Morey, C. C., AuBuchon, A., Attwood, M., Castelain, T., Cowan, N., Crepaldi, D., Fjerdingstad, E., Fredriksen, E., Jarrold, C., Koch, C., Lelonkiewicz, J., Lupyan, G., Mendenhall, W., Moreau, D., Schonberg, C., Tamnes, C., Vlach, H., & Elliot, E. (2024). Is verbal rehearsal strategic? An investigation into overt rehearsal of nameable pictures in 5- to 10-year-old children. *Journal of Cognition and Development*. Advance online publication. https://doi.org/10.1080/15248372.2024.2389123

Journal of Cognition and Development

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hjcd20

Is Verbal Rehearsal Strategic? An Investigation into Overt Rehearsal of Nameable Pictures in 5- to 10-Year-Old Children

Candice C. Morey, Angela M. AuBuchon, Meg Attwood, Thomas Castelain, Nelson Cowan, Davide Crepaldi, Emilie Fjerdingstad, Eivor Fredriksen, Chris Jarrold, Chris Koch, Jaroslaw R. Lelonkiewicz, Gary Lupyan, Whitney Mendenhall, David Moreau, Christina Schonberg, Christian K. Tamnes, Haley Vlach & Emily M. Elliott

To cite this article: Candice C. Morey, Angela M. AuBuchon, Meg Attwood, Thomas Castelain, Nelson Cowan, Davide Crepaldi, Emilie Fjerdingstad, Eivor Fredriksen, Chris Jarrold, Chris Koch, Jaroslaw R. Lelonkiewicz, Gary Lupyan, Whitney Mendenhall, David Moreau, Christina Schonberg, Christian K. Tamnes, Haley Vlach & Emily M. Elliott (22 Aug 2024): Is Verbal Rehearsal Strategic? An Investigation into Overt Rehearsal of Nameable Pictures in 5- to 10-Year-Old Children, Journal of Cognition and Development, DOI: 10.1080/15248372.2024.2389123

To link to this article: https://doi.org/10.1080/15248372.2024.2389123

Published online: 22 Aug 2024.

JOURNAL OF COGNITION AND DEVELOPMENT
https://doi.org/10.1080/15248372.2024.2389123

Is Verbal Rehearsal Strategic? An Investigation into Overt Rehearsal of Nameable Pictures in 5- to 10-Year-Old Children

Candice C. Morey[a], Angela M. AuBuchon[b], Meg Attwood[c], Thomas Castelain[d], Nelson Cowan[e], Davide Crepaldi[f], Emilie Fjerdingstad[g], Eivor Fredriksen[g], Chris Jarrold[c], Chris Koch[h], Jaroslaw R. Lelonkiewicz[f,i], Gary Lupyan[j], Whitney Mendenhall[h], David Moreau[k], Christina Schonberg[j], Christian K. Tamnes[g], Haley Vlach[j], and Emily M. Elliott[l]

[a]Cardiff University, UK; [b]Boys Town National Research Hospital; [c]University of Bristol, UK; [d]Serra Húnter Fellow, University of Girona; [e]University of Missouri; [f]Scuola Internazionale Superiore di Studi Avanzati (SISSA); [g]University of Oslo; [h]George Fox University; [i]Instituto Tecnológico y de Estudios Superiores de Monterrey; [j]University of Wisconsin; [k]University of Auckland, New Zealand; [l]Louisiana State University

ABSTRACT

Though verbal rehearsal is a frequently endorsed strategy for remembering short lists among adults, there is ambiguity around when children deploy it, and what circumstantial factors encourage them to rehearse. We recoded data from a recent multilab replication of a serial picture memory task in which children were observed for evidence of task-related speech or lip movements to extract finer-grained detail about how children spoke during the task. With these data, we aimed to better understand the manner in which children rehearse and the task scenarios which elicit overt rehearsal. Children in several countries from 5 to 10 years old were tasked with remembering 2–5 nameable pictures in serial order across a 15-second delay. Coders categorized children's speech or lip movements as reflecting fixed rehearsal of the last-presented item only, cumulative rehearsal of all the items presented so far, or some attempt at cumulative rehearsal. We found that most children, regardless of age, did not overtly rehearse at all during presentation of the objects or during the delay period. However, children who sometimes overtly rehearsed recalled longer lists of items than children who did not. Though rare, cumulative rehearsal was most frequently observed for list lengths close to the participant's demonstrated maximum recall length. Critically, on the trials where overt rehearsal was observed, recall improved. This evidence supports previous suggestions that rehearsal strategy, and possibly also its effectiveness, changes with task difficulty, and raises further questions about how verbal rehearsal affects serial recall.

Ask someone how they remember a short list of words and they will likely tell you that they say it to themselves, inside their head, perhaps repeatedly (Dunlosky & Kane, 2007). The vividness of this experience of verbal rehearsal makes it natural for theorists of short-term or "working" memory to give repetitive verbal rehearsal a prominent role in explanations of

CONTACT Candice C. Morey moreyc@cardiff.ac.uk School of Psychology, Cardiff University, 70 Park Place, CF10 3AT

remembering. How does rehearsal feature in theories of working memory? Depending on who you ask, rehearsal may serve as a key bulwark against decay: without rehearsal or some other strategic intervention, information is believed to eventually decay beyond recall (Barrouillet & Camos, 2012), or to be displaced by more recently encountered information (Baddeley et al., 1984). The gradual improvement of memory across childhood has been explained by the onset of a capability to verbally rehearse (Gathercole, 1998): after this rehearsal function switches on, we presumably rehearse by default. However, although adults endorse the strategy, it becomes increasingly difficult with age to observe clear evidence that rehearsal directly improves serial recall (Lewandowsky & Oberauer, 2015; Oberauer, 2019). Our experience of engaging something like rehearsal to aid memory could be epiphenomenal: perhaps we use it to describe a feeling of greater engagement or effort, but articulation itself (whether overt or internalized) may not cause any observed benefit. If so, then it is reasonable to question how consistently rehearsal is applied. Once we acquire the ability to verbally rehearse, does it become ubiquitous as some theories infer or is it occurring at key moments when extra insurance against forgetting may be needed?

Accumulated research and observation suggest that children are increasingly likely to rehearse as they get older (e.g., Elliott et al., 2021; Flavell et al., 1966; Lehmann & Hasselhorn, 2012). Flavell et al. (1966) observed children while they attempted a serial picture memory task, and recorded whether they spoke or mouthed words that might reflect conversion of the pictures to speech. The rough trajectory emerging from Flavell et al.'s work suggested rehearsal emerged around age 7 and by age 10 it was visible in nearly all children. If we furthermore make the reasonable assumption that rehearsal may also occur covertly, and that use of inner speech also increases with age (Vygotsky, 1962), then we may suppose that by adolescence, rehearsal occurs regularly. Experimental work with children may support this view, in that benchmark working memory effects of phonological coding, phonological similarity and word length, seem to affect recall in children under 7 years old less than children over 7 (Halliday et al., 1990; Henry et al., 2012; Hitch et al., 1989; Palmer, 2000), which would be consistent with the presumption that covert speech occurred less frequently in children under 7 (although this inference has been convincingly challenged; Jarrold & Hall, 2013).

One important factor to note is that Flavell et al.'s (1966) procedure focused primarily on discovering which children used speech during a memory task, not necessarily how frequently or under which circumstances they used it. Multiple influential models of working memory rely on the assumption that verbal information that is not rehearsed in time to prevent its decay will be forever lost. Baddeley (2012) suggests that information in the phonological loop must be repeated to prevent it being overwritten by incoming information or to prevent its eventual decay. The TBRS model (Barrouillet & Camos, 2012) likewise supposes that information decays if some maintenance process, either rehearsal or refreshing, is not marshaled to intervene. In either formulation, we might conclude that neglecting to rehearse could lead to rapid forgetting. If children are developing the ability to rehearse around age 7, then lapses in application as they hone their ability could lead to erratic memory performance. While we know that children under 7 who *sometimes* overtly rehearse recall more on average than children who *never* rehearse (Elliott et al., 2021; Flavell et al., 1966), we do not know whether the advantage of rehearsing can be localized to the particular trials on which rehearsal happened, or alternatively whether children seen to rehearse sometimes differ from non-rehearsers in some other way.

However, rehearsal does seem to be a strategy that can be chosen if circumstances suggest it (Henry et al., 2000; Keeney et al., 1967). Keeney et al. (1967) observed a sample of 6- and 7-year-olds to determine, prior to any intervention, which of them spontaneously rehearsed during a serial picture memory task. After a two-trial intervention in which participants were instructed to name each picture repetitively aloud, children who initially did not rehearse spontaneously rehearsed consistently and their recall improved (see also Miller et al., 2015). However, when allowed by the researcher to choose whether they rehearsed or not, the children who needed the intervention typically ceased rehearsing and lost the boost rehearsal brought them. By showing that children under 7 who do not spontaneously rehearse actually *can* rehearse with prompting, these findings cast doubt on the hypothesis that rehearsal depends on the maturation of a specialized phonological working memory component, thought to occur around age 7 (Gathercole, 1998). Even children who have practiced rehearsal and benefited from it do not always choose to deploy it. This sets rehearsal up as an optional tactic, rather than as memory theorists often see it (Watkins & Peynircioĝlu, 1982): a process constantly running in the background of our minds.

Rehearsal is the presumed process most frequently deployed to aid retention of verbal lists because it is believed to require little effort once it is initiated. Naveh-Benjamin and Jonides (1984) proposed that rehearsal occurs in two stages: effort is required to initiate the rehearsal process, but once rehearsal is underway, it proceeds automatically. If this were true then one might surmise that it is generally worthwhile to rehearse, and thus reasonable for theories to assume that rehearsal takes place whenever possible. However, recent evidence showing that rehearsal incurs a concurrent processing cost as much as 10 seconds after initiation suggests that rehearsal *does* require sustained effort. If more rehearsal is required before the effort of rehearsing decreases, then the notion that we generally do it regardless of whether it is essential for remembering becomes more difficult to support.

Once we accept that rehearsal entails a larger start-up cost than previously assumed, more nuanced predictions for its use emerge. Maybe rehearsal occurs less frequently than supposed, and if so, maybe it is deployed under predictable circumstances. One possibility is that rehearsal only benefits recall when the amount of information to be rehearsed is within the capacity of the individual to perform (Jarrold & Hall, 2013). If this is correct, participants might rehearse at these crucial moments. It may not be worth engaging rehearsal for lists that are impossibly long. Likewise, if starting rehearsal is cognitively costly, why bother with the effort of rehearsing for lists that are trivially easy? Assuming variability within individuals in how they apply rehearsal and what form of rehearsal they apply (Poloczek et al., 2019) fits well with the view that change in approaches to memory tasks occurs gradually as children accumulate more strategies in their repertoire (Siegler, 2016), as opposed to the view that once a superior method emerges, it dominates. While this nuanced view of rehearsal sounds plausible, it poses problems for models of working memory that propose that applying rehearsal is essential for preserving information from decay.

To further address this, we need to know more about how children rehearse, including at which moments in a task they deploy overt rehearsal. Poloczek et al. (2019) showed that children between 6 and 10 were more likely to cumulatively rehearse lists of nameable pictures that were a little longer than they could easily recall; for less challenging lengths, participants preferred simply labeling each item (Jarrold & Poloczek, 2024). Poloczek et al. measured rehearsal strategy using per-trial self-reports. The Elliott et al. (2021) multilab replication project we recently completed affords the possibility to verify Poloczek et al.'s conclusions with a different measure, namely whether overt speech observed during a picture memory task seems to reflect adoption of verbal rehearsal. Elliott et al.'s project aimed to replicate the classic study of Flavell et al. (1966), in which children were shown a series of nameable pictures and an observer recorded whether they spoke during presentation, retention, or recall of the items. Participants were never asked to speak, even for the purpose of recalling, so use of speech was taken as an indicator that children verbally recoded the pictures, and possibly rehearsed them. However, replicating Flavell et al. only required recording whether and when any overt speech occurred. Raters did not attempt to assess whether the speech observed constituted cumulative rehearsal or something else, only whether the speech was likely to be relevant to the task. While the coding performed for the Flavell et al. replication cannot directly be used to diagnose rehearsal strategies, many of the sessions from this project were video-recorded, and these recordings afford the opportunity to apply a fresh coding scheme addressing rehearsal strategy more specifically. Also, the replication analyses focused on whether individuals were observed speaking at all, not whether instances of speech behavior impacted trial-level performance.

Advancing beyond the replication described by Elliott et al. (2021), here we describe the frequency of overt rehearsal in our sample, considering at which list lengths and during which parts of the task rehearsal was deployed. We then successively consider person-level effects of overt rehearsal on memory performance. For instance, do children sometimes observed rehearsing recall more than children never observed rehearsing? As observed in the more lenient coding of Elliott et al. (2021), we found that children who sometimes overtly rehearsed reached higher maximum recall lengths than children never observed rehearsing. With this stricter coding of what constitutes rehearsal, we were able to identify sufficient numbers of rehearsers and non-rehearsers at all age groups to show that overt rehearsers achieved better recall across each age group in our sample. But when did these children deploy rehearsal, and did it specifically boost performance for the trials when they did it? If rehearsal is a selectively deployed strategy that children apply, then the specific trials on which overt rehearsal occurred should show selective improvement.

Method

Ten research groups from Elliott et al. (2021) "s replication of the Flavell et al. (1966) study video-recorded a substantial subset of their sessions and agreed to contribute further observational coding to this project. We will summarize the key details of the original data collection for readers' convenience before describing the novel coding procedure we applied.

Participants

Data from 522 children included in the Elliott et al. (2021) replication analyses between 5- and 10-years old with complete video records of their session and whose guardians consented to recording were included in this project. Ethical permissions for collecting and preserving the data for long enough to carry out this coding were obtained at each local laboratory. Sample sizes per laboratory and age group are given in Table 1.

Table 1. Participant demographics by lab.

University	Group	N	% male	Age in Months	
				Mean Age	SD (Age)
Boys Town National Research Hospital	5yo	15	47	65	3.96
	6yo	13	54	78	3.49
	7yo	15	60	91	3.66
	10yo	16	50	125	3.48
Cardiff University	5yo	13	46	66	2.06
	6yo	11	45	79	3.27
	7yo	15	47	88	2.89
	10yo	18	56	124	3.52
George Fox University	5yo	6	67	69	1.93
	6yo	6	33	78	2.99
	7yo	3	0	89	1.41
	10yo	6	67	126	3.93
Louisiana State University	5yo	7	43	65	2.86
	6yo	3	67	78	4.50
	10yo	3	0	126	2.07
Scuola Internazionale Superiore di Studi Avanzati (SISSA)	5yo	8	50	66	3.80
	6yo	13	54	79	3.33
	7yo	22	41	89	3.72
	10yo	12	50	125	2.88
University of Auckland	5yo	18	56	64	3.31
	6yo	18	56	78	2.76
	7yo	20	50	91	2.41
	10yo	20	50	127	3.87
University of Costa Rica	5yo	14	43	65	3.42
	6yo	17	47	77	3.60
	7yo	18	50	89	3.26
	10yo	15	40	127	2.24
University of Missouri	5yo	17	53	65	3.36
	6yo	23	35	78	3.69
	7yo	17	41	88	3.27
	10yo	19	53	125	3.88
University of Oslo	5yo	14	50	64	3.52
	6yo	10	60	76	3.74
	7yo	15	47	90	3.64
	10yo	12	58	123	2.69
University of Wisconsin	5yo	7	71	65	2.97
	6yo	12	25	76	3.87
	7yo	19	58	89	3.43
	10yo	12	50	123	4.30

Note. All included participants met the inclusion criteria for the Elliott et al. replication project. In addition, their sessions were video-recorded with guardians' permission.

Stimuli and materials

Participants were given nameable pictures to remember in order. A set of seven pictures of common objects (apple, comb, flag, flower, moon, owl, and pencil) was chosen from the Bank of Standardized Stimuli (BOSS; Brodeur et al., 2010). Presentation of the stimuli and response collection were controlled with custom software written in lab.js (Henninger et al., 2021) using personal computers. The lab.js software made use of standard web browsers to ensure stimulus presentation was identical at each lab site regardless of local apparatus, but data collection was always local (as opposed to online).

Participants took part in both immediate and delayed recall tasks, with task order counterbalanced. Within both tasks, 2 trials each at span lengths 2, 3, 4, and 5 were administered to all participants, regardless of whether recall of shorter lists was accurate.

During the 15-second retention interval of the delayed recall task, participants were instructed to cover their eyes with a fun prop so that a researcher could observe and record their speech behavior without awkwardness. In the original Flavell et al. (1966) study, the prop was a “space helmet” with an obstructed visor that could be lowered or lifted. In our replication, labs variably used helmets with adjustable visors, sunglasses, safety goggles, or masks, each with the visor or lenses obstructed using tape or paper. Because the 15-second retention interval affords a long opportunity for overt rehearsal, we focus our analyses only on data from the delayed recall task. Observations of rehearsal during the presentation period of the immediate recall task were comparable to those presented in our analysis of the delayed recall task below.

Experimental procedure

After consent from the guardian and assent from the child were obtained, participants began by completing two practice trials without a 15-second delay. Participants were then introduced to the prop they were to use to obscure their eyes during the delay periods. Participants practiced this for periods of 5, 10, and then 15 seconds so that they would get used to wearing the prop for the full length of the delay period in the delayed recall trials.

In each experimental trial, the set of pictures was displayed onscreen in a central, horizontal row, with order varying randomly per trial and run. The first to-be-remembered picture was highlighted in a red frame for 2 seconds, after which the frame disappeared and the second to-be-remembered picture was highlighted for 2 seconds, and so on for longer lists. In the delayed recall procedure, a 15-second delay was imposed between offset of the last to-be-remembered picture and the appearance of the response screen. At the start of the delay, participants were instructed to cover their eyes with their prop. At the end of the 15-second period, participants were instructed to remove the prop. Instructions appeared onscreen and the experimenter read them aloud. On the recall screen, the seven pictures were arranged in a randomly ordered horizontal row. Participants were instructed to indicate the pictures they saw in order. The experimenter operated the mouse for the participant, moving it to the picture indicated with the child’s response. Participants could recall by pointing or by speaking the names of the pictures, but they were not explicitly told to speak.

Behavioral coding procedures

We developed a coding protocol for categorizing what kind of rehearsal (if any) the child’s speech or lip movements represented during the presentation period and delay periods of the memory task. Coders were instructed to note the words spoken by the child by typing them in order in the specified cell of the coding template, separated by commas. Coders also made a judgment about whether the speech reflected fixed rehearsal (labeling each picture as it was presented without any repetitions), cumulative rehearsal (reiterating earlier list items as each new item was presented, appending the latest item to the end of the accumulating list), or partial cumulative rehearsal (more speech than fixed rehearsal, but short of the complete cumulative pattern). When no speech or discernible lip movements were detected, coders indicated that no overt rehearsal was detected. Coded data are available at https://osf.io/m8ks4/. Coders all had access to lip-reading training developed specifically for these picture stimuli by colleagues at the Boys Town National Research Hospital (https://osf.io/36ayh/).

We also considered participants' non-speech behaviors, particularly behaviors that may indicate distraction. We adapted five descriptors from the Restricted Academic Situations Task (RAST) (Milich et al., 1982) to refer more specifically to this task setting. For each of five descriptions (off-task behavior, fidgeting, vocalizing, getting up from their seat, playing with objects), coders entered 1 if the behavior was observed during the coded period and 0 otherwise. For both the rehearsal category and the RAST variables, coders entered a value for each presentation and delay trial period.

At least two coders worked with videos from each lab group. Procedures for dual-coding differed per lab group depending on availability of suitable personnel and whether sharing video materials externally was permitted. For some groups, dual coding was undertaken for only a small subset of the sample, and if inter-rater agreement was acceptable, the initial coder's ratings were accepted. For other groups, complete dual-coding was undertaken with a third rater introduced to assess disagreements. For rehearsal category, agreement ranged from 83.6% to 100% ($M = 93.86\%$, $SD = 6.37$). For the RAST variables, agreement was quite variable and sometimes poor (8.2–100%; $M = 77.51\%$, $SD = 21.51$). Where agreement was low, we approached this by outputting two versions of the coding. In the "strict" version, occurrence of a behavior was only assumed when two raters both observed it. In the "lenient" version, occurrence was assumed if either rater reported it. Any analyses reported involving the RAST variables were performed on both versions of the coding to check for robustness. We decided to omit analyses of the RAST variables from this report because interpretation given the agreement levels would be problematic, but these exploratory analyses are available at our OSF page, https://osf.io/m8ks4/. Generally, we did not observe many distracted behaviors from participants on average, and incidences of distraction decreased with age.

Results

Our inferential analyses are carried out using Bayesian methods, using the *BayesFactor* R package (Morey et al., 2018; Rouder et al., 2012). Bayesian methods afford several crucial advantages. First, in a data set this large, significant results are likely in a classical test simply because the amount of data affords the detection of tiny effects. The Bayes factor statistics accompanying our analyses tell us not *if* a pattern was discernible, but how strongly that pattern is discernible. In other words, whereas a *p*-value must be interpreted as reflecting presence of an effect or uncertainty, the Bayes factor tells us about the degree of certainty we might express in favor of a finding, which is a great advantage, especially when working with a large data set. Second, Bayes factors may be calculated in favor of null as well as alternative hypotheses, so the advantage of expressing degree of certainty in an outcome holds regardless of whether the data suggest there is or is not an effect. Bayesian ANOVA (the basis of most of our reported analyses) can be considered comparable to model selection techniques. In performing Bayesian ANOVAs, we look at each possible combination of main effects and interactions and select the combination that yields the highest Bayes factor (or the strongest evidence) as a baseline. We then consider the strength of evidence for including each factor in that model by comparing the best model with the combination that excludes the factor under consideration. Likewise, we can assess the strength of evidence for excluding a factor by comparing the best model with the version that includes the factor missing from the best model.

Who overtly rehearsed, and did it help?

We limited analyses of rehearsal style to the delayed recall task, where we could consider rehearsal during both presentation of the stimuli and during the 15-second delay period. Unlike in the Flavell et al. (1966) replication (Elliott et al., 2021), we did not consider spoken recall alone to count as evidence for rehearsal. While naming the items at recall indicates that participants have converted the presented images to speech at some point, we cannot tell from speech at recall alone when during the trial this conversion took place, nor could we classify a recalled response as reflecting a particular type of rehearsal.

Raters in this study, who were tasked with judging whether any observed speech corresponded to a type of rehearsal, were much less liberal in endorsing that rehearsal took place than the coders in the replication study were in endorsing that speech was observed. With the more detailed definitions of what counts as rehearsal, we saw many more instances of children in all age groups qualifying as "never" rehearsing. There are a few other factors that may have contributed to the lower rates of rehearsers we observed here than in the Elliott et al. replication analyses. Speech during the recall period counted in the Flavell et al. protocol and in the Elliott et al. replication for assigning children as speakers; here we have excluded the recall period completely. Also, even weak evidence of speech (e.g., indecipherable mumbling or unclear but repetitive lip movements that replication coders rated "2," which essentially meant "maybe") counted fully as incidences of speech, whereas in this analysis, coders would have required more certainty to assign incidences of speech to a particular rehearsal category. With these stricter definitions of what kind of speech constitutes rehearsal, we observed much less rehearsal than speech. Figure 1 shows the proportion of participants per age group who were ever observed rehearsing during the presentation and delay periods. The upper panels show proportions scored with the lenient scoring of Elliott et al. as applied during the presentation and delay periods, and the lower panels show the stricter coding implemented in this project. With the stricter scoring, the majority of participants at all age groups did not rehearse aloud. We will describe the rehearsal we observed using these stricter criteria further, but readers should keep in mind that overt behaviors that could be classified as reflecting a particular sort of rehearsal were far from ubiquitous.

Figure 2 shows the average maximum span length in which all items were recalled correctly in order achieved per age group for children who sometimes versus never rehearsed. Accuracy was scored strictly with respect to serial order: an item only counted as correct if it was recalled in the exact serial position in which it was presented. Table 2 reports the average number of items participants correctly recalled in order, per list length, to provide a more detailed account of overall performance. We analyzed maximum achieved recall with age group and rehearser status as possible factors. Models including all possible combinations of factors and interactions were compared, and the factors present in the option with the most support are considered as significantly influencing maximum number of items recalled. The best model included main effects of age group and rehearser status, $BF = 7.6 \cdot 10^{55}$. Excluding the interaction between the two factors was favored by $BF_{01} = 36.18$. Including both main effects was

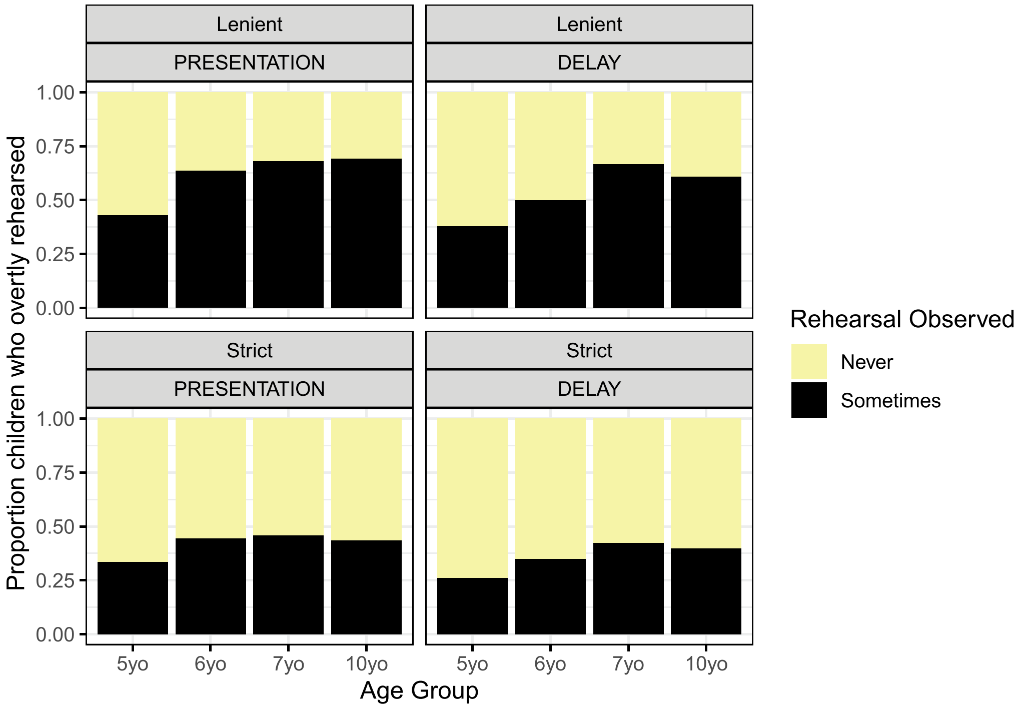

Figure 1. Proportions of children per age group who sometimes rehearsed. Upper panels use lenient scoring from Elliott et al. (2021), except that speech during the recall period was not considered to be rehearsal. Lower panels use the novel coding reported here, in which we strived to discern what sort of rehearsal children applied. Sample includes 119 5-year-olds, 126 6-year-olds, 144 7-year-olds, and 133 10-year-olds.

decisively favored, with BF_{10} of at least 105,000. Rehearsers' performance was better on average than non-rehearsers, for all age groups.

Even though our observational indicator of whether rehearsal took place is limited – of course participants may have fully internalized rehearsal, which our coding would not have detected – observed overt rehearsal seemed to pick up some signal that predicted how well children remember. At all ages, children seen to rehearse at least once achieved higher maximum span scores on average than children who were never observed rehearsing. Going further, an additional question we might ask is whether this overt behavior specifically affected performance on the trial in which it was observed, or whether it indicates a more general trait of the participant. Might the overt rehearsal we observe reflect processes also going on covertly, or did these instances of overt rehearsal reflect extra effort that paid off? Alternatively, some unmeasured trait (e.g., intelligence) may influence both engagement with rehearsal strategy and memory recall, rather than rehearsal directly impacting recall. If rehearsal is boosting recall, we should observe this advantage on trials where rehearsal is observed, and not necessarily more broadly.

When and how did children rehearse?

We carried out further analyses using the participants who sometimes rehearsed ($N = 249$), aimed at learning more about what exactly they did. Figure 3 shows

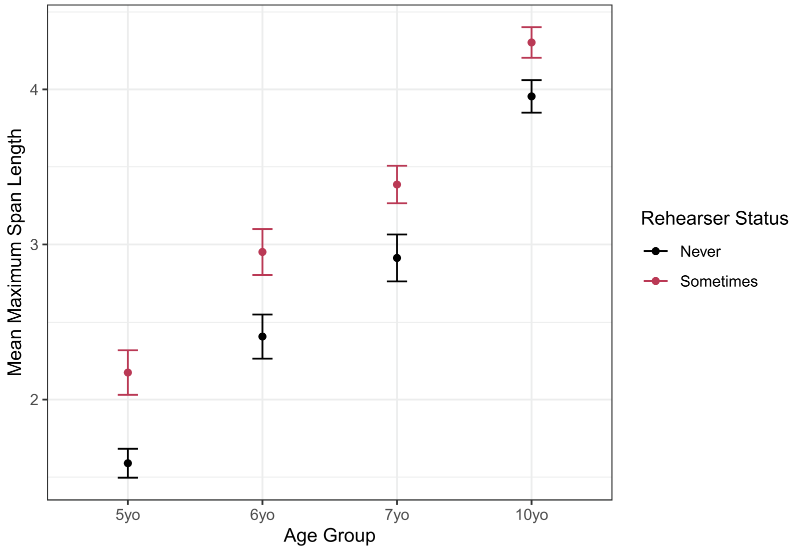

Figure 2. Average maximum span length with all items recalled correctly in order per age group, with standard errors of the mean.

Table 2. Average number reported correctly in order by rehearser status, age group, and list length.

		Never Rehearsed		Sometimes Rehearsed	
Group	Length	Mean	SD	Mean	SD
5yo	2	0.62	0.58	0.93	0.72
	3	0.83	0.79	1.14	0.90
	4	0.62	0.57	0.91	0.80
	5	0.58	0.59	0.80	0.73
6yo	2	0.96	0.73	1.33	0.67
	3	1.24	1.06	1.81	1.01
	4	1.31	0.98	1.64	1.00
	5	1.02	0.91	1.28	0.92
7yo	2	1.09	0.71	1.40	0.66
	3	1.70	1.03	2.11	0.98
	4	1.78	1.12	2.18	1.10
	5	1.46	1.15	1.35	1.04
10yo	2	1.55	0.57	1.75	0.38
	3	2.49	0.74	2.71	0.55
	4	2.97	1.00	3.01	0.95
	5	2.32	1.34	2.92	1.37

Note. In the Group column, 'yo' stands for 'years old'.

proportions of incidences of each type of rehearsal we coded for the participants who sometimes rehearsed, plotted per list length relative to the list length of their maximum recall. Zero in this plot is, per participant, the list length corresponding to their maximum span. Negative values are list lengths less than the individual's maximum recall, and positive values more than their maximum. This plot shows that even among children who sometimes rehearsed, there were many instances of trials with no overt sign of rehearsal. During presentation, there was no evidence of rehearsal on 50–60% of trials. During the 15-second delay, there was even less

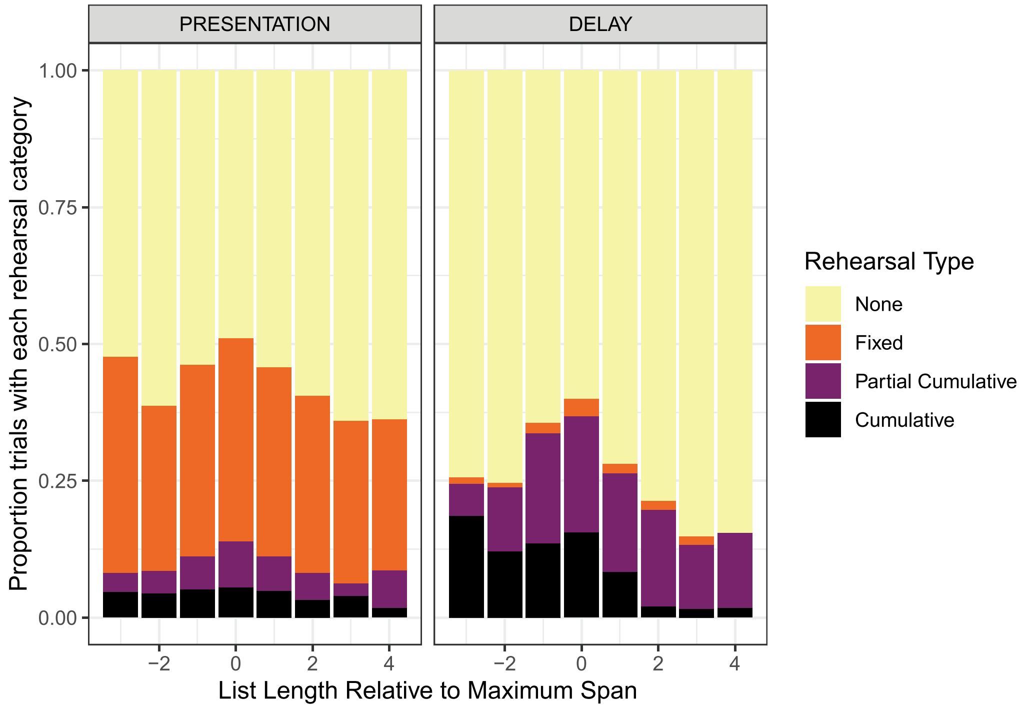

Figure 3. Proportions of rehearsal type observed per trial period, among children who sometimes rehearsed.

evidence of rehearsal, but it shifted from the mostly fixed rehearsal observed during presentation to some attempt at cumulative rehearsal. However, instances of rehearsal during the delay period do not appear to be distributed equally. We computed Bayes Factors on 2-way contingency tables categorizing relative span length and incidences of any versus no rehearsal separately for presentation and delay periods. Each individual trial counted as an independent observation. The underlying null assumption would be that observations of rehearsal would be independent, and so proportionally comparable, regardless of relative span length. Evidence favored the null assumption of independence during presentation (BF_{01} for the null > 27), consistent with the idea that rehearsal during presentation was equally likely regardless of list length. In contrast, during delay, evidence favored non-independence, BF_{10} >25 million. Rehearsal observations increased during the delay for list lengths approaching participants' achieved maximum span. This is consistent with the suggestion that overt rehearsal, particularly attempts at cumulative rehearsal which were more likely during the 15-second retention period, were more likely when participants had capacity for it, right at their limit. Rehearsal during the retention period was less likely to occur when the task was too easy or difficult.

Did overt rehearsal boost performance on individual trials, as well as maximums achieved?

Children observed overtly rehearsing on some trials may also have been covertly rehearsing on others. Perhaps they were usually rehearsing, but only on some trials were observed to be rehearsing overtly. If so, then their performance on the trials where we observed their rehearsal may not differ much from trials on which we did not observe rehearsal. However, if their overt rehearsal reflected an unusual effort, performance should have benefited from that exertion specifically on these trials. To learn whether overt rehearsal specifically assisted the trials on which children were observed rehearsing, we carried out a Bayesian ANOVA on mean number correct per trial for span levels (either at, sub, or supra their individual span) and age group, contrasted by whether overt rehearsal was observed. Because the number of sub- and supra-span levels with representative amounts of data differed per age age group (i.e., 5-year-olds' typical maximum span was much lower than 10-year-olds,' so they would have few observations sub-span, but many observations supra-span, and vice versa), we averaged all span lengths beneath and above each individual's span into single values. Only participants who were observed rehearsing at least sometimes were

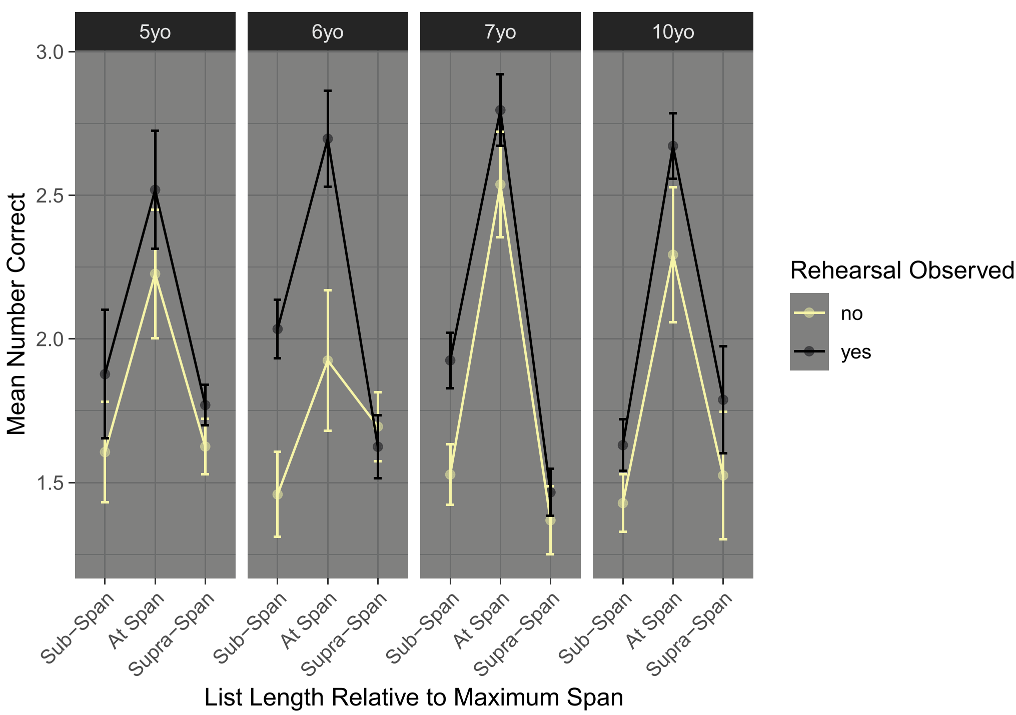

Figure 4. Average number correct per trial, in order, per span level and age group for children who overtly rehearsed, with within-participant standard errors of the mean (Morey, 2008).

included. Figure 4 depicts these values, with age groups presented in separate panels. The ANOVA confirmed that individual participants recalled more during the trials when they rehearsed overtly, and this interacted with span length relative to maximum span. The best model included main effects of age group, relative span length, and whether the participant was observed rehearsing on the trial, plus interactions between relative span length and observed rehearsal and relative span level and age group, $BF = 2.1 \cdot 10^{83}$. Including the interactions was favored by a factor of only $BF_{10} = 4.49$, and the next-best models were identical to the favored model except that they left out one interaction.[1] This suggests modest evidence that something about the relationships between observed rehearsal, span length, and age is fluctuating, but even with this much data it is difficult to pinpoint precisely what. However, evidence that overt rehearsal impacts the amount remembered on a trial-by-trial basis is decisively favored; inclusion of this factor was preferred $BF_{10} >$ 11 million. Inclusion of age group and span level were favored by even larger margins. Excluding an interaction between age group and observed rehearsal (which might mean that rehearsal only matters at some ages, or differentially matters per age group) was favored by a factor of $BF_{01} = 29.79$. We would advise some caution interpreting the absence of this interaction however, given that there are only two trials per participant and span length to work with.

Finally, we also considered how overt rehearsal might impact recall of items, regardless of whether they were given in the original correct order. We carried out the same Bayesian ANOVA described above, except with mean number of correct items reported regardless of order. Figure 5 depicts these values. The ANOVA confirmed that individual participants recalled more during the trials when they rehearsed overtly but unlike with ordered scoring, we no longer see evidence that rehearsal benefits performance most "at-span." The best model included main effects of age group, relative span length, and whether the participant was observed rehearsing on the trial, plus an interaction between relative span length and age group, $BF = 5.9 \cdot 10^{115}$. Including the interaction was barely favored, $BF_{10} = 2.95$. Excluding the interaction between relative span and observed rehearsal was likewise only slightly favored, $BF_{01} = 0.87$, but note from examination of Figure 5 that interpretation of this interaction would necessarily differ between this analysis and that of order-based scores: there is not a numerically greater boost for scores at-span here compared to scores below span. Inclusion of the main effects were all favored overwhelmingly, $BF_{10} \geq 5$ million. These patterns suggest that overt rehearsal consistently improved recall of items (when scored without regard to order), regardless of relative span length.

Because rehearsal was not observed most of the time, and because when it was, participants typically engaged in fixed rehearsal during stimulus presentation and some attempt at cumulative rehearsal during the delay period, we did not consider that there were sufficient trials uniquely representing each rehearsal type to test whether rehearsal types impacted recall differently.

Discussion

Our coding of video-recorded sessions from Elliott et al. (2021)'s many-labs replication of Flavell et al. (1966) yielded new observations that reinforce existing evidence and intuitions about the potential benefits of rehearsal for recall. During a delayed recall task, most

[1]When rehearsal is leniently considered using the scoring from Elliott et al., including an interaction between relative span and rehearsal per trial is favored $BF_{10} > 690$. It is still best to consider this effect as consistent across age groups, BF_{01} for excluding the relevant interaction with age group = 21.89.

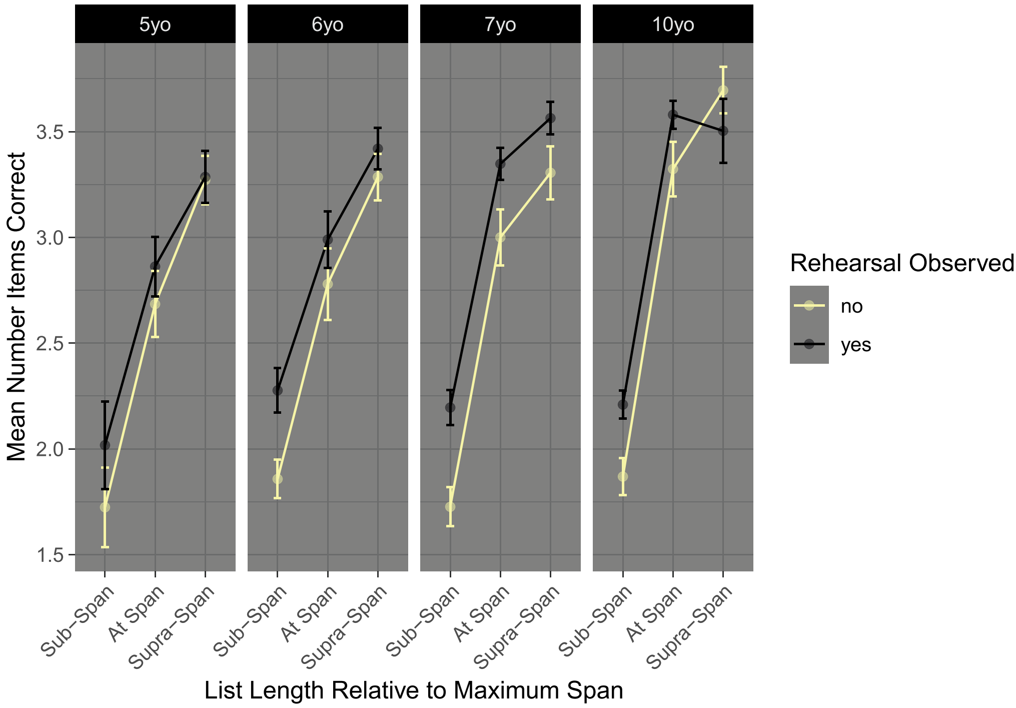

Figure 5. Average number correct per trial regardless of order per span level and age group for children who overtly rehearsed, with within-participant standard errors of the mean (Morey, 2008).

children 5–10 years old did not overtly rehearse in any observable way. However, the children observed rehearsing achieved higher maximum spans than the children who never overtly rehearsed. When they rehearsed, children usually adopted a fixed pattern during presentation of the memory items in which they labeled the items as they were presented, whereas during the 15-second delay period attempts at cumulative rehearsal increased. However, even children who rehearsed did not overtly rehearse most of the time. Cumulative rehearsal attempts were most likely on lists within one item of the child's demonstrated maximum recall length. Overt rehearsal also did not seem to be merely an indicator that the child usually covertly rehearsed, because performance was better on trials where overt rehearsal was observed. These observations are consistent with self-reported strategies described by Poloczek et al. (2019). Children in their sample usually reported deploying a variety of strategies, but cumulative rehearsal was most likely for challenging list lengths. Altogether, this evidence is consistent with the idea that rehearsal is applied selectively and strategically, perhaps when it is most likely to make an impact. Researchers aiming to investigate rehearsal should take note of this finding and ensure that they have designed stimuli likely to induce participants to adopt rehearsal in scenarios where it might have impact, such as titrating span lengths based on demonstrated ability (Doherty et al., 2019; Rhodes et al., 2019).

These analyses further bolster the view expressed by Jarrold and Hall (2013) that the success of rehearsal depends on whether the amount of to-be-remembered information is within the bounds of what the individual is capable of achieving. Jarrold and Hall emphasized the challenging end of this continuum; we would add that cumulative rehearsal may also be neglected when the amount to remember seems trivial, possibly because they prefer not to expend the additional effort that rehearsal entails (Thalmann et al., 2019). Application of overt rehearsal under each of these scenarios is described in Figure 6. The possibility that rehearsal is neglected when the task seems easy challenges theories of working memory that posit that rehearsal (or something comparable) must be performed to prevent memory loss. One caveat is that we cannot be certain that children did not rehearse simply because they did not rehearse aloud. Our observational method cannot access covert rehearsal attempts. However, the apparent decline in attempts at overt cumulative rehearsal for lists below span suggest the possibility that children are not expending as much effort to remember at list lengths they do not perceive as challenging. Furthermore, the improved performance on trials where overt rehearsal was observed suggests that rehearsing aloud indeed reflects some extra exertion. Whether that exertion is "extra" compared with doing nothing or compared to covert rehearsal, or whether it might signal the addition of another system to the memorization process (Barrouillet et al., 2021) we cannot be sure.

One piece of evidence that strongly suggests that we are usually comparing overt rehearsal to doing nothing rather than to internalized rehearsal is that the advantage of overt rehearsal was statistically similar across age groups: this would be unexpected under the assumption that 5-year-olds generally do not covertly rehearse whereas 10 year-olds generally do, but our analyses were inconsistent with the contention that 10 year-olds benefited less from overt rehearsal than other age groups. Moreover, there is good prior reason to believe that neither overt nor covert rehearsal, at least as typically modeled, can explain serial recall patterns alone (Lewandowsky & Oberauer, 2015). We therefore already need another means of explaining how lists are retained, and we already know that rehearsal as a counteraction of decay cannot universally work as

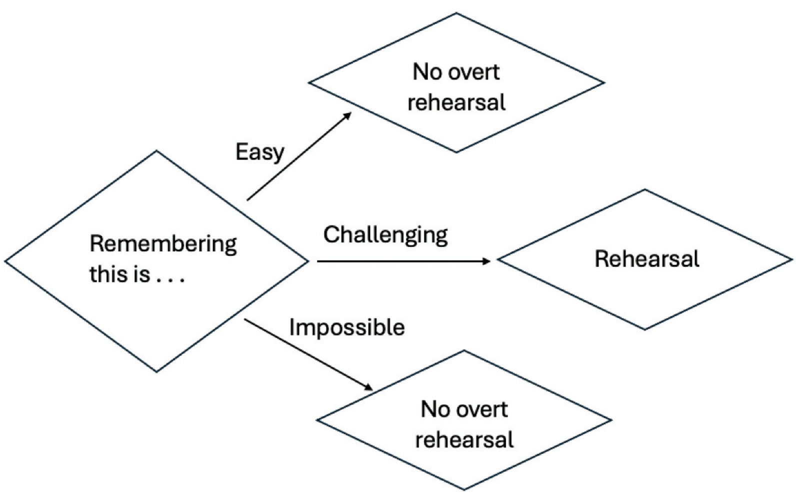

Figure 6. Flow chart describing circumstances under which children are likely to engage overt rehearsal, with easy and very difficult lists failing to elicit effort needed for overt rehearsal. Note that this process is description only. Further research and modelling could better flesh out why overt rehearsal is attempted at particular levels of difficulty, as well as what other factors affect adoption of rehearsal and how rehearsal affects mnemonic representations.

a general explanation for remembering. Possibly, rehearsal only needs to be invoked in much more limited circumstances; for short lists, perhaps primary memory is sufficient for successful maintenance as long as interference is minimal (Jarrold et al., 2010). Alternatively, overt rehearsal may provide additional benefits above and beyond covert rehearsal (e.g., re-perceiving the items), or may signal the application of some co-occurring process (e.g., increased task focus) that instead accounts for the trial-level improvements in performance we observed. It is also plausible that overt rehearsal may additionally distinguish children with certain traits that also help recall (e.g., the brighter children may be the ones more likely to adopt these effortful strategies, as well as being more likely to remember more). We do not have independent measures of intellectual or personality traits that might enable us to say how much rehearsal strategy adoption depends on these factors. However, we think that the observation that overt rehearsal benefits the specific trials on which that rehearsal occurred suggests that the benefit occurs at least in part because of the rehearsal action, rather than because of other traits of the individual performing it.

Memoranda were presented at a rate of one item every 2 seconds, as in Flavell et al. (1966). This pace is probably not ideal for encouraging spontaneous rehearsal during stimulus presentation, because it leaves little time for participants to do more than verbalize the most recent item. Children's lexical access speed varies within the ranges of ages we studied, and it is likely that many children in our sample, especially the younger children, needed more time to initiate verbalization of each new item (AuBuchon et al., 2022). Studying adults, Tan and Ward (2008) found that even with 5-second per item presentation rates, participants could only cumulatively rehearse the first few list items. This finding tends to support the suggestion that rehearsal cannot generally proceed in the idealized way it is described, particularly when lists exceed a comfortable length. Our participants were also observed during a 15-second post-list retention interval, which would give ample time to attempt further rehearsal of the items. However, if ongoing rehearsal during list presentation is crucial for eventual recall accuracy, children may have had difficulty remembering the items and their order to initiate rehearsal after list presentation. Future investigations of overt rehearsal in children should consider using slower paces of presentation to ensure that children would have enough time to cumulatively rehearse. This investigation arose from a multilab replication initiative (Elliott et al., 2021), and was naturally limited by the focus of that research. Nonetheless, undertaking the confirmatory multilab replication study afforded the possibility to conduct the exploratory analyses reported here on a very large, diverse sample, showing that these large-scale replications can add value beyond their primary aims (cf. AuBuchon et al., 2022).

In our large, diverse sample of children from several countries, 10-year-olds did not perform much better than 7-year-olds in terms of average number of items recalled in correct order per list, though their average maximum recall lengths were considerably higher. We suspect that though 10-year-olds were quite far from ceiling in our delayed recall task, this contrast between their maximum performance (more than 4 out of a maximum span of 5 on average) and their average number correct per list likely reflects variability of effort. If rehearsal is effective, but effortful and applied strategically, then 10-year-olds might have neglected to engage it for short lists, perhaps because they were overconfident about their ability to remember short lists. As list length increased, they apparently deployed rehearsal to achieve a high maximum score. Ten-year-olds would also be more likely to covertly rehearse, which would mean that our sample of overt rehearsers may have missed out many 10-year-old participants who might have been covertly rehearsing at least as much. However, note that on average the 10-year-olds who sometimes overtly rehearsed out-performed the 10-year-olds who never overtly rehearsed at each list length (Refer to Table 2).

Regardless of age, children who sometimes overtly rehearsed performed better than children who never overtly rehearsed. According to our inferential analyses, the magnitude of the boost that overtly rehearsing gave to maximum performance was comparable in all ages. This is inconsistent with the idea that verbal rehearsal is unavailable to children younger than 7 or so, as previously surmised (Gathercole, 1998). Possibly, the age of rehearsal acquisition is younger than previously supposed, and a younger age must be nominated for when rehearsal capability emerges. The absence of an interaction between age and rehearsal on performance in our data should be treated with caution: although we have a large sample, there were few trials per participant per list length, which will make individual values volatile. But also, our focus on observable overt rehearsal behavior naturally means that any fluent covert rehearsers were classed as “never” rehearsing, which one would expect would *decrease* evident differences between rehearsers and non-rehearsers. That we observed a measurable benefit while overtly rehearsing across the sample despite these limitations at the least confirms that the rehearsal strategy is available to children as young as 5 years old.

Taking this further, it is also reasonable to assume that the strategy of using speech to preserve information is available quite early, but is more consistently chosen and deployed as children mature. As shown in our analyses and those of Poloczek et al. (2019), children throughout the ranges of 5–10 years old report using a variety of strategies in picture memory tasks, including doing nothing at all, which may have been likely on trials in our study where overt rehearsal was not observed. This is consistent with the possibility that effortful strategies such as cumulative rehearsal are not always deployed, even by older children who adopt them more frequently and remember more. If we assume rehearsal requires effort, it stands to reason that one may decline to do it if it seems unnecessary or futile. Overt rehearsal could be seen as a specific application of proactive attentional control (Braver, 2012; Chevalier, 2015), in which a response is prepared before it is prompted. We would expect older children over 7 years old to use any proactive strategy more spontaneously and frequently, but would also expect that children under 7 could engage a proactive strategy when encouraged to do so (Chevalier et al., 2015). Verbal rehearsal may not require a special status, but may be simply considered one of a range of strategies that children can adopt, with older children more likely to spontaneously adopt it than younger children. Crucially though, when younger children do adopt a proactive strategy in an appropriate context, such as verbally rehearsing object names in a picture memory task, their performance improves.

Overall, our analysis suggests that children do not always rehearse, but when they overtly rehearse they are doing something that benefits serial memory. They are more likely to engage in rehearsal under circumstances when something extra is needed to boost their memory. The evidence we present here is consistent with the position that rehearsal is an effective means of preserving information, but not with the assumption that rehearsal is constantly occurring in the background with minimal effort. As such, these findings bolster existing evidence about strategic remembering during childhood, and also present challenges to prominent theoretical assumptions about how memory functions.

Disclosure statement

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

Data availability statement

Coded data and analysis scripts are available at https://osf.io/m8ks4/. The manuscript was generated using the papaja R package (Aust & Barth, 2023). The original replication project data can be found at https://osf.io/pn4rk/.

Open Scholarship Statement

This article has earned the Center for Open Science badge for Open Data. The data are openly accessible at https://osf.io/4b92a/

References

AuBuchon, A. M., Elliott, E. M., Morey, C. C., Jarrold, C., Cowan, N., Adams, E. J., & Voracek, M. (2022). Lexical access speed and the development of phonological recoding during immediate serial recall. *Journal of Cognition and Development*, *23*(5), 624–643. https://doi.org/10.1080/15248372.2022.2083140

Aust, F., & Barth, M. (2023). *Papaja: Prepare reproducible APA journal articles with R markdown* [manual]. https://github.com/crsh/papaja

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

Baddeley, A. D., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology A, Human Experimental Psychology*, *36*(2), 233–252. http://search.ebscohost.com/login.aspx?direct=truedb=psyhAN=1985-11343-001loginpage=Login.aspsite=ehost-livescope=site

Barrouillet, P., & Camos, V. (2012). As time goes by: Temporal constraints in working memory. *Current Directions in Psychological Science*, *21*(6), 413–419. https://doi.org/10.1177/0963721412459513

Barrouillet, P., Gorin, S., & Camos, V. (2021). Simple spans underestimate verbal working memory capacity. *Journal of Experimental Psychology General*, *150*(4), 633–665. https://doi.org/10.1037/xge0000957

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, *16*(2), 106–113. https://doi.org/10.1016/j.tics.2011.12.010

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank Of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLOS ONE*, *5*(5), e10773. https://doi.org/10.1371/journal.pone.0010773

Chevalier, N. (2015). The development of executive function: Toward more optimal coordination of control with age. *Child Development Perspectives*, *9*(4), 239–244. https://doi.org/10.1111/cdep.12138

Chevalier, N., Martis, S. B., Curran, T., & Munakata, Y. (2015). Metacognitive processes in executive control development: The case of reactive and proactive control. *Journal of Cognitive Neuroscience*, *27*(6), 1125–1136. https://doi.org/10.1162/jocn_a_00782

Doherty, J. M., Belletier, C., Rhodes, S., Jaroslawska, A., Barrouillet, P., Camos, V., & Logie, R. H. (2019). Dual-task costs in working memory: An adversarial collaboration. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *45*(9), 1529–1551. https://doi.org/10.1037/xlm0000668

Dunlosky, J., & Kane, M. J. (2007). The contributions of strategy use to working memory span: A comparison of strategy assessment methods. *The Quarterly Journal of Experimental Psychology*, *60*(9), 1227–1245. https://doi.org/10.1080/17470210600926075

Elliott, E. M., Morey, C. C., AuBuchon, A. M., Cowan, N., Jarrold, C., Adams, E. J., & Voracek, M. (2021). Multilab direct replication of flavell, beach, and Chinsky (1966): Spontaneous verbal rehearsal in a memory task as a function of age. *Advances in Methods and Practices in Psychological Science*, *4*(2), 25152459211018187. https://doi.org/10.1177/25152459211018187

Flavell, J. H., Beach, D. R., & Chinsky, J. M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development*, *37*(2), 283–299. https://doi.org/10.2307/1126804

Gathercole, S. E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *39*(1), 3–27. https://doi.org/10.1111/1469-7610.00301

Halliday, M. S., Hitch, G. J., Lennon, B., & Pettipher, C. (1990). Verbal short-term memory in children: The role of the articulatory loop. *Journal of Cognitive Psychology*, *2*(1), 23–38. https://doi.org/10.1080/09541449008406195

Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2021). Lab.js: A free, open, online study builder. *Behavior Research Methods*, *54*(2), 556–573. https://doi.org/10.3758/s13428-019-01283-5

Henry, L. A., Messer, D., Luger-Klein, S., & Crane, L. (2012). Phonological, visual, and semantic coding strategies and children's short-term picture memory span. *The Quarterly Journal of Experimental Psychology*, *65*(10), 2033–2053. https://doi.org/10.1080/17470218.2012.672997

Henry, L. A., Turner, J. E., Smith, P. T., & Leather, C. (2000). Modality effects and the development of the word length effect in children. *Memory*, *8*(1), 1–17. https://doi.org/10.1080/096582100387678

Hitch, G., Halliday, M., Dodd, A., & Littler, J. (1989). Development of rehearsal in short-term-memory - Differences between pictorial and spoken stimuli. *British Journal of Developmental Psychology*, *7*(4), 347–362. https://doi.org/10.1111/j.2044-835X.1989.tb00811.x

Jarrold, C., & Hall, D. (2013). The development of rehearsal in verbal short-term memory. *Child Development Perspectives*, *7*(3), 182–186. https://doi.org/10.1111/cdep.12034

Jarrold, C., & Poloczek, S. (2024). *The development of the adaptive use of different forms of rehearsal in verbal serial recall tasks. A multi-method study*. OSF. https://doi.org/10.31234/osf.io/f6emh

Jarrold, C., Tam, H., Baddeley, A. D., & Harvey, C. E. (2010). The nature and position of processing determines why forgetting occurs in working memory tasks. *Psychonomic Bulletin & Review*, *17*(6), 772–777. https://doi.org/10.3758/PBR.17.6.772

Keeney, T., Cannizzo, S., & Flavell, J. (1967). Spontaneous and induced verbal rehearsal in a recall task. *Child Development*, *38*(4), 953–953 &. https://doi.org/10.2307/1127095

Lehmann, M., & Hasselhorn, M. (2012). Rehearsal dynamics in elementary school children. *Journal of Experimental Child Psychology*, *111*(3), 552–560. https://doi.org/10.1016/j.jecp.2011.10.013

Lewandowsky, S., & Oberauer, K. (2015). Rehearsal in serial recall: An unworkable solution to the nonexistent problem of decay. *Psychological Review*, *122*(4), 674–699. https://doi.org/10.1037/a0039684

Milich, R., Loney, J., & Landau, S. (1982). Independent dimensions of hyperactivity and aggression: A validation with playroom observation data. *Journal of Abnormal Psychology*, *91*(3), 183–198. https://doi.org/10.1037/0021-843X.91.3.183

Miller, S., McCulloch, S., & Jarrold, C. (2015). The development of memory maintenance strategies: Training cumulative rehearsal and interactive imagery in children aged between 5 and 9. *Frontiers in Psychology*, *06*, 6. https://doi.org/10.3389/fpsyg.2015.00524

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial for Quantitative Methods in Psychology*, *4*(2), 61–64. https://doi.org/10.20982/tqmp.04.2.p061

Morey, R. D., Rouder, J. N., Jamil, T., Urbanek, S., Forner, K., & Ly, A. (2018). *BayesFactor: Computation of bayes factors for common designs*. https://CRAN.R-project.org/package=BayesFactor

Naveh-Benjamin, M., & Jonides, J. (1984). Maintenance rehearsal: A two-component analysis. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *10*(3), 369–385. https://doi.org/10.1037//0278-7393.10.3.369

Oberauer, K. (2019). Is rehearsal an effective maintenance strategy for working memory? *Trends in Cognitive Sciences*, *23*(9), 798–809. https://doi.org/10.1016/j.tics.2019.06.002
Palmer, S. (2000). Working memory: A developmental study of phonological recoding. *Memory*, *8*(3), 179–193. https://doi.org/10.1080/096582100387597
Poloczek, S., Henry, L. A., Messer, D. J., & Büttner, G. (2019). Do children use different forms of verbal rehearsal in serial picture recall tasks? A multi-method study. *Memory*, *27*(6), 758–771. https://doi.org/10.1080/09658211.2018.1563615
Rhodes, S., Jaroslawska, A., Doherty, J., Belletier, C., Naveh-Benjamin, M., Cowan, N., & Logie, R. (2019). Storage and processing in working memory: Assessing dual task performance and task prioritization across the adult lifespan. *Journal of Experimental Psychology General*, *148*(7), 1204–1227. https://www.research.ed.ac.uk/portal/en/publications/storage-and-processing-in-working-memory(f8ca650e-a008-4006-bf65-2214d5f420c2).html
Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001
Siegler, R. (2016). How does change occur? In D. J. Foss, R. J. Sternberg, & S. T. Fiske (Eds.), *Scientists making a difference: One hundred eminent behavioral and brain scientists talk about their most important contributions* (pp. 223–227). Cambridge University Press. https://doi.org/10.1017/CBO9781316422250.050
Tan, L., & Ward, G. (2008). Rehearsal in immediate serial recall. *Psychonomic Bulletin & Review*, *15*(3), 535–542. https://doi.org/10.3758/PBR.15.3.535
Thalmann, M., Souza, A. S., & Oberauer, K. (2019). Revisiting the attentional demands of rehearsal in working-memory tasks. *Journal of Memory and Language*, *105*, 1–18. https://doi.org/10.1016/j.jml.2018.10.005
Vygotsky, L. (1962). *Thought and language* (E. Hanfmann, G. Vakar Eds.). MIT Press. https://doi.org/10.1037/11193-000
Watkins, M. J., & Peynircioĝlu, Z. F. (1982). A perspective on rehearsal. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 16, pp. 153–190). Academic Press. https://doi.org/10.1016/S0079-7421(08)60549-6