Journal of Cognitive Psychology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/pecp21

The structure of working memory during childhood: a systematic review

Santiago Vernucci , Lorena Canet-Juric , Eliana V. Zamora & María M. Richard's

JOURNAL OF COGNITIVE PSYCHOLOGY
https://doi.org/10.1080/20445911.2021.1887199

Routledge
Taylor & Francis Group

REVIEW

The structure of working memory during childhood: a systematic review

Santiago Vernucci, Lorena Canet-Juric, Eliana V. Zamora and María M. Richard's

Instituto de Psicología Básica, Aplicada y Tecnología (IPSIBAT), Universidad Nacional de Mar del Plata (UNMDP), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Mar del Plata, Argentina

ABSTRACT

There are many working memory (WM) models, generally formulated and developed in adults. Controversies arise as to whether such models are adequate conceptualizations of WM in children. The aim of the present study is to perform a systematic review of studies that evaluate the structure of WM during childhood. Databases (PubMed, Scopus) and article reference lists were reviewed, identifying 264 studies and including 14 in the review. These include participants between 4 and 15 years of age, with typical development, and they evaluated the structure of WM using confirmatory factor analysis. Results show that from 4 to 6 years onwards a structure composed of a domain-general executive component together with two domain-specific storage components (verbal, visuospatial) is identified, generally being the best fitting model. Limitations and potential contributions of the reported results are discussed.

ARTICLE HISTORY
Received 28 August 2020
Accepted 2 February 2021

KEYWORDS
Working memory; children; structure; development

Working memory (WM) allows simultaneously maintaining and processing information for the execution of complex cognitive activities (Baddeley, 2012; Conway et al., 2007). During childhood it has great importance due to its relationship with processes and skills such as fluid intelligence (Dehn, 2017), reading comprehension (Borella & de Ribaupierre, 2014), mathematical skills (Cragg et al., 2017) and for its contribution to self-regulation (Hofmann et al., 2012). Likewise, children's WM capacity affects cognitive development, learning and academic performance; in particular, children with low WM capacity face difficulties in acquiring and consolidating these skills (Alloway et al., 2009).

Throughout childhood, WM undergoes a sustained increase in capacity (Best & Miller, 2010; Canet Juric & Burin, 2016). Both its variation throughout development as well as its role in complex activities have been studied based on many theoretical models. These models differ from each other by conceiving WM as a unitary construct (i.e. composed of a single component) or a system of multiple interacting components. Further, theoretical models differ regarding the number and specificity of the components that comprise WM, the sources of capacity limits, and how best to measure its capacity (see Conway et al., 2007; Miyake & Shah, 1999; Oberauer et al., 2018).

The most widespread model is the multiple-component model, initially formulated by Baddeley and Hitch (1974) and successively developed (Baddeley, 1996, 2000, 2012; Baddeley & Logie, 1999). It proposes that WM is composed of different specialised components, thus dividing attentional control mechanisms and temporary storage of information by its modality. The phonological loop provides temporary domain-specific retention of phonological information, through passive storage of information and active rehearsal. The visuospatial sketchpad provides temporary domain-specific retention of visuospatial information, by means of passive retention as well as reactivation of stored information. A domain-general central executive is in charge of the control and regulation of the system. It is responsible for task switching, focusing and dividing attention, and coordinating the storage components (Baddeley, 1996, 2012; Baddeley et al., 2019; Baddeley & Logie, 1999). Although the central executive was initially supposed to direct the retrieval of information from long-term

CONTACT Santiago Vernucci santiagovernucci@conicet.gov.ar Instituto de Psicología Básica, Aplicada y Tecnología, Facultad de Psicología, Universidad Nacional de Mar del Plata, CONICET, Funes 3250, Cuerpo V, Nivel III, B7602AYJ, Mar del Plata, Buenos Aires, Argentina

memory, a later modification no longer assigned this function, postulating the existence of an episodic buffer in charge of temporarily retaining multidimensional chunks, integrating information from different domains (Baddeley, 2000, 2012).

In addition to this model, there are several influential alternative proposals. Cowan (1999, 2001, 2005, 2014) proposes an embedded-processes model, in which WM is understood as the small amount of information that can be kept in an accessible state for its use in cognitive tasks. To achieve this, part of the contents of long-term memory becomes active, comprising a set of representations of different modalities (e.g. sensory, phonological, visual, spatial, semantic), functioning as a short-term memory store (STM). A portion of this information is within the attentional focus, being accessible to conscious processing. In addition, a central executive is in charge of voluntarily directing the attentional focus towards the relevant stimuli, ignoring distracting information. Engle, Kane et al. (Engle, 2018; Engle & Kane, 2004; Engle et al., 1999) consider that WM has different components with specific properties: short-term storage of information (i.e. STM), procedures and strategies for maintaining activation, and executive attention. Short-term storage consists of highly activated representations of different modalities. The central executive is responsible for controlled attention, which is essential in conditions of interference or conflict. Although WM functioning comprises the interaction of these components, WM capacity refers exclusively to controlled or executive attention; therefore, a greater WM capacity makes it possible to keep a higher amount of information in an active state, through better attentional control.

It has also been proposed that WM is composed of different domain-specific components, identifying two subsystems (i.e. verbal WM, visuospatial WM), each of which is in charge of both storing and processing information from different domains (Friedman & Miyake, 2000; Shah & Miyake, 1996). Furthermore, it has been suggested that the structure that best reflects WM is determined by a single component, without considering distinctions based on the domain of information (i.e. verbal, visuospatial) or function (i.e. storage, processing) (e.g. Colom et al., 2006; Unsworth & Engle, 2007). Although this conception is not considered widespread (Conway et al., 2007; Miyake & Shah, 1999), it has been argued that during development it is possible that a gradual differentiation in the structural organisation of WM takes place, from a single component present at an early age (Alloway & Alloway, 2013).

Structural organisation of WM during childhood

WM models are generally developed and tested in adult populations, which is why their relevance has been discussed when it comes to basing on them the study of the development and functioning of WM in children. One of the main controversies refers to the structure of WM (Alloway et al., 2006; Gray et al., 2017; Injoque-Ricle et al., 2012). Given its role in learning, academic performance and execution of numerous complex cognitive abilities, it becomes crucial to identify properly how WM structural organisation presents itself. That is, what components are part of WM and which specific functions they fulfil (e.g. Gathercole et al., 2004). This would allow the identification of relationships between the different components of WM and complex cognitive abilities, evaluating each component with appropriate tasks based on age and the functional organisation of the WM system, as well as developing intervention studies, knowing which component is being specifically targeted according to the demands of the tasks that are implemented (e.g. Rapport et al., 2013).

This leads to two related issues: the distinction between WM and STM, and the use of appropriate evaluation tasks. It is not uncommon to find confusion between WM and STM. The relationship between these constructs has been understood in different ways: considering them as equivalent, affirming they are components with different functions that are part of the same system, as well as conceiving them as completely different (Aben et al., 2012). An interchangeable use of STM and WM is sometimes found, showing that the difference between these constructs is not completely clear (Aben et al., 2012; Cowan, 2008, 2017). While STM involves the passive short-term storage of a limited amount of postsensory typically categorically encoded information (Cowan, 2017; Gathercole, 1999), WM not only involves the short-term storage of information, but also control mechanisms that enable the simultaneous processing of stored information. Possibly contributing to the confusion is that the main WM models propose at least one component of short-term storage of information, together with other mechanisms that intervene in processing this information (Aben et al., 2012; Cowan, 2008), and WM finds its direct antecedent in STM (Baddeley, 2012; Baddeley et al., 2019).

The second issue refers to the use of appropriate tasks to assess the different components of WM. Although many tasks have been proposed to evaluate WM functioning (Oberauer et al., 2018), span tasks are the most widely used (Conway et al., 2005; Dehn, 2008; Redick & Lindsey, 2013). Briefly, simple span tasks evaluate passive short-term retention of information, thus requiring the operation of domain-specific storage components of WM (Aben et al., 2012; Bayliss et al., 2003; Conway et al., 2005). Complex span tasks involve simultaneous storage and controlled processing of information, therefore requiring short-term domain-specific information storage and domain-general executive resources (Alloway & Alloway, 2013; Bayliss et al., 2003; Conway et al., 2005; Dehn, 2008). Despite this distinction, simple and complex tasks are sometimes used interchangeably by researchers and practitioners (Aben et al., 2012; Dehn, 2008). It should be noted that complex spans measure executive aspects of WM in addition to storage, so when referring to WM assessment, these tasks are generally being considered. It then becomes important to know if throughout development these tasks measure the same components of WM, or if at some point there are changes with respect to the components that are being measured.

An additional problem arises since complex span tasks are not the only ones that have been proposed to evaluate storage and processing in WM. For example, running memory tasks present a series of stimuli in a list of unpredictable length, and the participant must remember the last *n* stimuli presented, or else as many as possible from the end of the list, requiring continuous updating of information (Bunting et al., 2006; Oberauer et al., 2018). *N*-back tasks are also argued to require WM updating, as they involve a fast sequential presentation of stimuli and the participant must decide if the present stimulus is the same as the one presented *n* stimuli back in the sequence. Here, WM updating involves the continuous replacement in memory of previously presented elements by more recent ones (Szmalec et al., 2011). Although these tasks require controlled processing as well as temporary information storage (Conway et al., 2005), they are not necessarily adequate to be used interchangeably with complex span tasks. *N*-back and complex span tasks present correlations that are much lower than what would be expected if they measured the same underlying construct – leading to question over *n*-back tasks' validity as a WM measure (Gajewski et al., 2018; Kane et al., 2007; Miller et al., 2009; Redick & Lindsey, 2013). In short, evaluating the structural organisation of WM during childhood requires to adequately distinguish between WM components and their functions in order to use tasks that are adequate and valid for their measurement.

A statistical method that allows testing of the structural organisation of WM is confirmatory factor analysis (CFA), a type of structural equation model that addresses the relationships between indicators (i.e. observed measures) and factors (i.e. latent variables). One of its main characteristics is that it assigns a main role to the theory proposed by the researcher, since it requires the conceptual model to be specified prior to the evaluation of its fit to the data. To do this, the number of factors, the indicators associated with each one, as well as their relationships must be clearly specified. CFA has been used for the validation of theoretical constructs, as it allows evidence to be obtained regarding the structure of a given construct, as well as the interrelationships between its factors. In this sense, the proposed model is evaluated with respect to its adequacy in reproducing the relationships between the evaluated variables, using different indices to estimate its fit, allowing to compare the fit of different models to the same data set (Brown, 2015; Harrington, 2009; Kline, 2010).

Based on these considerations, different studies have addressed the adequacy of the main theoretical models of WM at different ages during childhood, through CFA (e.g. Alloway et al., 2006; Gathercole et al., 2004; Injoque-Ricle et al., 2012). This analysis allows one to not only to identify if the structure specified by a WM model presents a good fit to the data, but also to compare the fit of two or more models to identify which is the factor structure that best represents WM. However, some questions have not yet been fully addressed. Which is the best fitting WM model in children? Is WM a unitary or fractioned construct? And if it is fractioned, how many components can be identified? How are they organised according to domain and function? Does the structural organisation of WM vary throughout development? Therefore, the debate regarding the structural organisation of WM in childhood is still open (Alloway & Alloway, 2013; Gray et al., 2017;

Injoque-Ricle et al., 2012). Due to the importance of identifying the structural organisation of WM at this stage of the life cycle, added to the fact that there is no systematization of the evidence in this regard in the literature, the present narrative synthesis aims to systematically review the studies that have evaluated the structure of WM in typically developing children, using CFA, both throughout development and at specific ages.

Methods

Design

This review is based on the criteria proposed in the PRISMA-P Declaration, which consists of a checklist designed to facilitate the preparation and reporting of protocols of systematic reviews (Moher et al., 2015; Shamseer et al., 2015).

Eligibility criteria

Studies had to meet the following criteria: (a) include participants aged between 3 and 16 years; (b) participants must be of typical development, or they must not have been selected based on a specific pre-existing disorder or deficit; (c) evaluate the structure of WM through CFA, comparing at least two WM models; (d) be published in peer-reviewed journals; (e) be written in English or Spanish; (f) have been published until 10 February 2020 (last search date).

Information sources and search strategy

The information search was carried out through two sources: (a) electronic databases and (b) reference lists of published articles. The search for articles in electronic databases (PubMed Central and Scopus) was performed using the following search terms: *Working Memory* AND *Children* AND *Factor* AND *Structure*. These terms were searched in title, abstract and keywords of each record.

Study selection and data extraction

The study selection process was performed in accordance with the following steps. First, the search in electronic databases was conducted considering the specified terms. Once the resulting records were retrieved, they were assessed as to whether they met the proposed eligibility criteria. To do this, the first author reviewed the title and abstract of each record. In the event that it could not be determined whether or not it adequately met the proposed criteria, the record was independently reviewed by the other authors. The remaining studies were read in full text, to define the adequacy of their inclusion in the synthesis. To carry out the comparisons between the studies that met the eligibility criteria, the following information was considered: (a) number of participants, (b) age range, and/or age groups (if any), (c) country in which the study was conducted, (d) language of the participants, (e) WM tasks used, (f) WM models that were tested, (g) WM model(s) identified as the better-fitting according to data from the total sample or from each age group considered.

Results

The search in electronic databases resulted in 264 studies (Scopus $n = 204$, PubMed $n = 60$). From those, 58 duplicate records were removed. Then, studies that were identified in reference lists of published articles ($n = 5$) that had not been found in the previous search were included. The available studies ($n = 211$) were evaluated regarding compliance with the eligibility criteria, reading its title and abstract. After excluding 194 studies for not meeting these criteria, a total of 17 studies were identified as to be included in the review. After reading these studies full-text, 3 studies were excluded because they did not meet the criterion of evaluating the structural organisation of WM using CFA. The remaining studies ($n = 14$) met all the eligibility criteria and were finally included in the review (see Figure 1).

To synthesise results, the included studies were grouped into four areas of interest: (a) characteristics of the sample (number of participants, ages, country of origin, language), (b) WM tasks used, (c) WM models that were tested, (d) WM model(s) identified as the better-fitting. Table 1 presents the synthesis of the main characteristics of the reviewed studies.

Sample characteristics

Sample size

The studies present samples with a wide range of participants, from a minimum of 58 (Vugs et al., 2016) to a maximum of 1,699 (Michalczyk et al.,

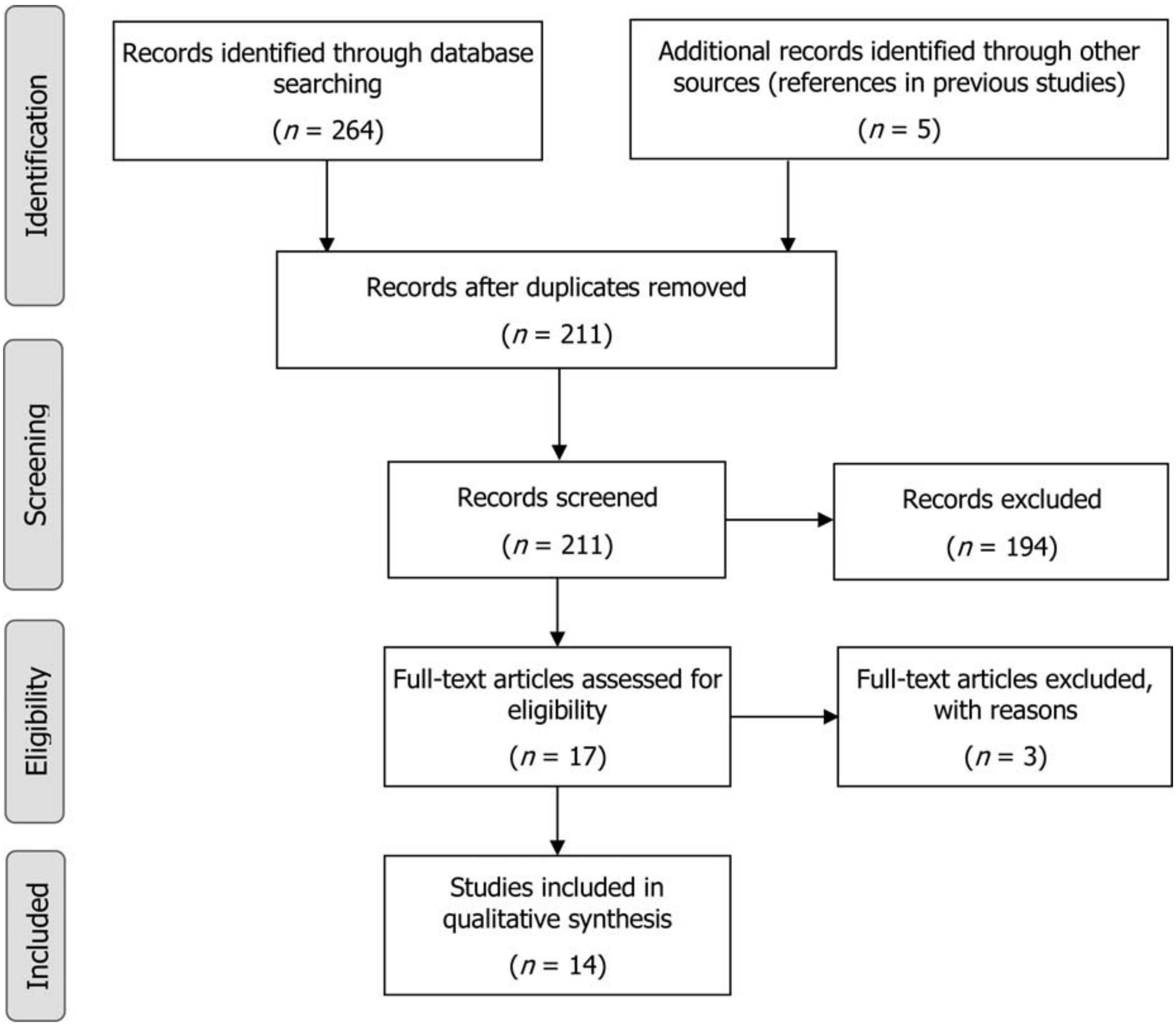

Figure 1. Flow diagram of the study selection. The different phases of the selection process are represented, detailing the number of records initially identified, those excluded in the different stages, as well as the number of articles that met the eligibility criteria and were included in the synthesis (Moher et al., 2009).

2013). Half of the studies ($n = 7$) have a relatively limited number of participants, with sample sizes equal to or less than 180 (Campos et al., 2013; Cockcroft & Milligan, 2019; Giofrè et al., 2013; Gray et al., 2017; Hornung et al., 2011; Injoque-Ricle et al., 2012; Vugs et al., 2016). Regarding those that evaluate the theoretical models of WM across age groups, 3 studies have groups of less than 200 cases each (n range = 60–184; Gathercole et al., 2004; Injoque-Ricle et al., 2012; Swanson et al., 2019) whereas 2 studies present larger groups (n range = 210–695; Alloway et al., 2006; Michalczyk et al., 2013).

Age

The studies consider a wide age range, extending from early childhood to adolescence. Although the ages of the participants overlapped in several studies, none of them fully coincided. The minimum age considered is 4 years (Alloway et al., 2004, 2006; Vugs et al., 2016), while the maximum is 15 years (Gathercole et al., 2004). Half of the studies focus on a relatively narrow age range, comprising 2 years ($n = 3$; Campos et al., 2013; Cockcroft & Milligan, 2019; Vugs et al., 2016) or 3 years ($n = 4$; Alloway et al., 2004; Giofrè et al., 2013; Gray et al., 2017; Hornung et al., 2011), while the remaining studies focus on a wider age range ($n = 7$; Alloway et al., 2006, 2017; Gathercole et al., 2004; Injoque-Ricle et al., 2012; Kuhn, 2016; Michalczyk et al., 2013; Swanson et al., 2019).

Country of origin and language

The studies come from the United Kingdom ($n = 3$), United States ($n = 2$), Germany ($n = 2$), Argentina (n

Table 1. Synthesis of the main characteristics of the reviewed studies analysing the structure of WM in children.

Study	Sample: *N*, age range, country (language)	Tasks (*n*)	WM models tested (components)	Better fitting WM models(s)
Alloway et al. (2004)	*N* = 604, 4–6 yrs, UK (English).	Verbal simple span (3) Verbal complex span (3) Sentence repetition (2)	Unitary Multiple-component (PL, EB, CE) 2-factors (post hoc): grouping simple span and sentence repetition, and complex span	Multiple-component (PL, EB, CE)
Gathercole et al. (2004)	*N* = 575, 6–15 yrs; groups: 6–7 yrs (*n* = 184), 8–9 (*n* = 154), 10–12 (*n* = 132), 13–15 (*n* = 105); UK (English).	Verbal simple span (3) Visuospatial simple span (3) Verbal complex span (3)	2-factors, by domain (verbal, visuospatial) Multiple-component (PL, VS, CE)	In all age groups: multiple-component (PL, VS, CE)
Alloway et al. (2006)	*N* = 708, 4–11 yrs; groups: 4–6 yrs (*n* = 285), 7–8 (*n* = 210), 9–11 (*n* = 213); UK (English).	Verbal simple span (3) Visuospatial simple span (3) Verbal complex span (3) Visuospatial complex span (3)	2-factors, by domain (verbal, visuospatial) 2-factors, by function (WM, STM) Multiple-component (PL, VS, CE) 4-factors (verbal WM, visuospatial WM, verbal STM, visuospatial STM)	In all age groups: multiple-component (PL, VS, CE)
Hornung et al. (2011)	*N* = 161, 5–7 yrs, Luxembourg (Luxembourgish or Portuguese as first language, evaluation in Luxembourgish).	Verbal simple span (2) Visuospatial simple span (2) Verbal complex span (2)	Unitary 2-factors, by domain (verbal, visuospatial) 2-factors, by function (WM, STM) Multiple-component (PL, VS, CE) Embedded-processes (attentional focus, verbal, visuospatial) Executive attention (verbal STM, visuospatial STM, executive)	Multiple-component = Embedded-processes = Executive attention
Injoque-Ricle et al. (2012)	*N* = 180, 6–11 yrs; groups: 6 yrs (*n* = 60), 8 (*n* = 60), 11 (*n* = 60); Argentina (Spanish).	Verbal simple span (3) Visuospatial simple span (3) Verbal complex span (3) Visuospatial complex span (3)	Unitary 2-factors, by domain (verbal, visuospatial) Multiple-component (PL, VS, CE) Multiple-component, 2-factors (PL, VS + CE) 4-factors (verbal WM, visuospatial WM, verbal STM, visuospatial STM)	Total sample: multiple-component (PL, VS, CE) = 4-factors (verbal WM, visuospatial WM, verbal STM, visuospatial STM) 6 yrs: none 8 yrs: multiple-component (PL, VS, CE) = Multiple-component, 2-factors (PL, VS + CE) 11 yrs: 4-factors (verbal WM, visuospatial WM, verbal STM, visuospatial STM)
Campos et al. (2013)	*N* = 103, 8–9 yrs, Portugal (Portuguese).	Verbal simple span (4) Visuospatial simple span (2) Verbal complex span (3)	Unitary 2-factors, by domain (verbal, visuospatial) Multiple-component (PL, VS, CE) Multiple-component, 2-factors (PL, VS + CE)	Multiple-component (PL, VS + CE) > multiple-component (PL, VS, CE)
Giofrè et al. (2013)	*N* = 176, 8–10 yrs, Italy (Italian).	Verbal simple span (2) Visuospatial simple span (2) Verbal complex span (2) Visuospatial complex span (1)	Unitary 2-factors, by domain (verbal, visuospatial) 2-factors, by function (WM, STM) Multiple-component (PL, VS, CE)	Multiple-component (PL, VS, CE)
Michalczyk et al. (2013)	*N* = 1669, 5–12 yrs; groups: 5–6 yrs (*n* = 284), 7–9 (*n* = 690), 10–12 (*n* = 695); Germany (German).	Verbal simple span (4) Visuospatial simple span (2) Verbal complex span (3)	Unitary Multiple-component (PL, VS, CE) Multiple-component,	Multiple-component (PL, VS, CE)

(*Continued*)

Table 1. Continued.

Study	Sample: *N*, age range, country (language)	Tasks (*n*)	WM models tested (components)	Better fitting WM models(s)
		Visuospatial complex span (1) Response inhibition (2)	4-factors (PL, VS, CE, inhibition)	
Kuhn (2016)	*N* = 275, 8–13 yrs, Germany (German).	Verbal free-recall (1) Verbal recollection (1) Visuospatial recollection (1) Visual array comparison task (1) Verbal complex span (2) Visuospatial complex span (1)	Unitary 2- factors, by function (visual array comparison task loading on WM factor) 2-factors, by function (visual array comparison task loading on STM factor) 3-factors (WM, STM, scope of attention) 3-factors (visuospatial WM, verbal WM, STM)	2- factors, by function (visual array comparison task loading on WM factor)
Vugs et al. (2016)	*N* = 58, 4–5 yrs, the Netherlands (Dutch).	Verbal simple span (3) Visuospatial simple span (3) Verbal complex span (3) Visuospatial complex span (3)	2-factors, by domain (verbal, visuospatial) Multiple-component (PL, VS, CE) 4-factors (verbal WM, visuospatial WM, verbal STM, visuospatial STM)	None
Alloway et al. (2017)	*N* = 1237, 5–10 yrs, Argentina (Spanish), Brazil (Portuguese), Canada (English), Italy (Italian), UK (English).	Verbal simple span (3) Visuospatial simple span (3) Verbal complex span (3) Visuospatial complex span (3)	Multiple-component (PL, VS, CE) 4-factors (verbal WM, visuospatial WM, verbal STM, visuospatial STM)	Multiple-component (PL, VS, CE) > 4-factors (verbal WM, visuospatial WM, verbal STM, visuospatial STM)
Gray et al. (2017)	*N* = 168, 7–9 yrs, USA (English).	Auditory N-back (1), Visual N-back (1), Verbal updating (1), Verbal simple span (2), Verbal running memory (1), Visuospatial simple span (2), Visuospatial running memory (2), Phonological binding span (1), Visuospatial binding span (1), Multimodal binding span (1)	Multiple-component (BF, AV, EC) Multiple-component, 4-factors (BF, AV, EC, BE) Embedded-processes (CE, attentional focus, STM)	Embedded-processes > Multiple-component (BF, AV, EC)
Cockcroft and Milligan (2019)	*N* = 92, 6–7 yrs, South Africa (African language as first language, English as second, evaluation in English).	Verbal simple span (3) Visuospatial simple span (3) Verbal complex span (3) Visuospatial complex span (3)	Unitary 2-factors, by domain (verbal, visuospatial) 2-factors, by function (WM, STM) Multiple-component (PL, VS, CE) 4-factors (verbal WM, visuospatial WM, verbal STM, visuospatial STM)	4-factors (verbal WM, visuospatial WM, verbal STM, visuospatial STM)
Swanson et al. (2019)	*N* = 614, 6–10 yrs; groups: 6 yrs (*n* = 133), 7 (*n* = 130), 8 (*n* = 141), 9 (*n* = 94), 10 (*n* = 121); USA (English language learners, Spanish as first language).	Verbal simple span (3) Visuospatial simple span (2) Verbal complex span (2) Verbal updating (running memory) (1) (verbal tasks both in English and Spanish)	Both in English and Spanish: Unitary 2-factors, by function (WM, STM) Multiple-component (PL, VS, CE)	Multiple-component (PL, VS, CE) English: at 6, 8, 9 and 10 yrs. Not specified at 7 yrs. Spanish: at 6, 7, 8 and 10 yrs. Not specified at 9 yrs.

Note: CE = central executive, EB = episodic buffer, PL = phonological loop, STM = short-term memory/storage, VS = visuospatial sketchpad, WM = working memory.

= 1), Italy (*n* = 1), Luxembourg (*n* = 1), the Netherlands (*n* = 1), Portugal (*n* = 1) and South Africa (*n* = 1). One study compared the fit of different WM models cross-nationally, with children from Argentina, Brazil, Canada, Italy and the United Kingdom (Alloway et al., 2017). Nine studies have been carried out exclusively with children from European countries. Regarding the languages, children were English speakers (*n* = 4), German speakers (*n* = 2); had Spanish as their first language and were English learners (*n* = 1); were speakers of an African language and had English as their second language ($n = 1$); had Luxembourgish or Portuguese as first language, being evaluated in Luxembourgish ($n = 1$). The remaining studies included children who spoke Spanish, English, Dutch or Portuguese as their native language (in each case, $n = 1$). Finally, the study by Alloway et al. (2017) includes children that speak either English, Italian, Spanish or Portuguese.

WM tasks used

Verbal simple span tasks are used in 13 of the 14 studies reviewed (with the exception of Kuhn, 2016), while visuospatial simple span tasks are used in 12 studies (except Alloway et al., 2004; Kuhn, 2016). Short-term storage is widely regarded as a component of WM. Regarding complex span tasks, all studies except Gray et al. (2017) use at least one task, and all of these studies include at least one verbal task, while visuospatial complex span tasks are used to a lesser extent ($n = 8$ studies). In addition, some studies use verbal but not visuospatial complex span tasks to evaluate executive aspects of WM (Alloway et al., 2004; Gathercole et al., 2004; Hornung et al., 2011; Campos et al., 2013; Swanson et al., 2019). Likewise, half of the studies ($n = 7$) use verbal and visuospatial span tasks, both simple and complex, which allows for the assessment of short-term storage, as well as simultaneous storage and processing of information from different domains (Alloway et al., 2006, 2017; Cockcroft & Milligan, 2019; Giofrè et al., 2013; Injoque-Ricle et al., 2012; Michalczyk et al., 2013; Vugs et al., 2016). Tasks that require updating of information in WM (e.g. running memory tasks, updating, *n*-back) have been scarcely used in studies assessing the structure of WM in children. Swanson et al. (2019) used an updating task in which the children had to remember the last three digits presented of sequences of different length, while Gray et al. (2017) used a number updating task, in which the children had to carry out two running totals simultaneously. In addition, this study used verbal, visual and spatial running memory tasks, and fundamentally, it is the only one reporting the use of *n*-back tasks (auditory and visual).

WM models tested

The unitary model is tested in 9 studies, while 2-factor models, which group the tasks by domain (i.e. verbal, visuospatial) or by function (i.e. WM, STM) are also frequently considered ($n = 8$ and $n = 6$, respectively). The most tested is the multiple-component model (e.g. Baddeley, 2000, 2012; Baddeley & Hitch, 1974), given that all but one study ($n = 13$) report having tested the fit of its theoretical structure. It should be noted that there are some differences between the studies in how they specify such model. The most widespread version, which proposes the existence of a domain-general central executive and two domain-specific storage components, is the most frequently evaluated ($n = 12$). However, a later formulation of the model that includes executive and storage components along with an episodic buffer has also been tested, although to a lesser extent ($n = 2$). In these cases, Alloway et al. (2004) included the episodic buffer but did not include a visuospatial storage component, while Gray et al. (2017) is the only study that evaluates the fit of the model as specified by Baddeley (2000).

Also, by assuming that the storage of visuospatial information is highly demanding in terms of attentional control resources, some authors pose that the relationship between the visuospatial sketchpad and the central executive should be strong during childhood. Therefore, two studies (Campos et al., 2013; Injoque-Ricle et al., 2012) have proposed a further version of the multiple-component model, by specifying its structure with a verbal storage component and another component that groups visuospatial storage (i.e. visuospatial sketchpad) together with attentional control (i.e. central executive).

Finally, a model that assumes the existence of four components that are distinguished by domain and function (i.e. verbal WM, visuospatial WM, verbal STM, visuospatial STM) is tested in a third of the studies ($n = 5$; Alloway et al., 2006, 2017; Cockcroft & Milligan, 2019; Injoque-Ricle et al., 2012; Vugs et al., 2016). Other influential WM model such as Cowan's (2005) embedded-processes, has been scarcely evaluated ($n = 3$; Gray et al., 2017; Hornung et al., 2011; Kuhn, 2016). Also one study specifically tests the fit of a version of Engle's WM model consisting of three factors (Hornung et al., 2011).

WM models evidencing the better fit in each study

In each of the 14 studies, the better-fitting WM model was identified based on the fit indices reported by the authors. Since studies differ in the age range of the participants, as well as with respect to whether they analyse changes in the adequacy of WM models as a function of age, results are presented distinguishing those studies that do not analyse changes in development ($n = 9$) and those that do analyse such changes ($n = 5$).

Studies that do not analyse developmental changes in the structure of WM

Studies involve children aged 4–5 (Vugs et al., 2016), 4–6 (Alloway et al., 2004), 5–7 (Hornung et al., 2011), 6–7 (Cockcroft & Milligan, 2019), 7–9 (Gray et al., 2017), 8–9 (Campos et al., 2013), 8–10 (Giofrè et al., 2013) and 8–13 years (Kuhn, 2016). Alloway et al. (2017) present a sample with a broader age range of children aged 5–10 years. Taken together, these nine studies analyse the structural organisation of WM between 4 and 13 years of age.

The results show that the multiple-component model composed of a central executive and two domain-specific storage components presents the best fit in 3 studies (Alloway et al., 2017; Giofrè et al., 2013; Hornung et al., 2011). In particular, Alloway et al. (2017) reported that both the 3-factor multiple-component model as well as a 4-factor model distinguishing components by domain and function presented an adequate fit to the data. Although the latter resulted in a slightly better fit, the authors indicate that due to the strong relationship (.90) between verbal WM and visuospatial WM factors it is possible to assume the presence of multicollinearity, noting that "it may be a more parsimonious choice to rely on a three-factor model" (Alloway et al., 2017, p. 347). Regarding the multiple-component model including the episodic buffer, Alloway et al. (2004) indicate that it is the better-fitting model among those tested. In contrast, Gray et al. (2017) show that the 4-factor multiple-component model (as in Baddeley, 2000) fits the data poorly. Additionally, despite the fact that the 3-factor multiple-component model presents a good fit, the structure specified on the basis of Cowan's model was the one that showed the better fit. For their part, Hornung et al. (2011) find that their specification of Cowan's model, as well as that of Engle and the 3-factor multiple-component model show an adequate fit and it is not possible to clearly identify a superior model.

Campos et al. (2013) found that the 3-factor model presents an adequate fit to the data. However, given the strong correlation (.91) between the central executive and the visuospatial sketchpad factors they proposed a 2-factor structure (i.e. phonological loop, visuospatial sketchpad + central executive), which fitted the data well and could be considered as more parsimonious than the 3-factor model (Campos et al., 2013, p. 9).

Two studies show an inadequate fit of the 3-factor multiple-component model. Vugs et al. (2016) compared the fit of a series of models both in children with specific language impairment and in typically-developing children. Regarding the latter subsample, none of the models show adequate fit indices (see Vugs et al., 2016, p. 14). However, they indicate that a 4-factor model (specifying components by domain and function) provides the best fit in the total sample, and after testing the factor loadings of this model for invariance across both groups they found no significant differences. Cockcroft and Milligan (2019) report a 4-factor structure distinguishing components by domain and function as the one that best represented the organisation of WM in children aged 6–7 years.

Kuhn (2016) report a 2-factor model distinguishing components by function (i.e. domain-general executive and short-term storage) shows the best fit, being the only one supporting this structure, and also the only study not evaluating the multiple-component model. Finally, no study reported either the unitary model or 2-factors model distinguishing components by domain (i.e. verbal WM, visuospatial WM) as the one better representing the structure of WM.

Studies that analyse developmental changes in the structure of WM

These five studies investigate the existence of variations in the structural organisation of WM throughout childhood, by comparing the fit of theoretical WM models in different age groups. These studies include participants aged 4–11 (Alloway et al., 2006), 5–12 (Michalczyk et al., 2013), and 6–10 (Swanson et al., 2019), 6–11 (Injoque-Ricle et al., 2012) and 6–15 years old (Gathercole et al., 2004).

Results show that in most of the studies ($n = 4$), the 3-factor multiple-component model presented the best fit to the data compared to the rest of the tested models, in each of the age groups that were considered. That is, from 4–6 years studies report that a structure composed of a domain-general executive component and two domain-specific storage components is in place (Alloway et al., 2006; Michalczyk et al., 2013; Swanson et al., 2019). It should be noted that the study by Swanson et al. (2019), with a sample of English language learners whose first language was Spanish, shows that when assessed with tasks in their first language the 3-factor model presents the best fit at 6, 7, 8 and 10 years, while when assessed in English the model shows the best fit at 6, 8, 9 and 10 years. In both cases, this model provided the best fit in 4 of the 5 age groups considered, while in the ages in which the fit was not satisfactory, no alternative model could be identified.

The study by Injoque-Ricle et al. (2012) found a different pattern of results. At 6 years, none of the models that were tested evidence an adequate fit. At age 8, the 3-factor model is the better-fitting, however, the authors state that the strong correlation (.93) between the executive component and the visuospatial storage component suggests a 2-factor model (i.e. verbal storage, and visuospatial storage + central executive) may result in a more parsimonious structure. In fact, this formulation evidenced slightly better fit indices, although not significantly different from those of the 3-factor model. At age 11, a 4-factor model, with components divided by domain and function, is the one with the best fit.

Finally, in the same sense as in the previous section, none of the considered studies reported that either the unitary model, 2-factors distinguishing components by domain (i.e. verbal WM, visuospatial WM), or 2-factors distinguishing components by function (i.e. WM, STM) was the one that better represented the structure of WM in any age group.

Discussion

The present study aimed to systematically review the studies that have evaluated the structural organisation of WM in typically developing children using CFA, both throughout development and at specific ages. To do this, 14 studies that met the proposed eligibility criteria were identified. Taken together, the evidence systematized in this review shows a structure composed of different specialised components as the one best representing WM during childhood. This structure is generally specified as comprising domain-specific storage components according to the information domain (verbal, visuospatial) and a domain-general executive component, which is consistent with the most widespread formulation of the multiple-component model of WM (Baddeley, 2012; Baddeley & Hitch, 1974; Baddeley & Logie, 1999). Importantly, these components seem to be in place as early as 4–6 years of age, and the structure they integrate is comparatively the one that best represents WM both when relatively narrow age ranges are considered, as well as at different ages throughout development. However, some important aspects need to be discussed to understand this in greater depth.

With respect to the WM models that were tested in the reviewed studies, not surprisingly the multiple-component model (Baddeley, 2012; Baddeley & Hitch, 1974), is the one that has been most frequently tested in children from 4 to 15 years of age, since it is considered the most influential model in WM research (e.g. Conway et al., 2007; Cowan, 2017; Logie, 2011). All but one study (Alloway et al., 2004) tested the classic 3-factor version of this model (i.e. Baddeley & Logie, 1999). However, some variations to this proposal have been tested, such as the inclusion of a component that represents the episodic buffer (Alloway et al., 2004; Gray et al., 2017), or the specification of a 2-factor structure taking into consideration the involvement of attentional control resources in the execution of visuospatial simple span tasks in children (Campos et al., 2013; Injoque-Ricle et al., 2012). In addition to the multiple-component model, the unitary model (e.g. Colom et al., 2006) is frequently evaluated, whereas other influential WM models such as Cowan's (2005) embedded-processes model or Engle's executive attention model (Engle et al., 1999) have been less evaluated.

Studies that do not analyse developmental changes in the structure of WM indicate that the multiple-component model shows the best fit by 4 years of age (Alloway et al., 2004), and that between 5 and 10 years of age the 3-component structure is in place (Alloway et al., 2017; Campos et al., 2013; Giofrè et al., 2013; Hornung et al., 2011). However, Campos et al. (2013) suggest that the central executive and the visuospatial sketchpad may not be sufficiently differentiated at 8–9 years of age. Gray et al. (2017) report that Cowan's model (i.e. central executive, attentional focus based on visuospatial tasks, and storage and rehearsal of phonological information) showed a slightly better fit in children aged 7–9 years than the multiple-component model. It should be noted that both these models propose verbal and visuospatial domain-specific storage components, as well as a domain-general attentional control component (see also Hornung et al., 2011).

Vugs et al. (2016) and Cockcroft and Milligan (2019) findings seem more contrasting, indicating that the multiple-component model showed an inadequate fit in children aged 4–5 and 6–7 years, respectively. Since none of the models tested by Vugs et al. showed an adequate fit (see p. 14), it may lead one to think that the structure of WM is not clear at that age and it will be defined throughout development. In this sense, Cockcroft and Milligan (2019, pp. 16–17) suggest that the fully fractionated 4-factor structure they found may be subject to change due to development, and that the multiple-component model may represent mature WM in children.

When considering studies that analyse developmental changes in the structure of WM, evidence favours more clearly the 3-factor multiple-component model. This structure is in place between 4 and 6 years of age, and is the better fitting across childhood (Alloway et al., 2006; Gathercole et al., 2004; Michalczyk et al., 2013; Swanson et al., 2019). Injoque-Ricle et al. (2012) indicate that this 3-factor structure is adequate when considering the total sample of children from 6 to 11 years of age, but the better fitting structure differs at 6, 8 and 11 years of age. They indicate that their results suggest a progressive differentiation of the WM system with age, and a further specialisation of processing resources according to the modality of information. These findings do not match those from other studies that obtained evidence of a domain-general executive component of WM at 11 years (Alloway et al., 2006; Michalczyk et al., 2013). Therefore, more evidence is needed to clarify this issue.

Importantly, with the exception of Kuhn (2016) who reports a two-component structure (i.e. central executive and short-term storage) no study shows that either the unitary model, or models that propose structures solely distinguishing two components by domain (i.e. verbal WM, visuospatial WM), or two components by function (i.e. WM, STM) evidenced an adequate fit. Thus, results support the claim that as early as 4–6 years of age and throughout childhood evidence is not in favour of a unitary model, or models that propose structures distinguishing components by domain (i.e. verbal WM, visuospatial WM) or by function (i.e. domain-general WM, domain-general STM). In line with what Miyake and Shah (1999, pp. 448–449) stated more than 20 years ago, WM must not be conceived of as a completely unitary construct, and a comprehensive theory of WM should account for domain-specific effects. These effects, at least from the results of the reviewed studies, seem limited to short-term storage of information, while processing of information in WM demanding controlled attention is better characterised as domain-general.

Results should be assessed taking into account the tasks used to evaluate WM, as these may not only help defining the models that are being tested, but also the structure of WM that is considered as the one better representing the data. Not surprisingly, simple and complex span tasks are the most used measures. Firstly, results show that throughout development the interchangeable use of simple and complex span tasks to assess WM is not adequate (Aben et al., 2012), as the tasks are tapping different components of WM. Simple span tasks were found to load consistently on either verbal or visuospatial domain-specific short-term storage components. Complex span tasks, requiring information storage and processing, load mainly on a single domain-general component that can be thought of as representing the central executive or controlled attention (Baddeley, 2012; Engle et al., 1999). Thus, researchers and practitioners wanting to assess WM functioning must carefully choose and interpret results from span tasks, since from an early age and throughout childhood simple and complex span tasks are measuring specific components.

Secondly, as Schmiedek et al. (2014) indicate, there would be a lot of profit if researchers were attempting to test their theories not only on their preferred paradigms but in the entire domain of tasks that define a construct. Hence, it may be important to consider how structures proposed by different WM models could be specified, not only using span tasks. This would allow one to evaluate whether tasks such as *n*-back, that have face validity as measures of executive aspects of WM but whose construct validity has been questioned (Gajewski et al., 2018; Kane et al., 2007; Miller et al., 2009; Redick & Lindsey, 2013), load on which specific components of WM. Do tasks that require updating load on the same factors as simple or complex span tasks? Evidence obtained from the reviewed studies is scarce. Swanson et al. (2019) use a single verbal updating task that loads on a factor representing the central executive, while Gray et al.

(2017) report that visual and auditory *n*-back tasks, and a verbal updating task load on Cowan's model central executive (which is specified identically in the 3-factors multiple-component model; see Gray et al., 2017, p. 193). However, complex span tasks were not used in this study so it is not possible to know if these load on the same factor. It would be of interest to study throughout development if span tasks, and updating and *n*-back tasks load on the same factor, or if these load on different factors at different ages.

Thirdly, a better understanding of the structure of WM during childhood needs to consider how tasks selected by authors help define such structure. The widespread use of span tasks allows testing relatively well the fit of a series of WM structures: single factor, 2-factors distinguishing components by domain (i.e. verbal WM, visuospatial WM) or by function (i.e. WM, STM), 3-factors multiple-component model (i.e. domain-general central executive, and two domain-specific storage components), as well as a 4-factor fully fractioned structure. However, when testing structures corresponding to other models, some differences are observed. For example, some authors assessed Cowan's embedded processes model using: span tasks (Hornung et al., 2011); complex span tasks, free-recall, recollection and visual array comparison tasks (Kuhn, 2016); *n*-back, updating, simple span, running memory and binding span tasks (Gray et al., 2017). Based on the previous considerations regarding WM tasks, it is difficult to know if the resulting structures, even those with a good fit, are comparable. In addition, the exclusive use of simple and complex span tasks could limit future studies from advancing and deepening the understanding of the structure of WM in childhood, not only testing the fit of these models, but of others that have been evaluated to a lesser extent. Thus, future studies could deepen the understanding of the structure of WM throughout development, testing models with measures that allow assessing different aspects of WM.

Another important matter that requires discussion in studies using CFA is sample size, since it affects the statistical power and precision of the model's parameter estimates. As sample size increases, the loadings have smaller standard errors; therefore, results become more precise estimates of population factor loadings and are also more stable across repeated sampling (Brown, 2015; Kline, 2010; MacCallum et al., 1999). However, there is no general agreement on how large a sample must be. A number of recommendations regarding sample size in factor analysis have been proposed, such as the minimum necessary sample size, or the ratio of the minimum sample size to the number of variables (MacCallum et al., 1999). Besides this, Kline (2010) indicates that factor analysis is a large-sample technique, which is required to achieve sufficient precision. What is specifically considered large enough may vary with respect to the complexity of the models being tested. Typical sample size should be around 200 cases, but this may still be small if a model of great complexity is being tested. Also, with less than 100 cases, most structural equation models would not be feasible. Taking this into account, the studies that were reviewed tend to present relatively small sample sizes. Half of the studies have less than 200 participants. Studies that include larger samples are generally those that evaluate changes in the structural organisation of WM throughout development, comparing the fit of the models in different age groups – although only 2 studies have $n > 200$ in each group. Therefore, given most age groups have a relatively small number of participants, findings should be interpreted with caution. In order to advance in the understanding of the structure of WM throughout development, and to generalise the results to the population of interest (in this case, typically developing children), future studies would benefit from having samples of sufficient size to allow a more precise estimate of WM components and their organisation during childhood.

With respect to the age of the children that were part of the reviewed studies, a greater number of studies that investigate ages that are less represented, such as preschool children, is needed. Moreover, evidence of developmental changes in the structure of WM is less compared to studies that consider specific ages. This suggests that more studies addressing the fit of different structures throughout childhood are needed. This is an important issue, because it may provide valuable evidence to better understand if the structure of WM changes from one point in development to another, since some studies suggest this (e.g. Injoque-Ricle et al., 2012) while others indicate that the components and their relationships are stable (e.g. Michalczyk et al., 2013).

Regarding the country of origin and the language in which the children were evaluated, the results show that the majority of the studies were carried out with children from European countries, while other regions such as South America or Africa have been scarcely represented. The languages in which the evaluations were carried out are quite diverse, English being the most frequent. Considering that various contextual variables can influence cognitive functioning (Last et al., 2018; Lipina & Segretin, 2015) it becomes crucial to evaluate the structure of WM in different contexts, especially those that are less represented in studies available in the literature. This would allow one to investigate in more detail if the structural organisation of WM is common throughout development in different contexts, or if factors such as language or sociocultural context have an impact on the way WM components develop and relate to each other. In this sense, the study by Swanson et al. (2019) is of interest, reporting that the 3-factor multiple-component model is the one that best represents WM structure regardless of whether it is tested in different languages, such as English or Spanish. Likewise, the study by Alloway et al. (2017) is relevant insofar it includes children from South America (Argentina, Brazil), North America (Canada) and Europe (Italy, United Kingdom), who are speakers of English, Italian, Spanish and Portuguese. The authors conclude that the 3-factor multiple-component model of WM does not appear to be exclusive to samples from mainly Western or English-speaking countries (Alloway et al., 2017, p. 348). Although it may be difficult to carry out a study that has data obtained in children from different countries, it is valuable to provide evidence regarding whether the results of a study with children of a specific context and age, testing theoretical models that are specified in the same way, are maintained in samples from diverse contexts, especially from less represented regions.

The results of this review should be evaluated taking into account some existing limitations. Many studies have a small sample size (or age groups with a small *n*) and this suggests a cautious interpretation of the findings. Likewise, studies are usually from Western developed countries, and the evidence from other contexts is comparatively less, hindering generalizability of the results. Furthermore, there are few studies that test the theoretical structure of WM in preschool-age children, considering the possibility of changes due to development. Thus, it is desirable that future studies address these limitations, obtaining samples of a sufficient size, with children from different contexts and taking into consideration ages that are less represented. In addition, only articles written in English or Spanish were included in the review. This constitutes a limitation since studies that were published in other languages may be missing. Also only two electronic databases were searched, and it could be the case that a particular study has not been considered to be reviewed.

Despite these limitations the present review is expected to contribute to a better understanding of the structural organisation of WM during childhood. Although there are issues that need to be deepened in future studies, some answers emerge from this review to those questions that had not been fully addressed regarding he structure of WM in children. Overall, evidence suggests that a 3-factor structure, with domain-specific verbal and visuospatial storage components and a domain-general executive component, consistent with the most widespread formulation of the multiple-component model of WM (Baddeley, 2012; Baddeley & Hitch, 1974; Baddeley & Logie, 1999) is the one better representing WM. These components are in place from preschool age throughout childhood.

Finally, the results from the present review are not expected to be taken as a definitive conclusion, but encourage researchers to carry out a greater number of studies addressing some of the unresolved issues that were identified. This will surely contribute to a better understanding of the organisation and functioning of WM during childhood.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina.

Data availability statement

The data supporting the results of this review are available from the corresponding author upon reasonable request.

ORCID

Santiago Vernucci http://orcid.org/0000-0003-1595-3106

Lorena Canet-Juric http://orcid.org/0000-0003-4147-4889
Eliana V. Zamora http://orcid.org/0000-0002-6278-6665
María M. Richard's http://orcid.org/0000-0002-7394-5967

References

References marked with an * indicate studies included in the systematic review.

Aben, B., Stapert, S., & Blokland, A. (2012). About the distinction between working memory and short-term memory. *Frontiers in Psychology*, *3*, 301. https://doi.org/10.3389/fpsyg.2012.00301

Alloway, T. P., & Alloway, R. G. (2013). Working memory across the lifespan: A cross-sectional approach. *Journal of Cognitive Psychology*, *25*(1), 84–93. https://doi.org/10.1080/20445911.2012.748027

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, *80*(2), 606–621. https://doi.org/10.1111/j.1467-8624.2009.01282.x

*Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development*, *77*(6), 1698–1716. https://doi.org/10.1111/j.1467-8624.2006.00968.x

*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 of Experimental Child Psychology*, *87*(2), 85–106. https://doi.org/10.1016/j.jecp.2003.10.002

*Alloway, T. P., Moulder, R., Horton, J. C., Leedy, A., Archibald, L. M., Burin, D., Injoque-Ricle, I., Passolunghi, M. C., & Dos Santos, F. H. (2017). Is it a small world after all? Investigating the theoretical structure of working memory cross-nationally. *Journal of Cognition and Culture*, *17*(3-4), 331–353. https://doi.org/10.1163/15685373-12340010

Baddeley, A. D. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 5–28. https://doi.org/10.1080/713755608

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

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

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1.

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2019). From short-term store to multicomponent working memory: The role of the modal model. *Memory and Cognition*, *47*(4), 575–588. https://doi.org/10.3758/s13421-018-0878-5

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). Cambridge University Press.

Bayliss, D. M., Jarrold, C., Gunn, D. M., & Baddeley, A. D. (2003). The complexities of complex span: Explaining individual differences in working memory in children and adults. *Journal of Experimental Psychology: General*, *132*(1), 71–92. https://doi.org/10.1037/0096-3445.132.1.71

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development*, *81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Borella, E., & de Ribaupierre, A. (2014). The role of working memory, inhibition, and processing speed in text comprehension in children. *Learning and Individual Differences*, *34*, 86–92. https://doi.org/10.1016/j.lindif.2014.05.001

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Press.

Bunting, M., Cowan, N., & Saults, J. S. (2006). How does running memory span work? *Quarterly Journal of Experimental Psychology*, *59*(10), 1691–1700. https://doi.org/10.1080/17470210600848402

*Campos, I. S., Almeida, L. S., Ferreira, A. I., & Martinez, L. F. (2013). Working memory as separable subsystems: A study with Portuguese primary school children. *Spanish Journal of Psychology*, *16*, 1–10. https://doi.org/10.1017/sjp.2013.6

Canet Juric, L., & Burin, D. I. (2016). La memoria de trabajo: escritorio, pizarra, energía mental. In I. Introzzi & L. Canet Juric (Eds.), *¿Quién dirige la batuta? Funciones Ejecutivas: herramientas para la regulación de la mente, la emoción y la acción* (pp. 25–35). EUDEM.

*Cockcroft, K., & Milligan, R. (2019). Working memory structure in atypical development: HIV-infected and HIV-exposed, uninfected school beginners. *Developmental Neuropsychology*, *44*(2), 248–272. https://doi.org/10.1080/87565641.2018.1564309

Colom, R., Rebollo, I., Abad, F. J., & Shih, P. C. (2006). Complex span tasks, simple span tasks, and cognitive abilities: A reanalysis of key studies. *Memory & Cognition*, *34*(1), 158–171. https://doi.org/10.3758/BF03193395

Conway, A. R. A., Jarrold, C., Kane, M. J., Miyake, A., & Towse, J. N. (Eds.). (2007). *Variation in working memory*. Oxford University Press.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, *12*(5), 769–786. https://doi.org/10.3758/BF03196772

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 University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity.

Behavioral and Brain Sciences, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Cowan, N. (2005). *Working memory capacity*. Psychology Press. https://doi.org/10.4324/9780203342398

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? In W. Sossin, J.-C. Lacaille, V. F. Castellucci, & S. Belleville (Eds.), *Progress in brain research: The essence of memory* (Vol. 169, pp. 323–338). Elsevier. https://doi.org/10.1016/S0079-6123(07)00020-9.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*(2), 197–223. https://doi.org/10.1007/s10648-013-9246-y

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin and Review*, *24*(4), 1158–1170. https://doi.org/10.3758/s13423-016-1191-6

Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition*, *162*, 12–26. https://doi.org/10.1016/j.cognition.2017.01.014

Dehn, M. J. (2008). *Working memory and academic learning. Assessment and intervention*. John Wiley & Sons.

Dehn, M. J. (2017). How working memory enables fluid reasoning. *Applied Neuropsychology: Child*, *6*(3), 245–247. https://doi.org/10.1080/21622965.2017.1317490

Engle, R. W. (2018). Working memory and executive attention: A revisit. *Perspectives on Psychological Science*, *13*(2), 190–193. https://doi.org/10.1177/1745691617720478

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation*, (Vol. 44, pp. 145–199). Elsevier. https://doi.org/10.1016/S0079-7421(03)44005-X.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). Cambridge University Press.

Friedman, N. P., & Miyake, A. (2000). Differential roles for visuospatial and verbal working memory in situation model construction. *Journal of Experimental Psychology: General*, *129*(1), 61–83. https://doi.org/10.1037/0096-3445.129.1.61

Gajewski, P. D., Hanisch, E., Falkenstein, M., Thönes, S., & Wascher, E. (2018). What does the n-back task measure as we get older? Relations between working-memory measures and other cognitive functions across the lifespan. *Frontiers in Psychology*, *9*, 2208. https://doi.org/10.3389/fpsyg.2018.02208

Gathercole, S. E. (1999). Cognitive approaches to the development of short-term memory. *Trends in Cognitive Sciences*, *3*(11), 410–419. https://doi.org/10.1016/S1364-6613(99)01388-1

*Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

*Giofrè, D., Mammarella, I. C., & Cornoldi, C. (2013). The structure of working memory and how it relates to intelligence in children. *Intelligence*, *41*(5), 396–406. https://doi.org/10.1016/j.intell.2013.06.006

*Gray, S., Green, S., Alt, M., Hogan, T., Kuo, T., Brinkley, S., & Cowan, N. (2017). The structure of working memory in young children and its relation to intelligence. *Journal of Memory and Language*, *92*, 183–201. https://doi.org/10.1016/j.jml.2016.06.004

Harrington, D. (2009). *Confirmatory factor analysis*. Oxford University Press.

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, *16*(3), 174–180. https://doi.org/10.1016/j.tics.2012.01.006

*Hornung, C., Brunner, M., Reuter, R. A. P., & Martin, R. (2011). Children's working memory: Its structure and relationship to fluid intelligence. *Intelligence*, *39*(4), 210–221. https://doi.org/10.1016/j.intell.2011.03.002

*Injoque-Ricle, I., Barreyro, J. P., & Burin, D. I. (2012). Estructura de la memoria operativa: comparando diferentes modelos en la infancia. *Escritos de Psicología / Psychological Writings*, *5*(2), 27–38. https://doi.org/10.5231/psy.writ.2012.1904

Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the n-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 615–622. https://doi.org/10.1037/0278-7393.33.3.615

Kline, R. B. (2010). *Principles and practice of structural equation modelling* (3rd ed.). Guilford Press.

*Kuhn, J. T. (2016). Controlled attention and storage: An investigation of the relationship between working memory, short-term memory, scope of attention, and intelligence in children. *Learning and Individual Differences*, *52*, 167–177. https://doi.org/10.1016/j.lindif.2015.04.009

Last, B. S., Lawson, G. M., Breiner, K., Steinberg, L., & Farah, M. J. (2018). Childhood socioeconomic status and executive function in childhood and beyond. *PloS One*, *13*(8), e0202964. https://doi.org/10.1371/journal.pone.0202964

Lipina, S. J., & Segretin, M. S. (2015). 6000 días más: Evidencia neurocientífica acerca del impacto de la pobreza infantil. *Psicología Educativa*, *21*(2), 107–116. https://doi.org/10.1016/j.pse.2015.08.003

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current Directions in Psychological Science*, *20*(4), 240–245. https://doi.org/10.1177/0963721411415340

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4*(1), 84–99. https://doi.org/10.1037/1082-989X.4.1.84

*Michalczyk, K., Malstädt, N., Worgt, M., Könen, T., & Hasselhorn, M. (2013). Age differences and measurement invariance of working memory in 5- to 12-year-old children. *European Journal of Psychological Assessment*, *29*(3), 220–229. https://doi.org/10.1027/1015-5759/a000149

Miller, K. M., Price, C. C., Okun, M. S., Montijo, H., & Bowers, D. (2009). Is the *n*-back task a valid neuropsychological measure for assessing working memory? *Archives of Clinical Neuropsychology*, *24*(7), 711–717. https://doi.org/10.1093/arclin/acp063

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge University Press.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, *4*(1), 1–9. https://doi.org/10.1186/2046-4053-4-1

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A. R. A., Cowan, N., Donkin, C., Farrell, S., Hitch, G. J., Hurlstone, M. J., Ma, W. J., Morey, C. C., Nee, D. E., Schweppe, J., Vergauwe, E., & Ward, G. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*(9), 885–958. https://doi.org/10.1037/bul0000153

Rapport, M. D., Orban, S. A., Kofler, M. J., & Friedman, L. M. (2013). Do programs designed to train working memory, other executive functions, and attention benefit children with ADHD? A meta-analytic review of cognitive, academic, and behavioral outcomes. *Clinical Psychology Review*, *33*(8), 1237–1252. https://doi.org/10.1016/j.cpr.2013.08.005

Redick, T. S., & Lindsey, D. R. B. (2013). Complex span and n-back measures of working memory: A meta-analysis. *Psychonomic Bulletin & Review*, *20*(6), 1102–1113. https://doi.org/10.3758/s13423-013-0453-9

Schmiedek, F., Lövdén, M., & Lindenberger, U. (2014). A task is a task is a task: Putting complex span, n-back, and other working memory indicators in psychometric context. *Frontiers in Psychology*, *5*, 1475. https://doi.org/10.3389/fpsyg.2014.01475

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*(1), 4–27. https://doi.org/10.1037/0096-3445.125.1.4

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *BMJ*, *349*(jan02 1), g7647. https://doi.org/10.1136/bmj.g7647

*Swanson, H. L., Kudo, M. F., & Van Horn, M. L. (2019). Does the structure of working memory in EL children vary across age and two language systems? *Memory*, *27*(2), 174–191. https://doi.org/10.1080/09658211.2018.1496264

Szmalec, A., Verbruggen, F., Vandierendonck, A., & Kemps, E. (2011). Control of interference during working memory updating. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(1), 137–151. https://doi.org/10.1037/a0020365

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. https://doi.org/10.1037/0033-295X.114.1.104

*Vugs, B., Knoors, H., Cuperus, J., Hendriks, M., & Verhoeven, L. (2016). Interactions between working memory and language in young children with specific language impairment (SLI). *Child Neuropsychology*, *22*(8), 955–978. https://doi.org/10.1080/09297049.2015.1058348