HHS Public Access

Author manuscript

Exp Aging Res. Author manuscript; available in PMC 2022 May 01.

Published in final edited form as:

Exp Aging Res. 2021 ; 47(3): 203–219. doi:10.1080/0361073X.2021.1884472.

Guilty, Innocent, or Just Not Proven? Bayesian Verdicts in the Case of Inhibitory Deficits

Joel Myerson[a,*], **Kyle G. Featherston**[a], **Cynthia Flores**[b], **Lindsey Lilienthal**[c], **Young Bui**[d], **Sandra Hale**[a]

[a]Washington University in St. Louis, St. Louis, MO 63130

[b]Stanford University, Stanford, CA 94305

[c]Penn State Altoona, Altoona, PA 16601

[d]Exponent, Inc., Los Angeles, CA 90066

Abstract

***Background*:** This study addresses two issues: Whether age-related differences in working memory (WM) can be studied in online samples, and whether such differences reflect an inhibitory deficit. Currently, the evidence is mixed, but the playing field was not level because traditional statistics cannot provide evidence for the null hypothesis.

Experiment 1: MTurk workers (ages 19–74) performed simple and complex visuospatial WM tasks to determine whether a secondary task affected the rate of age-related decline. Performance on both tasks replicated previous laboratory studies, establishing that age-related differences in WM can be studied online. Bayesian analyses revealed it is ten times as likely that there is no inhibitory deficit on visuospatial WM tasks than that there is.

Experiment 2: The effects of irrelevant location information on visuospatial WM were examined in older (M_{age}=64.0) and younger (M_{age}=25.0) MTurk workers. Irrelevant locations produced interference, but both groups were equally affected. Bayesian analyses provided support for the null hypothesis of no age difference.

Conclusions: The results of both experiments on working memory not only revealed equivalent visuospatial inhibitory function in older and younger adults, they also demonstrated that age-related differences in visuospatial WM can be effectively studied online as well as in the laboratory.

Keywords

age-related differences; visuospatial working memory; inhibitory deficit; Bayesian analysis

---

The present study addresses two important issues in cognitive aging research, one methodological and one theoretical. The first of these concerns whether age-related differences in working memory (WM) can be effectively studied in online samples, the

*Address correspondence to: Joel Myerson, Department of Psychological & Brain Sciences, Washington University in St. Louis, St. Louis, MO 63130, myerson.joel@gmail.com.

second whether such differences, which are robust in laboratory experiments, reflect an inhibitory deficit that results in older adults experiencing greater interference by irrelevant information (Hasher & Zacks, 1988; Lustig, et al., 2007).

Collecting data online is becoming increasingly common in many areas of psychology as researchers take advantage of the fact that the internet provides a convenient and relatively cheap source of large, diverse samples (e.g., Chandler & Shapiro, 2016; Paolacci et al., 2010). However, the growing popularity of this approach has led to concerns about the validity of data obtained from online samples (e.g., Necka, Cacioppo, Norman, & Cacioppo, 2016). These concerns seem quite reasonable, particularly with respect to data from older participants online, whose cognitive ability may be higher than that of the older adult population in general. Fortunately, there is a way to address these concerns, and that is to try and determine whether benchmark findings of laboratory research can be replicated online (e.g., Bui et al., 2015; Crump et al., 2013; Goodman et al., 2013) before proceeding to examine more novel experimental paradigms, and this is the approach used here.

Whereas the efficiency of online recruitment and data collection is an advantage in many research areas, it can be especially so in research with older adults. Currently, coming into the laboratory for testing poses special risks for older adults who may not be more likely than younger adults to become infected by the Covid-19 virus, but for whom infection may have more serious consequences., In St. Louis this year, for example, older adults have been around 50 times as likely as younger adults to die from a Covid-19 infection. Risks aside, older adults are typically paid for participating in behavioral research, unlike the undergraduate participants in many studies, but MTurk workers, regardless of age, are willing to participate for much less than the older adults who participate in laboratory studies. In addition, whereas collecting data in the laboratory often requires research assistants to schedule appointments and administer the experimental tasks, data collection on MTurk requires only that someone monitor the process at their convenience.

Online testing has other advantages as well, but it would all be for naught if the data collected were either unreliable or unrepresentative of the population of interest (e.g., if the cognitive abilities of online samples of older adults exceeded those of laboratory samples). Moreover, online testing is typically unsupervised, which poses special problems for memory studies because, regardless of age, participants could conceivably write down to-be-remembered items to aid them in later recall. Accordingly, the present study of age-related differences in WM focused on the visuospatial domain, and to-be-remembered locations were not presented in a grid because if they were, a participant might record them as easily as they could record a list of words.

We would note, however, that the present study of WM and the putative inhibitory deficit did not focus on the visuospatial domain for purely methodological reasons, important as they are here. The ability to process visuospatial information and maintain it in WM declines much more rapidly with age than the corresponding verbal abilities (Hale et al., 2011; Jenkins, Myerson, Joerding, & Hale, 2000; ; Lawrence, Myerson & Hale, 1998) performance that provide an excellent opportunity to assess Hasher and Zacks' (1988) inhibitory-deficit hypothesis. Notably, almost all previous research on aging, inhibition, and

WM has focused on verbal tasks (cf. Rowe et al., 2008), and thus the present study also addresses the important theoretical issue of the extent to which the aging of WM, and older adults' susceptibility to interference with WM in particular, involves domain-specific versus domain-general processes.

Hasher and Zacks (1988), in a landmark book chapter that has been cited well over 3,000 times, proposed that older adults suffer from a general inhibitory deficit that allows irrelevant information access to WM and hobbles efforts to delete it, which in turn has consequences for many other aspects of cognition as well. Nearly 20 years later, they concluded that their inhibitory-deficit perspective was "alive and well" (Lustig et al., 2007). Others, however, have long disagreed with this view (e.g., Burke, 1997; Rey-Mermet & Gade, 2017; Rey-Mermet et al., 2018; Verhaeghen, 2011).

How can researchers continue to hold opposite views for decades? One contributing factor may be the unlevel playing field on which proponents and opponents of the inhibitory-deficit hypothesis compete. Whereas proponents can point to statistically significant differences between younger and older samples, the best that opponents can say is that the null hypothesis could not be rejected. This situation is analogous to that in the American criminal justice system. One can be found *guilty* or *not guilty*, but not *innocent*, unlike in the Scottish system, which distinguishes between *not guilty* and *not proven*. As Leipold (2000) wrote regarding the American system, "When it comes to acquittals, it would be hard to devise a verdict that is less informative than the one currently in use," and the same could be said of the statistical verdict, "failure to reject the null hypothesis." (p. 1302)

Bayesian statistics has changed this situation, even going the Scottish system one better by quantifying evidence for and against competing hypotheses (for discussion of the profound implications, see Morey et al., 2016). Recently, researchers in gerontology have noted the applicability of Bayesian analyses to the question of whether or not there are age differences on particular measures because unlike traditional null hypothesis tests, such analyses allow one to compare support for the null hypothesis with that for hypothesized age differences (Brydges & Bielak, 2019; Isaacowitz, 2018; Lakens et al., 2018). A recent meta-analysis of age-related differences on eleven tasks commonly assumed to measure inhibition (e.g., the Stroop task) by Rey-Mermet and Gade (2017) was the first to apply this approach to the inhibitory-deficit hypothesis. Whereas Bayesian analyses indicated that older adults were 'guilty' of having an inhibitory deficit in two cases (stop-signal and go/no-go tasks, the verdict was 'not proven' in four cases and 'not guilty' (no deficit) in five. Although these findings are inconsistent with a *general* deficit, they increase the theoretical importance of determining under exactly what conditions inhibitory deficits are and are not observed.

The tasks examined by Rey-Mermet and Gade (2017) all assessed processing speed, however, whereas the inhibitory-deficit hypothesis originated with consideration of inhibition's role in working memory (WM). Subsequently, the hypothesis has been applied more broadly, largely because of WM's role in so many laboratory and nonlaboratory tasks (for a review, see Lustig et al., 2007). Given the evidence that although age-related differences in inhibitory ability may be observed, they are not general in nature, it seems appropriate to extend the search for boundary conditions by examining WM tasks, adapting

Rey-Mermet and Gade's (2017) two-pronged analytic approach. This approach combines null hypothesis statistical tests (NHSTs) of regression models with Bayesian model selection in order to determine in which cases the evidence supports the inhibitory-deficit hypothesis, in which it supports the null (no-deficit) hypothesis, and in which cases there is insufficient evidence to decide.

The current investigation consists of two experiments. Experiment 1 focused on a well-established WM phenomenon, interference from secondary tasks (Engle et al., 1999), and the potential role of such interference in the steep decline of visuospatial WM with age (Hale et al., 2011). In contrast, Experiment 2 involved a relatively novel, but reliable phenomenon: the benefit to visuospatial WM provided by environmental support for visuospatial rehearsal (e.g., Lilienthal et al., 2014; 2016). Specifically, support for rehearsal in visuospatial memory is provided by having an array of possible locations remain visible during the inter-item intervals between presentation of to-be-remembered locations. Experiment 2 contrasted the benefits of such support with the interference caused by presenting conflicting spatial information during inter-item intervals. By examining various forms of interference and using Bayesian analysis to obtain a balanced assessment of age-related differences in each case, we sought to determine the extent to which the putative inhibitory deficit is truly general in nature. More specifically, we examined whether the hypothesized deficit generalizes to visuospatial WM and whether it does so regardless of the manipulation leading to a need to exercise inhibitory control.

Experiment 1

Recent experiments have validated studying cognitive aging online, and this approach has many advantages (Bui et al., 2015). Visuospatial WM is well-suited for online research because, unlike with verbal WM, the memory items typically are difficult for unsupervised participants to keep a written record of. Accordingly, we recruited participants using Amazon's Mechanical Turk (MTurk) and compared rates of age-related decline on simple and complex versions of the same visuospatial memory span task, where the complex version interleaved a secondary task between to-be-remembered items. The difference between scores on the simple and complex versions indicates how much interference the secondary task creates. The primary issue here is whether or not that difference increases with age, reflecting increased interference by irrelevant secondary task information, as predicted by the inhibitory-deficit hypothesis.

Method

Participants—We recruited 107 participants ranging in age from 19–74 (69 female; M_{age} = 42.9 years, SD = 14.3; $M_{education}$ = 15.1 years, SD = 1.9) using a procedure like that in Experiment 2 of Bui et al. (2015) to obtain a relatively flat age distribution. More specifically, we initially recruited adults of all ages, keeping track of how many participants there were in each 10-yr band beginning at 18 (i.e., 18–27, 28–37, etc.), and as soon as there were a minimum of 20 volunteers in a band, we stopped accepting further volunteers in that age range. The present sample size was more than three times that needed for adequate power to detect a significant correlation between age and memory span at least the size of the correlations reported by Hale et al. (2011) for simple and complex versions of a visuospatial WM task similar to that used here. In fact, our sample provided adequate power to detect a weak correlation of .27.

Bayesian analysis to assess whether a nonsignificant difference in slopes (i.e., changes in span with age) reflected evidence so weak that it should be considered merely anecdotal, or whether the results actually supported the null hypothesis. The use of *BF*s does not eliminate the benefits of power analysis for planning experiments (Lakens et al., 2018). However, *BF*s provide quantitative comparisons of the support, or lack thereof, for the hypothesized outcomes, thereby allowing one to distinguish nonsignificant results that might result from a lack of power from results that although nonsignificant, actually support the null hypothesis (see the following *Analysis* subsection for more detail).

Materials and Procedure—In one session approximately 20-min long, participants performed Simple and Complex conditions of a visuospatial memory span task similar to the Dot Span task (Hale et al., 2011). Online stimulus presentation was programmed in Adobe ActionScript by Y. B. and presented using Adobe Flash Player. Each condition began with four practice trials, followed by eight test trials on which a subset of the circles in a 30-circle array turned red, one at a time, and participants were instructed to remember their locations. Each to-be-remembered circle turned red for 1.5 s, followed by a 1.5-s inter-item interval during which the array remained on the screen (see the diagram labelled Same-Array in Fig. 1, which depicts the procedure for the Simple condition). In the Complex condition only, lines appeared on each side of the array during inter-item intervals, and participants had to decide whether or not they were the same length. A different array of 30 circles, arranged to appear unstructured, was used for each trial.

On each trial, participants recalled the locations of the circles that had turned red during that trial by clicking on them using the computer mouse. List lengths were four and six to minimize the likelihood of ceiling effects, especially in younger participants in the Simple condition of the task, and floor effects, especially in older participants in the Complex condition. Participants completed eight trials in each condition, four at each length.

Analysis—Scores for each condition were calculated as the total number of correctly recalled locations at both list-lengths (Max = 40). One 53-year-old participant's data were removed because their score for the Simple condition was a significant outlier ($z = -3.97$, $p < .01$). Regression analyses, with Age as the independent variable, were conducted using SigmaPlot's Regression Wizard (Systat Software) to compare a full model that instantiated the inhibitory-deficit prediction with different slopes and intercepts for the Simple and Complex conditions (Ms=32.5 and 20.7, respectively; SDs=4.7 and 5.4) and a reduced model that instantiated the null hypothesis, with two intercepts but just one slope, reflecting the fact that according to the reduced model, scores in both conditions decline at the same rate. Results were assessed using both Null Hypothesis Significance Tests (NHSTs) and Bayesian analyses.

Following Rey-Mermet and Gade (2017), approximations of the Bayes Information Criteria for the two models (Wagenmakers, 2007) were used to calculate *BF*s representing the relative likelihoods of the two models given the data (Jarosz & Wiley, 2014). In the current study, for example, a $BF_{1,0} = 4.0$ would indicate that, given the data, it is four times as likely as not that there is an inhibitory deficit; its reciprocal (indicated by the reversal in the subscripts), $BF_{0,1} = 0.25$, restates that result as indicating that the null hypothesis of no deficit is one-fourth as likely given the data. In contrast, *BF*s between 3.0 and 0.33 and would be considered merely anecdotal evidence for either hypothesis (Jeffreys, 1961), analogous to a verdict of 'not proven.' Importantly, the same evidentiary standards apply to both hypotheses, thus providing for a balanced assessment of the truth of the two alternative hypotheses.

Results and Discussion

When online research is initiated that uses a new procedure or that is in a new area, it is usually validated by replicating benchmark phenomena, and as expected based on previous laboratory studies (e.g., Engle et al., 1999; Hale et al., 2011), the secondary task in Experiment 1 consistently produced interference. In fact, every participant but one recalled more locations in the Simple condition than the Complex condition. These effects may be seen in Figure 2, which depicts age-related declines in WM. Both the full and reduced regression models were significant: $F(3, 208) = 88.31$, $p <. 01$, and $F(2, 209) = 132.60$, $p < .01$). Importantly, however, the additional slope parameter in the full model was not significant ($t < 1.0$), indicating that the evidence for different slopes in the Simple and Complex conditions, as predicted by the inhibitory-deficit hypothesis, was insufficient (see Table 1).

Failure to reject the null hypothesis (i.e., the slopes are not different), of course, does not qualify as evidence for the null, but Bayesian analyses revealed a $BF_{0,1}$ of 11.61, indicating that given the data, it is more than ten times as likely that there is no inhibitory deficit on visuospatial WM tasks than that there is. According to guidelines originally suggested by Jeffreys (1961) and now widely adopted (Jarosz & Wiley, 2014), these results represent strong evidence for the null hypothesis that there is no age-related inhibitory deficit in visuospatial WM. Taken together with the fact that Hale et al. (2011), who examined three pairs of simple and complex visuospatial WM tasks, also found no consistent difference between the slopes for simple and complex versions, the results argue strongly against the hypothesis that as people age, their visuospatial WM becomes increasingly susceptible to interference.

Experiment 2

After validating online assessment of age differences in visuospatial WM, we examined a relatively new WM phenomenon, the effect of environmental support on visuospatial rehearsal of to-be-remembered locations (Lilienthal et al., 2014; 2016) as well as one not studied previously, the effect of incongruent visuospatial information. Lilienthal and colleagues showed that visuospatial WM performance is better when the array of possible locations remains visible during inter-item intervals (see Figure 1). Experiment 2 sought to determine whether, if the array changes during inter-item intervals, there are age-related differences in the ability to inhibit this incongruent, potentially interfering, location information.

To answer this question, we examined performance in three conditions of a visuospatial WM task similar to that in Experiment 1 (see the Same Array condition in Fig. 1). The Different Array condition presented incongruent visuospatial information that was expected to interfere with WM, and interference was measured as the difference between performance in this condition and that in two baseline conditions, an Array Absent condition as well as the Same Array condition. At issue was whether participants' age, their baseline performance, or both, would predict how much interference was observed.

Method

Participants—We recruited 32 younger (19 female; $M_{\text{age}} = 25.0$, $SD = 2.5$) and 32 older (22 female; $M_{\text{age}} = 64.0$, $SD = 3.8$) MTurk workers. The two age groups did not differ in level of education (high school, some college, college, postgraduate): $X^2(3) = 1.81$, $p = .612$. Given the novelty of the paradigm used, an expected effect size is difficult to estimate, and therefore a failure to find significant differences could be interpreted as due to lack of power. This issue was addressed, however, by calculating *BF*s that compare the relative likelihood of the obtained data given each of the two models under consideration, one based on the inhibitory deficit hypothesis and the other based on the null hypothesis.

Materials and Procedure—In a single session lasting approximately 40 minutes, participants performed three conditions of a visuospatial span task similar to the Simple condition in Experiment 1 (see Fig. 1). One condition (Same-Array) was identical to the Simple condition in Experiment 1 except that the inter-item intervals, during which the array of 30 possible locations remained visible, were 4-s long to provide more time for rehearsal. In a second condition (Array-Absent), the screen went blank during inter-item intervals. In the third condition (Different-Array), the 30-circle array seen during inter-item intervals was different from that when the to-be-remembered locations were presented. Each condition began with four practice trials followed by eight test trials, half with four to-be-remembered locations and half with seven. The order of conditions was randomized across participants. As in the first experiment, online stimulus presentation was programmed in Adobe ActionScript by Y. B. and presented using Adobe Flash Player.

Analysis—Scores for each condition were calculated as the number of correctly recalled locations summed across both list lengths (Max.= 44). Data from one younger and one older participant were excluded from the analyses because of very low scores in the Array-Absent condition (both *z*s < -2.8). A mixed-design Analysis of Variance (ANOVA) was conducted first, followed by multiple regression and Bayesian analyses comparing full and reduced state-trace regression models. The ANOVA was conducted using SPSS and as in Experiment 1, regression analyses were conducted using SigmaPlot.

The reduced models assumed that regardless of age, interference decreased participants' scores from their baseline scores by some amount, I, that was a linear function of their baseline scores (i.e., $I = mX + b$) in order to accommodate the suggestion that individuals with higher spans are more resistant to proactive interference (Hasher & Zacks, 1988; Lustig et al., 2007). Thus, the reduced model was Y = X − I = X − (*m*X + *b*) = (1 − *m*)X + *b*, where X is the baseline score and Y is the score in the Different-Array condition. The full model added a second intercept parameter, *db*, to capture the hypothesis that, after statistically controlling for differences in baseline, younger and older adults differed in the amount of interference.

Results and Discussion

Planned comparisons of Different-Array scores with Same-Array and Array-Absent scores, conducted as manipulation checks, confirmed that viewing a different array interfered with visuospatial WM, regardless of which baseline condition was used to measure interference, ts(61) > 5.0, ps < .001. A 2(Age) × 3(Condition) ANOVA revealed effects of Age, $F(1, 60) = 22.1$, $p < .001$, and Condition, $F(2, 120) = 182.5$, $p<.001$, but no interaction, $F(2, 120)= 0.245$, $p = .783$, indicating that, contrary to the inhibitory-deficit hypothesis, age-related differences were not significantly larger in the Different-Array condition (see Table 2). The current Same-Array and Array-Absent conditions represent conceptual replications of previous laboratory studies of this phenomena (e.g., Lilienthal et al., 2014; 2016), and as expected, providing time and environmental support for rehearsal increased participants' visuospatial WM.

State-trace regressions of scores in the Different-Array condition on scores in the Same-Array and Array-Absent conditions confirmed these conclusions (see Fig. 3): Contrary to the inhibitory-deficit hypothesis, there were no significant age-related differences in the intercepts (i.e., *db* was not significant; see Table 3). Finally, Bayesian analyses comparing full and reduced state-trace regression models provided substantial evidence for the null hypothesis that there is no age-related inhibitory deficit in visuospatial WM. Regardless of whether the Same-Array or the Array-Absent condition was the baseline against which interference was measured, reduced models in which the amount of interference was independent of age were more than three and six times as likely to be correct, respectively, than the corresponding full models ($BF_{0,1}$ = 3.56 and 6.66 for the Array-Absent and Same-Array conditions, respectively).

General Discussion

The primary goals of the present study were first, to evaluate the effectiveness of online samples for the purpose of studying age-related differences in WM, and second, to provide a balanced assessment of Hasher and Zack's hypothesis of a general, age-related inhibitory deficit (Hasher & Zacks, 1988; Lustig et al., 2007) by supplementing NHSTs with Bayesian analyses of age effects (Rey-Mermet & Gade, 2017). Two online experiments using different manipulations that interfered with visuospatial WM first replicated results originally obtained in the laboratory and then used Bayesian analyses to compare the strength of the evidence for the null hypothesis with that for the inhibitory-deficit hypothesis. These analyses showed that, in both cases, the evidence supported the null and not the inhibitory-deficit hypothesis. To revisit the Scottish law metaphor, older adults were tried and found *innocent* of a general inhibitory deficit, a far stronger conclusion than *not proven*, the strongest negative verdict possible with NHSTs.

Efficacy of Online WM Research

In a previous laboratory study, Hale et al. (2011) examined three pairs of simple and complex visuospatial WM tasks and found no consistent difference between the rates of age-related decline in performance on the simple and complex versions. Despite the presence of interference from a secondary task in the complex version but not the simple version, the rates of age-related decline did not differ significantly, contrary to the predictions of the inhibitory-deficit hypothesis. Not only was that finding replicated in Experiment 1, but the rates of decline were equivalent to those observed in previous studies with large samples.

With WM scores converted to z-scores in order to facilitate comparisons, the average rate of decline in the Simple and Complex conditions of the present online study was −0.027 z-scores/year, whereas in the laboratory study by Hale et al. (2011; $N = 388$), the average rate of decline on the analogous Simple and Complex tasks (Dot Span and Position Span; see their Fig. 2) was −0.026 z-scores/year. These rates are similar to the −0.025 z-scores/year observed with Spatial Span in the WMS III standardization sample (Myerson, Emery, White, & Hale, 2003; $N = 1050$). The correspondence between the rates of decline on the WMS III Spatial Span measure, the Hale et al. visuospatial WM tasks, and the present Simple and Complex tasks both validates the present online approach and demonstrates the robustness of the high rate of decline in visuospatial WM: Spans decreased approximately one full z-score from age 35 to 75.

Further support for the effectiveness of online WM research is provided by the replication in Experiment 2 of previous laboratory findings regarding the effects of environmental support for visuospatial rehearsal in younger adults (Lilienthal et al., 2014) as well as the finding that older adults' visuospatial WM performance also benefits from environmental support (Lilienthal et al., 2016). Further research is needed that assesses age-related differences in even older online samples. Currently, very old MTurk workers are lacking, but as might be expected, this is already changing. Indeed, one older adult in Experiment 2 was 80 years old, and consistent with Lilienthal et al.'s (2016) laboratory study of age-related differences in visuospatial WM, their performance also improved when environmental support was provided.

Although we consider the present study a demonstration of successful online data collection, one should be mindful of the different circumstances under which people participate in online research. Their devices likely differ, the internet systems to which they are connected likely differ as well, and the conditions under which they provide data, from bedrooms to Starbucks, will likely differ as well. How can one have confidence in data collected under such diverse circumstances?

We have three suggestions for those doing online research, all of which were followed here. The first is to begin with online replications of benchmark laboratory findings. This strategy has been followed by those pioneering online research in various research areas (e.g., Bui et al., 2015; Paolacci, Chandler, & Ipeirotis 2010), and it may reveal which phenomena are sufficiently robust that they can be studied online despite the added noise provided by diverse data collection circumstances and which are not (Crump, McDonnell, & Gureckis, 2013). Whereas Experiment 1 was a systematic replication of part of Hale et al.'s large laboratory study of WM relevant to the question of inhibitory deficits, Experiment 2 not only replicated laboratory findings by Lilienthal et al. (2014; 2016), but with our concerns about online WM research largely abated, added a novel condition not previously studied in the laboratory that specifically represented a test of the inhibitory deficit hypothesis.

Our second suggestion is that designs that expose different participants to different experimental conditions designs will be more affected by differences in the circumstances under which people participate than designs that compare a participant's behavior under different experimental conditions, and both experiments of the present study used such designs. Finally, because the circumstances under which participants provide data online may make them at risk of distraction, it is best to make experiments brief (e.g., compare the WM procedures in Experiment 1 with those in Hale et al., 2011). The briefer the experiment, the lower the likelihood of distraction, and more experiments with different samples and less conditions (e.g., Bui et al., 2015) may produce higher quality data and thus be more efficient in the long run.

Implications for the Inhibitory-Deficit Hypothesis

Although the present findings leave issues regarding an inhibitory deficit affecting verbal WM unresolved, they argue against a strong form of the inhibitory-deficit hypothesis in which deficits are *domain-general*, affecting both verbal and visuospatial WM. Instead, the results of Experiment 1 indicate that the steep declines in visuospatial WM with age (e.g., Hale et al., 2011) are not attributable to an increasing inhibitory deficit. Scores on a WM span task with a secondary task interleaved between to-be-remembered items (i.e., the Complex condition) and scores on the same span task without a secondary task (i.e., the Simple condition) both declined at the same rate, contrary to the prediction that a secondary task would lead to a faster decline because of a growing deficit in the ability to inhibit secondary-task information. The present results replicate previous results for three different visuospatial span tasks (Hale et al., 2011) but go beyond the previous finding that the rates of decline on simple and complex WM tasks did not differ significantly. Instead, Bayesian analysis revealed that the present data provide strong evidence for the null hypothesis of equivalent rates of decline, at least until age 75.

The results of the second experiment are consistent with those of Experiment 1. In Experiment 2, the irrelevant spatial stimuli interleaved between to-be-remembered items produced significant interference with visuospatial WM. Although the inhibitory deficit hypothesis predicted older adults would suffer more interference due to a decline in their ability to initially ignore or subsequently delete the irrelevant spatial information, no significant differences were observed between older and younger groups. However, two separate Bayesian analyses, each using a different condition as the baseline against which interference was measured, revealed substantial evidence for the null hypothesis.

Taken together, the present experiments, which not only required inhibition of different kinds of information, but which also used different kinds of designs (e.g., age as a continuous variable versus age as a dichotomous variable) provided converging evidence for the null hypothesis. Although the findings may still be specific to the situations examined, any situations in which equivalent interference with visuospatial WM is observed in younger and older adults are evidence that the hypothesized inhibitory deficit is not truly general in nature as previously claimed (e.g., Lustig et al., 2007). In particular, given that older adults' visuospatial abilities, including not just their WM (Hale et al., 2011) but also their information-processing speed (Lawrence et al., 1998) and visuospatial learning ability (Jenkins et al. 2000) typically decline much more rapidly than their verbal abilities, the finding that interference is not responsible for the rapid decline of visuospatial WM abilities argues strongly against a general explanation of age-related cognitive differences based on the assumption of an inhibitory deficit.

Boundary Conditions

This issue of whether the inhibitory-deficit hypothesis provides an adequate basis for a general theory of cognitive aging (Lustig et al., 2007) is, of course, not limited to the question of whether the putative deficit includes the visuospatial as well as the verbal domain. A recent meta-analysis (Rey-Mermet & Gade, 2017) using the same two-pronged analytic strategy (NHSTs supplemented by Bayesian analyses) found multiple cases in which younger and older adults showed equivalent interference effects on speeded tasks, including on the iconic Stroop task. This evidence against a general deficit is perhaps not surprising, as correlational analyses have long suggested that inhibition is not a general construct (Kramer, Humphrey, Larish, Logan, & Strayer, 1994), and even measures of the Stroop effect may not correlate if obtained using different tasks (Shilling, Chetwynd, & Rabbitt, 2002).

Most recently, structural equation modeling of data from multiple speeded tasks identified inhibition of prepotent responses and resistance to distracter interference as separate factors. Contrary to the inhibitory-deficit hypothesis, however, older adults actually were better at resisting distracter interference than younger adults although they were worse at the inhibition of prepotent responses (Rey-Mermet, Gade, & Oberauer, 2018). Not only was inhibition not a single factor, but even a two-factor inhibitory model had low explanatory power, reflecting the fact that many correlations between inhibition measures were not significant, replicating the findings of Kramer et al.'s multi-task study (1994). Rey-Mermet et al. (2018) concluded that, taken together, their findings regarding speeded tasks "challenge the hypothesis of a general inhibition deficit in older age."

It should be noted that whereas proving a general claim requires showing that it is true in multiple, diverse situations, disproving a general claim can require only evidence from one case – a single talking dog would disprove the claim that speech is uniquely human. With respect to WM, the original focus of Hasher and Zach's (1988) inhibitory deficit hypothesis, each experiment in the present study presents evidence of a 'talking dog.' The extent to which a secondary task interfered with WM was the same regardless of participants' age (Expt. 1), and irrelevant location information had equivalent effects on both younger and older adults' WM (Expt. 2). The results of the first experiment may represent stronger evidence in that they replicate those of a laboratory study that had a larger sample (N=388) of adults who varied more widely in age (from 20 to 89 years) that also examined the effects of secondary tasks on performance on three different visuospatial WM tasks (Hale et al., 2011). Taken together, these findings argue strongly that the effects of the hypothesized age-related deficit in inhibitory function on visuospatial WM are not general in nature.

Lustig, Hasher, and Zachs distinguish three inhibitory functions related to access, deletion, and restraint (Lustig et al., 2007). The first experiment of the present study focuses on deletion – complex span tasks like that in Experiment 1 require first allowing secondary task information temporary access to working memory and then deleting the information once it becomes irrelevant. The second experiment, on the other hand focuses on the access function in that inhibition is used to restrict the access of irrelevant information in the first place. In neither experiment did we see evidence of impaired inhibitory function. Notably, however, questions concerning the third putative inhibitory function, restraint as evidenced by inhibition of prepotent responses (e.g., Butler & Zacks, 2006), were not addressed in either experiment, and age-related differences in the effects of the restraint function on visuospatial WM tasks may well exist. However, even if they do, this would not alter our conclusions regarding the hypothesis of a *general* age-related inhibitory deficit, with the emphasis here on the word *general*. This is because our point is not that age-related inhibitory deficits do not exist, but rather that such deficits are not observed in all situations where inhibition is involved, as the present findings regarding visuospatial WM demonstrate, and the hypothesis of a general deficit would predict.

Given that age-related deficits in inhibitory function do not appear to be a truly general phenomenon, research is needed to establish the boundary conditions for putative age-related differences in inhibitory functions. The potential promise of this line of research is suggested by recent meta-analyses (Rey-Mermet & Gade, 2017; Verhaeghen, 2011) examining age differences on speeded tasks, analyses that go beyond testing a general null hypothesis against a general inhibitory-deficit hypothesis. Instead, they suggest that what is needed theoretically is a hypothesis about when age-related deficits will be observed and when they will not, and perhaps most importantly, why.

Efforts to determine the task characteristics that make older adults more vulnerable to interference may benefit from the present approach of supplementing NHSTs with Bayesian analyses (Rey-Mermet & Gade, 2017). Establishing the boundary that defines a class (e.g., task situations that differentially disadvantage older adults) requires not just evidence that a deficit is observed when certain conditions are met, but also evidence that no deficit is observed when those conditions are not met. The strength of Bayesian analyses is that, unlike NHSTs, they are equally capable of providing evidence in either case and of quantifying the strength of that evidence.

Conclusion

The long-standing issue of the generality of age-related differences in inhibition and their effects on WM, which was the original focus of the inhibitory-deficit hypothesis, is similar to the issue concerning age-related differences in inhibition on speeded tasks (e.g., Rey-Mermet & Gade, 2017; Rey-Mermet et al., 2018); Verhaeghen, 2011; Verhaeghen & Cerella, 2002). In both cases, inhibitory deficits have been reported, but they appear to be task- and process-specific. Consistent with this view, Rowe et al. (2008) reported that older adults' WM for locations was more susceptible to proactive interference than that of younger adults, in contrast to the equivalent interference with visuospatial WM observed in younger and older adults when the interfering events were secondary tasks (Expt. 1) and incongruent spatial information (Expt. 2). The present results converge with those from previous laboratory experiments, providing evidence that conducting research in visuospatial WM online with older adults is possible using both familiar and novel WM experimental paradigms. Given that visuospatial abilities typically decline much more rapidly with age than verbal abilities (Hale et al., 2011; Lawrence et al., 1998), strategies for minimizing interference with visuospatial WM could prove especially helpful to older adults, and an understanding of the boundary conditions on age-related inhibitory deficits like those demonstrated in the present online study could suggest such strategies.

Acknowledgements

Preparation of this manuscript was supported in part by the National Institute on Aging of the National Institutes of Health under Award Number R01AG058885.

References

Brydges CR, & Bielak AAM (2019). A Bayesian Analysis of Evidence in Support of the Null Hypothesis in Gerontological Psychology (or Lack Thereof). The Journals of Gerontology: Series B. doi:10.1093/geronb/gbz033

Bui DC, Myerson J, & Hale S (2015). Age-related slowing in online samples. The Psychological Record, 65, 649–655. doi:10.1007/s40732-015-0135-2

Burke DM (1997). Language, aging, and inhibitory deficits: Evaluation of a theory. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 52B, P254–P264. 10.1093/geronb/52B.6.P254

Butler KM, & Zacks RT (2006). Age deficits in the control of prepotent responses: evidence for an inhibitory decline. Psychology and Aging, 21, 638–643. 10.1037/0882-7974.21.3.638 [PubMed: 16953726]

Chandler J, & Shapiro D (2016). Conducting clinical research using crowdsourced convenience samples. Annual Review of Clinical Psychology, 12, 53–81. doi: 10.1146/annurev-clinpsy-021815-093623

Crump MJ, McDonnell JV, & Gureckis TM (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. PloS ONE, 8, e57410. 10.1371/journal.pone.0057410 [PubMed: 23516406]

Engle RW, Tuholski SW, Laughlin JE, & Conway ARA (1999). Working memory, short term memory and general fluid intelligence: A latent variable approach. Journal of Experimental Psychology: General, 128, 309–331. doi:10.1037/0096-3445.128.3.309 [PubMed: 10513398]

Goodman JK, Cryder CE, & Cheema A (2013). Data collection in a flat world: The strengths and weaknesses of Mechanical Turk samples. Journal of Behavioral Decision Making, 26, 213–224. 10.1002/bdm.1753

Hale S, Rose NS, Myerson J, Strube MJ, Sommers M, Tye-Murray N, & Spehar B (2011). The structure of working memory ability across the adult life span. Psychology & Aging, 26, 92–110. doi:10.1037/a0021483 [PubMed: 21299306]

Hasher L, & Zacks RT (1988). Working memory, comprehension, and aging: A review and a new view. In Bower GH (Ed.), Psychology of learning and motivation (Vol. 22, pp. 193–225). San Diego, CA: Academic Press. doi:10.1016/S0079-7421(08)60041-9

Isaacowitz DM (2018). Planning for the future of psychological research on aging. The Journals of Gerontology: Series B, 73, 361–362. doi:10.1093/geronb/gbx142

Jarosz AF, & Wiley J (2014). What are the odds? A practical guide to computing and reporting Bayes factors. The Journal of Problem Solving, 7, 2–9. doi:10.7771/1932-6246.1167

Jeffreys H (1961). Theory of probability (3rd ed.). Oxford: Oxford University Press, Clarendon Press.

Jenkins L, Myerson J, Joerding JA, & Hale S (2000). Converging evidence that visuospatial cognition is more age-sensitive than verbal cognition. Psychology and Aging, 15, 157. doi: 10.1037//0882-7974.15.1.157. [PubMed: 10755297]

Kramer AF, Humphrey DG, Larish JF, Logan GD, & Strayer DL (1994). Aging and inhibition: beyond a unitary view of inhibitory processing in attention. Psychology and Aging, 9, 491–512. doi:10.1037/0882-7974.9.4.491 [PubMed: 7893421]

Lakens D, McLatchie N, Isager PM, Scheel AM, & Dienes Z (2018). Improving Inferences about Null Effects with Bayes Factors and Equivalence Tests. The Journals of Gerontology: Series B, 75, 45–57. doi:10.1093/geronb/gby065

Lawrence B, Myerson J, & Hale S (1998). Differential decline of verbal and visuospatial information processing speed across the adult life span. Aging, Neuropsychology & Cognition, 5, 129–146. doi:10.1076/anec.5.2.129.600

Leipold AD (2000). The problem of the innocent, acquitted defendant. Northwestern University Law Review, 94, 1297–1356.

Lilienthal L, Hale S, & Myerson J (2014). The effects of environmental support and secondary tasks on visuospatial working memory. Memory & Cognition, 42, 1118–1129. doi:10.3758/s13421-014-0421-2 [PubMed: 24874509]

Lilienthal L, Hale S, & Myerson J (2016). Effects of age and environmental support for rehearsal on visuospatial working memory. Psychology and Aging, 31, 249–254. doi:10.1037/pag0000077 [PubMed: 26950223]

Lustig C, Hasher L, & Zacks RT (2007). Inhibitory deficit theory: Recent developments in a "new view." In Gorfein DS & MacLeod CM (Eds.), Inhibition in cognition (pp. 145–162). Washington, DC: American Psychological Association. doi:10.1037/11587-008

Morey RD, Romeijn JW, & Rouder JN (2016). The philosophy of Bayes factors and the quantification of statistical evidence. Journal of Mathematical Psychology, 72, 6–18. doi:10.1016/j.jmp.2015.11.001

Myerson J, Baumann AA, & Green L (2014). Discounting of delayed rewards: (A)theoretical interpretation of the Kirby questionnaire. Behavioural Processes, 107, 99–105. 10.1016/j.beproc.2014.07.021 [PubMed: 25139835]

Myerson J, Emery L, White DA, & Hale S (2003). Effects of age, domain, and processing demands on memory span: Evidence for differential decline. Aging, Neuropsychology, and Cognition, 10, 20–27. doi:10.1076/anec.10.1.20.13454

Necka EA, Cacioppo S, Norman GJ, & Cacioppo JT (2016). Measuring the prevalence of problematic respondent behaviors among MTurk, campus, and community participants. PloS one, 11(6), e0157732. 10.1371/journal.pone.0157732 [PubMed: 27351378]

Paolacci G, Chandler J, & Ipeirotis PG (2010). Running experiments on Amazon Mechanical Turk. Judgment and Decision Making, 5, 411–419.

Rey-Mermet A, & Gade M (2017). Inhibition in aging: What is preserved? What declines? A meta-analysis. Psychonomic Bulletin & Review, 25, 1695–1716. doi:10.3758/s13423-017-1384-7

Rey-Mermet A, Gade M, & Oberauer K (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. Journal of Experimental Psychology: Learning, Memory, and Cognition, 44, 501–526. doi:10.1037/xlm0000450

Rowe G, Hasher L, & Turcotte J (2008). Age differences in visuospatial working memory. Psychology and Aging, 23, 79–84. doi:10.1037/0882-7974.23.1.79 [PubMed: 18361657]

Shilling VM, Chetwynd A, & Rabbitt PMA (2002). Individual inconsistency across measures of inhibition: An investigation of the construct validity of inhibition in older adults. Neuropsychologia, 40, 605–619. doi:10.1016/S0028-3932(01)00157-9 [PubMed: 11792402]

Tremblay S, Saint-Aubin J, & Jalbert A (2006). Rehearsal in serial memory for visual-spatial information: Evidence from eye movements. Psychonomic Bulletin & Review, 13, 452–457. 10.3758/BF03193869 [PubMed: 17048730]

Verhaeghen P (2011). Aging and executive control: Reports of a demise greatly exaggerated. Current Directions in Psychological Science, 20, 174–180. doi:10.1177/0963721411408772 [PubMed: 25866452]

Verhaeghen P, & Cerella J (2002). Aging, executive control, and attention: A review of meta-analyses. Neuroscience & Biobehavioral Reviews, 26, 849–857. doi: 10.1016/s0149-7634(02)00071-4. [PubMed: 12470697]

Wagenmakers E-J (2007). A practical solution to the pervasive problems of *p* values. Psychonomic Bulletin & Review, 14, 779–804. doi:10.3758/BF03194105 [PubMed: 18087943]

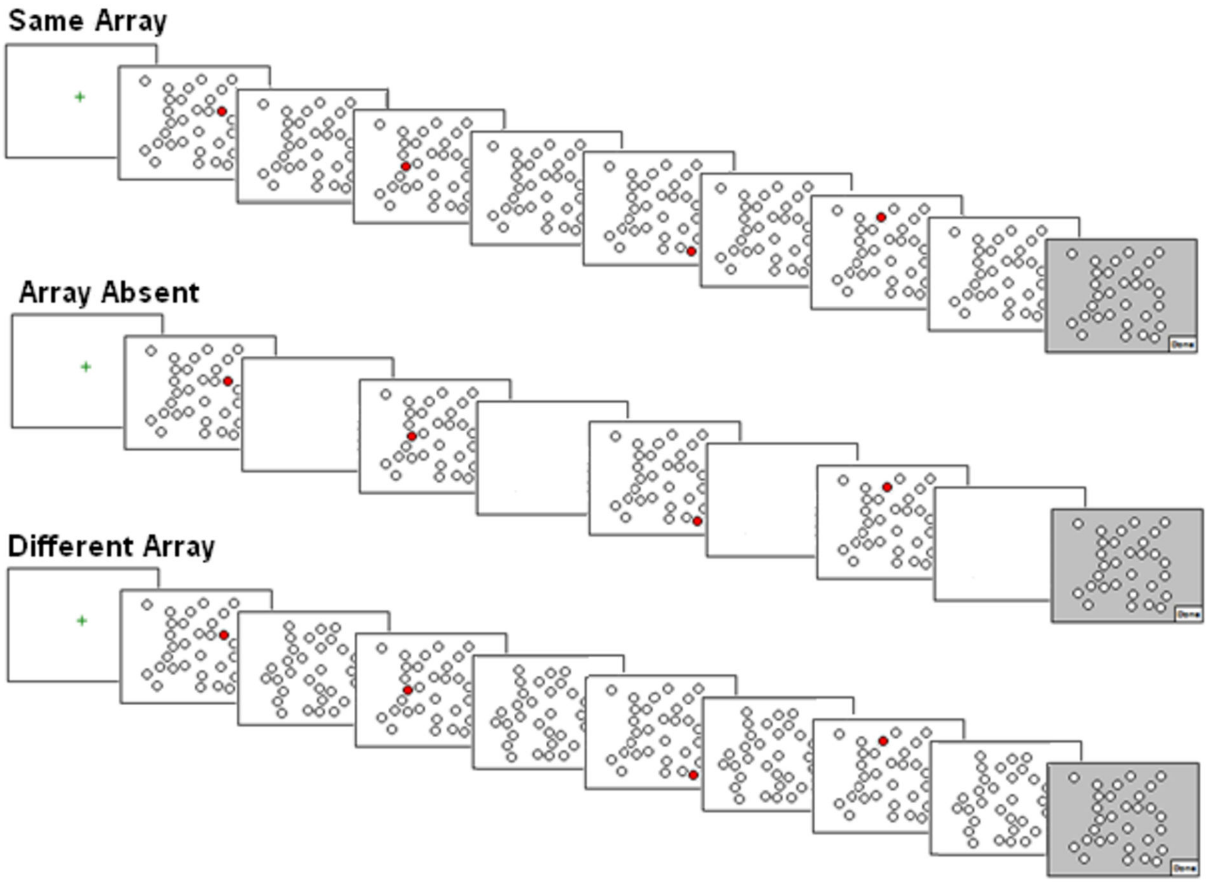

Figure 1.
Diagrams of memory span procedures used in Experiment 1 (Same Array procedure) and Experiment 2 (all three procedures: Same Array, Array Absent, and Different Array). The examples shown are for lists of four to-be-remembered locations (red circles). Participants indicated which locations they recalled by clicking on the corresponding circles in the final array (dark background).

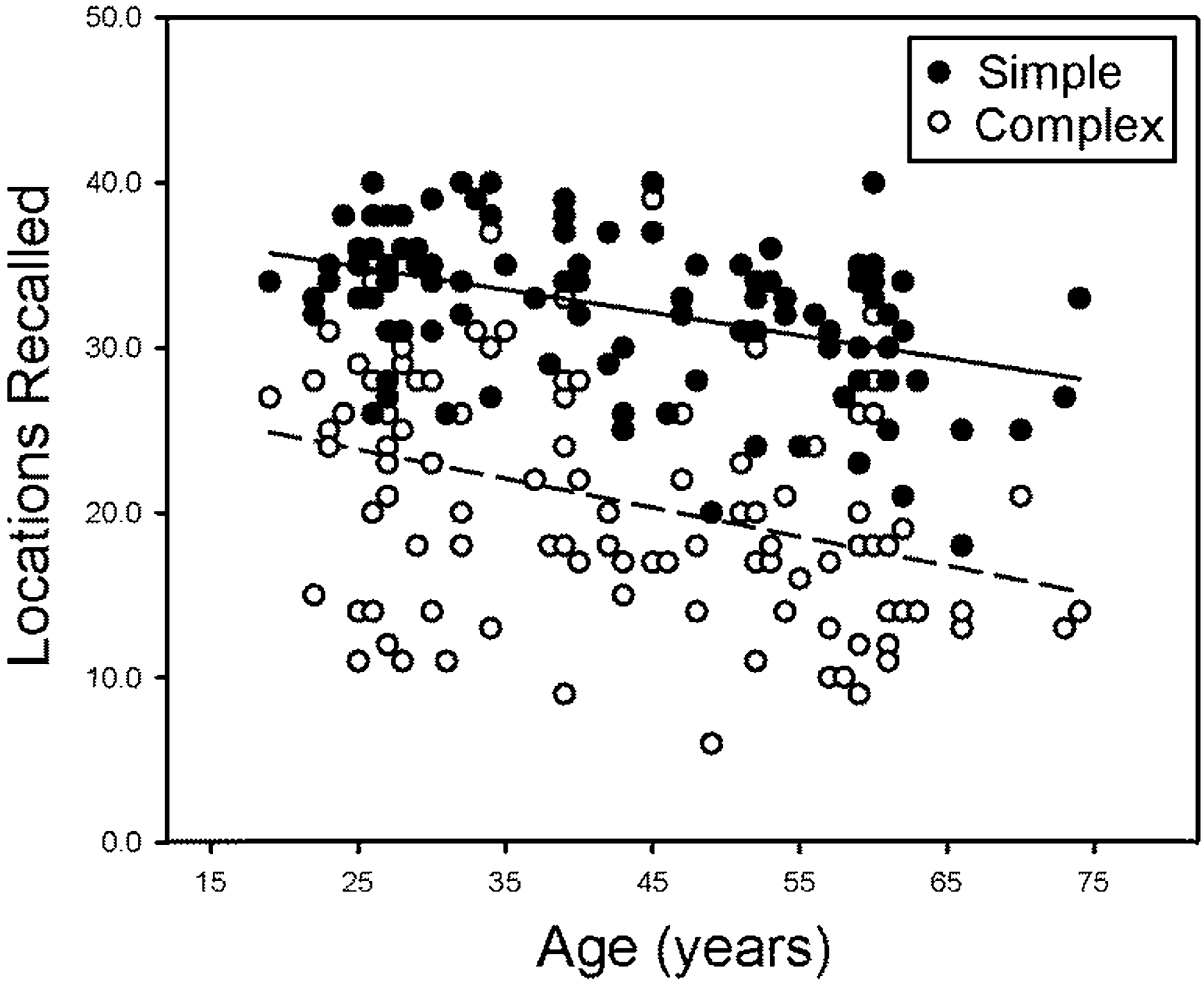

Figure 2.

Total number of locations recalled on simple and complex span tasks as a function of age in Experiment 1. The lines represent the regression of scores on age determined separately for each condition (as in the Full Model). If there was no interference, the points in the Simple and Complex conditions would fall along a single regression line; the vertical distance between points in the Simple and Complex conditions represents the amount of interference by the secondary task.

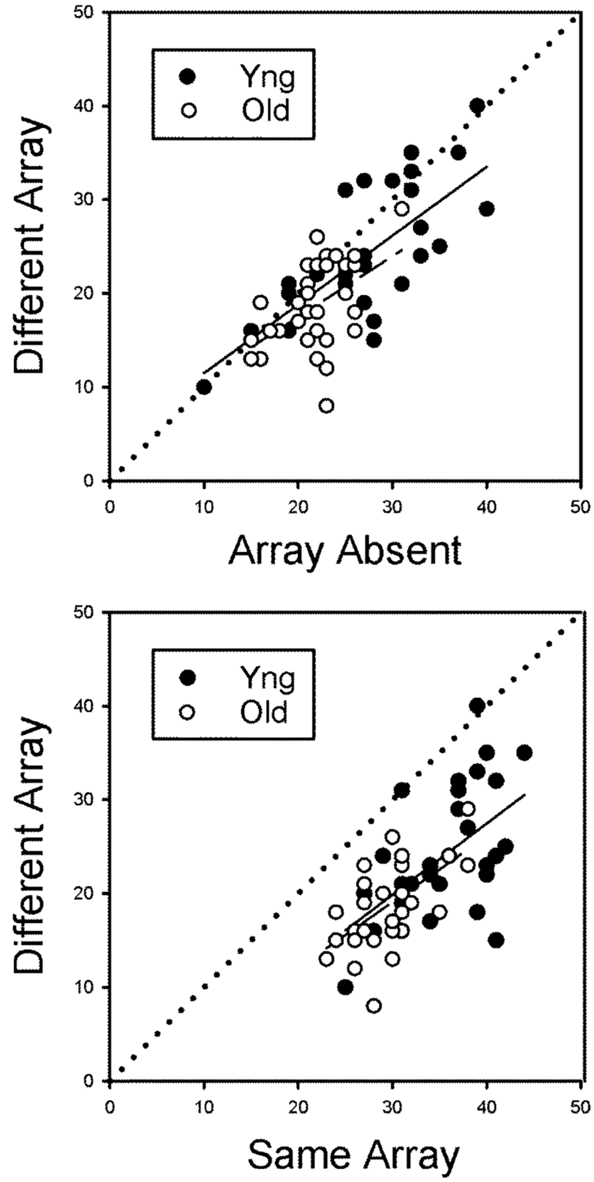

Figure 3.

Number of locations recalled in the Different-Array condition of Experiment 2 as a function of locations recalled in the baseline condition. In the top panel, the baseline is the Array-Absent condition; in the bottom panel, the baseline is the Same-Array condition. If there was no interference, scores would fall along the dotted diagonal lines; distance below that line measures the degree of interference. Regression lines for the younger adult group are solid; those for the older adult group are dashed.

Table 1

Regression Models of Age-related Decline for Experiment 1

Model	Parameter	Coefficient (SE)	*t*
Reduced model of age-related decline (no inhibitory deficit):			
WM = *m* * Age + *b* + *db* * Task	*m*	−00.16 (0.03)	5.79**
R^2 = .559	*b*	39.19 (1.28)	30.57**
	db	−11.75 (0.77)	15.22**
Full model (age-related inhibitory deficit):			
WM = *m* * Age + *dm* * Task * Age + *b* + *db* * Task	*m*	−00.14 (0.04)	3.61**
R^2 = .560	*dm*	−00.04 (0.05)	0.68
$BF_{0,1}$ = 11.61	*b*	38.40 (1.73)	22.18**
	db	−10.16 (2.45)	4.15**

Note. In the Reduced model, the parameter *m* is the rate of decline in scores in both the Simple and Complex conditions, Task is a categorical variable that distinguishes between the conditions (Simple=0, Complex=1), *b* is the intercept for Simple condition scores, and *db* is the difference between the intercepts for Simple and Complex scores (if *db* < 0, the complex intercept is lower). The Full model differs from the Reduced model in that *m* is the rate of decline in Simple scores only, and *dm* is the difference between the rates of decline in the two conditions (if *dm* < 0, Complex scores decline more steeply).

** $p < .01$

Table 2

Mean working memory scores (SDs) for Incongruent, None, and Same array conditions and younger and older age groups in Experiment 2.

	Incong	None	Same	Mean
Older	18.71	21.81	29.45	23.32
	(4.75)	(3.63)	(4.07)	(6.14)
Younger	24.10	27.16	35.55	28.94
	(7.01)	(6.98)	(5.20)	(8.03)
Mean	21.40	24.48	32.50	
	(6.53)	(6.14)	(5.56)	

Table 3

State-Trace Regression Models Predicting Different-Array Performance for Experiment 2

Model	Parameter	Coefficient (SE)	*t*
With Array-Absent as Baseline			
Reduced state-trace model (no age-related inhibitory deficit):			
Different = (1−*m*) * Absent + *b*	*m*	0.23 (0.09)	2.43 *
R^2 = .526	*b*	2.53 (2.38)	1.06
Full state-trace model (age-related inhibitory deficit):			
Different = (1−*m*) * Absent + *b* + *db**Age	*m*	0.29 (0.11)	2.73 **
R^2 = .538	*b*	4.69 (2.96)	1.59
	db	−1.56 (1.28)	1.22
$BF_{0,1}$ = 3.56			
With Same-Array as Baseline			
Reduced state-trace model (no age-related inhibitory deficit):	*m*	0.22 (0.11)	1.92
Different = (1−*m*) * Same + *b*	*b*	−4.05 (3.72)	1.09
R^2 = .445			
Full state-trace model (age-related inhibitory deficit):			
Different = (1−*m*) * Same + *b* + *db**Age	*m*	0.26 (0.14)	1.92
R^2 = .448	*b*	−2.17 (4.92)	0.44
	db	−0.88 (9.12)	0.59
$BF_{0,1}$ = 6.66			

Note. The first pair of reduced and full models compares performance in the Different-Array condition with that in the Array-Absent condition; the second pair of models compares performance in the Different-Array condition with that in the Same-Array condition. In all models, for reasons given in the text, the regression slope equals (1−*m*), and *m* measures the extent to which the slope differs from 1.0. In reduced models, *b* is the intercept for the single state-trace regression line; in full models, *b* is the intercept of the younger adult group's line, *db* is the difference between the intercepts for the two groups, and Age is a categorical variable: Younger=0; Older=1.

* $p < .05$,

** $p < .01$.