Are Age Differences in Recognition-based Retrieval Monitoring an Epiphenomenon of Age Differences in Memory?

Christopher Hertzog[1], Taylor Curley[1], & John Dunlosky[2]
1: Georgia Institute of Technology
2: Kent State University

1

Metamemory Monitoring

- Metamemory is a multi-faceted construct domain that includes
  - knowledge & beliefs about memory
    - (e.g., memory complaints)
  - monitoring ongoing memory processes
  - adaptive use of strategies to effectively control encoding and retrieval
- Today's talk focuses on age differences in retrieval monitoring

2

2

Older Adults' Elevated False Alarm Rates

- Widely known that OA more likely to generate FAs in recognition memory
  - Not necessarily consistent w/ single process SDT (Koen & Yonelinas, 2014)
- Several studies indicate OAs produce more HIGH CONFIDENCE FAs, in which they are confident in incorrect answers on yes/no recognition tasks
  - May not be true of forced-choice tasks
- Is this a evidence of deficient memory monitoring?

3

3

- OAs show no deficits in monitoring encoding
- There are mixed reports of age differences in feeling-of-knowing (FOK) accuracy
- More widespread reports of possible issues with OA's retrospective confidence judgments (RCJs) in recognition tasks
- But not all studies find it
- Example, newly published study finds age differences in associative recognition d' but not monitoring it with RCJs as measured by meta-d' and 'efficiency' – meta d' / d'
  - Did find higher meta-d' for associative over item recognition in both age groups

4

4

- What about HCFAs?
- Important: cues for metacognitive judgments, like RCJs, derive from retrieved qualia about contents of memory (indirect access via retrieval mechanisms)
- Perfect & Stollery (1993) argued (for source memory) that one needed to control on age differences in memory to be certain that metamemory deficits weren't merely due to lack of access to metacognitive cues due to deficits in memory itself
- Our study addressed this issue

5

5

Hypotheses

- H1: Age differences in HCFAs (or resolution) when OA and YA have equivalent experimental conditions producing age differences in memory
- H2: When age differences in memory equated by differential delays, no age differences in HCFAs (or resolution)
- Note: H2 is consistent with the null hypothesis in standard Neyman-Pearson inference
- We used Bayes Factors to estimate likelihood of H0 (equal HCFAs) relative to H1: (age differences in HCFAs)
  - Brydges & Bielak, 2019

6

6

- Also manipulated distinctiveness at encoding, which is known to affect memory for both OA and YA
- Distinctiveness (Hunt, 2012): identifying differentiating features of stimuli in the context of similar (confusable) stimuli
- Our category task presented a set of 4 related exemplars from a given noun category (e.g., FISH)
- Identifying a distinctive feature of the designate target exemplar for a later memory test should aid memory and reduce HCFAs
- H3: HCFAs lower with distinctive vs shared encoding for both age groups

7

7

- Older adults: 2-day delay
  - *N* = 56
  - Age: *Range* = 60 – 80, *M* = 70.52, *SD* = 4.59
  - Years of Education: *M* = 16.13, *SD* = 2.80
  - Shipley Vocab (40 questions): *M* = 35.69, *SD* = 3.27
- Young adults: 7-day delay
  - N = 86
  - Undergraduates at Georgia Tech
- Young adults: 2-day delay
  - N = 34
  - Undergraduates at Georgia Tech

8

8

METHODS

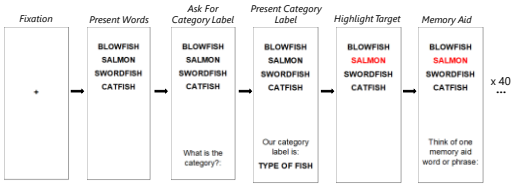
9

9

METHODS

First Session

- 4 concrete nouns drawn from taxonomic category presented in column
- Category queried, then explicitly presented exactly as cued in future
- One element highlighted in RED FONT to designate it is the target for future recall



10

10

METHODS

Between-Subjects Manipulation

- At final stage of item study individuals prompted to type in a feature that is shared by all 4 nouns (similar) or distinguishes the target from its companions (item distinctiveness)

Think of one memory aid word or phrase that you know is SHARED between the word in red and the other words:	***OR***	Think of one memory aid word that you know is DISTINCT about the word in red compared to the other words:

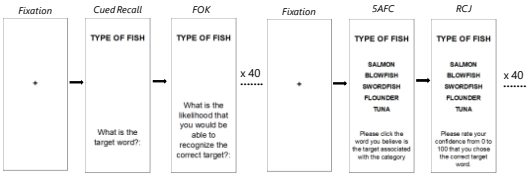
11

11

METHODS

Second Session

- Cued recall (category label as cue)
- FOK immediately after each category-cued recall attempt
- 5-Alternative Forced Choice recognition test (target, 2 OLD lures (presented at study), 2 NEW lures from category norms)
- RCJ collected after each recognition response



12

METHODS

Different Delays to Equate Memory

- Do not want age differences in memory to affect age differences in metamemory (Perfect & Stollery, 1993)
- Recognition is at ceiling for younger adults under distinctive encoding if tested immediately
- Different delays assigned to age groups:
- Young: 7-days
- Old: 2-days

13

13

RESULTS: Frequentist Statistics

- We first report ANOVAs examining age differences in memory performance, as well as if age differences are affected by processing during encoding.
- All 3 experimental conditions (i.e. OAs with a 2-day delay, YAs with a 2-day delay, and YAs with a 7-day delay) are included in these ANOVAs.
- However, we also separately examined when older and younger adults have the same delay between study and test versus when young adults have a longer delay.

14

14

RESULTS: Bayes Factors

- We also examined performance for each young adult group (2-day and 7-day delays) against the older adult group.
  - YA 2-day vs. OA 2-day and YA 7-day vs. OA 2-day
  - Do qualitative differences in recognition memory performance between YAs and OAs still exist after memory performance is matched?
- Following recommendations from Brydges & Bielak (2020), we quantified evidence for and against the null hypothesis (i.e. no age difference) using Bayes Factor.
- Bayes Factor (BF) "bench marks":
  - $1 < BF < 3$ – anecdotal evidence
  - $3 < BF < 10$ – moderate evidence
  - $BF > 10$ – strong evidence

15

15

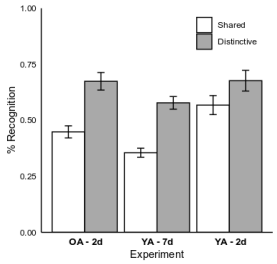
RESULTS: Matched Analysis

- A reviewer expressed concern that merely showing similar overall mean-level recognition performance is not sufficient for inferences about a lack of age deficit.
- We performed a matched-sample analysis by iteratively matching young adults from the 7-day delay condition and older adults from the 2-day delay condition with similar recognition memory performances.
- We used this procedure to match all 55 older adults.
- The matched sample (55 OAs and 55 YAs) was used to examine HCFAs and retrospective confidence judgment accuracy.

16

16

RESULTS: Average Recognition



2 (Encoding Condition: Shared vs. Distinctive) x 3 (Group: OA 2-day vs. YA 2-day vs. YA 7-day) ANOVA

Main effect of Encoding Condition:
- $F(1,169) = 45.25$, $n^2_p = 0.21$

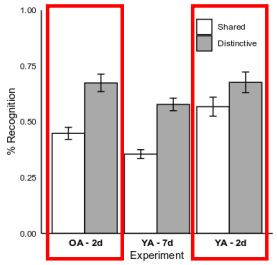
Main effect of Group:
- $F(2,169) = 11.90$, $n^2_p = 0.12$

Interaction: *n.s.*

17

17

RESULTS: Average Recognition (BF)



OA 2-day delay vs. YA 2-day delay

Bayesian null hypothesis: No age differences.

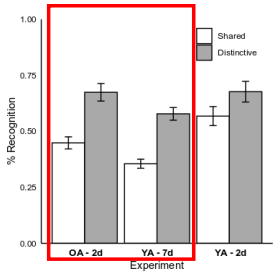
Alternative hypothesis: Group x Condition interaction effect.

Group x Condition:
- BF (H_1) = 1.20

18

18

RESULTS: Average Recognition (BF)



OA 2-day delay vs. YA 7-day delay

Bayesian null hypothesis: No age differences.

Alternative hypothesis: Group by Condition interaction effect.

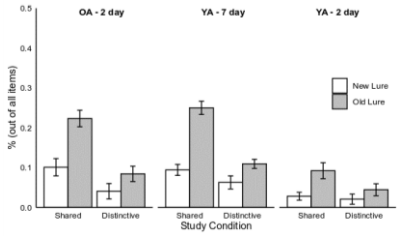
Effect of Group:
- BF (H_0) = 0.22

Group x Condition:
- BF (H_0) = 1.05
- BF (H_1) = 0.95

19

19

RESULTS: High-confidence false alarm rates



2 (Encoding Condition: Shared vs. Distinctive) x 3 (Group: OA 2-day vs. YA 2-day vs. YA 7-day) x 2 (Error Type: New vs. Old Lure) mixed GLM

3-way interaction (Encoding, Group, and Error Type):
- F(2,169) = 3.69, n^2_p = 0.04

Interaction (Condition and Error Type):
- F(1,169) = 11.21, n^2_p = 0.06

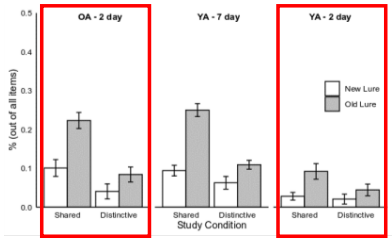
Interaction (Group and Error Type):
- F(2,169) = 6.38, n^2_p = 0.07

Interaction (Group and Condition):
- F(2,169) = 6.34, n^2_p = 0.07

20

20

RESULTS: High-confidence false alarm rates (BF)



OA 2-day vs. YA 2-day only

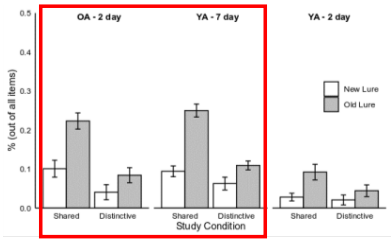
Bayesian null hypothesis: No age differences

Condition x Group
- BF (H_1) = 48.4

21

21

RESULTS: High-confidence false alarm rates (BF)



OA 2-day vs. YA 7-day only

Bayesian null hypothesis: No age differences

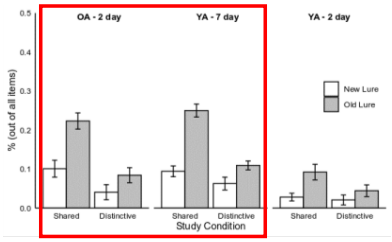
Effect of Group
- BF (H_0) = 4.35

Condition x Group
- BF(H_0) = 5.56

22

22

RESULTS: High-confidence false alarm rates (Matched)



OA 2-day vs. YA 7-day only: Matched Analysis

Bayesian null hypothesis: No age differences

Effect of Group
- BF (H_0) = 5.75

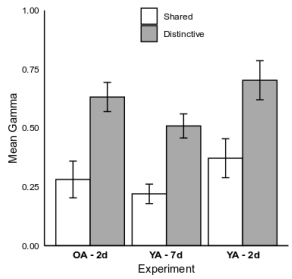
Condition x Group
- BF(H_0) = 4.72

Condition x Group x Error Type
- BF(H_0) = 6.37

23

23

RESULTS: RCJ Gammas



2 (Encoding Condition: Shared vs. Distinctive) x 3 (Group: OA 2-day vs. YA 2-day vs. YA 7-day) ANOVA

Main effect of Encoding Condition:
- $F(1,169) = 40.29$, $n^2_p = 0.19$

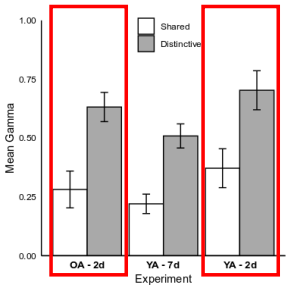
Main effect of Group:
- $F(2,169) = 9.56$, $n^2_p = 0.10$

Interaction: *n.s.*

24

24

RESULTS: RCJ Gammas (BF)



OA 2-day delay vs. YA 2-day delay

Bayesian null hypothesis: No age differences.

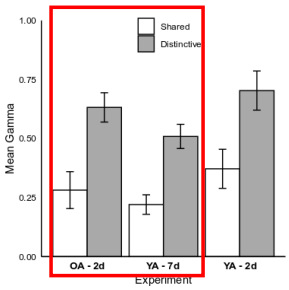
Alternative hypothesis: Group x Condition interaction effect.

Group x Condition:
- BF (H_1) = 0.32

25

25

RESULTS: RCJ Gammas (BF)



OA 2-day delay vs. YA 7-day delay

Bayesian null hypothesis: No age differences.

Alternative hypothesis: Group x Condition interaction effect.

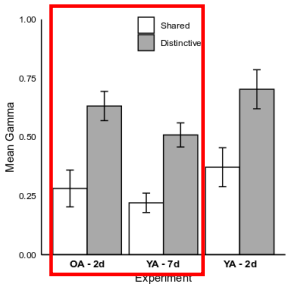
Group:
- BF (H_0) = 2.45

Group x Condition:
- BF (H_0) = 5.56
- BF (H_1) = 0.18

26

26

RESULTS: RCJ Gammas (Matched)



OA 2-day delay vs. YA 7-day delay: Matched

Bayesian null hypothesis: No age differences.

Group:
- BF (H_0) = 5.43

Group x Condition:
- BF (H_0) = 5.03

27

27

RESULTS: RCJ Accuracy (Alternative Measures)

- As an alternative to the gamma correlation, we also calculated d_a, an estimation of area under the curve in a type-2 ROC function (Benjamin & Diaz, 2008; Masson & Rotello, 2009), for each participant
  - The greater the d_a value, the greater the correspondence between a type-2 judgment and performance.
- High correspondence between gammas and d_a values.

	OA – 2 Day Delay			YA – 2 Day Delay			YA – 7 Day Delay		
	All	Shared	Distinctive	All	Shared	Distinctive	All	Shared	Distinctive
d_a	1.00 (0.10)	0.71 (0.14)	1.37 (0.11)	0.80 (0.13)	0.34 (0.13)	1.23 (0.16)	0.76 (0.09)	0.47 (0.12)	1.05 (0.12)

28

28

Conclusions

- Our data support the hypothesis that age differences in HCFA in this 5AFC memory experiment disappear when controlling on memory
- Use of BF statistics helpful in making this claim
- No age differences in RCJ resolution (as measured by gamma and meta-d')
  - Raises concerns about HCFA measure that has been used in literature
- Distinctiveness encoding repairs tendency for HCFA and increases resolution in both age groups
- See Hertzog et al. (in press), Psychology and Aging

29

29

QUESTIONS

30

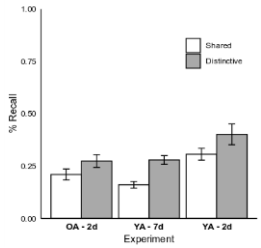
30

EXTRA SLIDES

31

31

RESULTS: Average Recall



2 (Encoding Condition: Shared vs. Distinctive) x 3 (Group: OA 2-day vs. YA 2-day vs. YA 7-day) ANOVA

Main effect of Encoding Condition:
- $F(1,169) = 16.34$, $n^2_p = 0.09$

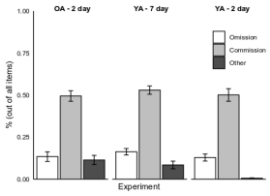
Main effect of Group:
- $F(2,169) = 11.37$, $n^2_p = 0.12$

Interaction: *n.s.*

32

32

RESULTS: Recall Errors



2 (Encoding Condition: Shared vs. Distinctive) x 3 (Group: OA 2-day vs. YA 2-day vs. YA 7-day) x 3 (Error Type: Omission vs. Commission vs. Other) ANOVA

Main effect of Encoding Condition:
- $F(1,169) = 16.46$, $n^2_p = 0.09$

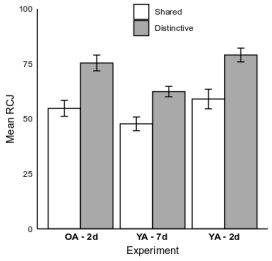
Main effect of Group:
- $F(2,169) = 9.98$, $n^2_p = 0.11$

Main effect of Error Type:
- $F(2,338) = 152.49$, $n^2_p = 0.47$

33

33

RESULTS: Average RCJ



2 (Encoding Condition: Shared vs. Distinctive) x 3 (Group: OA 2-day vs. YA 2-day vs. YA 7-day) ANOVA

Main effect of Encoding Condition:
- $F(1,169) = 40.29$, $n^2_p = 0.19$

Main effect of Group:
- $F(2,169) = 9.56$, $n^2_p = 0.10$

Interaction: *n.s.*

34