Journals of Gerontology: Psychological Sciences
cite as: *J Gerontol B Psychol Sci Soc Sci*, 2023, Vol. 78, No. 6, 969–976
https://doi.org/10.1093/geronb/gbac187
Advance Access publication December 5, 2022



Research Article

# Determining Whether Older Adults Use Similar Strategies to Young Adults in Theory of Mind Tasks

**Anne C. Krendl, PhD,* Willa Mannering, BSc, Michael N. Jones, PhD, Kurt Hugenberg, PhD, and Daniel P. Kennedy, PhD**

Department of Psychological and Brain Sciences, Indiana University, Bloomington, Indiana, USA.

*Address correspondence to: Anne C. Krendl, PhD, Department of Psychological and Brain Sciences, Indiana University, 1101 E. 10th St., Bloomington, IN 47405, USA. E-mail: akrendl@indiana.edu

Received: June 22, 2022; Editorial Decision Date: November 11, 2022

**Decision Editor:** Phoebe E. Bailey, PhD

## Abstract

**Objectives:** Theory of mind—the ability to infer others' mental states—declines over the life span, potentially due to cognitive decline. However, it is unclear whether deficits emerge because older adults use the same strategies as young adults, albeit less effectively, or use different or no strategies. The current study compared the similarity of older adults' theory of mind errors to young adults' and a random model.

**Methods:** One hundred twenty older adults ($M_{Age}$ = 74.68 years; 64 female) and 111 young adults ($M_{Age}$ = 19.1; 61 female) completed a novel theory of mind task (clips from an episode of the sitcom *The Office*®), and a standard measure of cognitive function (Logical Memory II). Monte Carlo resampling estimated the likelihood that older adults' error patterns were more similar to young adults' or a random distribution.

**Results:** Age deficits emerged on the theory of mind task. Poorer performance was associated with less similarity to young adults' response patterns. Overall, older adults' response patterns were ~2.7 million times more likely to match young adults' than a random model. Critically, one fourth of older adults' errors were more similar to the random distribution. Poorer memory ability contributed to this relationship.

**Discussion.** Age deficits in theory of mind performance may be driven by a subset of older adults and be related to disparities in strategy use. A certain amount of cognitive ability may be necessary for older adults to engage similar strategies to young adults' during theory of mind.

**Keywords:** Social cognitive aging, Theory of mind

Social connectedness is associated with improved physical, cognitive, and mental outcomes for older adults (Boss et al., 2015; Kuiper et al., 2015). Theory of mind—the ability to infer others' mental states (Frith & Frith, 2005)—plays a key role in developing and maintaining social connectedness across the life span (Bishop-Fitzpatrick et al., 2017; Krendl et al., 2022; Watson et al., 1999). Though healthy and pathological aging have both been associated with declines in theory of mind (e.g., Demichelis et al., 2020; Henry et al., 2013), efforts to characterize the magnitude and causes of these deficits have yielded mixed results (Fernandes et al., 2021). Methodological disparities and age-related cognitive decline have been commonly implicated as potential drivers of these disparities (Fernandes et al., 2021; Grainger et al., 2019; Laillier et al., 2019), but they have yet to provide a systematic understanding of how aging affects theory of mind. The current study thus explores the possibility that age-related cognitive decline affects the types of strategies that older adults use on theory of mind tasks, and those strategies predict performance.

Theory of mind engages multiple cognitive resources, such as maintaining multiple pieces of information in working memory, inhibiting the incorrect prediction, and episodic memory (Bottirolo et al., 2016; Fernandes et al., 2019; Laillier et al., 2019; Leslie et al., 2004; Scholl & Leslie, 2001; Wellman & Cross, 2001). Older adults may lose the ability to engage these systems effectively as their cognitive abilities decline. This could unfold in two possible ways. On the one hand, young and older adults may all use similar strategies to perform theory of mind tasks, but the cognitive decline may disrupt the effectiveness with which older adults do this. On the other hand, older adults who have acquired a certain level of cognitive decline may shift to using less demanding albeit less effective strategies than other older adults or young adults. If true, some older adults would use similar strategies (to varying degrees) as young adults, whereas others would use a different strategy. One way to disentangle these possibilities is to compare the *types of errors* committed by young and older adults. The current study leveraged a computational resampling approach and a novel theory of mind task to disentangle these possibilities.

Though many standard theory of mind measures are scored as being correct (the mental state was accurately inferred) or incorrect (the mental state was not accurately inferred), other tasks include multiple response options wherein only one is correct (e.g., Eckerly, 2021; Fretland et al., 2015; Frith, 1994). In some cases, error patterns can then be compared at the group level to assess similarities (or differences) in responses between groups (e.g., Eckerly, 2021). This approach provides the flexibility of retaining an agnostic approach to strategy-type, and focus on whether members of one group (e.g., older adults) all use similar strategies. We employed this approach in the current study to determine whether older adults are attracted to the same foils as young adults when they make errors. If so, it would suggest that they use similar strategies (which lead to similar missteps in theory of mind). However, if older adults tend to select dissimilar foils to young adults, it would suggest that their errors result from using either a different strategy or no strategy at all.

To investigate this question, we employed a Monte Carlo simulation approach. Here, two probability distributions can be compared: (a) a young adult probability distribution that is based on the frequency with which young adults endorse each response option for every item, and (b) a random probability distribution that simulates error responses uniformly across the foils. Because this simulation generates a value for each older adult that predicts the likelihood that their response profile is more similar to a young adults' or to a random distribution, we can answer two questions. First, we can determine whether older adults use similar strategies to each other (e.g., respond in a way that is more similar to young adults or more similar to chance responding) and to young adults. Second, we can isolate whether poorer performance is predicted by using a strategy that is more similar to chance performance instead of using a strategy that is more similar to young adults'.

An important benefit of using a random chance comparison is that it aligns theory of mind research with other cognitive aging research (e.g., memory) that commonly compares older adults' memory performance to chance. Though a chance comparison has not been integrated into theory of mind research, this practice has been adapted in other types of social cognitive aging research, including emotion recognition (Ruffman et al., 2009) and deception detection (Stanley & Blanchard-Fields, 2008). Indeed, prior work has shown that older adults' perform at chance in some types of deception detection (Stanley & Blanchard-Fields, 2008), which is a subdomain of theory of mind. Thus, to rule out the possibility that poor performance is driven by chance, we compared older adults' performance to young adults' and a random model. Of critical interest here is whether the type of strategy that older adults use is related to age-related cognitive decline.

Prior work has implicated a range of different cognitive abilities in older adults' theory of mind deficits, including memory (Fernandes et al., 2021; Laillier et al., 2019) and executive function (Bailey & Henry, 2008; Charlton et al., 2009; Wang & Su, 2013), with conflicting findings across studies (e.g., Bottirolo et al., 2016; Wang & Su, 2013). One contributing factor to these discrepant findings may be that recent work has shown that different cognitive abilities (episodic memory and executive function) relate to different types of theory of mind (e.g., cognitive, affective, and respectively; Fischer et al., 2017; Laillier et al, 2019). Because age deficits in theory of mind have been more commonly observed on cognitive than affective theory of mind tasks (Bottirolo et al., 2016, Li et al., 2013; Wang & Su, 2013; but see Henry et al., 2013), an exploratory goal of the current study was to determine whether age differences in episodic memory predicted older adults' strategy shifts.

An important consideration in the current study is the type of stimuli used. A limitation of prior work is that it has used different types of stimuli (e.g., pictures or words) to compare performance on affective and cognitive theory of mind (e.g., Henry et al., 2013). In addition to creating confounds in task performance, these stimuli do not capture older adults' real-world theory of mind (e.g., Grainer et al., 2019). These limitations have been addressed, at least in part, by recent work that used stimuli that are more ecologically valid (Grainger et al., 2019; Laillier et al., 2019). Such stimuli may provide unique insight into understanding why older adults' theory of mind deficits emerge because they may better capture the complexity of real-world social interactions (see also Hamilton et al., in press). The current investigation adopts a similar approach using a novel task in which participants viewed multiple brief clips of a television episode (the U.S. version of *The Office*®) and were asked multiple choice questions about people viewed in the clip (Byrge et al., 2015; Krendl et al., 2022). A recent study using this task found that it predicted the nature of older adults' real-world social relationships (e.g., Krendl et al., 2022), thus highlighting its ecological validity. Additional benefits of this method include that it allowed us to assess multiple subcomponents of theory of mind (understanding others' affective states, beliefs, thoughts, or intentions, and detecting deception) using the same types of stimuli. Prior work has often collapsed different subcomponents of theory of mind into a single measure of theory of mind (Fischer et al., 2017; Wang & Su, 2013), thereby possibly reducing measurement sensitivity. Critically, and central to the current study, this method allowed us to compare the *types of errors* made by younger and older adults during the task by introducing multiple (3–4) response options (i.e., foils), rather than relying strictly on binary "correct" or "incorrect" responses.

### Current Research

We used Monte Carlo resampling to compare error patterns between young and older adults on this theory of mind task. The behavioral data for the present investigation were drawn from Krendl et al. (2022), and the simulation data were created for the present investigation. We predicted that older adults' theory of mind deficits are *systematic* (e.g., the same error patterns as young adults, but amplified), not *random* (e.g., distinct from young adults' error patterns, reflecting an inability to decode or attribute mental states; Hypothesis 1). Hypothesis 2 predicted that older adults' error patterns would positively predict their task accuracy, and explored whether their similarity to young adults' response patterns differed across multiple domains of theory of mind (inferring intentions, understanding motivations, detecting deception, and understanding emotions). Finally, Hypothesis 3 was exploratory and examined the extent to which age deficits in episodic memory explained whether older adults' errors were systematic or random. Specifically, we examined whether the extent to which older adults' errors were systematic mediated the relationship between episodic memory and task performance (e.g., Fischer et al., 2017; Laillier et al., 2019).

## Method

One hundred twenty older adults ($M_{Age}$ = 74.68 years, standard deviation [*SD*] = 7.13; 64 female) were recruited from the Bloomington, Indiana community. They were well-educated (86.7% had a college degree or higher) and not cognitively impaired (as indicated by scoring >26 on the Mini-Mental State Exam; Folstein et al., 1975). They received monetary compensation for participating. A group of 111 younger adult undergraduates at Indiana University ($M_{Age}$ = 19.1, *SD* = 1.4; 61 female) participated in exchange for partial course credit.

Sensitivity analyses were conducted in G*Power 3.1 (Faul et al., 2007) to determine the power that could be detected for the repeated measures analysis of variance (ANOVA; Hypothesis 2) using the smallest *N* (75) and the lowest correlation observed between the theory of mind measures ($r$ = 0.34; as reported in Krendl et al., 2022). The analysis revealed that a power = 0.80 and five question types with 75 participants were sufficient to identify small effects ($f^2$ = 0.146) at $p < .05$.

During the testing session, young and older adults completed several social cognitive tasks as well as a social network interview as part of a study on social connectedness (see Krendl et al., 2022 for these results). Of interest in the current investigation was their performance on the dynamic theory of mind task, which was performed during this testing session, and their episodic memory (measured with the Logical Memory II; Wechsler, 2009). One older adult did not complete the theory of mind task and was excluded from the analyses.

### Measures

**Dynamic theory of mind task**

Cognitive and affective theory of mind were measured using a novel task that was adapted from previous research (Byrge et al., 2015), and used in a recent study on social connectedness and aging (Krendl et al., 2022). In the task, participants viewed brief (10–60 s) clips from an episode of the sitcom *The Office*®, and responded to questions about each clip to assess their *cognitive* and *affective* theory of mind. The questions assessed three distinct components of the cognitive theory of mind—inferring intentionality, inferring others' beliefs, and detecting deception—as well as affective theory of mind (understanding emotions). An example of a question that measured respondent's ability to infer the intentions of others was, "Why does Michael suggest having an ice cream cake?" (answer: "He wants an ice cream cake"). A question measuring belief inference was, "Will Meredith want an ice cream cake?" (answer: "no"). An example of a deception question was, "Why does Pam go downstairs?" (answer: "Pam is trying to fool Dwight"). Questions related to affective theory of mind focused on understanding the emotional responses of the characters, such as "Is Jim happy to see Pam's fiancé, Roy?" (answer: no). For all questions, a still image was presented on the screen depicting the face and name of the character(s) referenced in the question. At the conclusion of the task, participants were asked to indicate whether they had ever seen *The Office*® before and, if so, how familiar they were with the show (1 = very little, 9 = have seen the entire series).

There were 51 questions in total on the task: 11 control questions to assess basic comprehension, seven questions related to deception, 11 related to inferring beliefs, 13 related to inferring intentions, and nine related to emotions. Response options were either multiple choice or yes/no/do not know selection. The questions were developed by three of the authors (A. C. Krendl, D. P. Kennedy, and K. Hugenberg) and categorized on two domains: (a) answer accuracy (to determine that answer identified as the correct answer was unequivocally correct) and (b) theory of mind subcategory (whether the question captured *inferring intentions*, *inferring beliefs*, *detecting deception*, or *understanding emotions*). Full consensus had to be reached on both domains for the question to be retained. One question (in the *inferring intentions* category) was removed because consensus could not be reached and not included in the analyses. This left 50 total questions. See Supplementary Appendixes A and B for information on clips and the full list of questions.

Participants were given 30 s to read and respond to each question. If they did not respond in that time, that item was considered a "missed response" and excluded from the analyses. Performance was scored as proportion of items correct (#correct/#responded) within each theory of mind subcomponent. To minimize missed responses, participants first completed a practice trial with the experimenter. This also served to ensure they could hear the stimuli. Accuracy was calculated by subcomponent (inferring intentions, inferring beliefs, detecting deception, understanding emotions, and control) using the number of correct divided by the total number of questions to which the individual responded (i.e., excluding missed responses) for that subcomponent. See Supplementary Materials for performance results (also Krendl et al., 2022).

### Episodic memory measure

Episodic memory was assessed with the Logical Memory II from the Wechsler Memory Scale IV (Wechsler, 2009), a widely used measure of verbal episodic memory. Standard task administration and scoring were used. Consistent with standard protocols, an experimenter read two distinct passages detailing events about an individual, and told the participants they would be asked to recall them. Participants then completed an immediate and 30-min delayed recall for each passage in which they were instructed to retell the stories with as much detail as possible. Points were given for each detail correctly recalled. Because age deficits are most pronounced on delayed recall (Larry & Asenath, 2003), memory performance was operationalized from the Logical Memory II (delayed recall) using total number of items recalled from both stories.

## Results

*Hypothesis 1: Determining whether older adults' error patterns are systematic or random*

As reported in Krendl et al. (2022), young adults outperformed older adults on all channels of theory of mind task, with the most pronounced age-related deficits on detecting deception and understanding emotions (see Supplementary Results for full analyses, also Figure 1). Analyses on reaction times are provided in the Supplementary Materials; see also Supplementary Table 1.

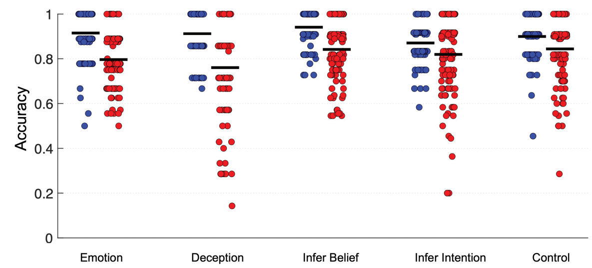


**Figure 1.** Older and young adults' accuracy (proportion correct) on each domain of the theory of mind task (emotion, deception, inferring belief, inferring intention, and control). Scatterplots show accuracy for each individual participant, with the respective group mean indicated by a black bar. Older adults are shown in red, while young adults are in blue.

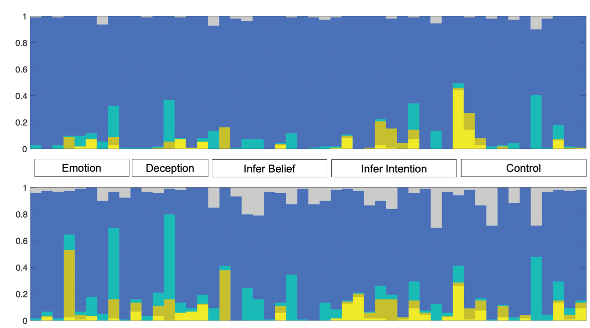


**Figure 2.** Young (top) and older (bottom) adults' response profiles. Each column on the *x*-axis shows the proportion of participants who endorsed a specific response item on each question. Questions are organized by theory of mind domain (emotion, deception, infer belief, infer intention, and control). For each question, blue represents the proportion of respondents who selected the correct response. Cyan, yellow, and gold indicate a specific foil. Foils are the same color for young and older adults. For example, for a given item, young and older adults who selected the same foil are indicated in the same color (e.g., gold), with higher proportions in one group indicating a greater propensity to select that foil. Gray is nonresponses.

Hypothesis 1 used young and older adults' response patterns on the task to test the prediction that older adults' theory of mind deficits are *systematic* (e.g., they have similar error patterns to young adults, but amplified), not *random* (e.g., they respond in a seemingly random manner). We employed a Monte Carlo simulation to test this prediction. Across items, every question has one single correct answer and two or three foils. We first identified the frequency with which young adults endorsed each response option for every item, creating a probability distribution of endorsement over responses per item. The goal of the simulation approach was to determine whether older adults were attracted to foils in a similar pattern as young adults (but endorsed the foils with higher frequency) or selected from the foils in a pseudorandom manner. See Figure 2 for a graphical depiction of the response patterns for young and older adults across all items.

We simulated predictions for each older adult's error data under two models. The first model simulated responses of foil endorsement using the young adult response distribution. The second model was a random baseline that simulated error responses uniformly across the foils (reflecting chance responding). Hence, both models predicted the likelihood of each foil being endorsed when an error is made. We then computed the likelihood of the observed data having been generated by each model. As is standard in cognitive modeling (Busemeyer & Diedrich, 2010), this likelihood was represented using the log-likelihood across all errors per participant. As the log of a probability is a negative value, it is common to reverse the sign and use the negative log-likelihood (NLL). Values of the NLL closer to zero indicate a better fit between the model and data.

Across all older adult participants, the NLL demonstrated that 89 were better fit by the young (versus random) distribution ($M_{NLL}$ = 3.93, *SD* = 3.73; range = 0.09–17.13), whereas the remaining 30 were better fit by the random (versus young) distribution ($M_{NLL}$ = 8.62, *SD* = 8.10; range = 0.16–28.87). Older adults whose data better fit the random model had lower performance on the Logical Memory II (*M* = 18.30, *SD* = 8.75) than those who had a better fit to the young adult model (*M* = 22.13, *SD* = 7.54, $t[117] = 2.31$, $p = .023$, $d = 0.49$), but the average ages of individuals in these two groups did not differ ($t[117] = 1.06$, $p = .29$, $d = 0.22$).

*Hypothesis 2: Strategic errors predict better performance than random errors for theory of mind overall and by domain*

Hypothesis 2 predicted that older adults' theory of mind performance would be positively associated with having greater similarity to young adults' error response patterns. We examined whether this was the case for performance overall, and then by the four domains of theory of mind measured in the current study (detecting deception, inferring intentions, understanding emotions, and understanding motivations), as well as control items. While NLL reports the likelihood of each observed data set having been generated by either the systematic or random model, it does not provide a direct measure of the extent to which one model is preferred over the other. However, this preference can be calculated through the Bayesian Information Criterion (BIC), which quantifies the posterior probability for the young adults' response profile over the random model given the data (Wagenmakers, 2007).

BIC was calculated as $(2NLL_{RND} + \log_n) - (2NLL_{YA} + \log_n)$ for each participant. Positive values for BIC indicate a stronger preference for the young adult model over the random model, whereas a negative value indicates the opposite. BIC was positively correlated with accuracy, $r(119) = 0.31$, $p < .001$, and was not significantly associated with familiarity with the show, $r(119) = 0.16$, $p = .08$. There was a wide range of scores (−36.95 to 18.81), highlighting the individual differences on this measure. Converting the total BIC to a Bayes Factor as exp(BIC/2) yielded a Factor of $3 \times 10^{-77}$—very strong evidence for the young adult model as having generated the observed data over the random model under Jeffreys' (1961) evidence scale for Bayes Factors.

We next examined whether performing the task with greater similarity to young adults predicted better performance for older adults across each domain of theory of mind. However, because BIC incorporates the number of errors in its calculations, data were only available for participants who had at least one error on the given domain (this ranged from $N = 75$ on deception items to $N = 92$ on control items). To include the whole sample, we also calculated similarity using the difference scores between NLLs for the young and random distribution. This approach yielded a similar patterns of results (see Supplementary Results).

Using BIC as the dependent variable, we conducted a repeated measures ANOVA with question type (control, detecting deception, inferring intentions, understanding emotions, and understanding motivations) as the independent variable. Results yielded a main effect of question type, $F(4,240) = 4.36$, $p = .002$, $\eta^2_{partial} = 0.07$. The main effect was driven by the fact that older adults' response profiles were most similar to young adults' for inferring intentions ($M_{BIC}$ = 3.00, *SD* = 5.44), followed by inferring beliefs ($M_{BIC}$ = 1.86, *SD* = 2.41). Specifically, older adults' response profiles were more similar to young adults when inferring intentions than when understanding emotion ($M_{BIC}$ = 1.07, *SD* = 2.99), $t(89) = 2.77$, $p = .007$, $d = 0.29$, detecting deception ($M_{BIC}$ = 1.41, *SD* = 3.96), $t(74) = 2.34$, $p = .022$, $d = 0.27$, or completing control tasks ($M_{BIC}$ = 1.00, *SD* = 4.38), $t(74) = 3.19$, $p = .002$, $d = 0.37$. BIC scores for inferring beliefs did not differ from any other domain, all *p*s > .09.

*Hypothesis 3: Strategy use mediates the relationship between older adults' episodic memory and theory of mind performance*

Hypothesis 3 was exploratory, and examined the extent to which the relationship between older adults' episodic memory and theory of mind performance was mediated by their BIC score. Older adults' performance on the memory task was positively related to their task performance, $r(119) = 0.44$, $p < .001$. We then used PROCESS with 5,000 bootstrap samples (Hayes, 2012) to determine whether BIC mediated this relationship. Results showed a small but the significant, indirect effect, $B = 0.001$, standard error [*SE*] = 0.0004, confidence interval [CI] 0.000, 0.002. See Figure 3.

## Discussion

Together, our results demonstrate several novel findings. First, though most older adults employed theory of mind strategies similar to young adults, about one quarter of older adults had seemingly random responses. Second, older adults' response profiles were most similar to young adults' for inferring intentions, and least similar for understanding emotion or detecting deception—the two domains on which age-related performance deficits were most

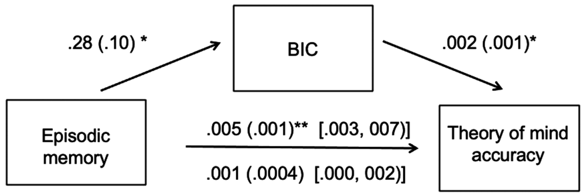


**Figure 3.** Results from a mediation analysis demonstrating that the extent to which older adults' response profiles were similar to young adults' (BIC) mediated the relationship between their episodic memory (performance on the Logical Memory II) and their theory of mind performance. ***p* < .001; **p* ≤ .01. BIC = Bayesian Information Criterion.

pronounced. Third, as noted in Figure 1, there was a great deal of variability in older adults' performance, with lower performance being more strongly related to responding in a seemingly random manner. This suggests that age deficits in theory of mind performance may be driven, at least in part, to disparities in strategy use. Finally, using a response strategy that was more similar to young adults' mediated the relationship between cognitive ability and theory of mind performance. Together, these findings suggest that observed age deficits in theory of mind may be driven, at least in part, but a subset of older adults who adopt more random response strategies.[1] If true, it could provide insights into the seemingly mixed findings in the extant research on age-related theory of mind deficits (e.g., Fernandes et al., 2021), by suggesting that the magnitude of observed age deficits in the literature may be driven, at least in part, by a subset of participants.

Also notable is that older adults' response profiles were least similar to young adults' on items related to understanding emotions or detecting deception, the two domains on which performance deficits were also the most pronounced. Given prior mixed findings regarding potential age deficits on affective theory of mind (Bottirolo et al., 2016, Li et al., 2013; Wang & Su, 2013; but see Henry et al., 2013), a potential interpretation of this finding is that strategy differences may contribute to disparate patterns across studies. Regarding why the performance deficits on understanding emotions and detecting deception were more pronounced, one possibility is that both may necessitate attending to and recognizing emotional displays (e.g., fear and worry). For example, understanding emotions and detecting deception may both rely on decoding emotional displays (e.g., Ruffman et al., 2008; Stanley & Blanchard-Fields, 2008). Numerous studies have shown that older adults are impaired in emotion recognition (e.g., Ruffman et al., 2008). Alternatively, understanding emotions or detecting deception may be more cognitively demanding than inferring intentions. Future work should examine what factors relate to older adults' response patterns being more similar or dissimilar to young adults.

The results of the current work may also provide insight into how cognitive decline affects older adults' theory of mind performance (e.g., Bailey & Henry, 2008; Cavallini et al., 2013; Fernandes, 2021; German & Hehman, 2006). Specifically, our results suggest that declines in memory ability may affect the strategies older adults engage during theory of mind tasks, which subsequently disrupts their performance. An important caveat to this finding, however, is that the mediation effects were relatively small. One reason for that might be our focus on episodic memory. Prior work has implicated a range of different cognitive abilities in older adults' theory of mind deficits, including memory (Fernandes et al., 2021; Laillier et al., 2019) and executive function (Bailey & Henry, 2008; Charlton et al., 2009; Wang & Su, 2013), with conflicting findings across studies (e.g., Bottirolo et al., 2016; Wang & Su, 2013). Though recent work has demonstrated that episodic memory consistently predicted age deficits in cognitive theory of mind performance (Fischer et al., 2017; Laillier et al, 2019), future work should disentangle the role of episodic memory and executive function in older adults' response patterns across different domains of theory of mind.

## Limitations and Future Directions

The limitations of the present work provide multiple avenues for future work. First, though our findings suggest that episodic memory plays a role in theory of mind performance, it is not possible to make conclusions about specificity. That is, without directly testing other potential mechanisms (e.g., processing speed and inhibition) in the same study (e.g., Fischer et al., 2017; Laillier et al., 2019), we cannot determine whether the strategies older adults use are specifically related to episodic memory. Moreover, our measure of episodic memory involved remembering information from passages. Though this is a well-validated measure of episodic memory, it does not capture memories for autobiographical events. This is important because older adults may utilize their own personal experiences to help infer and interpret other people's behaviors. Consistent with this assertion, theory of mind and autobiographical memory engage similar neural systems (Spreng et al., 2009), suggesting they may be intertwined. Because autobiographical memory relies on distinct neural systems from other episodic memories (Gilboa, 2004), our finding that episodic memory predicts theory of mind accuracy should not be interpreted as including autobiographical memory. Future work should explore this possibility.

Second, an important limitation of using *The Office*® in the current task is that young adults reported greater familiarity with the show than did older adults. Though our analyses controlled for participants' prior exposure to the show, we could not account for the possibility that

[1]Older adults whose data fit better with the random model had lower accuracy (M = .77, SD = .03) than those who had a better fit to the young adult model (M = .83, SD = .08, $t(117) = 3.15$, $p = .002$, $d = .67$).

the show might depict social behaviors that are more normative and/or common for young adults than they are for older adults. However, it is important to note that age deficits in performance and the extent to which older adults' response patterns were similar to young adults' differed across the different theory of mind domains. Thus, if young adults outperformed older adults because they could better relate to the material, this effect was not systematic. Moreover, our results also showed that nearly three-quarters of older adults had response patterns that were more similar to young adults', suggesting that, irrespective of the generational differences in familiarity, most older adults used similar strategies to young adults' to perform the task. Regardless of these findings, future work should leverage tasks that are similarly unfamiliar to young and older adults to control for potential confounds.

Finally, an important caveat to our results is though we used a random model for comparison, we do not mean to suggest that low-performing older adults respond in a random manner. Rather, the results suggest that these older adults' response patterns were more similar to a uniform (random) distribution than to the young adults' distribution. It is also possible that lower performing older adults used a nonrandom strategy, but one that was different from young adults. For example, prior work suggests that low-performing older adults may be more literal in their interpretations during theory of mind tasks (Wang & Su, 2013). However, it is worth noting that although the older adults' whose response patterns were more similar to the uniform model had poorer memory than those whose errors were more similar to young adults, the mean age of the two groups did not differ. Thus, it is unlikely that their response patterns were due to more literal responding. An alternate possibility is that missing a key piece of information early in the show (e.g., that Jim was only pretending to be in the alliance with Dwight) might have had downstream effects on performance. The benefit of this approach is that it parallels real-life interactions where information is temporally ordered and contingent. As this task was not designed to detect different potential strategies or lapses in understanding, future research should systematically investigate these possibilities.

## Conclusion

Together, these studies contribute to the growing literature on age deficits in theory of mind. Specifically, they show that, though age deficits emerged on the task, engaging similar strategies to young adults predicted better theory of mind performance for older adults. Moreover, memory decline predicted more random error response patterns for older adults, which affected performance accuracy. This work suggests that for many older adults, theory of mind interventions may be most effective if they target reducing errors, rather than changing strategies.

## Supplementary Material

Supplementary data are available at *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences* online.

## Funding

This project was supported in part by an R01 AG070931 from the NIA (PI: A. C. Krendl). This study was not preregistered, but data, analytic methods, and study materials will be made available to other researchers upon request.

## Conflict of Interest

None declared.

## References

Bailey, P. E., & Henry, J. D. (2008). Growing less empathic with age: Disinhibition of the self-perspective. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, **63**(4), P219–P226. doi:10.1093/geronb/63.4.p219

Bishop-Fitzpatrick, L., Mazefsky, C. A., Eack, S. M., & Minshew, N. J. (2017). Correlates of social functioning in autism spectrum disorder: The role of social cognition. *Research in Autism Spectrum Disorders*, **35**, 25–34. doi:10.1016/j.rasd.2016.11.013

Boss, L., Kang, D. H., & Branson, S. (2015). Loneliness and cognitive function in the older adult: A systematic review. *International Psychogeriatrics*, **27**(4), 541–553. doi:10.1017/S1041610214002749

Bottiroli, S., Cavallini, E., Ceccato, I., Vecchi, T., & Lecce, S. (2016). Theory of mind in aging: Comparing cognitive and affective components in the faux pas test. *Archives of Gerontology and Geriatrics*, **62**, 152–162. doi:10.1016/j.archger.2015.09.009

Busemeyer, J. R., & Diederich, A. (2010). *Cognitive modeling*. Sage.

Byrge, L., Dubois, J., Tyszka, J. M., Adolphs, R., & Kennedy, D. P. (2015). Idiosyncratic brain activation patterns are associated with poor social comprehension in autism. *Journal of Neuroscience*, **35**(14), 5837–5850. doi:10.1523/jneurosci.5182-14.2015. PMID: 22498897.

Cavallini, E., Lecce, S., Bottiroli, S., Palladino, P., & Pagnin, A. (2013). Beyond false belief: Theory of mind in young, young-old, and old-old adults. *International Journal of Aging and Human Development*, **76**(3), 181–198. doi:10.2190/ag.76.3.a

Charlton, R. A., Barrick, T. R., Markus, H. S., & Morris, R. G. (2009). Theory of mind associations with other cognitive functions and brain imaging in normal aging. *Psychology and Aging*, **24**(2), 338–348. doi:10.1037/a0015225

Demichelis, O. P., Coundouris, S. P., Grainger, S. A., & Henry, J. D. (2020). Empathy and theory of mind in Alzheimer's disease: A meta-analysis. *Journal of the International Neuropsychological Society*, **26**(10), 963–977. doi:10.1017/s1355617720000478

Eckerly, C. (2021). Answer similarity analysis at the group level. *Applied Psychological Measurement*, **45**(5), 299–314. doi:10.1177/01466216211013109

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, **39**(2), 175–191. doi:10.3758/bf03193146

Fernandes, C., Barbosa, F., Martins, I. P., & Marques-Teixeira, J. (2021). Aging and social cognition: A comprehensive review of the literature. *Psychology and Neuroscience*, **14**(1), 1. doi:10.1037/pne0000251

Fernandes, C., Gonçalves, A. R., Pasion, R., Ferreira-Santos, F., Barbosa, F., Martins, I. P., & Marques-Teixeira, J. (2019). Age-related decline in emotional perspective-taking: Its effect on the late positive potential. *Cognitive, Affective, and Behavioral Neuroscience*, **19**(1), 109–122. doi:10.3758/s13415-018-00648-1

Fischer, A. L., O'Rourke, N., & Loken Thornton, W. (2017). Age differences in cognitive and affective theory of mind: Concurrent contributions of neurocognitive performance, sex, and pulse pressure. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, **72**(1), 71–81. doi:10.1093/geronb/gbw088

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-Mental State Exam: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, **12**, 189–198. doi:10.1016/0022-3956(75)90026-6

Fretland, R. A., Andersson, S., Sundet, K., Andreassen, O. A., Melle, I., & Vaskinn, A. (2015). Theory of mind in schizophrenia: Error types and associations with symptoms. *Schizophrenia Research*, **162**(1–3), 42–46. doi:10.1016/j.schres.2015.01.024

Frith, U. (1994). Autism and theory of mind in everyday life. *Social Development*, **3**(2), 108–124. doi:10.1111/j.1467-9507.1994.tb00031.x

Frith, C., & Frith, U. (2005). Theory of mind. *Current Biology*, **15**(17), R644–R645. doi:10.1016/j.cub.2005.08.041

German, T. P., & Hehman, J. A. (2006). Representational and executive selection resources in "theory of mind": Evidence from compromised belief-desire reasoning in old age. *Cognition*, **101**(1), 129–152. doi:10.1016/j.cognition.2005.05.007

Gilboa, A. (2004). Autobiographical and episodic memory—one and the same?: Evidence from prefrontal activation in neuroimaging studies. *Neuropsychologia*, **42**(10), 1336–1349. doi:10.1016/j.neuropsychologia.2004.02.014

Grainger, S. A., Steinvik, H. R., Henry, J. D., & Phillips, L. H. (2019). The role of social attention in older adults' ability to interpret naturalistic social scenes. *Quarterly Journal of Experimental Psychology*, **72**(6), 1328–1343. doi:10.1177/1747021818791774

Hamilton, L. J., Gourley, A. N., & Krendl, A. C. (2022). They Cannot, They Will Not, or We Are Asking the Wrong Questions: Re-examining Age-Related Decline in Social Cognition. *Frontiers in psychology*, **13**, 894522.

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling.

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging*, **28**(3), 826–839. doi:10.1037/a0030677

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

Krendl, A. C., Kennedy, D. P., Hugenberg, K., & Perry, B. L. (2022). Social cognitive abilities predict unique aspects of older adults' personal social networks. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, **77**(1), 18–28. doi:10.1093/geronb/gbab048

Kuiper, J. S., Zuidersma, M., Voshaar, R. C. O., Zuidema, S. U., van den Heuvel, E. R., Stolk, R. P., & Smidt, N. (2015). Social relationships and risk of dementia: A systematic review and meta-analysis of longitudinal cohort studies. *Ageing Research Reviews*, **22**, 39–57. doi:10.1016/j.arr.2015.04.006

Laillier, R., Viard, A., Caillaud, M., Duclos, H., Bejanin, A., de La Sayette, V., Eustache, F., Desgranges, B., & Laisney, M. (2019). Neurocognitive determinants of theory of mind across the adult lifespan. *Brain and Cognition*, **136**, 103588. doi:10.1016/j.bandc.2019.103588

Larry, P., & Asenath, L. (2003). What does the WMS–III tell us about memory changes with normal aging? *Journal of the International Neuropsychological Society*, **9**(1), 89–96. doi:10.1093/geronb/gbw088

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in "theory of mind." *Trends in Cognitive Sciences*, **8**(12), 528–533. doi:10.1016/j.tics.2004.10.001

Li, X., Wang, K., Wang, F., Tao, Q., Xie, Y., & Cheng, Q. (2013). Aging of theory of mind: The influence of educational level and cognitive processing. *International Journal of Psychology*, **48**(4), 715–727. doi:10.1080/00207594.2012.673724

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience and Biobehavioral Reviews*, **32**(4), 863–881. doi:10.1016/j.neubiorev.2008.01.001

Ruffman, T., Sullivan, S., & Dittrich, W. (2009). Older adults' recognition of bodily and auditory expressions of emotion. *Psychology and Aging*, **24**(3), 614–622. doi:10.1037/a0016356

Scholl, B. J., & Leslie, A. M. (2001). Minds, modules, and meta-analysis. *Child Development*, **72**(3), 696–701. doi:10.1111/1467-8624.00308

Spreng, R. N., Mar, R. A., & Kim, A. S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience*, **21**(3), 489–510. doi:10.1162/jocn.2008.21029

Stanley, J. T., & Blanchard-Fields, F. (2008). Challenges older adults face in detecting deceit: The role of emotion recognition. *Psychology and Aging*, **23**(1), 24–32. doi:10.1037/0882-7974.23.1.24

Wagenmakers, E. J. (2007). A practical solution to the pervasive problem of *p* values. *Psychonomic Bulletin and Review*, **14**, 779–804. doi:10.3758/bf03194105

Wang, Z., & Su, Y. (2013). Age-related differences in the performance of theory of mind in older adults: A dissociation of cognitive and affective components. *Psychology and Aging*, **28**(1), 284–291. doi:10.1037/a0030876.

Watson, A. C., Nixon, C. L., Wilson, A., & Capage, L. (1999). Social interaction skills and theory of mind in young children. *Developmental Psychology*, **35**(2), 386–391. doi:10.1037//0012-1649.35.2.386

Wechsler, D. (2009). *Wechsler memory scale–fourth edition (WMS-IV)*. New York, NY: The Psychological Corporation.

Wellman, H. M., & Cross, D. (2001). Theory of mind and conceptual change. *Child Development*, **72**(3), 702–707. doi:10.1111/1467-8624.00309