comprehension and Working Memory's Executive Processes: An Intervention Study in Primary School Children. *Reading Research Quarterly*, *48*(2), 155-174.

García-Madruga, J.A, Gómez-Veiga, I. & Vila-Chaves, J.O. (2016). Executive functions and the improvement of thinking abilities: The intervention in reading comprehension. *Frontiers in Psychology. 7*, pp. 58 – 73.

García-Madruga, J. A., Gutiérrez, F., Carriedo, N., Luzón, J. M., & Vila, J. O. (2007). Mental models in propositional reasoning and working memory´s central executive. *Thinking and Reasoning, 13* (4), 370-393.

Gutiérrez, F., García-Madruga, J.A., Carriedo, N., Vila, J.O., & Luzón, J.M. (2005). Dos pruebas de Amplitud de Memoria Operativa para el Razonamiento [Two tests of Operating Memory Width for Reasoning]. *Cognitiva, 17* (2), 183-203.

Hannon, B., & Daneman, M. (2001). A new tool for measuring and understanding individual differences in the component process of reading comprehension. *Journal of Educational Psychology*, *93*(1), 103–128.

Hannon, B., & Daneman, M. (2004). Shallow semantic processing of text: An individual-differences account. *Discourse Processes*, *37*(3), 187–204.

Johnson, J.S., Hollingworth, A., & Luck, S.J. (2008). The role of attention in the maintenance of feature bindings in visual short-term memory. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 41–55.

Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension. *Psychological Review, 99*(1), 122-149.

Kane, M.J., & Engle, R.W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, *9*(4), 637–671.

Kane, M.J., Hambrick, D.Z., Tuholski, S.W., Wilhelm, O., Payne, T.W., & Engle, R.W. (2004). The generality of working-memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*(2), 189–217.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Miyake, A., & Shah, P. (1999). Toward unified theories of working memory: Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 442–481). New York: Cambridge University Press.

Oberauer, K., Schulze, R., Wilhelm, O., & Süß, H.-M. (2005). Working memory and intelligence—their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*(1), 61–65.

Raven, J.C., Court, J.H., & Raven, J. (1996). *Raven matrices progresivas. Escalas: Color (CPM), General (SPM), Superior (APM) [Raven progressive matrices]. Manual*. Madrid: TEA Ediciones S.A.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime (Version 2.0). [Computer software and manual]. Pittsburgh, PA Psychology Software Tools Inc.

Turner, M.L., & Engle, R.W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28,*127-154.

Tulving, E., & Patkau, J.E. (1962). Concurrent effects of contextual constraint and word frequency on immediate recall and learning of verbal material. *Canadian Journal of Psychology, 16*, 83-95.

Unsworth, N., & Engle, R.W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*(1), 104–132.

Vukovic, R.K., & Siegel, L.D. (2006). The role of working memory in specific reading comprehension difficulties. In T.P. Alloway & S.E. Gathercole (Eds.), *Working memory and neurodevelopmental disorders* (pp. 89–112). New York: Psychology.

Appendix

A. STRUCTURE OF THE NEW PROCEDURE

The new procedure is administered via computer (using the E-Prime software package). It presents all of the information through successive screens that the experimenter (or by the participant him/herself) controls by the keyboard (see below for an example of the sequence). Each begins by reading some instructions that explain how the two recall tasks are performed. A few practice trials are then presented in order to familiarize the participant with the procedure and to resolve any lingering doubts. After the practice phase ends the experimental phase begins and follows a structure that combines two measures of memory: a first obtained from the PAR—anl test and a second measure, contextualized byt the new PA-CONTEX test.

First memory (PAR-anl)

Participants must read aloud a set of incomplete verbal analogies and select the correct alternative from two presented. The presentation occurs serially, beginning with two analogies and increases in number up to a maximum of 5, with three trials at each level. At the end of each series, participants must remember the word solution they consider correct in each problem, and say these in the correct order. Further details on this test and its scoring criteria can be found by consulting *Gutiérrez-Martínez et al. (2005).*

Second memory –PA-Contex–

After having completed the first memory task for each series, as described above, a text appears with blanks the participant must fill out with the words contained in the solution to the analogies task previously administered. As in the previous test, the text must be read and completed aloud.

B. CHARACTERISTICS OF THE TEXTS USED.

The texts refer to daily and simple stories in order that the words to be remembered were familiar and well known. All of the words that made up the texts had high frequency use in the target language.
The texts begin with an introduction of several sentences (three lines on the screen comprised of 26 to 32 words in total). We found that participants took an average of 15 seconds to read this introduction, such that the memory of the first word never occurred before the first 20 seconds.
Every one of the sentences had a length of 12 words, and the blank that required completion through memory was always located in the last position of the sentence.

This table lists an approximate translation of the Spanish task

	1st Series	2nd Series	3ª Serie
1° Recall PAR-ANL	**Two element Series**		
	A glass it to drink, as a plate is to… - break - eat A suit is to a tailor, as a novel is to… - writer - offer as gift	Doctor is to patient, as teacher is to... - student - conceirge Cover is to sofa, as glove is to... – wool - hand	Happy is to sad, as skinny is to... - overweight - blond Navigate is to ship, as fly is to...- cloud - airplane
2°. Memroy PA-Contex	This morning, after waking, he received a telephone call. An acquaintance had invited him to a party. It was actually his own surprise party that his colleagues had put together. After giving a moving speech, he felt like something to _____. In the end they awarded him with a prize for best _____ of the year.	Some time ago everybody was talking about him. He used to be quite stubborn and ungrateful towards everyone. Little by little his possibilities for improvement were becoming limited. The professionals at the school considered him to be a complicated _____. Lately he acquired the habit of painting his _____ in watercolor.	Jesús worked in an administrative office. His colleagues suggested that he try to resolve his problem. He told them all that he would do so upon his return. He himself recognized the he was rather _____. While at times he had issues, this time he said nothing before boarding the _____.
1° Recall PAR-ANL	**Three element Series**		
	Eyelid is to eye, as blind is to ... - window - door Cellar is to wine, as garden is to ... - flowers - pond Up is to down, as left is to ... - arm - right	Television is to watch, as radio is to ... - listen - broadcaster Maestro es a escuela, como médico es a ... -hospital - operation Elections are to votes, as vintage is to ... - grapes - field	Plants are to sap, as humans are to... - lung - blood Attic is to chalet, as peak is to... - mountain - snowfall Daisy is to flower, as apple is to ... - red - fruit

2°. Memory PA-Contex	When we arrived we saw that it was already crowded. Whatever happened, we wanted to see her, even if it would not be easy. For a long time we were all yelling at once. After much prodding, she finally came to the _____. Although we could hardly see because of the _____. We had no choice but to go running to the _____.	The argument amongst families came from a land dispute. All indications were that it was a difficult case. Julian refused to be present for the resolution of the trial. As much as he opposed it, in the end he had to _____. Upon returning to the village and coming across the others, everybody ended up in the _____. It seems silly that everything started with some _____.	Eduardo went on holiday with friends. After arriving at the cottage, the prepared an excursion for the following day. Upon waking the next day, they prepared everything and started walking. After two hours on the trail, and while they had _____ on their feet, they continued walking towards the _____. Upon returning they observed that they only had one piece of _____ to eat.

Four element Series

1° Recall PAR-ANL	Potatoes are to edible, as trousers are to... - bag - clothes To love is to hate, as filling is to... -emptying -saving Drowsiness is to sleep, as sadness is to... - singing - crying To follow is to reach, as competing is to... - playing - winning	Breakfast is to morning, as having lunch is to... - evening - sandwich Fork is to silverware, as closet is to a... - wood - furniture Brick is to mason, as brush is to ... - painter - stainer To live is to die, as to forget is to... - dream - remember	Learning is to school, as fitness is to... - gym - strength Saturn is to planet, as elephant is to... - large - animal To clean is to scrub, as to pass is to... - jump - study Song is to a CD, as chapter is to... - book - story

2°. Memory PA-Contex	María won a trip to Cuba with her best friend. They had fun salsa dancing all day. Upon returning to Spain, they missed it. Half of what they brought back from their trip were _____. They realized this after _____ their suitcases. Every time they recalled their trip they got emotional and started _____. María had always dreamed of _____ a trip to Cuba and enjoying the week there.	Ana spends all day playing sports. Everybody says she is a great athlete. She is in the habit of running everyday. That very _____ she went to visit her brother Luis. He had built her some _____ in which she could display her medals. Now she would have to remodel her room, Luis convinced her to call a friend of his who was a _____. The problem is that he could not _____ his telephone number.	María is a nurse in the "Severo Ochoa" hospital. She is very happy as she enjoys her work, and being very active there pleases her. Every evening she spends two hours in the_____, because she does not like to be alone. She thought about buying a companion ____, but she still has not decided. Every evening she tries to _____ everything she can about them. Her work colleagues are thinking about buying her a _____ about them.

Five element Series

1° Recall PAR-ANL	In front is to behind, as before is to... - yesterday - after Ice is to cold, as light is to... - heat - sunny Glass es a fragile, as silk is to....- jersey - soft* Shovel is to dig, as needle is to... - sewing - ripping Rome is to city, as the Duero is to... - shore - river	Attentive is to careless, as strong is to... - weights - weak Light is a dark, as poor is to... - rich - hungry Old is to young, as tall is to... - ugly - short Painting is to art, as physics is to... - science - chemistry Roof is to house, as hat is to. – cap - head	Snow is to winter, as beach is to... -parasol - summer Tractor is to farmer, as plane is to... - pilot - air Leaf is to green, as sky is to... - star - blue Hemorrhage is to bleed, as hilarity is to... - please - laugh Wave is to water, as dune is to... - sand - palm

2°. Recuerdo PA-Contex	He is considered the most intelligent of his class. Ricardo is 16 years old and gets very high grades. He wants to study medicine and specialize in cardiology Every day _____ eating he goes to the library. Even on especially _____ days he studies. Even though he may do so in a _____* way, that still rather affects him. Ricardo's mother works as a _____ in order to save up and pay for his studies. The store where she sell the clothes she makes is located beside the _____.	Everything was in place for a good evening of bullfighting. There was a good poster and his fight was at a good time. When it started Manual felt nervous. He finished his performance with a return to the ring, but he felt rather _____. He knew that this afternoon's performance might fulfill his dream of becoming _____. He triumphed even though he was rather_____ for a bullfighter. Despite his dedication to bullfighting, he did not want to give up his pursuit of _____. He liked to train and his _____ was also always full of thoughts.	Summer arrived and they still did not know if they could take August off. She should have known whether or not she'd be going on holiday. In the end they were able to off in September . Because they prepared everything quickly they had forgotten to bring the _____. Owing to their work as a _____ they never had fixed dates to plan for their holiday. After arriving they found that the skies were _____. This allowed her to forget her initial anger they later started to _____ a lot. The children were very happy that they could now play all day in the _____.

* soft translates as "suave" in Spanish, which also can mean to study "lightly" (i.e., softly).