**Supplementary Materials**

Short- and Long-Delay Consolidation of Memory Accessibility and Precision across Middle Childhood and Young Adulthood

# Supplementary Methods

*1.0. Precision Normalization*

|  |  |
| --- | --- |
| *A) Identity Line (line of equality)* | *B) Bland-Altman Plot* |
|  |  |

*Figure S0***. Sensitivity Analysis for Memory Precision Normalization**. A) Identity line (lien of equality) reflects that all data points lie exactly on this line, rendering the two measures identical. It reflects that precision normalization using the precision range (i.e., minimal, and maximal values) based on aggregated data is a valid approach to normalized precision. B) The Bland-Altman Plot describes the consistency between the two methods (range calculation based either on item or aggregated level). The results show that two methods of normalization are consistent since all data points (about 95%) fall between the two blue dashed lines.

## Assessment of demographic and cognitive covariates

Other cognitive covariate tasks were also assessed, such as cognitive switching and object-location memory with immediate test but were not included in the current paper.

## Congruency Ratings



*Figure S1***. Congruency Ratings**. In the *Congruency Rating* *Phase*, participants were instructed to rate whether the object and scene fit together. RT – reaction time; s – second.

# Supplementary Results

*2.0 Eye-tacking measurement during encoding (Day 0).*

During encoding following eye-tracking measurements were collected (these are beyond the scope of this paper):

1. Number of fixations before fixating the region of interest (an object within a scene);
2. Duration of the first fixation on the region of interest (an object within s scene);
3. Total duration of fixations on the region of interest (an object within a scene) between entering and leaving the region of interest for the first time;
4. Total duration of fixations on the region of interest (an object within a scene) during the whole encoding time (10 seconds per item);
5. Number of entering and leaving the region of interest (an object within the scene) during the whole encoding time.

## Story quality rating (Day 0)

We first tested for potential group by congruency differences in the self-rating of stories during initial encoding of object-location associations in the learning phase (Fig. S2A). There was a significant main effect of Group*,* F(2,105) = 10.39, p <.001, w2 = .15, *Item Type,* F(1,161) = 49.78, p <.001, w2 = .23, and *Group* x *Item type* interaction, F(2,8057) = 36.82, p < .001, w2 = .008. There was no *Sex* effect (p = .31). Model-based Sidak post hoc comparisons revealed that (i) younger children (b = .32, z = 8.06, 95% CI [.22 – .43], p < .001) and older children (b = .29, z = 7.42, 95% CI [.19 – .39], p < .001) rated their stories for congruent pairs significantly higher than for incongruent pairs and did not differ in their ratings (p = .98), but their ratings differed significantly from that of young adults (all p < .001). No such difference in ratings was observed in young adults (p = .85). Taken together, both child groups, but not young adults, rated their stories to memorize the location of the objects within the scene significantly higher for congruent than incongruent object-scene pairs.

To control for possible effect of story quality ratings, we included story quality rating as a continuous covariate of no interest into all models to predict memory precision and accessibility. However, it was not significant in either model or did not improve the model fit (all p > . 587).

## Post hoc congruency ratings (Day 14 after retrieval)

In the following we characterize the post hoc congruency rating of object-scene pairs (Fig. S2B). The LME model’s total explanatory power was substantial (R2 = .95), and the part related to the fixed effects alone was of .94. Further, we observed a significant *Item Type x Group* interaction, F(2,285) = 28.44, p < .0001, w2 = .16; a significant *Item Type x Item Experience* interaction, F(1,285) = 7.08, p = .008, w2 = .02. Model-based Sidak post hoc comparisons showed that all age groups differentiated significantly between congruent and incongruent items: younger children, b = 2.01, t(284) = 46.61, 95% CI [1.9 – 2.1], p < .0001; older children, b = 2.08, t(284) = 48.14, 95% CI [2.0 – 2.2], p < .0001; and in young adults, b = 2.43, t(284) = 59.62, 95% CI [2.3 – 2.5], p < .0001. Further, younger children and older children differentiated between congruent and incongruent items similarly (p = .86). Young adults differentiated between congruent and incongruent items significantly more than younger children, b = -.41, t(284) = -6.96, 95% CI [-.57 – -.26], p < .0001, and then older children, b = -.35, t(284) = -5.84, 95% CI [-.50 – -.19], p < .0001. Further reports on model effects and comparisons can be found in Table S1. To sum it up, all age groups differentiated significantly between congruent and incongruent items and this differentiation was significantly more pronounced in young adults in comparison to both child groups. Experience attenuated this overall congruency-based differentiation.



*Figure S2***.** **(A) Encoding.** Self-rating of stories during initial encoding***.*** Participants rated the quality of their stories to memorize object locations during initial encoding on the scale from 1 to 4 (1-no story; 2-bad quality; 3-good quality; 4-excellent quality). **(B) Post hoc congruency ratings**. Self-experienced rating of congruency and incongruency between old and new object-scene pairs. All tests use Sidak correction for multiple comparisons.Error bars indicate standard error based on the underlying LME-model. YC – 6-8-year-old children; OC – 9-11-year-old children; YA – young adults. \**p* < .05; \*\**p* < .01; \*\*\**p* < .001 – significant difference; n.s. – non-significant difference.

Table S1

*Statistical overview of the man and interaction effects of the linear mixed effects model for learning on Day 0 (R2 = .23)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Fixed Effects*** | *F(DF)* | | *p* | *ω2* |
| Item Type | .02(1,378) | .904 | | .00 |
| Learning Cycle | **645.72(1,22839)** | **<.001\*\*\*** | | **.03** |
| Group | **48.97(2,168)** | **<.001\*\*\*** | | **.36** |
| Sex | 3.37(1,106) | .070 | | .02 |
| Group x Item Type | .72(2,22744) | .486 | | .00 |
| Group x Learning Cycle | **18.21(2,22840)** | **<.001\*\*\*** | | **.002** |
| Item Type x Learning Cycle | .26(2,22737) | .613 | | .00 |
| Group x Item Type x Learning Cycle | .98(2,22753) | .376 | | .00 |

*Notes.* The following reference levels where used: for Item Type – congruent; for Group – younger children; F – F-value; DF – degrees of freedom; p – p-value; R2 – amount of variance explained by the model; *ω2* – effect size. Type III Analysis of Variance Table with Satterthwaite's method. \*p < .05; \*\* <.01, \*\*\*<.001 (significant difference).

Table S2

*Statistical overview of the main and interaction effects of the linear mixed effects models for (A) overall memory accessibility and (B) precision.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***(A)Memory Accessibility (R2 = .47)*** | | | ***(B)Memory Precision (R2 = .22)*** | | |
| ***Fixed Effects*** | *χ2(df)* | *p* | *F(DF)* | *p* | *ω2* |
| Item Type | .76(1) | .385 | .47(1,164) | .495 | .00 |
| Session | 72.82(2) | **<.001\*\*\*** | 133.90(2,10651) | **<.001\*\*\*** | .02 |
| Group | 21.73(2) | **<.001\*\*\*** | 49.21(2,106) | **<.001\*\*\*** | .47 |
| Sex | 1.94(1) | .163 | 1.05(1,106) | .309 | .00 |
| Group x Item Type | .57(2) | .752 | .47(2,16216) | .624 | .00 |
| Session x Item Type | 2.07(2) | .356 | .71(2,10674) | .490 | .00 |
| Group x Session | 4.51(4) | .342 | 16.30(4,16228) | **<.001\*\*\*** | .003 |
| Group x Item Type x Session | 6/94(4) | .139 | .63(4,16209) | .638 | .00 |

*Notes.* The following reference levels where used: for Item Type – congruent; for Group – younger children; for Session – Day 0; for Delay – short delay. *χ2*– chi-squared; F – F-value; DF – degrees of freedom;p – p-value; *ω2* – effect size. R2 – amount of variance explained by the model; Type III Wald chi-squared test was used in memory accessibility model. Type III Analysis of Variance Table with Satterthwaite's method. \*p < .05; \*\* <.01, \*\*\*<.001 (significant difference).

Table S3

*Statistical overview of the main and interaction effects of the linear mixed models for delay-related difference in memory precision for items that were successfully accessed on Day 0 (R2 = .10).*

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | | |
| ***Fixed Effects*** | *F(DF)* | *p* | *ω2* |
| Item Type | .32(1,155) | .573 | .00 |
| Delay | **106.05(1,158)** | **<.001\*\*\*** | .40 |
| Group | **9.61(2,104)** | **<.001\*\*\*** | .14 |
| Sex | .19(1,104) | .667 | .00 |
| Group x Item Type | .11(2,6330) | .892 | .00 |
| Delay x Item Type | .17(1,157) | .680 | .00 |
| Group x Session | .12(2,6400) | .882 | .00 |
| Group x Item Type x Delay | .07(2,6375) | .932 | .00 |

*Notes.* The following reference levels where used: for Item Type – congruent; for Group – younger children; for Session – Day 0; for Delay – short delay. F – F-value; DF – degrees of freedom; p – p-value; R2 – amount of variance explained by the model; *ω2* – effect size. Type III Analysis of Variance Table with Satterthwaite's method. \*p < .05; \*\* <.01, \*\*\*<.001 (significant difference).

|  |  |
| --- | --- |
| Decrease  Increase |  |
|  |  |

*Figure S3***.** **Delay-related Differences in Memory Precision**. Delay-related differences were calculated as item-level differences in memory precision for short delay (Day 1– Day 0) and long delay (Day 14 – Day 0) based on only successfully accessed items on Day 0. Higher positive values of change indicate higher increase in memory precision. Higher negative values of change indicate higher decrease in memory precision. YC – 6-8-year-old children; OC – 9-11-year-old children; YA – young adults. \**p* < .05; \*\**p* < .01; \*\*\**p* < .001 – significant difference; n.s. – non-significant difference.

Table S4

*Statistical overview of the main and interaction effects of the linear mixed models for delay-related difference in memory accessibility for items that were not successfully accessed on Day 0 (R2 = .19).*

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | | |
| ***Fixed Effects*** | *F(DF)* | *p* | *ω2* |
| Item Type | .26(1,285) | .609 | .00 |
| Delay | **40.34(1,290)** | **<.001\*\*\*** | .12 |
| Group | .85(2,98) | .429 | .00 |
| Sex | .44(1,98) | .508 | .00 |
| Group x Item Type | .59(2,285) | .556 | .00 |
| Delay x Item Type | **21.78(1,287)** | **<.001\*\*\*** | .00 |
| Group x Session | .02(2,290) | .980 | .07 |
| Group x Item Type x Delay | .12(2,287) | .883 | .00 |

*Notes.* The following reference levels where used: for Item Type – congruent; for Group – younger children; for Session – Day 0; for Delay – short delay. F – F-value; DF – degrees of freedom; p – p-value; R2 – amount of variance explained by the model; *ω2* – effect size. Type III Analysis of Variance Table with Satterthwaite's method. \*p < .05; \*\* <.01, \*\*\*<.001 (significant difference).

|  |  |
| --- | --- |
| Decrease |  |
|  |  |

*Figure S4***.** **Delay-related Differences in Memory Accessibility**. Delay-related differences were calculated as item-level differences in memory precision for short delay (Day 0 – Day 1) and long delay (Day 0 – Day 14) based on only not successfully accessed items on Day 0. Higher negative values of differences indicate higher increase in memory accessibility. YC – 6-8-year-old children; OC – 9-11-year-old children; YA – young adults. \**p* < .05; \*\**p* < .01; \*\*\**p* < .001 – significant difference; n.s. – non-significant difference.





*Figure S5*. **A) Relation of Memory Accessibility to Age**. Memory accessibility significantly increases with age in younger 6-to-8-year-old children, in older 9-to-11-year-old children, but not in adults. **A) Relation of Memory Precision to Age**. Memory precision does not increase with age in younger 6-to-8-year-old children, significantly increases with age in older 9-to-11-year-old children and significantly decreases with age in young adults. All p-values of Spearman’s correlation were corrected with False Discovery Rate (FDR) or Benjamini-Hochberg procedure for multiple comparisons (six instances). \**p* < .05; \*\**p* < .01; \*\*\**p* < .001 – significant difference.

Table S5

*Statistical overview of Spearman’s correlation between memory accessibility or precision with age within a group over time.*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Memory Accessibility** | | | **Memory Precision** | | |
|  |  | *p-value* | *p-value(FDR corrected)* |  | *p-value* | *p-value(FDR corrected)* |
| Younger Children |  | .00013 | .0005 |  | .841 | .841 |
| Older Children |  | .027 | .036 |  | .0005 | .0009 |

*Notes.* Type III Analysis of Variance Table with Satterthwaite's method. \*p < .05; \*\* <.01, \*\*\*<.001 (significant difference).