
Article:

Forsberg, A., Adams, E.J. and Cowan, N. (2022) Why does visual working memory ability improve with age : more objects, more feature detail, or both? A registered report. Developmental Science, 26 (2). e13283. ISSN: 1363-755X

https://doi.org/10.1111/desc.13283

Received: 24 March 2020 | Revised: 30 March 2022 | Accepted: 29 April 2022

DOI: 10.1111/desc.13283

REGISTERED REPORT

Developmental Science WILEY

Why does visual working memory ability improve with age: More objects, more feature detail, or both? A registered report

Alicia Forsberg | **Eryn J. Adams** | **Nelson Cowan**

University of Missouri, Columbia, Missouri, USA

Correspondence
Alicia Forsberg, Department of Psychology at the University of Sheffield, Cathedral Court, 1 Vicar Ln, Sheffield S1 2LT, UK
Email: a.forsberg@sheffield.ac.uk

Funding information
National Institutes of Health, Grant/Award Number: R01 HD-21338

Abstract

We investigated how visual working memory (WM) develops with age across the early elementary school period (6–7 years), early adolescence (11–13 years), and early adulthood (18–25 years). The work focuses on changes in two parameters: the number of objects retained at least in part, and the amount of feature-detail remembered for such objects. Some evidence suggests that, while infants can remember up to three objects, much like adults, young children only remember around two objects. This curious, nonmonotonic trajectory might be explained by differences in the level of feature-detail required for successful performance in infant versus child/adult memory paradigms. Here, we examined if changes in one of two parameters (the number of objects, and the amount of detail retained for each object) or both of them together can explain the development of visual WM ability as children grow older. To test it, we varied the amount of feature-detail participants need to retain. In the baseline condition, participants saw an array of objects and simply were to indicate whether an object was present in a probed location or not. This phase begun with a titration procedure to adjust each individual's array size to yield about 80% correct. In other conditions, we tested memory of not only location but also additional features of the objects (color, and sometimes also orientation). Our results suggest that capacity growth across ages is expressed by both improved location-memory (whether there was an object in a location) *and* feature completeness of object representations.

KEYWORDS
cognitive development, feature memory, object memory, working memory

1 | INTRODUCTION

Working Memory (WM) is the system that holds mental representations available for processing for use in higher-level cognitive activities (e.g., Logie & Cowan, 2015). WM capacity is thought to be a crucial determinant of cognitive development in childhood (Bayliss et al., 2003; Holmes et al., 2010) and individual differences in intellectual abilities (Conway et al., 2003; Jarrold & Towse, 2006). Generally, WM performance improves as children grow older (e.g., Brockmole & Logie, 2013; Cowan et al., 2006; Cowan et al., 2010; Cowan et al., 2006; Gathercole et al., 2004; Isbell et al., 2015; Riggs et al., 2006; Riggs et al., 2011), and understanding this development has important consequences for educational settings. For instance, children's ability to follow instructions may be constrained by WM capacity (Jaroslawska et al., 2016). However, despite general consensus that WM abilities improve as we reach adulthood, it is unclear which aspect of WM drives this improvement. Numerous candidate processes have been proposed, tested, and rejected. For instance, WM development does not seem driven by improved ability to allocate attention effectively (Cowan et al., 2010), improved object knowledge (Cowan et al., 2015), or reduced memory encoding limitations (Cowan et al., 2011).

Here, we focus on two factors that may explain visual WM improvement as children grow older from the elementary school years to adulthood. The first factor on which we focus is the increase in the number of objects that can be retained in WM, and the second is the amount of feature detail retained for each object. Consider when someone asks a child to remember three animals: a bird, a fish, and a giraffe. To distinguish a bird from a fish, they need to rely on certain features of these objects (do they have wings, do they have a beak?). The animals may also differ in size, color, and other features. It is possible that retaining three separate objects (or animals) is too much, and children will forget one of the animals. Or, they may retain something about each animal, but not all the features. For instance, they may remember that one animal was yellow, but forget its other features, and recall that another animal was a bird, while forgetting its color.

A large body of research suggests that adults typically can remember three to four items when there is no way to combine the presented items into fewer, larger chunks (Cowan, 2001; Luck & Vogel, 1997). However, when item-complexity increases, featural detail is not complete (Cowan et al., 2013; Hardman & Cowan, 2015; Oberauer & Eichenberger, 2013, although see Luck & Vogel, 1997). Cowan et al. (2013) presented arrays of colored shapes and required memory of only colors, only shapes, or both. After a brief retention period, participants judged whether a visual array differed from a comparison probe item that was presented ("*change*") or did not differ from it ("*no change*"). Young adult participants remembered something about around three items on average in all conditions but, when responsible for both features, they often forgot either the shape or the color. Similar results were found for multifeatured objects with 4 – 6 features (Hardman & Cowan, 2015; Oberauer & Eichenberger, 2013). Although it is beyond the scope of the present work, making adult participants responsible for two features instead of just one may also reduce the precision of memory, such as its exact location on a circle representing possible orientations or colors (Fougnie et al., 2010).

We hypothesized that the number of objects and featural detail of those objects may follow separate developmental trajectories, based on an intriguing paradox in the literature on memory ability in infants, children, and young adults. While adults typically make errors if the number of items to hold in mind exceeds three to four items (Cowan, 2001; Luck & Vogel, 1998), preschoolers and children who just started school seem able to retain only about 2–2.5 items (e.g., Cowan et al., 2005; Cowan et al., 1999; Riggs et al., 2006; Simmering, 2012). Surprisingly, though, there is evidence seeming to suggest that 18-month old infants may remember around three objects (e.g., Ross-Sheehy et al., 2003; Zosh & Feigenson, 2015). This would lead to the unsettling conclusion that memory capacity decreases with age in young children. That conclusion would be unwarranted; however, inasmuch as infant studies use different paradigms than studies of WM in older children. For example, Feigenson and Carey (2003; 2005) found that 14-month-old infants searched for the correct number of items when up to three objects were hidden. In such infant research, participants are attributed a memory capacity of three simply by remembering that three items were there. In contrast, paradigms used with older children typically involve detecting changes to (or reproducing) items based on features such as color, shape, or orientation (e.g., Burnett Heyes et al., 2012; Cowan et al., 2006; Heyes et al., 2016; Riggs et al., 2011; Sarigiannidis et al., 2016), requiring participants to remember what they saw, instead of simply indicating that they saw something.

Indeed, when exploring infants' memory for item features, memory capacity estimates are lower. Zosh and Feigenson (2012) tested whether infants remembered item features by replacing hidden objects that the infant has seen with hidden objects that have not been seen. If infants remember feature-detail they should notice when one object has been switched out for another, and search for the missing item. If, in contrast, infants only remember that they saw some object, but not what it was (i.e., no feature-detail), they would not notice the switch, and therefore not search for the original object. Using this approach, Zosh and Feigenson found that 18-month-olds appeared to remember sufficient featural detail to distinguish between objects (i.e., noticing identity switches) when tasked with remembering one or two objects. The infants were allowed to retrieve the objects from a container and kept searching for the remembered objects when the new objects were found in the container instead, presumably thinking that the old objects must still be in there. However, when three objects were hidden, the infants no longer appeared to notice such switches inasmuch as they stopped searching after three objects. Thus, despite remembering the presence of three objects, they appeared to remember three-object arrays with less featural detail than one- and two-object arrays. Interestingly, when the identity change was more pronounced – the researchers replaced an object with a nonsolid substance – infants

RESEARCH HIGHLIGHTS

- We tested whether the developmental increase in working memory was driven by improved ability to remember more objects, more features of such objects, or both.
- We examined memory for whether an object was presented at a particular location, with the task including 0, 1, or 2 other object features.
- Simple working memory capacity (number of item locations known) and the richness of object representations (including objects' locations, colors, and orientations) seemed to develop.
- We incorporated innovative methods with Bayesian sample size estimation, performance level adjustments, testing of multiple object features, and inference based on recent Bayesian methods.

did seem to notice, even at set-size three. This indicates that while some feature detail was retained, the representation may be too weak to differentiate between two solid objects, but sufficient to distinguish between more distinctly different representations (i.e., a solid vs. a non-solid object). Similar results have been found in infants as young as six months old, who seemed to remember the categorical identity (ball vs. doll head) of a hidden object, but failed to remember its perceptual identity (e.g., its color; Kibbe & Leslie, 2019).

Thus, the marked increase in WM ability seen from toddlerhood to adolescence (e.g., Cowan et al., 2006; Cowan et al., 2010; Cowan et al., 2006; Riggs et al., 2006; Riggs et al., 2011) may not be driven by the ability to retain more items, but instead by increased feature-detail retained for the remembered items. Differences in what constitutes “remembering an object” in typical infant-paradigms – compared to in paradigms used with children and adults – may explain this counterintuitive U-shaped function of memory capacity with age. If, like infants in some procedures, children merely need to know if something was there (without remembering featural-detail), their estimated memory capacity (*k*), should be about three items, similar to estimates obtained for infants in the aforementioned procedures and for young adults in testing procedures like ours. If so, this would suggest that the the number of objects that can be held in mind is constant across the human lifespan, but the amount of detail per item may explain the memory improvement associated with development. Consistent with this possibility, children who are only able to retain one or two items in visual WM still typically judge that they have about three items in mind, when asked about an array of colors before an objective test (Blume, 2018). In these cases, the children may remember certain objects for which they do not realize that they no longer retain the critical feature to be tested (in Blume’s procedure, color).

As these examples suggest, measuring the act of “remembering an object” is not necessarily straightforward. Indeed, the relationship between features and objects, and the space they occupy in working memory, is a contentious issue. Some research had suggested that a specific number of items can be held in memory regardless of the number of features per item (Luck & Vogel, 1997; Luria & Vogel, 2011; Vogel et al., 2001). However, others have found that remembering additional features does impair memory (Cowan et al., 2013; Hardman & Cowan, 2015; Oberauer & Eichenberger, 2013). Hardman & Cowan (2015) used a similar paradigm to ours and found that both increases in the number of objects *and* feature load impaired memory, in adult participants. The model that was fits to all three of these more recent data sets is one in which there is a limit to about three objects, and also a limit to the number of features per object for briefly presented arrays. By including a baseline condition titrated to yield a constant performance level across age groups, we plan to examine whether adding features to be remembered to each object creates more difficulty for younger children than for older children or adults.

We approached measurement of objects remembered in two ways. Firstly, we measured object memory as remembering that something was present in a specific location. This’ “was-something-there” implementation hews closely to the concept of an “object file” (Kahneman & Treisman, 1984). In this view, visual events are likened to reports to a police station, where a new file is opened for each novel event, by its time and location. Then, more features can be added (such as details about the crime or, for a visual object in our study, its color and orientation). Our “was-something-there” question may be like asking whether an object file was created, given that the objects do not move within an array. Then, additional details that may have been added to the object file (color and orientation) were sometimes probed. This order of testing fits with the idea that location is used to access specific visual features (Nissen, 1985), and has a special status in binding visual features (e.g., Kahneman et al., 1992; Treisman & Zhang, 2006; Wheeler & Treisman, 2002). Nevertheless, even if the percept is created via such a location-specific object file, it is theoretically possible that, for some objects, location is subsequently forgotten while the color or orientation of the object is retained. Indeed, other research suggests that location and feature information are not necessarily integrated. While location appears crucial for initial perceptual binding, location’s special status might be lost once representations are formed in WM, which operates according to different principles than visual attention and perception (e.g., Hedayati & Wyble, 2020). We allowed for the possibility that exact location could be mis-remembered while other features were remembered in a second measurement of objects in memory, namely, the the number of objects for which at least one feature was remembered (in addition to using the first quantification, object locations correctly remembered).

This theoretical question that we asked may be orthogonal to some other questions one could ask about the development of WM. For example, it is possible that what develops is the speed of refreshing the representations of items in working memory (e.g., Gaillard et al., 2011). Even if that is the case, one can still ask whether the developing rate allows retention of more items, more features per item, or both. Similarly, there may be developmental increases in knowledge and strategies (e.g., Cowan, 2016) but their development would not settle the issue of whether the advances that occur affect the number of WM representations or their detail, the latter determining if the representations are sufficient to answer experimental test questions.

We tested if developmental changes from child- to adulthood are driven by remembering more objects, and/or remembering objects with richer feature detail, by asking participants to remember objects at different levels of featural complexity. Below, we outline how we conceptualized the completeness of object representations, define the terms we will use, and present the key questions we addressed.

1.1 | Experimental aims and hypotheses

In practical applications of knowledge about working memory, as in education, one potential way to work around WM limitations and facilitate learning is to adjust the presentation of materials by reducing the number of parts to be held in mind independently (see Cowan, 2014; Gathercole & Alloway, 2007). To do this, it is useful to know if the number of chunks or the amount of feature-detail – or both – tend to overload young children’s WM. Returning to the animal example above, we would be interested in whether young children’s WM limitations are caused by the number of objects (animals), or the number of features (the feature complexity of those animals). Using simpler three-feature objects, we aimed to answer two questions. First, do children remember fewer objects? Second, does increasing the complexity of the memory task (i.e., asking participants to remember more features per object), influence performance equally across development?

We tested the hypothesis that what improves with development is the completeness of object representations but not the number of objects in WM per se. According to this hypothesis, the reason that young children do more poorly than adults on remembering arrays of, say, colored squares (e.g., Cowan et al., 2005) is that the colored squares are actually two-featured objects with location and color as distinguishing features, and children may remember the locations of just as many objects as adults, while remembering fewer colors, or may remember at least one feature (location or color) of just as many objects as adults, while remembering fewer features overall.

This *feature enrichment* hypothesis of developmental change leads to two predictions: (1) Regardless of age, the the number of objects at least partially in WM should be about three, but (2) for such objects, younger children should be less able to remember features (i.e., their performance will be more impaired when asked to remember additional feature detail about these objects). We examined developmental changes in both of these parameters (number of objects and number of features within objects) using a version of methods previously used in adults with multifeatured objects (Cowan et al., 2013; Hardman & Cowan, 2015), adapted here to the study of child development. Importantly, we included a baseline' "was-something-there" condition, making our paradigm conceptually related to some used in infant research, to achieve some general comparability to capacity estimation methods in such studies. We used a task in which the kind of visual display is always the same, but in which the information required for perfect task performance varies. Sometimes participants were only responsible for remembering whether an object was present at a particular location on the computer screen (baseline); other times, for remembering object location and color (one added feature); and still other times, for remembering object location, color and orientation (two added features). The advantage of this design is that the perceptual complexity of memory items is identical in all conditions, which is important as more complex items may be harder to remember because they are more challenging to perceive in a limited time frame (Eng et al., 2005), a factor outside the scope of our study. However, participants may disregard the task-instructions or prefer to focus a certain feature regardless of task-instructions. Such preferential encoding should be especially noticeable in the "any one-feature trial block" (in which any one of the three features can be probed). We also included a control analysis to detect selective dropping of second feature (see Supplementary Materials, Section 1).

1.2 | Specific hypotheses and how they will be tested

Hypotheses are summarized in Table 1. There we state only hypotheses in which conditions differ, but for any of them the opposite, null hypotheses can also be demonstrated given our Bayesian methods of inference.

First, according to a *capacity increase* hypothesis of developmental growth, we might find that older participants retain more objects in the Location condition (Hypothesis H_{1A}) and that this advantage should extend to Location tests within every condition (Hypothesis H_{1B}), with no claim of developmental change in the feature detail for remembered objects. That would fit with suggestions that WM improvement during childhood is due to a discrete increase in visual WM capacity, that is, that the maximum the number of objects that can be held in visual WM increases (e.g., Cowan, 2016) while the level of feature-detail of each successfully encoded object remains constant with age. For instance, Riggs et al. (2011) compared memory performance for single- and multifeature objects across three age groups: 7-year-olds, 10-year-olds, and adults. While adults remembered more objects than young children, the multifeature condition did not incur additional performance deficits in any age group, compared to trials in which only one feature could change. This suggests that the number of integrated multifeature object representations in WM changed with age. However, in Riggs et al. (2011) single-feature condition only orientation could change, while in the multifeature condition either color or orientation could change. Memory performance for color is typically better than that for orientation (see Hardman & Cowan, 2015; Oberauer & Eichenberger, 2013; Peich et al., 2013). Indeed, Riggs et al.'s adults also performed equally well in the single- and multifeature conditions, in contrast to Hardman & Cowan (2015; see also Oberauer & Eichenberger, 2013). It is possible that a general boost for color memory masked the detrimental effect of remembering two features. Therefore, to rule out this possibility in this study, we systematically examined memory for some features while varying the demand to remember other features. For example, we examined memory for object location in three conditions: when it alone must be remembered, when location and color must be remembered, and when location, color, and orientation all must be remembered.

The second potential outcome accords with the *feature enrichment* hypothesis of developmental growth. According to that hypothesis, all participants could retain an equal the number of objects, but older participants will retain more feature detail for each object. For everyone, it was expected that performance on location memory will decline as the need to retain additional features are added (Hypothesis H_{2A}), but according to the feature enrichment hypothesis, this decline will be steeper for younger participants (Hypothesis H_{2B}). Similarly, memory for color should decline when memory for orientation is also required (Hypothesis H_{2C}), and according to the feature enrichment hypothesis, this decline should be steeper for younger participants (Hypothesis H_{2D}). Improved memory stemming from increasingly detailed representations of remembered objects, rather than an increase in the number of objects, fits with the literature suggesting that infants can remember about three objects at once (Oakes & Luck, 2013; Zosh & Feigenson, 2015), but with limited feature-detail (see Zosh & Feigenson, 2012). This account might also align with accounts of feature-binding deficits in young children, compared to conditions when only one feature is required (see Cowan et al., 2006; Lorsbach

TABLE 1 Analysis plan

Research questions	Hypotheses	Analyses
1. Does the number of objects held in memory increase with age?	H_{1A} If the number of objects in WM increases with age, older participants should have higher *k* (items in WM) estimates in the “was-something-there” condition.	1. Compare *k* estimates (DV) by Age Group for location-memory (“was-something-there”), using BayesFactor ANOVA.
	H_{1B} If the number of objects in WM increases with age, older participants should have higher *k* (items in WM) estimates for “was-something-there” probes, regardless of feature-load condition.	2. Compare *k* estimates (DV) by Age Group for location-memory (“was-something-there”), when participants also need to remember (a) color and (b) color and orientation, using BayesFactor ANOVAs.
	H_{1C} If the number of objects in WM increases with age, the estimated number of items for which at least one feature was remembered should be larger in older participants.	3. Based on the final testing block (*any one-feature*), compare the estimated number of objects for which at least one feature was known (DV), relying on the fairly-well-met assumption of independence of the features (e.g., Fougnie & Alvarez, 2011), by Age Group, using BayesFactor ANOVA.
2. Does memory for feature detail in terms of proportion correct increase with age?	If remembering additional features reduces memory compared to remembering fewer features, increases in feature load will affect retention of location. Specifically:	4. Compare proportion correct (correct vs. incorrect responses; DV) when remembering Location-only vs. Location (+ Color) and Location (+ Color + Orientation), by Age Group, using Bayesian Logistic Regression.
	H_{2A} Memory for location should decrease when color also must also be remembered and decrease further when both color and orientation must be remembered.	
	H_{2B} If the ability to retain more feature-detail develops with age, feature-load decreases for location memory should be more severe for younger children.	
	H_{2C} Specifically, memory for color should decrease when orientation also must also be remembered.	5. Compare proportion correct (correct vs. incorrect responses; DV) when remembering Color-only versus Color (+ Orientation), by Age Group, using Bayesian Logistic Regression.
	H_{2D} If the ability to retain more feature-detail develops with age, feature-load decreases for color memory should be more severe for younger children.	
	H_{2D} If the ability to retain more feature-detail develops with age, the number of features per known object should increase with age.	6. Based on the final testing block (*any one-feature*) estimate the number of known features within objects for which at least one feature is known (DV), relying on the fairly-well-met assumption of independence of the features (e.g., Fougnie & Alvarez, 2011). Compare by Age Group, using BayesFactor ANOVA.

& Reimer, 2005). If the feature enrichment hypothesis completely accounts for the development of working memory, the assumption is that previous findings of increasing capacity with age were obtained because younger children more often forgot the tested feature (e.g., color), while still retaining knowledge of where about three objects were located.

These opposing outcomes (*capacity increase* vs. *feature-enrichment*) can be checked in a different manner that is not dependent on the special status of any one feature or on the total feature load. Specifically, in our final testing block, each trial included only one probe, which could be Location, Color, or Orientation, unknown to the participant until the probe is presented. Based on this trial block, as explained later, we could use a recent WM model to estimate the number of trials for which at least one feature is known, a type of *k* that could increase with development (Hypothesis H_{1C}), and alternatively, we could also estimate whether the total number of features known for at-least-partly known objects increases with development (Hypothesis H_{2D}) (cf. Cowan et al., 2013; Hardman & Cowan, 2015; Oberauer & Eichenberger, 2013).

A third potential outcome is that older participants would retain more objects *and* more feature detail for those objects; both the capacity increase and the feature enrichment hypotheses could be correct. As children develop, both parameters may increase and contribute to improved WM ability. There are similar, though not identical, findings in the literature. In particular, recent work has shown that both the capacity (the number of objects in WM) and the precision with which such objects were remembered were greater in adults than children at a set-size of two objects (Sarigiannidis et al., 2016). Increases in both the number of items retained and precision with age were also found for memory for tone series (Clark et al., 2018). Similarly, children's memory precision when reproducing the orientation of one or three bars improved with age, using both cross-sectional (Burnett Heyes et al., 2012) and longitudinal data (Heyes et al., 2016). This age-benefit was significantly greater in a three-bar than the one-bar condition, which may reflect that the number of high-precision slots increased with age. We were not investigating precision and do not know if precision of any feature plays a comparable role in development to what we are investigating, the number of features per object.

Finally, children and young adults theoretically could retain equal objects and level of feature detail. This overall null hypothesis is unlikely given past research, as young adults typically outperform children and adolescents on visual WM tasks (e.g., Brockmole & Logie, 2013; Cowan et al., 2005, 2006; Gathercole et al., 2004; Isbell et al., 2015; Riggs et al., 2006).

2 | METHOD

2.1 | Proposed sample characteristics

We planned to recruit 40 children (6 – 7 years old), 40 early adolescents (11 – 13 years old), and 40 college-age adults (18 – 25 years old). This sample size was selected following Bayes Factor design analysis simulations, and simulation of Bayesian posteriors (see Section 2 in the Supplementary Materials; Figure S1). Moreover, if evidence for age differences in *k* between age groups (Hypothesis $H_{1A,}$ Analysis 1) were inconclusive (defined as a Bayes Factor between 0.33 and 3), we would recruit 10 more participants per age group and reanalyze, a maximum of two times (see Schönbrodt & Wagenmakers, 2018). Eligible participants reported having normal or corrected-to-normal vision and normal color vision and speak English fluently. The study has been approved by the local research ethics committee (Institutional Review Board) at the University of Missouri. All participants (or, for child participants, their legal guardians) provided informed consent prior to participation. We outline detailed exclusion (and replacement) criteria in the Supplementary materials (see Section 3), based on near-floor performance (below .55 proportion correct), near-ceiling (above .97), performance below 90% on the perceptual matching task, as well as failure to complete the task.

2.2 | Demographic information to be collected

Participants and their parents or guardians reported their age (measured in months) and gender (female, male, other/prefer not to say), and we will report mean age and gender ratios by age group. Optional demographic information about participant race and ethnic group was gathered for research participation monitoring purposes and federal funding requirements, but was not reported or analyzed in this study.

2.3 | Experimental procedure

2.3.1 | Overview

Our original plan was to collect data in person. However, due to the COVID-19 pandemic, data was collected virtually through an online video call with the experimenter. Therefore, participants did not receive books and stickers as originally planned. While undergoing written and oral consent procedures, participants learned of cash payment. The phases of the experiment, in order, included participant instructions and engagement, a perceptual matching task, titration of set sizes to adjust the difficulty to accommodate individual differences in ability level, and the experiment proper. The titration procedure was based on an array of multifeatured cats followed by a probe location at which a cat was or was not placed, with recognition of that location tested. The experiment proper included trials with set sizes equal to and one above the result of titration, divided into trial blocks with each trial including one probe (testing location), two probes (testing location and color), and three probes (testing location, color, and orientation) in that order or the reversed order. Finally, each participant will complete a block in which, on every trial, any one feature (Location, Color, or Orientation) is probed. We estimate that the total time of testing will last between 45 and 55 min including all procedures.

2.3.2 | Participant instructions and engagement

Before starting the experiment, the experimenter will use a cover story to improve task understanding and engagement. They will tell participants that we need help figuring out which cats were having fun at a birthday party (e.g., was a cat with a hat of this color at the party?). When the probe is the same as an item in the memory array ("the cats at the party"), party should press "YES," if it is different from all such items, they should press "NO." All participants will see on-screen task instructions before starting the experiment, while the experimenter reads them aloud. On-screen, written instructions will also appear before each new block of trials (see Supplementary Materials, Section 4, for details on participant instructions). Participants will receive feedback after each trial; a green tick-mark (✓) will indicate correct, a light red cross (✗) incorrect, responses. Also, at the end of each block, participants will see numerous green tick marks on the screen, which represent all correctly answered trials, as well as a moving progress bar indicating how much of the study they have completed. An experimenter was available online via a virtual communication software for questions and encouragement.

2.3.3 | Perceptual matching task

Participants completed a perceptual matching task in which they matched the probe color (presented in the center of the screen) to one of the eight cats, by clicking on the appropriate cat. Similarly, they matched the orientation probe to one of the cats. Participants performed two trials for each feature (2 × 8 colors and 2 × 8 orientations). We had planned to exclude and replace participants who got less than 90% of these correct, allowing no more than three errors for 32 trials. However, we deviated from this rule. As we moved the study online due to the pandemic, participants (especially young children) seemed to click through this task rapidly and seemed to sometimes accidentally select the same option twice, perhaps due to the screen-sharing lag. Thus, a total of 17 participants performed under the 90% cut-off point (thirteen of the youngest children, $M = .81$, $SD = .07$, one adolescent, $M = .88$, and three adults, $M = .76$, $SD = .05$). The lowest performance was .69. Given the changed circumstances, we included these participants in the analysis. Due to an error, data was missing from 12 of the adult participants for this measure, resulting in $N = 38$.

2.3.4 | Titration protocol

The purpose of our titration procedure was to find a set size individualized for each participant so that all would perform at about the same proportion correct in a baseline condition. It is important to do so in order to measure any differential deficit of one age group compared to another (cf. MacDonald, 2015). In psychometric terms, the purpose of titration is to enable performance-level matching in the baseline (Location-only tested) condition to be described below, to prevent non-removable interactions (see Loftus, 1978; Wagenmakers et al., 2012). Without this matching, statistical evidence for a significant interaction term could vary depending on whether the absolute or relative performance difference was considered. For example, if young adults remember 90% of items in Condition A but performance drops by 25% in Condition B, the performance difference would be 22.5 percentile units. If children only remember 70% in Condition A and also drop by 25% in Condition B, their performance difference would be 17.5 percentile units. Then, we might find statistical evidence for a larger difference in young adults than for children for absolute differences in proportion correct, but not for relative differences. This muddles the theoretical validity of statistical evidence for an Age Group × Condition interaction effect. Furthermore, since we wanted to test whether the level of feature detail differs for the objects that participants do remember, it was important that the number of memory items was adjusted to each individual's performance level. Titration should also be an efficient approach to obtain accurate, individual *k* estimates (estimates of the number of items in working memory), while reducing the length of the experimental session compared to gathering data across a wide range of set-sizes for each participant, some of which may either be too easy or overwhelming.

To obtain a measure of the set size at which each participant can respond successfully in around 80% of trials, this titration procedure was performed for the Location-only (baseline) condition, at the start of the study (see Figure 1 A). After four practice trials, all participants completed 40 trials of titration, starting at a set-size of one item. When participants responded correctly three times in a row, one item was added to the next trial's set-size (e.g., after three consecutive correct responses at set-size one, they moved on to two-item sets). When participants responded incorrectly, their set-size for the next trial was reduced by one, with minimum and maximum possible set-sizes of one and seven item(s), respectively. This *three-up, one-down* procedure estimates an experimental set-size for which participants respond correctly around 80% of the time (Tansley et al., 1982).

In half of the experimental trials, participants' set-size was equal to their average set-size in their 25 final titration trials (rounded to the nearest integer), with the exception that if a participant's final obtained set-size was one, their experimental set-size was two items, as a precaution against ceiling effects. In the other half of trials, the set size exceeded that determined in titration by one item, as a precaution against underestimation of actual memory capacity with this method. For instance, if a participant's obtained set-size was three items, they completed one block (20 trials) in each experimental condition (Location, Location + Color, Location + Color + Orientation) at that set-size, and another block in each condition with four items (three plus one). If a participant reached the maximum possible level of seven items, they would need to remember seven items for all trials. Underestimation could occur when set sizes are rounded down from the titration average to the nearest integer (e.g., from 3.40 to 3.00) as it is not possible to present fractions of an item. Additionally, underestimation could occur as participants improve with practice.

We used this set size determination procedure instead of a common set size because floor and/or ceiling performance in the baseline condition would prevent meaningful testing of the effect of additional feature-load. Based on previous literature, we expected age-differences in performance, suggesting that a set size appropriate for most adults (e.g., four items) would likely result in floor-performance in many children. Therefore, this titration procedure – while unlikely to be perfect – appeared more efficient for collecting useable data. Based on previous literature on working memory capacity, we expected that most individuals will have three-item sets for comparison, because the smallest possible set size we administered was two items, with span plus one (in that case three items) also presented, and adult capacity is typically around three items. However, all participants completed the 'any-one feature block' at set size three (see details below).

2.3.5 | Experimental task

In the titration procedure just described, which used a baseline, location-only testing procedure, and in all conditions of the consequent experimental testing, participants saw arrays (see Figure 1) of cat-faces wearing hats. The hats were of different colors (including eight prototypical colors: red, green, blue, purple, pink, yellow, orange, and turquois) and the cats were presented at different orientations (−70, −50, −30, −10, 10, 30, 50, or 70 degrees of tilt of cat-faces in cone-shaped hats) with eight colors and eight orientations in the stimulus pool, and no repetition of color or orientation within an array. Locations were selected randomly within an imaginary rectangle (width = 9.8, height = 7.3 degrees) at the center of the screen, separated by at least 2.5 degrees from one another. When location probes were different, they were also be presented within this rectangle, at least 2.5 degrees from any presented item. After seeing an array (500 ms) followed by a blank interval (1000 ms), a probe item was presented.

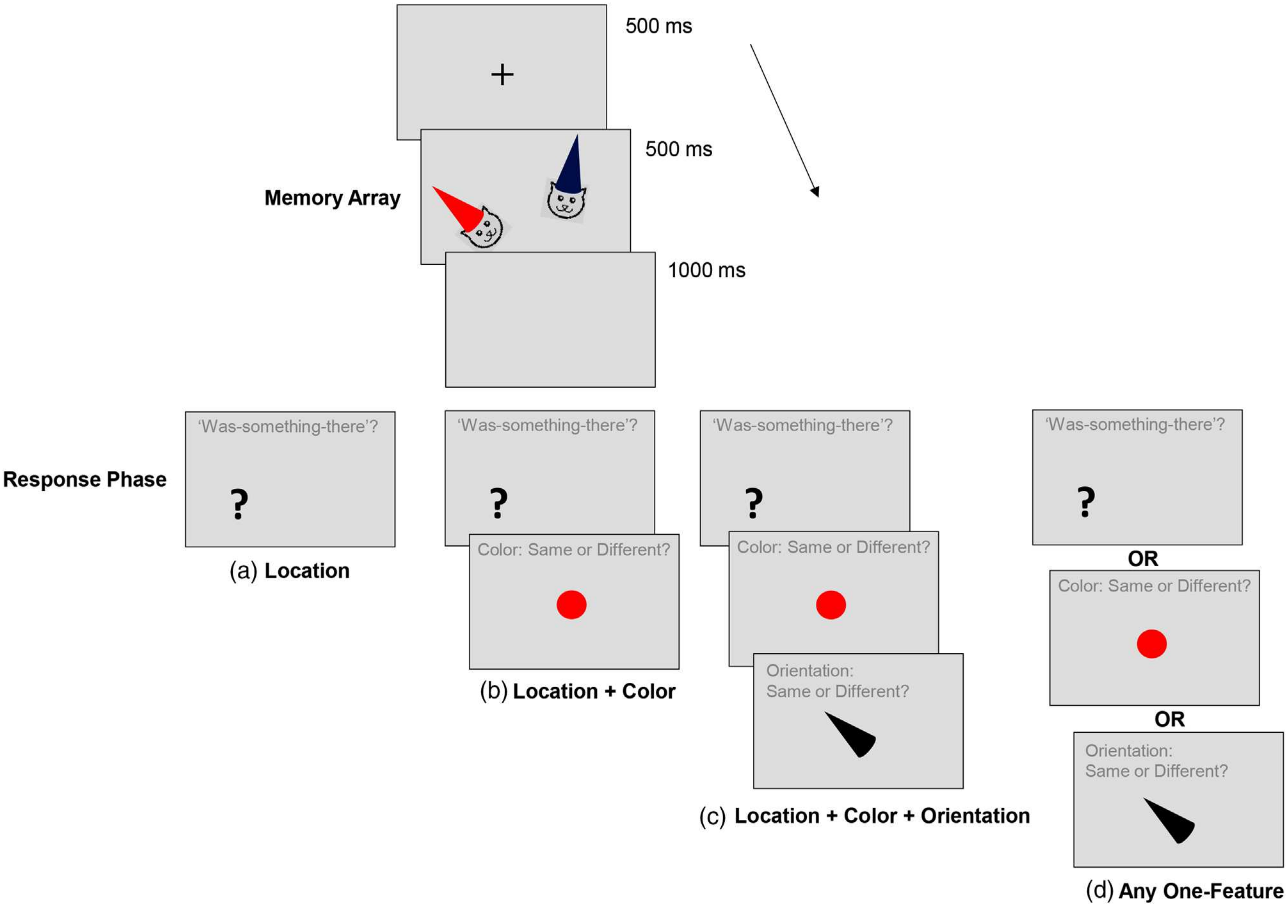

FIGURE 1 Outline of the experimental procedure at set size two, for the three different conditions. Participants respond to whether there was an item at the location of the question mark-probe, whether the probe color matches the color of the hat previously in that location, and/or whether the orientation of the hat matches the orientation in the memory array. In the main experimental segment (encompassing Parts a-c in the figure), features are always probed in this order. In this illustration, in the Location + Color + Orientation condition, one of the items is probed twice, because there are only two memory items. When there are three or more items, the location, color and orientation probes will all refer to different memory items. In the final experimental segment (Part d), any one feature is probed per trial. Stimuli are enlarged and text included for this illustration

Probe types

The probe items were (A) for location probes, a question mark, (B) for color probes, a circle filled with one of the eight study colors, or (C) for orientation probes, a black' "hat" pointing in one of the eight study angles. Participants responded to the probe in a change-detection procedure to indicate whether it was the same in the probed feature as a previously presented cat-object. Participants responded using keyboard keys (to indicate "YES" or "SAME" the "c" key, and to indicate "NO" or "DIFFERENT", the "m" key). They were advised to label the keys to help them remember these response options. The probe feature was the same as the to-be-remembered object half the time, and different from all array objects half the time (i.e., a new location or feature).

Trial block types

In a given trial block, there was only one probe per trial (Location, Figure 1 a), two probes per trial (Location, then Color, Figure 1b), or three probes per trial (Location, then Color, and then Orientation, Figure 1 c). When two or three features were probed, probes related to different items (i.e., the location of one item is probed, the color of another item and, when orientation is probed, the orientation of a third item). If there were fewer memory objects than features probed (which could occur at set-size two in Location + Color + Orientation trial blocks), one item was probed twice. Our original plan was to start each block with four practice trials. There would be a total of 40 Location ("was-something-there") trials, 40 Location + Color trials, and 40 Location + Color + Orientation trials, with each trial type presented within two blocks of 20 trials each. However, we found that young children did not complete the study, as the length caused frustration and boredom. We used our prespecified rule ("*if more than 4 of the 10 first participants in a given age group fail to complete the experiment, we will review the study length, reduce each experimental block to 16 instead of 20 trials, and recruit 8 additional participants in that age group (resulting in a total N of 48 in that age group)*") to facilitate data collection. We reduced trial numbers and increased participant numbers in all groups. We also reduced the number of practice trials in each block from four to two, as four practice trials appeared to make participants frustrated, and task understanding was generally excellent. Based on the noncompletion of the first child participants, we also decided to implement our second prespecified rule: "… *if participants want to end the session early (e.g., due to fatigue or boredom), they may opt to complete the rest of the study on a different day)*", and test all 6 –7 year old children in two sessions, to prevent an overly long session and to help ensure that they had an enjoyable experience. Adolescents and adults completed all their trials in one session.

Trial block order

For half of the participants in each age group, feature-load increased gradually (i.e., they started with Location-only (two blocks, at their two different set-sizes) followed by Location + Color (two blocks), then Location + Color + Orientation (two blocks). For the remaining participants, feature load gradually decreased (i.e., they start with Location + Color + Orientation (two blocks), followed by Location + Color (two blocks), and finally Location-only (two blocks). This difference in condition order allowed us to test whether it is difficult for participants to change their task set, adding or omitting features to be retained as the study progresses (see Supplementary Materials; Section 1, Table S1, Analysis 1). For each block type, all participants first received 16 trials of their lowest the number of objects per array or set size, based on each participants' performance level via a staircase titration procedure. The second set of 16 trials of the same block type included an additional item in each array.

Final, any one-feature trial block

Finally, each participant completed a block in which any one feature (Location, Color, or Orientation) was probed (Figure 1 d). In this final block, each feature was probed 10 times (resulting in a total of 30 trials), with randomized probe order, so that participants needed to retain all features. Each participants' higher set-size was used in this block. To allow comparison across age groups for a common set size, we also required that participants who did not complete this block with set-size three completed an extra block at this set-size.

2.4 | Proposed analysis pipeline

The purpose of our analyses was to compare (1) the number of objects held in working memory, *k*, between participants in the different age groups, and (2) the effect of increasing feature-load on performance (measured both by *k* and by more theory-neutral means), in the different age groups.

Figure 2(a) shows some key expectations according to the feature enrichment hypothesis as examined in the main part of our procedure. It depicts a situation in which older participants have richer representations of the objects in working memory. The titration procedure for location and the location-only (Load 0) procedure also yield tests of the capacity increase hypothesis. Figure 2(b) shows results that could be obtained in the final testing block, according to either hypothesis if age differences are general across different features when all three features must be retained. As this figure also shows, the latter results can be further analyzed to yield estimates of the number of objects for which at least one feature is known (relevant to the capacity increase hypothesis) and the number of features per known object (relevant to the feature enrichment hypothesis). Below, we explain how tests of the hypotheses are carried out.

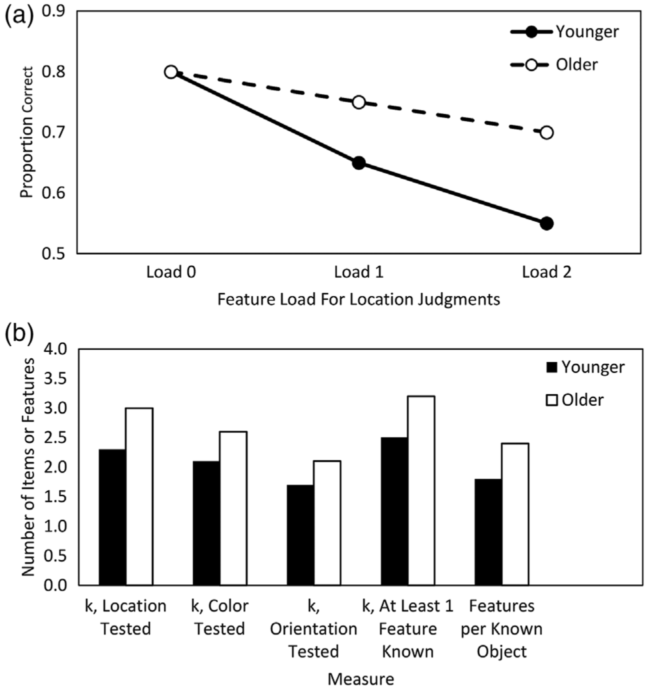

FIGURE 2 Depiction of some key possible results for any two age groups. (a) After the set size has been titrated to yield location-probe scores of about 80% correct for all individuals, an effect of adding the need to retain other features as well (Load 1, Color; Load 2, Color + Orientation). The graph depicts an effect of load that depends on age group. The results of the titration procedure and location-only test block also are expected to yield age effects in capacity for the location probes. (b) Results of the final testing block showing an age difference in estimated items in working memory, *k*, for the three tested features. The final two pairs of bars show derived measures to examine possible age differences in the estimated number of objects for which at least one feature is known and the estimated number of known features for such objects

2.4.1 | Items in working memory

The *k* parameter is based on simple logic in which the participant either answers correctly when the probe is presented because the corresponding array item is in WM, or guesses at a certain rate in the absence of such knowledge (Cowan, 2001; Cowan et al., 2013). To estimate *k* values for each age group, we implemented a hierarchical Bayesian model that uses all data to constrain each estimate (see Rhodes et al., 2018). The implementation used JAGS (Plummer, 2003) and the R package R2jags (R Core Su & Yajima, 2015; Team, 2015). See Supplementary Materials, Section 5, for details.

For our inferential statistical comparisons, we combined Bayesian estimation with Bayes factors, in line with suggestions that such approaches are complementary (e.g., see Rouder et al., 2018), and used different approaches to account for the distribution of our data. These analyses provide a measure of evidence for a model of an effect being non-null versus a model of it being null (see Dienes, 2019; Etz & Vandekerckhove, 2018; Morey et al., 2016). Specifically, we compared whether *k* estimates differ by age groups (see Table 1, Analysis 1) using a Bayesian ANOVA model comparison approach. We then explored differences between *k* values obtained for each age group, and for each feature-load condition, with a similar Bayesian ANOVA model comparison approach (see Table 1, Analysis 2). For further details, see the online supplement, Section 6. The key question here is whether the effect of increasing the number of features tested will produce a steeper decline in the tested features in younger children. For example, will younger children's *k* value for locations suffer a greater decline when color must also be retained, compared to older children and adults? Will still greater age differences be observed when orientation also must be retained? Will color memory suffer from the need to remember orientation, more in younger children?

As a second type of theoretical analysis convergent with the first, we also estimated the number of objects for which at least one feature was known for each participant (see Cowan et al., 2013; Hardman & Cowan, 2015; Oberauer & Eichenberger, 2013 for similar approaches. See the online supplement, Section 7, for details). For this estimation, we used data from the final trial block, in which only one feature (Location, Color, or Orientation) was probed in each trial, but participants did not know which feature. We assumed that the probability of remembering each feature is independent, consistent with the suggestion that separate memory stores with largely independent capacity limits exist for different features (see Bays et al., 2011; Wang et al., 2017; Wheeler & Treisman, 2002). This assumption is borne out by previous research, for example, Fougnie and Alvarez (2011). Then, we can estimate the proportion of trials in which at least one feature was known and also the number of known features within such objects (as shown in the online supplement, Section 8).

2.4.2 | Response distributions

Next, we used a Bayesian Logistic Regression to examine our main question in a theory-neutral manner using trial-level performance data. That method accounts for the binary distribution of our data (correct or incorrect), and include guessing rate in the model (50% correct with two-forced choice), see Table 1; contrasts 4 and 5. This model estimates the effect on the parameter η (*eta*; memory performance) by Age Group and Feature Load-condition, using a Bernoulli distribution. Participant identity was included as a random intercept, to account for individual variation. We used a normally distributed prior for η (*eta;* memory performance). For each model parameter, we report the parameter estimate, and its' 95% credible interval.

Specifically, to provide a theory-neutral answer our research question, we tested whether there was a main effect of feature-load condition, Age Group, and the crucial interaction between Feature-Load and Age Group (children vs. adults; children vs. adolescents). Feature-Load will be coded as a continuous variable (0, 1, 2 additional features to remember), and Age Group as categorical, since the three groups here are seen as distinct categories (childhood, adolescence, and adulthood) rather than continuous. To test whether increased feature-load is more detrimental for younger children's memory, we compared this model to a model with both the main effects and their interaction. We compared the strength of the evidence for including the interaction both by examining the confidence intervals and calculating a Bayes Factor for/against inclusion of the interaction parameter (see the online supplement, Section 6, for more details).

2.5 | Pilot data and simulations

2.5.1 | Sample size determination

Trying out some complex simulations with different sample sizes, we settled on 40 per age group[1], somewhat higher than the 24–30 participants per age group used in some previous studies in this area (e.g., Cowan et al., 2010, 2011). To determine an appropriate sample size, we used two separate procedures. Firstly, our most critical hypothesis is that the effect of increased feature-load will differ across age groups. Estimating power for interactions is notoriously difficult (e.g., see McClelland & Judd, 1993). In our analysis, we used a Bayesian logistic regression to account for the binary (correct vs. incorrect) nature of the data. This enables us to account for guessing rates and examine posterior distributions for each parameter. We simulated data for two imaginary populations, one with an Age Group × Feature Load interaction (H_1), and one without this interaction (H_0). We used some adult pilot data to estimate plausible feature-load differences (see Supplementary Materials for details). With 40 participants per group, 86.4% of our 500 simulations produced a 95% Bayesian credible interval that did not straddle 0 (indicating a differential effect of feature-load in children and adults), seeFigure S2. In contrast, in samples drawn from a population in which such an interaction effect was not present (H_0), in 94.6% of trials of our 500 simulations ($N = 40$ per group), the 95% Bayesian credible interval straddled 0, correctly rejecting the hypothesis of a differential effect of feature-load in children and adults. Based these simulations, we proposed 40 participants per age group as our predetermined, fixed sample size.

Next, for the comparisons of *k* values between age groups (see Table 1, Analysis 1, 2, 3 and 6) we used a Bayes Factor Design Analysis (BFDA; Schönbrodt & Wagenmakers, 2018). This method is based on the concept of Bayesian hypothesis testing and model comparison (Jeffreys, 1961; Kass & Raftery, 1995; Wagenmakers et al., 2010; Wrinch & Jeffreys, 1921). Using effect sizes in the field, with 40 participants per group, 97.9% of samples showed evidence for H_1 ($BF > 6$), 2.1% were inconclusive ($0.1667 < BF < 6$), and 0.0% showed evidence for H_0 ($BF < 0.1667$).

---

[1] Note that this was increased to 48 participants per age groups as trial numbers were reduced.

For further details on our sample size determination procedures see the Supplementary Materials, Section 2. In addition to procedures we have mentioned here, we include estimates for smaller effect sizes, taking into account possible publication bias, and we estimate the ability to find evidence for the null hypothesis.

2.6 | Outcome-neutral criteria that must be met for successful testing of the stated hypotheses

2.6.1 | Absence of near-floor and ceiling level performance

We excluded and replaced such participants (see exclusion criteria above for details).

Selective dropping of the second feature

Participants may strategically focus only on one feature in blocks in which they are asked to also remember color and orientation. If so, they may essentially perform the one-feature task again. It will be difficult to know whether this was driven by a decision to ignore a secondary feature, or inability to remember it. To err on the side of caution, we report the number of participants whose color memory performance is less than .55, and rerun the main analyses without such participants (see Supplementary Materials; Table S1, Analysis 2), to see whether it influences the main outcome.

Accounting for attentional lapses (which may be more prominent in children)

We examined the number of trials with Reaction Times over 5 s, which are presumably from lapses of attention. If they exceed 5% of total overall experimental trials, we would run a separate control analysis without such trials (see Supplementary Materials; Table S1, Analysis 3).

2.7 | Timeline for completion of the study and proposed resubmission date if Stage 1 review is successful

The completion date depends on official recommendations and university policy regarding social distancing due to the current global pandemic (COVID-19). Currently the university is closed but, once we can return to normal operations, we estimate the completion of data collection in 6 months and completion of analysis and writing in an additional 3 months.

2.8 | Online data collection

As this registered report received in-principle acceptance at the start of the COVID-19 pandemic, we were granted permission to collect the data virtually instead of in person. This led to some procedural changes, which we have outlined above. The preregistered protocol with in-principle acceptance was uploaded to the OSF prior to data collection (https://osf.io/59ekp/).

2.9 | Final sample of participants

One child participant turned eight before completing their second session and was excluded from all analyses. Following prespecified rules, five participants were excluded from the main analyses and replaced (location-only performance $< .55$: one child and adolescent, location-only performance $> .97$: one adolescent and two adults). However, these five participants performed in the specified range in the any-one-feature conditions and were included in those analyses. Overall, the study included data from 49 children ($M = 6.6$, $SD = 0.5$ years, 65.3% female, 36.7% male), 50 adolescents ($M = 11.9$, $SD = 0.9$ years, 46.0% female, 44.0% male) and 50 adults ($M = 19.7$, $SD = 1.8$ years, 72.0% female, 26.0% male, 2.0% nonbinary). All participants resided in the United States, except for two participants who resided in the United Kingdom.

3 | RESULTS

To compare the capacity increase and the feature enrichment hypotheses of WM development, we included various measures of both object and feature memory. We will discuss our results in the following order: First, we will focus on our different estimates of WM capacity, including standard estimations based on performance for location memory probes, as well as estimations of the number of objects for which at least one feature was known. Then, we discuss feature memory – specifically, the number of known features within objects for which at least one feature was known. Finally, we discuss how asking participants to remember additional features impacted location and color memory, respectively.

3.1 | WM capacity estimates (k)

First, we explored developmental changes in object memory (WM capacity). WM capacity estimates (k) for location memory varied by age group in the location-only ("was-something-there") condition ($BF_{10} = 6.1 \times 10^{20}$, children: $M = 1.7$, $SD = 0.5$, early adolescents: $M = 3.2$, $SD = 0.9$, adults: $M = 3.7$, $SD = 1.0$. Children vs. adults, $d = -2.67$). Similarly, location k-estimates also varied by age group when participants also need to remember color, and/or orientation (age group effect $BF_{10} = 3.7 \times 10^{70}$) with inconclusive evidence against a load effect ($BF_{01} = 2.9$), and evidence against an age group × load interaction ($BF_{01} = 32.0$). These values are shown in Figure 3. For details about how we obtained these hierarchical k-estimates see the Supplement, Section 9. Overall, these WM capacity analyses suggested that older participants tended to remember more objects than younger participants.

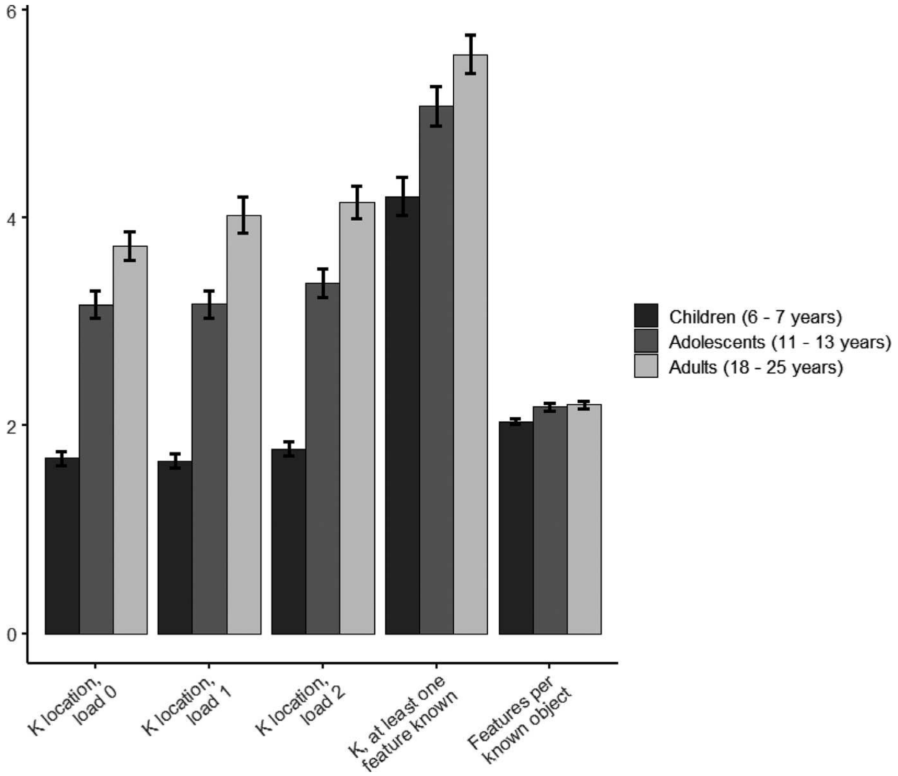

FIGURE 3 Different estimates of working memory capacity and feature knowledge by age group. Error bars represent Standard Error. *Note.* Plotted data for *K*, *at least one feature known* and *Features per known object* is from the "any-one-feature" trials in which the number of cat participants remembered was adjusted based on their performance in the titration block. The maximum features per known object were three (location, color, and orientation)

3.2 | The estimated number of objects for which at least one feature was known

In this alternative measure of object memory capacity, an object was considered remembered if at least one of its three features – location, color, or orientation – was known. Older participants once more tended to remember more objects than younger participants. In the block in which each participant worked in their estimated span set size + 1 item, the estimated the number of objects for which at least one feature was known was higher in older participants (BF_{10} = 1.9 × 10^3, Children: *M* = 4.2, *SD* = 1.1, Adolescents: *M* = 5.1, *SD* = 1.3, Adults: *M* = 5.6, *SD* = 1.2, Children vs. Adults, *d* = −1.15). When set size was fixed at three items for all participants, age differences were again observed (BF_{10} = 1.3 × 10^4, Children: *M* = 2.93, *SD* = 0.06, Adolescents: *M* = 2.98, *SD* = 0.04, Adults: *M* = 2.98, *SD* = 0.04. Children vs. Adults, *d* = −1.12). This reliable result was obtained despite near-ceiling-level performance on the latter metric.

3.3 | The number of known features within objects for which at least one feature is known

Finally, the number of *known features within objects for which at least one feature is known* appeared to differ between age groups when participants were asked to remember their assigned set size + one item (BF_{10} = 10.4, Children: *M* = 2.0, *SD* = 0.17, Adolescents: *M* = 2.2, *SD* = 0.23, Adults: *M* = 2.2, *SD* = 0.24. Children vs. Adults, *d* = −0.76). A similar pattern was observed when all participants had to remember three objects (BF_{10} = 2.1 × 10^{10}, Children: *M* = 2.1, *SD* = 0.18, Adolescents: *M* = 2.4, *SD* = 0.25, Adults: *M* = 2.5, *SD* = 0.23, Children vs. Adults, *d* = −1.92). This suggests that, *within known objects*, older participants remembered more object features than did younger participants.

3.4 | Memory for feature detail under memory load: Location

Next, we tested whether memory load (i.e., the number of features participants were told to remember and subsequently tested on) affected WM location performance. Location was the feature that was tested with a load of one, two, or three features. Using hierarchical Bayesian logistic regression, we found credible evidence that overall memory performance decreased as load increased (η = −0.23; SE = 0.08, 95% Bayesian Credible Interval; BCI [−0.39, −0.08]). There was also credible evidence that the children were outperformed both by the adolescents (η = 0.82; SE = 0.17, 95% BCI [0.49, 1.16]), and adults (η = 1.18; SE = 0.17, 95% BCI [0.86, 1.51]). There was no credible evidence that the load effect differed between children and adolescents (η = 0.14; SE = 0.10, 95% BCI [−0.06, 0.34]). However, the load effect

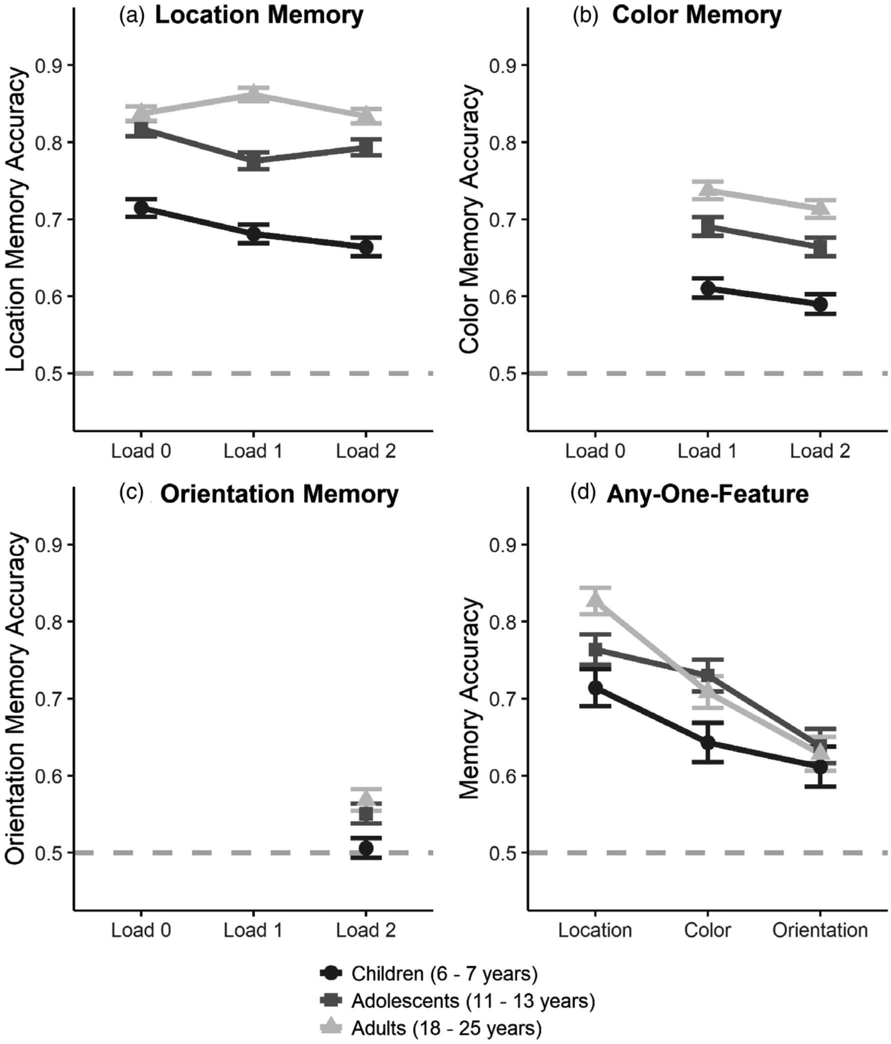

FIGURE 4 Accuracy in different feature load conditions. Panel (a). Location ("was-it-there") memory accuracy by the instructional load. (b). Color memory accuracy by the instructional load. (c). Orientation memory accuracy, only examined when the two other features were also to be remembered. (d). Memory accuracy in the any-one-feature condition, first-probed feature, in which any one of the three features could be probed first. In this condition, only one feature was probed. Error bars represent Standard Error. The dashed horizontal line represents chance level performance (50% accuracy). *Note.* Plotted data for the *Any-One-Feature* condition is from the "any-one-feature" trials in which the number of cat participants were asked to remember was adjusted based on their performance in the titration block

seemed to differ between children and adults ($\eta = 0.22$; SE = 0.10, 95% BCI [0.03, 0.42]), see Figure 4. Yet, the BF in favor of the model not including the load by age group interaction was 220.4 over a model including this interaction.

3.4.1 | Exploratory analysis: Children versus adults

Since this BF reflects a comparison with the full model (also including contrasts between children and adolescents), we did an exploratory analysis including only the children and adults. The BF in favor of the model not including this interaction was 3.9 over a model including it. Thus, the evidence is inconclusive regarding whether increasing feature load affected children's memory performance more than adults'. Interestingly, children's average *k*-estimates appeared quite consistent across loads (see Figure 3) compared to the decrease observed for location memory accuracy (see Figure 4). These differences might indicate shifts in the proportion of hits and false alarms across load conditions, which influences *k*-estimates differently than accuracy scores.

3.4.2 | Decreasing or increasing feature-load

Half of the participants experienced an increasing feature-load, the others a decreasing feature load. A planned control analysis suggested that feature-load order did not affect memory performance (decreasing feature-load: *M* = .78, *SD* = .12, increasing feature-load: *M* = .77, *SD* = .12, evidence against including the block order factor, $BF_{10} = 34.9$). See Supplement, Section 10 for the complete model output.

3.4.3 | Selective dropping of color memory

Next, in our second planned control analysis, we excluded participants who performed less than < .55 in the color condition, which might indicate "selective dropping" of that feature, which would be likely if participants decided to focus only on location memory, because it was probed first. This this excluded sixteen children, five adolescents, and two adults. We reran the analyses and the patterns were similar to those in the primary analysis (see Supplement, Section 11 for details).

3.5 | Memory for feature detail under memory load: Color

Next, we tested whether an additional orientation memory impaired color memory load. We conducted a similar Bayesian regression analysis for color memory (comparing color memory in the Location + Color condition with color memory with Location + Orientation) by age group. As per the prespecified exclusion rule, 23 participants (16 children, five adolescents, and two adults) were excluded due to performing below < .55 in the color memory condition. We found credible evidence that memory performance decreased as load increased ($\eta = -0.62$; SE = 0.23, 95% Bayesian Credible Interval; BCI [−1.06, −0.19]). While there were credible performance differences between children and adolescents ($\eta = .62$; SE = 0.28, 95% BCI [0.07, 1.18]) and children and adults ($\eta = 0.99$; SE = 0.27, 95% BCI [0.45, 1.51]) in the model without the interaction, this was not the case in the model including the interaction (children vs. adolescents: $\eta = 0.01$; SE = 0.48, 95% BCI [−0.92, 0.97] and children vs. adults: $\eta = 0.30$; SE = 0.47, 95% BCI [−0.60, 1.26]), likely reflecting how the exclusion of 16 of the poorest performing children reduced age differences. However, the load effect seemed equal in both younger children and adolescents ($\eta = 0.42$; SE = 0.28, 95% BCI [−0.12, 0.98]), and younger children and adults ($\eta = 0.48$; SE = 0.27, 95% BCI [−0.03, 1.00]). The BF in favor of the model not including this interaction was 85.9 over a model including it.

3.5.1 | Exploratory analysis: Children versus adults

Contrasting children and adults only, the BF in favor of the model not including this interaction was 3.5 over a model including it. Similar to the analysis above, the evidence that increasing feature load affected children's memory performance more than it affected adults' memory performance appears inconclusive.

3.6 | Reaction time

The number of trials with Reaction Times over 5 s did not exceed 5% of total overall experimental trials for either feature (Location, 4.1%, Color: 3.6%, Orientation: 3.3%). Thus, no planned control analysis excluding higher RT trials was needed.

4 | DISCUSSION

We sought to address whether developmental improvements in visual WM ability are expressed by location-memory (whether there was an object in a probed location, taken to reflect the presence of an object file as in the concept of Kahneman & Treisman, 1984), by feature completeness of object representations, or by both qualities. Overall, our results suggested that older participants retain more objects *and* more feature detail for those objects, supporting both the capacity increase and the feature enrichment hypotheses. Thus, as children develop, both the number of objects, and the number of features remembered within those objects, appear to increase and contribute to improved WM ability. This aligns with previous research using different methods and materials (e.g., Clark et al., 2018; Sarigiannidis et al., 2016). We discuss the observed age differences in these two parameters in more detail next.

4.1 | Do children remember fewer objects?

We observed strong evidence that younger participants remembered fewer objects than adolescents and adults. Average child WM capacity (*k*) in our "was-it-there" condition was just under two items, while adolescents' capacity was just over three, and adults' capacity was closer to four items. Thus, WM capacity doubled between the early school years (6-7 years) and adulthood. This suggests that WM improvement during childhood is partially driven by a discrete increase in the maximum number of object locations that children can hold in visual WM (e.g., Cowan, 2016) and provides evidence for the *capacity increase* hypothesis of WM development.

4.2 | Does increasing the complexity of the memory task (i.e., asking participants to remember more features per object) influence performance equally across development?

The answer to this question was less straightforward. We observed some evidence that the additional load requirement was more detrimental for children than adults for location memory. However, this effect was supported by credible intervals but was "inconclusive" when using a Bayes Factor model comparison approach. The color memory analysis provided similar weak evidence against differential effects of additional load in the different age groups.

For these contrasts, we manipulated load demand through the instructions, while stimuli were consistent (i.e., each cat had a location, a color, and an orientation) across trials, keeping the perceptual load consistent. Our procedure differs from previous research in several ways. For example, we focus on location memory and used a titration procedure, so that participants performed the task at different set sizes, based on their individual memory ability. Some previous research has suggested that additional feature load was detrimental for young adults' memory performance (Hardman & Cowan, 2015; Oberauer &

Eichenberger, 2013). However, others found that increased feature load did not impact either adults' or children's memory performance (Riggs et al., 2011). Our analyses do not provide clear evidence for or against age-differences in this instructional load effect. Next, we discuss our final contrasts, which were devised to directly compare the capacity increase versus feature-enrichment hypotheses in a manner that is not dependent on the load instruction.

4.3 | Capacity increase versus feature-enrichment

4.3.1 | The number of objects for at which at least one feature was known

We estimated the *number of objects for which at least one feature was known* based on an experimental condition where either location, color, or orientation could be probed. We observed age differences in this measure, further supporting the *capacity increase hypothesis*. Notably, these estimates were higher than the traditional *k*-estimates – children remembered at least one feature for just over four objects, adolescents for just over five, and adults for 5.6 items (see Figure 3). These results provide a potential explanation for the "infant paradox," that is, evidence that infants seem able to hold more objects in WM (up to three objects, Ross-Sheehy et al., 2003; Zosh & Feigenson, 2015) than school-aged children (less than two items; for example, 1.7 objects in this study). As discussed above, infant memory is measured using simpler paradigms which allow noticing a change in one of many features (e.g., one item might be a doll, the other a ball, Kibbe & Leslie, 2019). Our measure of the *the number of objects for which at least one feature was known* may provide a fairer comparison to infant measures, as it credits broader feature knowledge.

This more lenient measure of capacity also may explain discrepancies between how many colored squares participants believe they can hold in mind and their actual WM capacity, estimated using color-probes (Blume, 2018; Forsberg, Blume, et al., 2021). Indeed, subjective memory ratings may capture a sense of knowing *something* about the object rather than remembering the probed feature specifically. Finally, these values appear higher than traditional *k* estimates of 3–4 objects in young adults (c.f. 5.6 objects for which at least one feature was known observed in this study). This suggests that perhaps standard capacity estimates obtained by probing memory for one feature (e.g., color) may underestimate the the number of objects for which *something* is known – be that the location, color, or orientation. However, arguably, location may have a special role in visual processing (see Tsal & Lavie, 1988; Yousif et al., 2021), and our location test may allow spatial perceptual grouping, which could potentially' "inflate" these capacity estimates (see Brady & Alvarez, 2015; Morey, 2019).

4.3.2 | The estimated features known per object

Within the known objects, older participants appeared to remember more features. Specifically, adults and adolescents remembered 2.2 out of three features per known object, compared to 2.0 features in children. This supports the feature-enrichment hypothesis of visual WM development. Notably, the development of feature knowledge appeared more pronounced at earlier developmental stages (i.e., the difference between childhood and adolescence appeared more prominent than the difference between adolescence and adulthood).

4.4 | How do these findings relate to fundamental theories of cognitive development?

Different theories of cognitive development propose different mechanisms for developmental increases in WM capacity (for a review, see Cowan, 2022). In an Empiricist framework, WM changes are likely attributed to a child's learning history while a Nativist framework would emphasize the role of the brain's biological growth. A Cognitivist perspective would focus on identifying subprocesses that change with age, based on both learning and biological maturation (Cowan, 2016). Cognitivists also emphasize the *interaction* between the development of obligatory and voluntary processes that contribute to performance on WM capacity tests. For example, developmental capacity increases may be driven by increased use of grouping strategies, attentional refreshing of information, and verbal rehearsal. Increases in these processes may all be driven by developmental growth in a more general process, such as self-directed use of attention (see Cowan, 2022). This idea seems aligned with *Dynamic Systems Theory* (Spencer, 2020), which suggests that development in different processes may be captured more parsimoniously by considering more general developments in how the brain implements sustained activation. In such *Dynamic Systems Theory* accounts, basic changes in neural functioning, perhaps combined with environmental input (see Witherington & Margett, 2011), are believed to produce developmental capacity growth (e.g., Perone et al., 2021). In the context of our findings, the same improved ability for sustained activation – explained by the same biological brain maturation – may improve the ability to hold both object and feature memory in mind, without necessarily considering these processes as separate. Based on our data, we cannot discern whether the observed developmental increases in object and feature memory are driven by improvements in the same underlying developmental process, or separate processes. However, our findings of developmental increases in both parameters suggest that a shared process is theoretically plausible, and may be a parsimonious explanation.

4.5 | Developmental increases in both object and feature-memory: Theoretical and practical implications

Understanding the mechanisms of developmental WM constraints has implications for successful problem-solving in the classroom, as well as long-term learning (see Forsberg, Adams, et al., 2021; Forsberg, Blume, 2021; Forsberg, Guitard, 2021). Our findings make an important theoretical contribution as they provide evidence for developmental increases in *both* object and feature-memory – rather than one or the other. This suggests that as children mature, they are able successfully hold: (1) an increased number of informational chunks, and (2) increasingly complex informational chunks, in WM. Our results also provide a likely explanation for the "infant paradox" (i.e., seemingly better object memory in infants than in school-aged children), by demonstrating how children's (and adults') object memory capacity may actually be higher than typical estimates, if broader feature knowledge is included in the memory estimate. Finally, our results align with conceptualizations of WM as being limited by both the number of objects "slots" (Adam et al., 2017) and the "precision" (or featural detail) of memory representations (Ma et al., 2014).

To conclude, as children develop, both the the number of objects *and* the number of features remembered within those objects, appear to increase and contribute to improved WM ability – supporting both the *capacity increase* and the *feature enrichment* hypotheses of WM development. While the mechanisms driving these improvements are yet to be determined, one parsimonious explanation would be developmental growth in a more general process, such as self-directed use of attention (see Cowan, 2022), perhaps underpinned by developments in how the brain implements sustained activation (e.g., Spencer, 2020). Our findings have important implications for the theoretical understanding of visual WM capacity increases across childhood, as they suggest suggesting that developmental improvements cannot be explained exclusively by either increases in the number of informational slots, or the ability to remembering richer representations. Instead, developmental WM capacity growth is characterized by improvements in both these cognitive parameters.

ACKNOWLEDGMENTS

This research is supported by NIH Grant R01 HD-21338 to Cowan. We thank Bret Glass, Cayden Lawrence, Daniel Fridman, and Reese Lavers for assistance with data collection, and Nathaniel Greene, Stephen Rhodes, Yu Li, and other members of the Working-Memory Laboratory at the University of Missouri for helpful comments and advice.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The original data, analysis code and the experimental script is available online at https://osf.io/wskcf/.

ORCID

Alicia Forsberg https://orcid.org/0000-0002-6635-6160

REFERENCES

Adam, K. C., Vogel, E. K., & Awh, E. (2017). Clear evidence for item limits in visual working memory. *Cognitive Psychology*, *97*, 79–97.

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.

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia*, *49*(6), 1622–1631.

Blume, C. L. (2018). *School-aged children's awareness of their own working memory contents. Doctoral dissertation,*, Columbia, MO: University of Missouri.

Brady, T. F., & Alvarez, G. A. (2015). No evidence for a fixed object limit in working memory: Spatial ensemble representations inflate estimates of working memory capacity for complex objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 921.

Brockmole, J. R., & Logie, R. H. (2013). Age-related change in visual working memory: A study of 55753 participants aged 8-75. *Frontiers in Psychology*, *4*, 12.

Burnett Heyes, S., Zokaei, N., van der Staaij, I., Bays, P. M., & Husain, M. (2012). Development of visual working memory precision in childhood. *Developmental Science*, *15*(4), 528–539.

Clark, K. M., Hardman, K. O., Schachtman, T. R., Saults, J. S., Glass, B. A., & Cowan, N. (2018). Tone series and the nature of working memory capacity development. *Developmental Psychology*, *54*(4), 663.

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, *7*(12), 547–552.

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. (2014). Working memory underpins cognitive development, learning, and education. *Educational Psychology Review*, *26*(2), 197–223.

Cowan, N. (2016). Working memory maturation: Can we get at the essence of cognitive growth? *Perspectives on Psychological Science*, *11*(2), 239–264.

Cowan, N. (2022). Working memory development: A 50-year assessment of research and underlying theories. *Cognition*, *224*, 105075.

Cowan, N., AuBuchon, A. M., Gilchrist, A. L., Ricker, T. J., & Saults, J. S. (2011). Age differences in visual working memory capacity: Not based on encoding limitations. *Developmental Science*, *14*(5), 1066–1074.

Cowan, N., Blume, C. L., & Saults, J. S. (2013). Attention to attributes and objects in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 731–747.

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42–100.

Cowan, N., Fristoe, N. M., Elliott, E. M., Brunner, R. P., & Saults, J. S. (2006). Scope of attention, control of attention, and intelligence in children and adults. *Memory & Cognition*, *34*(8), 1754–1768.

Cowan, N., Morey, C. C., AuBuchon, A. M., Zwilling, C. E., & Gilchrist, A. L. (2010). Seven-year-olds allocate attention like adults unless working memory is overloaded. *Developmental Science*, *13*(1), 120–133.

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-span development of visual working memory: When is feature binding difficult? *Developmental Psychology*, *42*(6), 1089–1102.

Cowan, N., Nugent, L. D., Elliott, E. M., Ponomarev, I., & Saults, J. S. (1999). The role of attention in the development of short-term memory: Age differences in the verbal span of apprehension. *Child Development*, *70*(5), 1082–1097.

Cowan, N., Ricker, T. J., Clark, K. M., Hinrichs, G. A., & Glass, B. A. (2015). Knowledge cannot explain the developmental growth of working memory capacity. *Developmental Science*, *18*(1), 132–145.

Dienes, Z. (2019). How do I know what my theory predicts?. *Advances in Methods and Practices in Psychological Science*, *2*(4), 364–377.

Eng, H. Y., Chen, D., & Jiang, Y. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review*, *12*(6), 1127–1133.

Etz, A., & Vandekerckhove, J. (2018). Introduction to Bayesian inference for psychology. *Psychonomic Bulletin & Review*, *25*(1), 5–34.

Feigenson, L., & Carey, S. (2003). Tracking individuals via object-files: Evidence from infants' manual search. *Developmental Science*, *6*(5), 568–584.

Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition*, *97*(3), 295–313.

Forsberg, A., Adams, E. J., & Cowan, N. (2021). The role of working memory in long-term learning: Implications for childhood development. In: *Psychology of Learning and Motivation* (Vol. 74, pp. 1–45). Academic Press.

Forsberg, A., Blume, C. L., & Cowan, N. (2021). The development of metacognitive accuracy in working memory across childhood. *Developmental Psychology*, *57*(8), 1297.

Forsberg, A., Guitard, D., & Cowan, N. (2021). Working memory limits severely constrain long-term retention. *Psychonomic Bulletin & Review*, *28*(2), 537–547.

Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision*, *11*(12), 3–3.

Fougnie, D., Asplund, C. L., & Marois, R. (2010). What are the units of storage in visual working memory? *Journal of Vision*, *10*(12), 27–27.

Gaillard, V., Barrouillet, P., Jarrold, C., & Camos, V. (2011). Developmental differences in working memory: Where do they come from? *Journal of Experimental Child Psychology*, *110*, 469–479.

Gathercole, S. E., & Alloway, T. P. (2007). *Understanding working memory: A classroom guide*. Lontoo: Harcourt Assessment.

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.

Hardman, K. O., & Cowan, N. (2015). Remembering complex objects in visual working memory: Do capacity limits restrict objects or features? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 325–347.

Hedayati, S., & Wyble, B. (2020). Memories of visual events can be formed without specific spatial coordinates. *Journal of Cognition*, *3*(1), 13.

Heyes, S. B., Zokaei, N., & Husain, M. (2016). Longitudinal development of visual working memory precision in childhood and early adolescence. *Cognitive Development*, *39*, 36–44.

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2010). Poor working memory: Impact and interventions. In *Advances in Child Development and Behavior*, *39*, 1–43. JAI.

Isbell, E., Fukuda, K., Neville, H. J., & Vogel, E. K. (2015). Visual working memory continues to develop through adolescence. *Frontiers in Psychology*, *6*, 696.

Jaroslawska, A. J., Gathercole, S. E., Logie, M. R., & Holmes, J. (2016). Following instructions in a virtual school: Does working memory play a role? *Memory & Cognition*, *44*(4), 580–589.

Jarrold, C., & Towse, J. N. (2006). Individual differences in working memory. *Neuroscience*, *139*(1), 39–50.

Jeffreys, H. (1961). *Theory of ProbabilityTheory of probability*. Oxford: UK Oxford University Press.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In: R. Parasuraman & Davies DR (Eds). *Varieties of attention*. Academic Press Inc.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*(2), 175–219.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795.

Kibbe, M. M., & Leslie, A. M. (2019). Conceptually rich, perceptually sparse: Object representations in 6-month-old infants' working memory. *Psychological Science*, *30*(3), 362–375.

Loftus, G. R. (1978). On interpretation of interactions. *Memory Cognition*, *6*(3), 312–319.

Logie, R. H., & Cowan, N. (2015). Perspectives on working memory: Introduction to the special issue. *Memory & Cognition*, *43*(3), 315–324.

Lorsbach, T. C., & Reimer, J. F. (2005). Feature binding in children and young adults. *The Journal of Genetic Psychology*, *166*(3), 313–328.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 279–281.

Luck, S. J., & Vogel, E. K. (1998). Response from Luck and Vogel (Response to Commentary by Nelson Cowan). *Trends in Cognitive Sciences*, *2*, 78–80.

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia*, *49*(6), 1632–1639.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356.

MacDonald, A. W. III (2015). Differential deficit. In R. L. Cautin & S. O. Lilienfeld (Eds.), *Encyclopedia of clinical psychology*. Wiley online library.

McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin*, *114*(2), 376–390.

Morey, C. C. (2019). Perceptual grouping boosts visual working memory capacity and reduces effort during retention. *British Journal of Psychology*, *110*(2), 306–327.

Morey, R. D., Romeijn, J. W., & Rouder, J. N. (2016). The philosophy of Bayes factors and the quantification of statistical evidence. *Journal of Mathematical Psychology*, *72*, 6–18.

Nissen, M. J. (1985). Accessing features and objects: Is location special. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and Performance* (Vol. 11, pp. 205–219). Hillsdale, NJ: Erlbaum.

Oakes, L. M., & Luck, S. J. (2013). Short-term memory in infancy. *The Wiley handbook on the development of children's memory*. Wiley Blackwell, vol. 1, pp. 151–180.

Oberauer, K., & Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Memory & Cognition*, *41*(8), 1212–1227.

Peich, M. C., Husain, M., & Bays, P. M. (2013). Age-related decline of precision and binding in visual working memory. *Psychology and Aging*, *28*(3), 729–743.

Perone, S., Simmering, V. R., & Buss, A. T. (2021). A dynamical reconceptualization of executive-function development. *Perspectives on Psychological Science*, *16*(6), 1198–1208.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In Proceedings of the 3rd international workshop on distributed statistical computing, Vol., *124*(125), pp. 1–10.

R CoreTeam. (2015). *R: A language and environment for statistical computing*. Team, 3, p. 2.

Rhodes, S., Cowan, N., Hardman, K. O., & Logie, R. H. (2018). Informed guessing in change detection. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(7), 1023–1035.

Riggs, K. J., McTaggart, J., Simpson, A., & Freeman, R. P. (2006). Changes in the capacity of visual working memory in 5-to 10-year-olds. *Journal of Experimental Child Psychology*, *95*(1), 18–26.

Riggs, K. J., Simpson, A., & Potts, T. (2011). The development of visual short-term memory for multifeature items during middle childhood. *Journal of Experimental Child Psychology*, *108*(4), 802–809.

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development*, *74*(6), 1807–1822.

Rouder, J. N., Haaf, J. M., & Vandekerckhove, J. (2018). Bayesian inference for psychology, part IV: Parameter estimation and Bayes factors. *Psychonomic Bulletin & Review*, *25*(1), 102–113.

Sarigiannidis, I., Crickmore, G., & Astle, D. E. (2016). Developmental and individual differences in the precision of visuospatial memory. *Cognitive Development*, *39*, 1–12.

Schönbrodt, F. D., & Wagenmakers, E. J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, *25*(1), 128–142.

Simmering, V. R. (2012). The development of visual working memory capacity during early childhood. *Journal of Experimental Child Psychology*, *111*(4), 695–707.

Spencer, J. P. (2020). The development of working memory. *Current Directions in Psychological Science*, *29*(6), 545–553.

Su, Y.-S., & Yajima, M. (2015). R2jags: Using R to Run 'JAGS' [Computer software manual]. Retrieved from: https://CRAN.R-project.org/package=R2jags (R package version 0.5-7)

Tansley, B. W., Regan, D., & Suffield, J. B. (1982). Measurement of the sensitivities of information-processing channels for frequency change and for amplitude change by a titration method. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *36*(4), 723–730.

Treisman, A., & Zhang, W. (2006). Location and binding in visual working memory. *Memory & Cognition*, *34*(8), 1704–1719.

Tsal, Y., & Lavie, N. (1988). Attending to color and shape: The special role of location in selective visual processing. *Perception & Psychophysics*, *44*(1), 15–21.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 92.

Wagenmakers, E. J., Krypotos, A. M., Criss, A. H., & Iverson, G. (2012). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition*, *40*(2), 145–160.

Wagenmakers, E. J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. *Cognitive psychology*, *60*(3), 158–189.

Wang, B., Cao, X., Theeuwes, J., Olivers, C. N., & Wang, Z. (2017). Separate capacities for storing different features in visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(2), 226.

Wheeler, M. E., & Treisman, A. M. (2002). Binding in short-term visual memory. *Journal of Experimental Psychology: General*, *131*(1), 48–64.

Witherington, D. C., & Margett, T. E. (2011). How conceptually unified is the dynamic systems approach to the study of psychological development? *Child Development Perspectives*, *5*(4), 286–290.

Wrinch, D., & Jeffreys, H. (1921). XLII. On certain fundamental principles of scientific inquiry. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *42*(249), 369–390.

Yousif, S. R., Rosenberg, M. D., & Keil, F. C. (2021). Using space to remember: Short-term spatial structure spontaneously improves working memory. *Cognition*, *214*, 104748.

Zosh, J. M., & Feigenson, L. (2012). Memory load affects object individuation in 18-month-old infants. *Journal of Experimental Child Psychology*, *113*(3), 322–336.

Zosh, J. M., & Feigenson, L. (2015). Array heterogeneity prevents catastrophic forgetting in infants. *Cognition*, *136*, 365–380.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Forsberg, A., Adams, E. J., & Cowan, N. (2022). Why does visual working memory ability improve with age: More objects, more feature detail, or both? A registered report. *Developmental Science*, e13283. https://doi.org/10.1111/desc.13283