Evolution of the Chunking Concept in Working Memory: Theoretical Models

Authors: Chen Yutian, Chen Rui, Li Peng, Li Peng

Date: 2022-06-22T00:00:00+00:00

Abstract

Following Miller' s proposal of "the magical number 7$±$2" , "chunk" has been adopted by numerous theoretical frameworks as a unit possessing stable structure in the processing of individual working memory for the measurement of memory capacity. However, as research on chunking has deepened, definitions of chunking have likewise evolved. Concurrently, substantial research has demonstrated that individuals' developmental stages correspond to the hierarchical levels of chunking they predominantly employ; however, it remains unclear whether transitions between chunking levels are confined to fixed age intervals, and explanations of the underlying mechanism remain contested. Accordingly, this article provides a comprehensive discussion addressing three facets: the development and evolution of chunking definitions, developmental stage characteristics, and its mechanism. Future research should further investigate the role of long-term memory in the operational mechanisms of working memory chunking, elucidate the chunking characteristics of individuals across different developmental stages, and how to exploit the advantages of rehearsal strategies and the "less is more" principle in the chunking process.

Following Miller' s proposal of "the magical number 7$±$2" , "chunk" has been adopted by numerous theoretical frameworks as a unit possessing stable structure in the processing of individual working memory for the measurement of memory capacity. However, as research on chunking has deepened, definitions of chunking have likewise evolved. Concurrently, substantial research has demonstrated that individuals' developmental stages correspond to the hierarchical levels of chunking they predominantly employ; however, it remains unclear whether transitions between chunking levels are confined to fixed age intervals, and explanations of the underlying mechanism remain contested. Accordingly, this article provides a comprehensive discussion addressing three facets: the development and evolution of chunking definitions, developmental stage characteristics, and its mechanism. Future research should further investigate the role of long-term memory in the operational mechanisms of working memory

chunking, elucidate the chunking characteristics of individuals across different developmental stages, and how to exploit the advantages of rehearsal strategies and the "less is more" principle in the chunking process.

Full Text

The Evolution of the "Chunking" Concept and Theoretical Models in Working Memory

CHEN Yutian, CHEN Rui, LI Peng
(Faculty of Education, Yunnan Normal University, Kunming 650500, China)

Abstract: Following Miller' s proposal of "the magical number 7$±$2," the "chunk" has been adopted by many theories as a stable structural unit in individual working memory processing that can be used to measure memory capacity. However, as researchers have deepened their investigation of chunking, their definitions of chunks have continued to evolve. Meanwhile, numerous studies have found that an individual' s developmental stage corresponds to their primary chunking hierarchy, yet it remains unclear whether transitions between chunking hierarchies occur within fixed age intervals, and explanations of the chunking mechanism remain contested. This paper provides a comprehensive discussion of the development and evolution of chunking definitions, age-stage characteristics, and underlying mechanisms. Future research should further explore the role of long-term memory in the operation of working memory chunking, refine the chunking characteristics of individuals at different ages, and examine how to leverage the advantages of rehearsal strategies and the "less is more" principle in the chunking process.

Keywords: chunking, working memory, model structure, capacity

1 Introduction

In daily life, when you need to memorize a large amount of information within a short period (e.g., a phone number), what strategy would you employ? Chunking information represents an effective strategic choice. Research demonstrates that during interpersonal interactions, people can chunk groups of people based on social relationships to increase working memory capacity, thereby enhancing their ability to identify social groups (Stahl & Feigenson, 2014). In foreign language learning, chunking can improve the efficiency of vocabulary and grammar acquisition (Alluhaybi & Witzel, 2020). Therefore, investigating the chunking mechanism contributes to exploring better methods for storing and retrieving information. Some studies suggest that due to working memory capacity limitations, individuals tend to reorganize recurring or interrelated information through recoding into chunked information to expand storage capacity (Chekaf et al., 2016). In other words, this process involves chunking information. Early research on chunking originated from short-term memory studies, which posited that the primary function of chunking was information storage (Miller, 1956).

However, other research argues that compared to short-term memory, chunking in working memory involves not only storage but also temporary processing functions (Cowan & Albarracín, 2019). Consequently, research on working memory inevitably involves the process of information chunking. Studies using chess players as participants have found that the difference between experts and novices lies in experts' ability to remember more chess piece configurations, as they can utilize richer experience to form more piece combinations (Sala & Gobet, 2017). This indicates that chunked storage of information can improve individual processing efficiency. Furthermore, in cognitive psychology research, the degree of chunking has become one of the standards for measuring individual working memory capacity when processing complex tasks (Sala et al., 2017).

2 The Definition and Evolution of Chunking

James (1891) discovered in memory capacity research that linking temporally associated items could improve individual memory capacity and efficiency, a finding that can be considered a description of the prototype of chunking. Since then, chunking has entered researchers' field of vision. However, some domestic studies have used the term "chunking" with conceptual confusion, applying it to both the integration process and the result of information integration. A review of the literature reveals that "chunk" emphasizes the result of information integration, while "chunking" emphasizes the process of dividing scattered information according to certain characteristics (Gobet et al., 2001; Peng, 2019). Both share the common goal of enhancing memory capacity, and this paper focuses on the concept of chunking. The definition of chunking has continuously evolved, likely because different researchers hold varying understandings of the individual information chunking process. A summary of previous research indicates that individuals primarily chunk information through two approaches: strategic reorganization and identification of common features.

The strategic reorganization approach originates from Miller' s (1956) proposal that chunking is a process whereby individuals, when facing complex memory tasks, integrate fragmented information based on semantic connections to improve memory capacity. When semantic information is continuously input, individuals organize information through structured accumulation, with each accumulation forming a chunk that contains more information than the previous one. Miller termed this process recoding. Building on Miller' s concept, Baddeley (2003) proposed that working memory, as an independent memory system, contains an episodic buffer that serves as a temporary storage system, binding information into episodes. Chunking, in this view, involves the activation of old information while simultaneously transforming original information under working memory operations. Additional research has also demonstrated that recoding of associated information across different memory systems significantly influences chunking (Atkinson & Shiffrin, 1968; Malmberg et al., 2019). These definitions emphasize that individuals activate and employ reorganization strategies when chunking to-be-remembered information.

The common features identification approach posits that while information reorganization strategies can continuously increase chunk capacity, they simultaneously interfere with calculating pure working memory capacity. Cowan (2001) argued that chunking should be a process of associating common features across different pieces of information. Subsequently, Gobet et al. (2001) further emphasized that the amount of shared features between pieces of information is key to chunking. As the chunking concept has evolved, its hierarchical structure has also changed, corresponding to different age stages.

3 Hierarchical Structure and Age-Related Developmental Characteristics

Factors influencing information chunking include the degree of association between information (including external and internal connections), semantic knowledge, and past experience. The common function of these factors in information chunking is information compression, differing in the degree and manner of compression. When individuals chunk information based on the degree of association, the primary method involves identifying patterns, which does not alter the original appearance of information (e.g., when memorizing the letter sequence "a a a a," individuals compress it into "4 a' s" as one chunk). Another example: when memorizing the four word chunks "ship, car, banana, apple," individuals categorize them into two chunks based on intrinsic attributes: "ship, car" and "banana, apple." When individuals chunk information based on semantic knowledge and past experience, they recode and compress information, changing its original form (e.g., when presented with the four digits "1 9 4 9," individuals recode them as "the founding year of the PRC" as one chunk) (Chekaf et al., 2016). These influencing factors also result in chunks having different hierarchical structures.

Previous research has described various chunk types. Some researchers suggest chunks can be divided into three types: (1) chunks formed by grouping items with completely identical external forms and properties; (2) chunks formed by grouping items with inconsistent external forms but internal associations; and (3) chunks formed by reintegrating existing information using past experience (Mclean et al., 1967). Other researchers propose that chunk types include: (1) chunks formed by single independent items; (2) chunks formed by combining single items; and (3) super chunks formed by interlocking multiple chunks (Rosenberg & Feigenson, 2013). Research suggests that the first two types of chunks in Mclean' s classification have no qualitative difference in formation and should be classified as two different chunk types at the same hierarchical level. Therefore, based on previous classifications and chunk complexity, chunks can be divided into three hierarchical levels: (1) chunks formed by single, unrelated independent items; (2) chunks formed by classifying items based on external or internal feature connections; and (3) super chunks formed by recoding items based on semantic knowledge and past experience. A review of previous literature reveals that chunk hierarchical structure is closely related to individual age and cognitive development. Huang et al. (2017) proposed that individuals have a fundamental cognitive system for representing external objects, and the development of this system influences the development of individual chunking functions.

Research indicates that individuals in infancy before 14 months primarily exhibit first-level chunking characteristics. Infants at this stage can only identify single objects and cannot chunk items based on external or internal connections. Using a manual search task paradigm, Kingo and Krøjgaard (2012) first showed 10- to 12-month-old infants one ball, hid it in a box for the infant to search for, then simultaneously showed the infant four identical balls and again had them search. Results from both searches showed that infants could only retrieve one ball each time, with similar search durations (Wang & Feigenson, 2019), demonstrating that children cannot chunk and remember four identical balls. These findings are supported by research evidence from infants aged 6-7 months (Kibbe & Leslie, 2019), 10-12 months (Applin & Kibbe, 2021), and 12-14 months (Stahl & Feigenson, 2018). The reason infants cannot chunk multiple balls may lie in their limited understanding of singular/plural concepts (Peng et al., 2018).

After infants reach 14 months, with the rapid development of language abilities, they gradually form chunks with second-level characteristics. However, due to large individual differences among infants at this stage, existing research has not reached consensus on when infants can form stable second-level chunks, though the age range is generally identified as approximately 14 to 22 months. Previous research presents two viewpoints. One perspective suggests that the transition from first-level to second-level chunking occurs around 14 months. Research has found that starting from 14 months, infants become aware of similarities between objects, but their ability to chunk objects remains unstable and requires adult assistance (Stahl & Feigenson, 2018). The alternative viewpoint places the transition period around 22 months, supported by two lines of evidence: (1) From a brain mechanism development perspective, the lateral prefrontal cortex and ventral temporal regions, which are crucial for object identification and working memory storage, only mature relatively after 22 months (Guida et al., 2020). Subsequently, infants shift from identifying features of single objects to identifying common features across multiple objects, establishing the physiological foundation for second-level chunking ability. (2) From a language development perspective, infants around 22 months can grasp singular/plural concepts, enabling conceptual differentiation between single and multiple objects (Peng et al., 2018). After 22 months, infants begin second-level chunking, independently extracting common features to form chunks. However, infants at this stage lack knowledge and experience, making it difficult to access shared attributes and relying primarily on perception for chunking memory (Forsberg et al., 2020; Smalle et al., 2015).

Upon reaching adulthood, with rapid development of semantic knowledge and past experience, individuals begin to master third-level chunking. Feigenson et al. (2013) found that compared to infants and young children, adults more easily form super chunks based on semantics and past experience. For example, marathon runners can recode the four digits "5 1 2 4" as "5 hours, 12 minutes, 4 seconds" for memorization. In adult chunking research, Ding et al. (2017) discovered in studies of interactive actions, individual actions, and non-interactive actions that during memory processing of interconnected, interactive actions, participants retrieved semantic knowledge from long-term memory to chunk actions. In letter memory research, similar to digit encoding, adult participants could also recode different letters for memory (Stahl & Feigenson, 2018). For instance, in chessboard layout memory research, chess players reintegrated and recoded piece positions into super chunks based on positional relationships, assigned linguistic labels, stored them in long-term memory, and facilitated rapid retrieval (Kibbe & Feigenson, 2014). Summarizing these studies reveals that with proficient mastery of linguistic concepts, adults gradually grasp common abstract attributes of objects for chunking. Current research predominantly focuses on infancy and adulthood, with limited studies on childhood and adolescence. Consequently, no evidence indicates when the transition from second-level to third-level chunking occurs, representing a direction for future research.

4.1 Chunking Theory

Building upon Miller' s theory, Chase and Simon (1973) more systematically proposed chunking theory and, based on the "Elementary Perceiver and Memorizer" (EPAM) computer program, hypothesized that learning development and short-term memory storage occur through interconnections among nodes in a discrimination network. EPAM is a hierarchical discrimination network model that simulates the interconnections of different external stimuli, composed from top to bottom of multiple discriminating nodes and connecting lines. Individuals can classify different types of stimuli through the network to form new chunks for storage (Gobet, 2005). Subsequently, Simon and Barenfeld (1969) further linked the PERCEIVER program (used to collect chess players' eye movement data) with EPAM to propose the "Memory-aided Pattern Perceiver—MAPP" model. This model comprises two components: a learning component and a performance component. The learning component uses EPAM' s discrimination network learning mechanism to simulate long-term memory storage patterns for different chess layouts. The performance component contains three operational steps: (1) detecting significant chess layouts; (2) identifying chess layouts stored in long-term memory from the learning component and retrieving them into short-term memory for temporary storage; and (3) recoding chess layouts stored in short-term memory and overwriting old layouts for position learning and chunk size expansion (Pförtner & Hristova, 2021) (see MAPP flowchart in Figure 1 [Figure 1: see original paper]). For example, when memorizing the six-letter chunk "B B C F B I," the learning component uses past experience stored in long-term memory to encode the six letters as two letter chunks: "BBC" and "FBI." When individuals need to correctly identify the positions of these six letters among randomly placed letters, the performance component first detects the layout of all letters, then identifies and retrieves the correctly positioned

"BBC" and "FBI" chunks stored in long-term memory into short-term memory, and finally retrieves the two temporarily stored chunks to exclude extraneous letters and correctly select the six letters.

Although chunking theory proposes a complete system for extracting and encoding chunk information, it has three major defects: (1) It fails to distinguish between short-term memory and working memory. From a functional perspective, short-term memory primarily serves a storage function, while information encoding occurs in working memory. Although chunking theory was proposed based on short-term memory, the short-term memory encoding subsystem in its performance component operates more similarly to working memory mechanisms. Therefore, this theoretical model should be viewed as a primary working memory chunking operation mechanism. (2) Chase and Simon (1973) proposed that creating a new chunk requires 8 seconds, while adding a new chunk to existing data blocks requires 2 seconds. This slow chunk creation speed cannot explain experts' rapid storage and immediate recall of large amounts of information. (3) It cannot explain how experts can perform tasks using complex new information when working memory capacity is limited (Gobet & Simon, 1998). Consequently, Gobet proposed template theory based on Simon' s work.

4.2 Template Theory

Gobet proposed template theory based on the CHREST (Chunk Hierarchy and Retrieval Structures) computational model composed of short-term memory. The CHREST model includes three sequentially arranged subsystems. Using Go board layout processing as an example: first, individuals input board layout information through the visual perception system; second, the input information is preserved in an intermediate device similar to visual short-term memory storage; and third, frequently encountered piece information in short-term memory is further recoded into chunks stored in long-term memory. These chunks serve as fixed templates that can be retrieved into short-term memory when needed later (Bennett et al., 2020). Compared to the MAPP model, template theory adds two new mechanisms: (1) New parallel nodes—original nodes encode sensory information, while new nodes encode action information, which can connect to form new structures to some extent. (2) Retrieval structures—used to automatically acquire rapidly encoded chunks. Frequently encountered chunks develop into templates with slots, allowing new information to be automatically added to templates. Slots function to judge and test whether new information fits existing templates, improving chunking efficiency while enriching existing templates (Gobet et al., 2001) (see Figure 2 [Figure 2: see original paper]). For example, individuals are first asked to memorize the four-digit chunk "2 0 1 8." Visual information temporarily preserves the four digits in short-term memory, which are then recoded as the character chunk "year 2018" stored in long-term memory. Subsequently, if asked to memorize the three-digit chunk "8 3 1," individuals will connect these four digits as new nodes with the original "year 2018" character chunk to form a new connection, which is then recoded as the larger character chunk "August 31, 2018." This character chunk can be continuously rehearsed to become a template in long-term memory and can be retrieved holistically when recalling the seven digits "2 0 1 8 8 3 1."

Template theory' s slot structure explains individuals' rapid chunking and immediate recall of large amounts of complex information within short timeframes, with its operation process similar to working memory information processing. Researchers believe template theory provides a more comprehensive description of working memory chunking mechanisms than chunking theory. However, this model has two shortcomings: (1) It still does not treat working memory as an independent chunking processing system. (2) It lacks an attention system. Due to limitations on individual attentional resources, multiple chunks cannot be processed in parallel, and connections between working memory information and long-term memory slots also require attentional involvement (Guida et al., 2012). It was not until Cowan' s "embedded-processes model of working memory" that working memory was regarded as a relatively independent chunking processing system.

4.3 Embedded-Processes Model of Working Memory

Cowan (1998) integrated long-term memory and working memory, proposing that working memory is not an independent memory system but consists of a capacity-limited focus of attention, an attentional system, and temporarily activated regions in long-term memory. Its operation proceeds as follows: information first activates brief sensory memory, which then further activates the focus in long-term memory (Cowan & Albarracín, 2019). The focus is controlled by different environments and individual attention scope size (Cowan, 2016), with numerous sensory functions, especially semantic functions, capable of activating the focus. Subsequently, individuals can automatically access information contained in the focus, while other information not in the focus but closely connected to it can be indirectly accessed by expanding the attentional scope (Cowan et al., 2018). The attentional scope is flexible and can be adjusted by individuals to adapt to different task requirements (Cowan et al., 2020). The attentional scope can simultaneously accommodate up to four simple chunks; when encountering more complex chunks, adding relevant information can assist in processing them (Verhaeghen et al., 2015). Although there is no limit to the amount of focus activation in long-term memory, only attended foci can be easily activated (Cowan et al., 2021). Activated foci are recoded and integrated with new information to form super chunks that are retained as part of long-term memory, a process that both improves chunking efficiency and increases chunk memory capacity (Cowan, 2010).

This model systematically elaborates working memory chunking processing mechanisms, manifested in the integration and extraction of attentional foci in long-term memory. Although Hitch et al. (2019) did not oppose the view that working memory is part of long-term memory, they argued that working memory' s unique role cannot be denied and that the two systems should be distinguished while maintaining attention to their interconnections. Some studies have found that new information can still be chunked and retrieved without forming a focus, indicating that the information is not stored in long-term memory. The embedded-processes model cannot explain where such information is stored, suggesting the existence of a temporary storage system separate from long-term memory (Baddeley et al., 2018). Meanwhile, Norris and Albarracín (2017) argued that important evidence that long-term memory cannot replace working memory comes from patients with severe working memory impairment, who can recall familiar, well-learned words but struggle to recall newly learned words. This proves that long-term memory and working memory are two independent systems, emphasizing the importance of working memory' s own functions.

All three models acknowledge long-term memory' s influence on chunking, but template theory' s slots enable it to process more information fragments faster than chunking theory, while the embedded model' s attentional focus limits the effect of rehearsal on chunk quantity. Therefore, the following discussion addresses the different chunk capacities proposed by the three theories.

4.4 Chunk Capacity

The key reason for the lack of a unified standard for chunk capacity in working memory is whether strategies can be used in calculating chunk capacity. Neither chunking theory nor template theory opposes strategy use during chunking. Chunking theory proposes that learners continuously use rehearsal strategies to construct various types of chunks into a large database stored in long-term memory (Gobet & Simon, 1998). Experts store at least approximately 50,000 data chunks in long-term memory. Once learning and recall occur, data chunks are rapidly accessed and integrated with new information through short-term memory (including working memory) to form larger chunks (Chase & Simon, 1973). Therefore, although similarly limited by short-term memory capacity, experts possess far greater chunk capacity than novices (Pluss et al., 2019). Regarding short-term memory (including working memory) capacity size, Chase and Simon (1973) proposed that during initial recall, participants preserve larger, coarser chunks, and in a second phase, reconstruct chunk structures based on information connections. Consequently, chess players can recall approximately seven chunks. Similarly, template formation in template theory also results from individual rehearsal strategies; more experienced individuals form more stable templates, allowing new information to combine more quickly with template slots and be recoded into larger chunks. Although template theory suggests short-term memory (including working memory) template capacity is approximately three chunks (Smith et al., 2021), it also proposes that rapid encoding only occurs during slot filling; otherwise, its information encoding speed and chunk quantity are no different from chunking theory (Gobet & Clarkson, 2004). In contrast, other researchers argue that strategies should not be used when calculating working memory chunk capacity.

Cowan (2021) proposed that due to attentional resource limitations, chunk capacity in long-term memory foci cannot exceed four chunks. For example, the digits 12345678 can be quickly recoded into four large chunks: 12, 34, 56, 78. Miller' s measurement method can integrate objectively unrelated information into chunks through rehearsal strategies when short-term memory capacity reaches its limit, failing to truly reflect pure working memory chunk quantity. This leads to more experienced individuals showing larger working memory capacity, violating the premise of fixed capacity limitations (Cowan, 2020). Therefore, Miller' s calculation of short-term memory (including working memory) chunk capacity is not recognized (Bhandari & Badre, 2018). Mathy and Feldman (2012) also proposed that seven chunks represent composite capacity under rehearsal strategies, while four chunks represent true memory capacity after compression without any mnemonic aids. Therefore, it is necessary to exclude strategy effects on chunk formation in experiments to correctly identify working memory capacity limits and obtain pure chunk quantity (Bayliss et al., 2003). Consequently, Cowan proposed four methods to accurately measure working memory chunk capacity limits: (1) When individuals split chunks into single items during memory tasks, this proves capacity limits have been reached. (2) Employ alternative methods to prevent chunk recoding. (3) When efficiency decreases due to capacity limitations during chunking, this proves maximum capacity limits have been reached. (4) When working memory capacity limits produce other indirect effects on storage and processing efficiency, this proves the upper limit has been reached (Cowan, 2016).

Calculating working memory capacity limits has become an unavoidable topic in chunking research. Some researchers have also proposed that overloaded chunk capacity slows information processing flexibility and accuracy, preventing individuals from simultaneously managing multiple important tasks (Perfors, 2012). Therefore, the question of how to maintain a reasonable number of chunks in working memory also requires attention in future research.

5.1 The Relationship Between Long-Term Memory and Chunking Mechanisms

Current controversies in chunking mechanism research primarily concern the relationship between long-term memory and chunking mechanisms. Researchers propose the following two directions: (1) The role of long-term memory in working memory chunking mechanisms cannot be ignored. First, Miller only treated chunks as simple information-containing modules, briefly describing the encoding, storage, and retrieval processes of chunks in short-term memory without explicitly proposing a chunking mechanism theory or mentioning long-term memory' s role in short-term memory (including working memory) chunking processes. However, Chase and Simon (1973), in their chunking theory based on Miller' s research, proposed that long-term memory can provide necessary chunking pattern information for short-term memory (including working memory), offering past experience to assist chunking mechanism operation. Second, some researchers argue that chunk formation results from connecting past experience with current information, with chunks more likely stored in long-term memory and working memory serving as the operating site for chunking mechanisms (Gobet et al., 2001). Therefore, long-term memory cannot be excluded from working memory chunking operation mechanisms. (2) The degree of influence long-term memory exerts on chunk encoding and storage processes in working memory requires continued investigation.

First, Baddeley et al. (2018) argue that long-term memory only provides past experiential knowledge for working memory chunking processes, while chunk encoding and storage occur only in working memory. The episodic buffer, as an interface between the central executive system and long-term memory, is where information chunking and storage occur, operating not under attentional resource influence but purely under episodic buffer capacity constraints. According to fMRI testing, working memory' s three subsystems—the phonological loop, central executive system, and visuospatial sketchpad—have independent brain regions, and working memory chunk encoding processes can operate in these three systems without long-term memory influence (Hitch et al., 2019). Similarly, some researchers believe visual and auditory information are constructed and integrated in working memory, with long-term memory only providing experiential knowledge in later chunk construction stages (Darolia & Varshney, 2016). However, Cowan and Albarracín (2019) argue that long-term memory participates in chunk encoding processes in working memory, providing important attentional resources to assist chunk integration. Moreover, after chunk construction, chunks are stored in long-term memory and only temporarily stored in working memory when needed. Additionally, Ericsson et al.' s (1995) concept of long-term working memory also agrees that chunk encoding and storage occur in long-term memory. Experts' superior skill primarily results from numerous mature chunks stored in long-term memory; new information is encoded and integrated with existing chunks in long-term memory to form larger chunks, enabling experts to skillfully use external cues to retrieve chunks into working memory for current tasks. Other researchers have demonstrated through empirical studies that when information in working memory is activated, associated information in long-term memory is simultaneously activated, and individual information encoding cannot occur without long-term memory assistance (Liu & Guo, 2013).

5.3 Perfecting Age-Stage Chunking Characteristics

Current research on chunking characteristics across different age stages focuses mainly on infancy and adulthood. Future research should investigate chunking development trends and characteristics across all life stages. Two aspects warrant deeper exploration: (1) Do chunk level transitions occur within several fixed age intervals? As age increases, does the proportion of influence from physiological and environmental factors on chunk level transitions show dynamic changes? (2) Research on chunk characteristics in childhood, adolescence, and old age remains scarce. A few studies have found that when elderly individuals learn concepts about chemicals contained in different vegetables and are tested on their learning outcomes through conceptual classification tasks, they show that although working memory declines with age, elderly individuals use more past experience to conceptually integrate new knowledge, creating larger working memory storage space to compensate for working memory decline (Soederberg Miller, 2011). However, this study did not clarify what differences exist between elderly and adult use of past experience for chunking. Therefore, research on chunk characteristics across different age stages needs to examine chunk encoding, storage, and retrieval processes to reveal age-related features of chunking.

5.2 The Influence of "Rehearsal Strategy"and "Less is More" Principle on Chunk Capacity

No consensus exists regarding whether a definite numerical range exists for individual working memory chunk capacity. Future research can consider two aspects: (1) Is rehearsal strategy the key cause of large variations in working memory chunk capacity range? Although Cowan (2019) consistently affirms this, Berry et al.' s (2018) research showed that under both suppressed and unsuppressed rehearsal strategy conditions, whether presenting to-be-remembered items slowly or rapidly, participants' core memory capacity remained at 3-4 chunks, similar to Cowan' s observations. Researchers propose that working memory contains a fixed number of chunks, and rehearsal strategies may only affect the chunking process, causing overall chunk connection patterns to be in dynamic flux and making memory more flexible to meet specific information needs in different contexts, without affecting working memory chunk capacity. Unsworth and Engle (2007) proposed that working memory comprises a primary memory system and a secondary memory system. The primary memory system, maintained by attention, permanently stores four chunks, while the secondary memory system can retrieve additional information needed by the primary system from long-term memory to update chunks. Research has shown that in sports, experts outperform novices because, compared to novices who only use procedural memory, experts use more declarative memory and can consciously employ strategies to continuously transform connection patterns between chunks in working memory, enabling more flexible responses to different competition formats (Christensen et al., 2019). Therefore, rehearsal strategies may be a key factor causing changes in chunk connection patterns but may not affect chunk capacity. (2) Expanding working memory chunk capacity as much as possible should not be the fundamental goal for improving working memory efficiency. Future research should not only explore the capacity range of chunks in working memory but also investigate a capacity range that maximizes chunking function to ensure maximum quality of cognitive activities. Although increasing working memory capacity can promote and improve efficiency and quality in advanced cognitive activities such as problem-solving (Sang, 2017) and reading (Ni, 2017), some scholars propose that one hypothesis for why children can become better second language learners than adults is that although children' s chunk quantity is far lower than adults' , children can better concentrate attentional resources on linguistic information fragments contained within each chunk. Adults' working memory contains numerous chunks, leading them to focus more on the whole while neglecting details in language learning, resulting in lower language learning efficiency and accuracy compared to children (Gordon et al., 2020). Newport (1990) proposed the "less is more"hypothesis, suggesting that immature language learners can more easily grasp small amounts of linguistic information in each learning episode, which helps decompose linguistic information into many fine structures in complex language learning environments, making learning more precise. Therefore, researchers argue that current academia mostly focuses only on the upper and lower limits of chunk quantity in working memory, neglecting the characteristic that each chunk comprises multiple information fragments. Both chunk storage and retrieval and information fragment analysis require conscious resources; excessive chunks occupy conscious resources, causing people to neglect some information details within single chunks. Even when numerous chunks are retrieved, chunk usage quality decreases.

References

Huang, F., He, M., & Luo, J. (2017). The cognitive neural mechanism of chunking in insight problem solving. *Chinese Science Bulletin*, 62(31), 11.

Liu, Z., & Guo, C. (2013). ERP evidence for shared information representation between working memory and long-term memory. *Acta Psychologica Sinica*, 45(03), 285-293.

Ni, J. (2017). The effect of interactive memory reading strategies on second language working memory capacity. *Foreign Language Teaching*, 38(5), 49-53.

Peng, D. (2019). *General psychology* (5th ed.). Beijing Normal University Press.

Sang, N. (2017). The brain neural mechanism of visual working memory content maintenance. *Chinese Science Bulletin*, 63(27), 13.

Alluhaybi, I., & Witzel, J. (2020). Letter connectedness and Arabic visual word recognition. *Quarterly Journal of Experimental Psychology*, 73(10), 1660-1674.

Applin, J. B., & Kibbe, M. M. (2021). Young children monitor the fidelity of visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(5), 789-800.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of Learning and Motivation*, 2, 89-195.

Baddeley, A. D., Hitch, G. J., Allen, R. J., Malmberg, K. J., Raaijmakers, J. G. W., & Shiffrin, R. M. (2018). From short-term store to multicomponent working memory: The role of the modal model. *Memory & Cognition*, 47(4), 575-588.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829-839.

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.

Bennett, D., Gobet, F., & Lane, P. (2020, August). Forming concepts of Mozart and Homer using short-term and long-term memory: A computational model based on chunking. Paper presented at the Proceedings of the 42nd Annual Meeting of the Cognitive Science Society, Virtual.

Berry, E. D. J., Waterman, A. H., Baddeley, A. D., Hitch, G. J., Allen, R. J., & Dubow, E. F. (2018). The limits of visual working memory in children: Exploring prioritization and recency effects with sequential presentation. *Developmental Psychology*, 54(2), 240-253.

Bhandari, A., & Badre, D. (2018). Learning and transfer of working memory gating policies. *Cognition*, 172, 89-100.

Chase, W. G., & Simon, H. A. (1973). Perceptions in chess. *Cognitive Psychology*, 4(1), 55-81.

Chekaf, M., Cowan, N., & Mathy, F. (2016). Chunk formation in immediate memory and how it relates to data compression. *Cognition*, 155, 96-107.

Christensen, W., Sutton, J., & Bicknell, K. (2019). Memory systems and the control of skilled action. *Philosophical Psychology*, 32(5), 692-718.

Cowan, N. (1998). *Attention and memory: An integrated framework*. Oxford Psychology Series (No. 26). New York: Oxford 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.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, 19(1), 51-57.

Cowan, N. (2016). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, 24(4), 1158-1170.

Cowan, N. (2016). *Working memory capacity: Classic edition* (1st ed.). New York: Routledge.

Cowan, N. (2020). Why and how to study working memory development. *L' Année Psychologique*, 120(2), 135.

Cowan, N., AuBuchon, A. M., Gilchrist, A. L., Blume, C. L., Boone, A. P., & Saults, J. S. (2021). Developmental change in the nature of attention allocation in a dual task. *Developmental Psychology*, 57(1), 33-46.

Cowan, N., Belletier, C., Doherty, J. M., Jaroslawska, A. J., Rhodes, S., Forsberg, A., Naveh-Benjamin, M., Barrouillet, P., Camos, V., & Logie, R. H. (2020). How do scientific views change? Notes from an extended adversarial collaboration. *Perspectives on Psychological Science*, 15(4), 1011-1025.

Cowan, N., Li, Y., Glass, B. A., & Scott Saults, J. (2018). Development of the ability to combine visual and acoustic information in working memory. *Developmental Science*, 21(5), e12635. https://doi.org/10.1111/desc.12635

Cowan, N., & Albarracín, D. (2019). Short-term memory based on activated long-term memory: A review in response to Norris (2017). *Psychological Bulletin*, 145(8), 822-847.

Cowan, P. A., & Pape Cowan, C. (2019). The role of parental relationships in children' s well-being: A modest set of proposals for improving the lives of children. *Human Development*, 62(4), 171-174.

Darolia, C. R., & Varshney, N. (2016). Role of long-term memory activation in working memory. *Indian Journal of Health and Wellbeing*, 7(1), 153.

Ding, X., Gao, Z., & Shen, M. (2017). Two equals one: Two human actions during social interaction are grouped as one unit in working memory. *Psychological Science*, 28(9), 1311-1318.

Ericsson, K. A., Kintsch, W., & Bjork, R. A. (1995). Long-term working memory. *Psychological Review*, 102(2), 211-245.

Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links between the intuitive sense of number and formal mathematics ability. *Child Development Perspectives*, 7(2), 74-79.

Forsberg, A., Guitard, D., & Cowan, N. (2020). Working memory limits severely constrain long-term retention. *Psychonomic Bulletin & Review*, 28(2), 537-547.

Gobet, F. (2005). Chunking models of expertise: Implications for education. *Applied Cognitive Psychology*, 19(2), 183-204.

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, 5(6), 236-243.

Gobet, F., & Clarkson, G. (2004). Chunks in expert memory: Evidence for the magical number four or is it two? *Memory*, 12(6), 732-747.

Gobet, F., & Simon, H. A. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory*, 6(3), 225-255.

Gordon, R., Smith-Spark, J. H., Newton, E. J., & Henry, L. A. (2020). Working memory and high-level cognition in children: An analysis of timing and accuracy in complex span tasks. *Journal of Experimental Child Psychology*, 191, 104736.

Guida, A., Fartoukh, M., & Mathy, F. (2020). The development of working memory spatialization revealed by using the cave paradigm in a two-alternative spatial choice. *Annals of the New York Academy of Sciences*, 1477(1), 54-70.

Guida, A., Gobet, F., Tardieu, H., & Nicolas, S. (2012). How chunks, long-term working memory and templates offer a cognitive explanation for neuroimaging data on expertise acquisition: A two-stage framework. *Brain and Cognition*, 79(3), 221-244.

Hitch, G. J., Allen, R. J., & Baddeley, A. D. (2019). Attention and binding in visual working memory: Two forms of attention and two kinds of buffer storage. *Attention, Perception, & Psychophysics*, 82(1), 280-293.

James, W. (1891). The moral philosopher and the moral life. *International Journal of Ethics*, 1(3), 330-354.

Kibbe, M. M., & Feigenson, L. (2014). Developmental origins of recoding and decoding in memory. *Cognitive Psychology*, 75, 55-79.

Kibbe, M. M., & Leslie, A. M. (2019). Conceptually rich, perceptually sparse: Object representations in 6-month-old infants' working memory. *Psychological Science*, 30(3), 362-376.

Kingo, O. S., & Krøjgaard, P. (2012). Object function facilitates infants' object individuation in a manual search task. *Journal of Cognition and Development*, 13(2), 152-173.

Malmberg, K. J., Raaijmakers, J. G. W., & Shiffrin, R. M. (2019). 50 years of research sparked by Atkinson and Shiffrin (1968). *Memory & Cognition*, 47(4), 561-574.

Mathy, F., & Feldman, J. (2012). What' s magic about magic numbers? Chunking and data compression in short-term memory. *Cognition*, 122(3), 346-362.

Mclean, R. S., Gregg, L. W., & Grant, D. A. (1967). Effects of induced chunking on temporal aspects of serial recitation. *Journal of Experimental Psychology*, 74(4p1), 455-459.

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.

Newport, K. T. (1990, July). Incorporating survivability considerations directly into the network design process. Paper presented at the Infocom 90, Ninth Joint Conference of the IEEE Computer & Communication Societies: The Multiple Facets of Integration, San Francisco, CA, USA: IEEE.

Norris, D., & Albarracín, D. (2017). Short-term memory and long-term memory are still different. *Psychological Bulletin*, 143(9), 992-1009.

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., Dardick, W., Tao, S., & Albarracín, D. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin*, 144(1), 48-76.

Perfors, A. (2012). When do memory limitations lead to regularization? An experimental and computational investigation. *Journal of Memory and Language*, 67(4), 486-506.

Pförtner, P., & Hristova, P. (2021). Thematic bootstrapping: Performance differences between expert chess players and novices. *Cognition, Brain, Behavior:*

An Interdisciplinary Journal, 25(1), 55-68.

Pluss, M. A., Bennett, K. J. M., Novak, A. R., Panchuk, D., Coutts, A. J., & Fransen, J. (2019). Esports: The chess of the 21st century. *Frontiers in Psychology*, 10, 156.

Rosenberg, R. D., & Feigenson, L. (2013). Infants hierarchically organize memory representations. *Developmental Science*, 16(4), 610-621.

Sala, G., Burgoyne, A. P., Macnamara, B. N., Hambrick, D. Z., Campitelli, G., & Gobet, F. (2017). Checking the "Academic Selection" argument: Chess players outperform non-chess players in cognitive skills related to intelligence: A meta-analysis. *Intelligence*, 61, 130-139.

Sala, G., & Gobet, F. (2016). Do the benefits of chess instruction transfer to academic and cognitive skills? A meta-analysis. *Educational Research Review*, 18, 46-57.

Sala, G., & Gobet, F. (2017). Does far transfer exist? Negative evidence from chess, music, and working memory training. *Current Directions in Psychological Science*, 26(6), 515-520.

Simon, H. A., & Barenfeld, M. (1969). Information-processing analysis of perceptual processes in problem solving. *Psychological Review*, 76(5), 473-483.

Smalle, E. H. M., Bogaerts, L., Simonis, M., Duyck, W., Page, M. P. A., Edwards, M. G., & Szmalec, A. (2015). Can chunk size differences explain developmental changes in lexical learning? *Frontiers in Psychology*, 6, 1925.

Smith, E. T., Bartlett, J. C., Krawczyk, D. C., & Basak, C. (2021). Are the advantages of chess expertise on visuo-spatial working-memory capacity domain specific or domain general? *Memory & Cognition*, 49(8), 1600-1616.

Soederberg Miller, L. M., Gibson, T. N., Applegate, E. A., & de Dios, J. (2011). Mechanisms underlying comprehension of health information in adulthood: The roles of prior knowledge and working memory capacity. *Journal of Health Psychology*, 16(5), 794-806.

Stahl, A. E., & Feigenson, L. (2014). Social knowledge facilitates chunking in infancy. *Child Development*, 85(4), 1477-1490.

Stahl, A. E., & Feigenson, L. (2018). Infants use linguistic group distinctions to chunk items in memory. *Journal of Experimental Child Psychology*, 172, 149-167.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" words. *Psychological Review*, 69(4), 344-354.

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.

Verhaeghen, P., Cerella, J., & Basak, C. (2015). A working memory workout: How to expand the focus of serial attention from one to four items in 10 hours or less. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(6), 1322-1337.

Wang, J. J., & Feigenson, L. (2019). Infants recognize counting as numerically relevant. *Developmental Science*, 22(6), e12805. https://doi.org/10.1111/desc.12805

Note: Figure translations are in progress. See original paper for figures.