

Running head: Age-related inhibitory deficits

Episodic memory and age-related deficits in inhibitory effectiveness

Malcolm D. MacLeod¹, School of Psychology & Neuroscience,

University of St Andrews

&

Jo Saunders, School of Psychological Sciences & Health,

University of Strathclyde

Word Count: 10516 (including references)

Now at: ¹Faculty of Natural Sciences, University of Stirling

Address for Correspondence:

Malcolm D. MacLeod
School of Natural Sciences
University of Stirling
Stirling FK9 4LA

Email: malcolm.macleod@stir.ac.uk

Tel: +44 (0) 1786 467007

Abstract

While an age-related deficit in inhibitory control is well-established in some areas of cognition, the evidence for such a deficit in episodic memory remains inconclusive. Two novel retrieval practice studies were conducted to clarify this issue via the extent to which a loss in inhibitory effectiveness – as measured by the extent of retrieval-induced forgetting (RIF) - is only detectable in (1) the very old, and (2) that a failure to control for non-inhibitory mechanisms can lead to the misinterpretation of intact inhibition in episodic memory in the very old. In Study 1, we chose not to cue practised stronger items at final test and employed independent cues throughout in order to provide as clean a measure of inhibitory functioning as possible. Three groups of older adults were tested: younger-old (60-64 years), old (65-69 years), and older-old (70-74 years). RIF effects emerged in all age groups except in the older-old. In Study 2, we directly manipulated the contribution of output interference (a non-inhibitory mechanism) in a group of young adults (18-34 years), and two older age groups (61-69 years and 70-85 years). Forgetting effects consistent with intact inhibition emerged in young adults and in older adults under the age of 70 years but not in adults over 70 years. In conditions where output interference was promoted, RIF effects also emerged in adults over 70 years thereby giving the misleading impression that inhibitory functioning remains intact. Implications for memory retrieval, measurement, and the modelling of cognitive inhibition in older age are considered.

249 words

Key words: retrieval-induced forgetting; age-related inhibitory deficit; episodic memory retrieval

Episodic memory and age-related deficits in inhibitory effectiveness

The notion that a deficit in inhibitory effectiveness accompanies old age continues to gain acceptance, particularly for its ability to explain age-related deficits in working memory performance (e.g., Collette, Schmidt, Scherrer & Salmon, 2009; Hasher, 2007; Hasher, Stoltzfus, Zacks & Rypma, 1991; Hasher & Zacks, 1988; Healey, Campbell & Hasher, 2008; Lustig, Hasher & Tonev, 2001; Lustig, Hasher & Zacks, 2007; Persad, Abeles, Zacks & Denburg, 2002; Verhaeghen & Salthouse, 1997; Wilson & Kipp, 1998; although see McDowd, 1997 for an alternative view). One of the attractions of such an inhibitory deficit model is that it provides an intuitive means of explaining why, as we grow older, we seem less able to keep distracting stimuli, thoughts, and memories from coming to mind. What remains unclear, however, is whether such an inhibitory deficit is inevitable, and whether such a deficit extends to spheres of cognitive functioning beyond working memory.

From a theoretical point of view, inhibitory resources are also considered to play a pivotal role in resolving unwelcome competition from related but irrelevant information accessed by common cues during the retrieval of episodic memories (see Anderson, 2003; Anderson, Bjork & Bjork, 1994; Bjork, 1989; M. D. MacLeod & Macrae, 2001; Macrae & MacLeod, 1999; Storm & Levy, 2012). By resolving unwanted competition, the retrieval of target memories can be promoted. By the same token, the inability to deal with unwanted competition during retrieval should result in less successful remembering of target material. If we assume this to be the case, any age-related deficit in inhibitory effectiveness in episodic memory should mean that older adults would be less able to deal with interference from competing information during retrieval.

Despite the research effort in this field, an ambivalent picture has emerged regarding the role of inhibition in age-related decrements in episodic memory. Evidence consistent with an age-related inhibitory deficit in episodic memory comes from a range of paradigms including list-method directed forgetting (e.g., Aslan & Bäuml, 2013; Titz & Verhaegen, 2010; Zellner & Bäuml, 2006), and Think/No-Think tasks (e.g., Anderson, Reinholz, Kuhl & Mayr, 2011) - both of which provide participants with explicit instructions to forget or suppress a sub-set of previously learned material. This act of inhibition conveys benefits for one's ability to retrieve other related material; that is, the greater one's ability to inhibit interfering information, the greater one's ability to retrieve target material on demand because of the consequent reduction in interference. In general, older adults have been found to be less able to inhibit or suppress information effectively following explicit instructions to do so which means that recall performance for target material tends to be poorer.

In principle, there is no reason to expect any different pattern of effects when older adults are tested under conditions which promote interference and where there is no explicit instruction to forget. The reality, however, has produced a more complicated picture. This particular kind of non-directed forgetting has been extensively explored using the retrieval practice paradigm (Anderson et al., 1994). In young healthy adults, this paradigm produces a retrieval-induced forgetting effect (RIF) – the magnitude of which is generally taken as a proxy for the effectiveness of inhibition; basically, the bigger the RIF effect observed, the more effective inhibition is assumed to be (although see Anderson & Levy, 2011; Noreen & MacLeod, 2015 for some important caveats to this principle). In the current article, we set out to demonstrate why we need to exercise further caution about how such RIF effects are

interpreted – especially when the retrieval practice paradigm is used with elderly populations.

In the initial stage of the retrieval practice paradigm, participants are typically presented with a series of category-exemplar pairs to study (e.g., loud – jackhammer, siren, traffic, cannon, grenade, gun; green – dollar, emerald, lawn, artichoke, lettuce, pepper). This is followed by a selective retrieval practice phase in which half of the exemplars from half of the categories are practised usually up to three times (e.g., loud – ja_____; si_____; tr_____.). This procedure creates three item types which differ in retrieval status: Rp+ items are practised items from practised categories (e.g., jackhammer, siren, traffic); Rp- items are unpractised items from practised categories (e.g., cannon, grenade, gun); and Nrp items are unpractised items from unpractised categories (e.g., items from the ‘green’ category). Following a distractor task, participants are tested for all the items originally presented.

Two patterns of retrieval performance typically emerge. First, Rp+ items are better remembered than Nrp items, thereby confirming the facilitatory effect of retrieval practice (e.g., Karpicke & Roediger, 2007; Morris & Fritz, 2000). Second – and of much more theoretical interest - a RIF effect emerges whereby Rp- items are remembered less well than Nrp items, despite both item types having been treated as functionally equivalent (i.e., neither receive any retrieval practice). According to inhibitory theory (Anderson, 2003; Anderson & Spellman, 1995), Rp- items have been inhibited or suppressed by virtue of the fact that they represent unwelcome competition due to their relatedness to Rp+ items (i.e., as Rp- items belong to the same category as Rp+ items, they are also likely to be accessed by common cues and, therefore, compete for retrieval with Rp+ items).

One of the problems for an inhibitory account of RIF, however, is that non-inhibitory mechanisms can also contribute to such forgetting effects (see Anderson & Levy, 2011; Anderson & Spellman, 1995; Jakab & Raaijmakers, 2009; C. M. MacLeod, 2007; C. M. MacLeod, Dodd, Sheard, Wilson & Bibi, 2003; Raaijmakers & Jakab, 2013). Thus, to increase our confidence that any observed forgetting in the retrieval practice paradigm is likely to be a function of inhibition, novel retrieval cues can be employed at final test; that is, cues that have not previously been used in either the presentation or retrieval practice phases (Anderson, 2003; Anderson & Spellman, 1995; Anderson & Levy, 2007; Huddleston & Anderson, 2012; Hulbert, Shivde & Anderson, 2012; Veling & van Knippenberg, 2004). In doing so, the provision of novel cues at final test should prove sufficient to circumvent any forgetting of Rp-items caused by other associated items interfering with retrieval. Thus, if the forgetting of Rp-items is due to such interference, the use of independent cues should circumvent this problem and therefore no forgetting effect should be observed. If, on the other hand, forgetting occurs despite the use of such novel cues, inhibitory theorists have argued (see Anderson et al, 1994; Anderson & Spellman, 1995) that the item itself has been inhibited; that is, it is no longer available to conscious inspection (but see Camp, Pecher & Schmidt, 2007; Camp, Pecher, Schmidt & Zeelenberg, 2009; Raaijmakers & Jakab, 2013; Tomlinson, Huber, Rieth & Davelaar, 2009, for alternative theoretical views).

The complexities associated with the interpretation of RIF effects has particular relevance for our understanding of inhibitory deficits and whether they exist. Moulin, Perfect, Conway, North, Jones and James (2002), for instance, reported the presence of intact inhibitory functioning in patients who had been diagnosed with Alzheimer's disease and also in a matched non-clinical control group of healthy older

adults following a selective retrieval practice procedure. Retrieval-induced forgetting effects were observed in both groups, consistent with intact inhibitory functioning in older age.

The difficulty with such an interpretation, however, is the possibility that the observed RIF effects may have been a product of non-inhibitory mechanisms. Specifically, participants in Moulin and colleagues' study were required to engage in a free recall procedure during the final test phase of the retrieval practice paradigm and therefore we cannot eliminate the possibility that the RIF effects observed may have been a function output interference. Participants may have tended to output stronger Rp+ items first which may have subsequently interfered with the retrieval of weaker Rp- items.

Storm and White (2010) had previously drawn attention to this potential problem when they demonstrated that inhibitory deficits in people suffering from attention-deficit/hyperactivity disorder (ADHD) only became apparent where output interference had been adequately controlled during the final test phase. Specifically, they showed that where ADHD participants were presented with category only cues, RIF effects emerged. In contrast, where category plus stem cues were employed – and therefore the order in which items were retrieved could be controlled – no RIF effect emerged. Thus, where output interference is not adequately controlled for at final test, confounds can arise which, in turn, prevent an adequate test of the inhibitory deficit hypothesis.

Recognizing the need to disambiguate between inhibitory and non-inhibitory accounts of RIF in older adults, Aslan, Bäuml and Pastötter (2007) reported the results of two studies, the first of which established that the magnitude of RIF in older adults was comparable to that of young adults. In their study, Aslan and colleagues

not only employed the independent probe procedure (cf. Anderson & Spellman, 1995) but also took the additional precaution of *not* cueing participants for Rp+ items during final recall. The rationale for this novel modification was that, if participants were not required to recall Rp+ items at final test, any observed forgetting effects could not be a function of stronger practised items interfering with the retrieval of weaker items (Rundus, 1973). Having adopted these careful controls, Aslan and colleagues confirmed that there was no evidence of an age-related inhibitory deficit in episodic memory retrieval. In fact, the extent of RIF in older adults was found to be roughly equivalent to that in younger adults (-7% vs. -8%, respectively).

More recently, however, Aslan and Bäuml (2012) sought to provide a more nuanced approach by looking at older-old adults in addition to old adults. Retrieval-induced forgetting effects emerge in old adults (aged 60-75 years, $M = 70$ years), thereby confirming their earlier findings, but failed to emerge in very old adults (above 75 years of age, $M = 84$ years). This latter finding is also broadly consistent with Ortega and colleagues (2012, Study 2) who found that, on using a divided attention manipulation during the selective retrieval practice phase, young adults required a relatively demanding task before RIF effects could be eliminated. Older adults, in contrast, required only a moderately demanding task to eliminate RIF. Taken together, these studies would indicate that there is a decrease in the effectiveness of inhibition in later life, but that this loss of inhibition may only become apparent in the very old – at least, as measured by this paradigm.

Despite the additional clarity provided by such research, the majority of retrieval practice studies published to date continue to challenge the notion of an inhibitory deficit in episodic memory in older adults (see Barber & Mather, 2012; Gómez-Ariza, Pelegrina, Lechuga, Suárez & Bajo, 2009; Hogge, Adam & Collette,

2008; Koutstall, Schacter, Johnson & Galluccio, 1999; Ortega et al., 2012, Study 1).

Given this to be the case, it would seem particularly important to clarify whether the apparent disparity in the literature is simply a function of the age ranges of older adults sampled. Specifically, the mean age of older adults in the latter studies where no inhibitory deficits were reported was 69.3 years whereas the mean age of older-old adults in Aslan and Bäuml's (Study 2) was considerably higher at 84 years.

In the current article, we report two studies which attempt to address these various issues. In our first study we sought to establish the robustness of RIF in groups of younger-old, old, and older-old adults. This also allowed us to establish whether any marked differences in inhibitory effectiveness existed between these three older age groups. Following Aslan and colleagues (2007), we used independent cues. Also, Rp+ item retrieval cues were not presented at final test in order to minimize the possibility that stronger practised items could interfere with the retrieval of related but unpractised Rp- items. Based on Aslan and Bäuml's (2012) and Ortega and colleagues' (2012) work, we could expect RIF effects to emerge for younger-old and old adults, but to be absent for older-old adults – consistent with an inhibitory deficit account of episodic memory in the very old.

In our second study, we sought to provide a direct test of the possible contribution of non-inhibitory interference to the production of RIF effects in young adults, old adults, and older-old adults. In order to accomplish this, we systematically manipulated whether practised items were cued early (i.e., early Rp+ condition), or late (i.e., early Rp- condition), during final recall. In other words, we not only wished to consider whether RIF effects could be eliminated in the early Rp- condition but whether they could be facilitated in the early Rp+ condition. We also looked at the pattern of second-order inhibition effects; that is, the retrieval performance for those

Nrp items which are semantically related to Rp- items in comparison to Nrp items which are not semantically linked to Rp- items. The rationale here is that, if inhibition is responsible for forgetting in the retrieval practice paradigm, then systematic forgetting effects should be evident for those Nrp items related to Rp- items (as well as Rp- items) by virtue of their semantic relatedness to Rp+ items (Anderson & Spellman, 1995; M. D. MacLeod & Saunders, 2005, 2008; Saunders & MacLeod, 2006). Thus, if an inhibitory deficit exists only in very old adults, we could expect first-order (RIF) and second-order forgetting effects to emerge in young adults and old adults, irrespective of whether Rp+ items are cued early or late. Older-old adults, in contrast, could be expected to show evidence of first- and second-order forgetting effects only when Rp+ items are cued early. In other words, any forgetting effects in very old adults would be a consequence of non-inhibitory interference rather than a function of inhibition per se.

STUDY 1

Method

Participants and design

Our study had a 3 (age: younger-old, old, older-old) x 2 (item type: Rp- and Nrp) mixed design with repeated measures on the latter factor. In this study, we considered episodic memory performance in three age groups: younger-old (60-64 years), old (65-69 years), and older-old (70-74 years) adults. Sixty participants volunteered to participate in this study, with 20 participants assigned to each condition. All participants were recruited from the southern region of Wales. In order to determine comparability in intellectual functioning between groups, we used participants' mean error score on the National Adult Reading Test (NART) in order to

calculate the WAIS-R predicted FSIQ (Davis, Bradshaw & Szabadi, 1999). This provided an estimate of premorbid IQ. We also used the Mini Mental State Exam (MMSE) as a means of screening for dementia with a score below 27 as the cut-off (see Table 1 for details). No differences were detected between the age groups on the MMSE, $F(2, 57) = .28$, $MSe = .01$, *ns*, or FSIQ, $F(2, 57) = 1.43$, $MSe = 1.89$, *ns*. No participant was recruited with a past history of head trauma. We did not collect any other subjective measures of health as previous research has indicated they are not predictive of cognitive performance (see e.g., Earles & Salthouse, 1995; Salthouse & Babcock, 1991; Salthouse, Kausler & Saults, 1988; Salthouse & Mitchell, 1990).

Materials and procedure

Given that we wished to replicate and extend Aslan and colleagues' (2007, Study 2) original finding, we employed the same procedures and experimental materials. The only exception to this was in respect of the cue-exemplar pairing 'green-police' which was replaced with 'green-emerald' as police officers in the UK do not wear green uniforms. In the study phase, participants were presented with a booklet containing 32 category cue-exemplar word pairs (see Appendix A). Each word pair was presented on a separate sheet of paper and presented for 5 secs. The order of word pairs in the booklet was randomized for each participant. Participants were paced through the booklet by the experimenter.

After studying the last word pair, participants engaged in selective retrieval practice where participants were cued to retrieve half of the exemplars (i.e., 2 items) from half of the categories (i.e., 4 categories). Participants were presented with the category cue plus the first two letters of the exemplar and given 5 secs in which to

respond. Participants practised each item twice. Following the retrieval practice phase, participants worked on sets of unrelated anagrams for a period of 60 seconds.

In the final phase, recall performance for Rp- and Nrp items was tested. Participants were presented with a series of independent probes, each of which consisted of a new (i.e., never-before-seen) cue plus the first two letters of the target word (see Appendix A). Participants were given the following instructions: “In this task we would like you to remember the words you had studied at the beginning of the experiment. To help you, we have given you some cues or category labels that describe the item. These cues differ from the ones you had originally studied at the beginning of the experiment. Please consider the cue and then think of the words that you originally studied and whether any of those items belong to that category. For example, if you studied ‘strawberry’ under the cue ‘FRUIT’ and are given the cue ‘RED’ at final test, then a ‘strawberry’ is also a ‘red thing’ and should be recalled in response to ‘RED’.” Following Aslan and colleagues (2007, Study 2), we chose not to cue Rp+ items at final test in order to eliminate stronger practised items interfering with the retrieval of weaker unpractised items. Participants were given 4 seconds in which to respond to each cue. This final recall task was paper-and-pencil based. On completion, participants were paid £3.00 (~\$4.73) for their participation, thanked and debriefed.

Results & Discussion

Retrieval practice success rates were 71.88%, 70.63%, and 71.25% for the younger-old, old, and older-old age groups, respectively. Note that facilitation effects (i.e., Rp+ versus Nrp baseline) are not reported here as Rp+ items were not cued at

final test. Holm's sequential Bonferroni approach was employed for all multiple comparisons.

Retrieval-induced forgetting

A 3 (age group: younger-old, old, or older-old) x 2 (item type: Rp- and Nrp) mixed analysis of variance (ANOVA) revealed an effect of item type, $F(1, 57) = 16.79$, $MSe = .01$, $p < .001$, $\eta^2_p = .21$, indicating that fewer Rp- items were reported than Nrp items (see Figure 1). An effect of age was also detected, $F(2, 57) = 10.46$, $MSe = .01$, $p < .001$, $\eta^2_p = .27$. Independent t-tests confirmed that the younger-old age group remembered more items than either the old group ($M_s = .55$ vs. $.40$, respectively; $t(78) = 5.23$, $p < .001$, Cohen's $d = .25$), or the old-old group ($M_s = .55$ vs. $.33$, respectively; $t(78) = 7.99$, $p < .001$, $d = .45$). Also, the old group remembered more items than the old-old group ($M = .40$ vs. $.33$), $t(78) = 2.48$, $p < .01$, $d = .07$. An interaction between item type and age was also present, $F(2, 57) = 4.27$, $MSe = .01$, $p < .05$, $\eta^2_p = .13$, thereby indicating that recall for Rp- and Nrp items differed across age groups. Mean inhibition scores were $-.14$ (younger-old), $-.01$ (old), and zero for the older-old age groups. A correlational analysis confirmed that RIF effect size was negatively related to the age of the participant, $r(60) = .35$, $p < .01$.

Paired samples t-tests revealed RIF in the younger-old age group ($M_s = .41$ vs. $.55$, for Rp- and Nrp items, respectively), $t(19) = 5.21$, $p < .001$, $d = .59$, and the old group ($M_s = .36$ vs. $.44$, Rp- and Nrp items, respectively), $t(19) = 3.04$, $p < .001$, $d = .33$, consistent with Aslan and colleagues (2007, Study 2). A paired samples t-test for Rp- versus Nrp item performance in the older-old age group, however, revealed no evidence of RIF ($M = .33$ vs. Nrp $M = .33$), $t(19) = .15$, ns . This is consistent with an

inhibitory deficit account in the older-old and provides a valuable replication of Aslan and Bäuml's earlier finding (2012, Study 2).

Given that we had employed independent cues at final test (Anderson & Spellman, 1995) and that no Rp+ items had been cued, we can be reasonably confident (according to inhibitory theory) that, when RIF effects emerged for younger-old and old adults, they were likely to have been a function of inhibition. Following this logic, we could also expect that, when no RIF emerged in the older-old group, this is likely to have been due to a deficit in inhibitory functioning – at least, as measured via this particular paradigm.

One of the complexities with this inhibitory deficit interpretation in the older-old group is that it involves the interpretation of a null effect. In our second study, therefore, we sought to provide a further test of the age-related inhibitory deficit hypothesis (see Levy & Anderson, 2008; Storm & White, 2010) by assessing directly the contribution of non-inhibitory interference to the production of RIF. As indicated previously, one of the difficulties in establishing the action of inhibition is the need to eliminate the possibility that the observed forgetting effects may be a function of non-inhibitory mechanisms. To date, this has been addressed by employing a number of modifications including the independent cue method (Anderson & Spellman, 1995), recognition test procedures at final test instead of recall (Aslan & Bäuml, 2012; Ortega et al., 2012, but see Verde & Perfect, 2011 for an alternative non-inhibitory account), asking participants to engage in retrieval practice for 'impossible' exemplars (Storm et al., 2006), or not cueing Rp+ items at final test (Aslan et al., 2007; Study 1 of the current article). In our second study, we decided to take a novel approach whereby we directly manipulated the likelihood of stronger practised items interfering with the retrieval of weaker unpractised items. In other words, could we

directly manipulate the magnitude of RIF by manipulating the prevailing recall conditions and, thereby, demonstrate the importance of controlling recall order when looking at older populations?

We set out to accomplish this by assigning participants to conditions in which they were required to retrieve Rp+ items early or where they had to retrieve Rp- items early during the final test phase of the retrieval practice paradigm. Our rationale was that, if there is no inhibitory deficit in older-old adults, we could expect RIF to emerge irrespective of whether Rp+ items are cued early or late. If Rp- items are cued early and a RIF effect still occurs, it is unlikely to be due to stronger practised items interfering with the retrieval of weaker unpractised Rp- items. Thus, where RIF effects emerge under these retrieval conditions, we can be reasonably confident that they are not due to non-inhibitory means and that we can infer that inhibition remains intact. In contrast, if there is a deficit in inhibitory functioning in older-old adults, we could expect RIF to emerge only where stronger Rp+ items are cued early at final test - based on the assumption that the early retrieval of Rp+ items can contribute to RIF by interfering with the subsequent retrieval of weaker Rp- items (i.e., output interference). It is worth noting in this regard that Aslan and colleagues (2007) reported a drop of 5% in the magnitude of RIF in older adults between their first study and second study where they had initially not controlled for output interference.

Given that, in Study 1, we found no significant difference in inhibitory functioning between young-old and old adults, we could expect that young adults and old adults (up to age 69 years) could be expected to produce RIF effects, irrespective of whether Rp+ items are cued early or not. In other words, RIF should be apparent in these two age groups in both retrieval conditions, consistent with there being no inhibitory deficit. If, in contrast, older-old adults (70+ years) have a deficit in

inhibition, we could expect RIF to emerge only when Rp+ items are cued early during the final retrieval phase (i.e., where the retrieval of stronger items can interfere with the retrieval of subsequent weaker Rp- items).

As an additional test of this inhibitory account of forgetting, we also constructed our materials in such a way that we could look at second-order inhibition effects. Specifically, we constructed our materials in such a way that some Nrp items were similar (i.e., semantically related) to Rp- items, or dissimilar to Rp- items (i.e., semantically unrelated). According to inhibitory theory, all items that are related to the practised target item should be subject to inhibition; that is, inhibition should affect not only Rp- items by virtue of their relatedness to Rp+ items but also any Nrp items that could compete for retrieval with either Rp+ items or Rp- items. Thus, as well as the inhibition of Rp- items (first-order forgetting effects), we could expect any Nrp-similar items to be inhibited relative to an Nrp-dissimilar baseline. Given this to be the case, we could expect such second-order effects to emerge for young adults and old adults, irrespective of whether Rp+ items are cued early or not as inhibition should remain effective. For older-old adults, however, second-order effects should emerge only when Rp+ items are cued early, consistent with an inhibitory deficit account.

STUDY 2

Method

Participants and Design

One hundred and eighty participants volunteered to participate in this study (60 young adults, 60 old adults, and 60 older-old adults). Young adults were included in order to provide an appropriate baseline (i.e., intact inhibition) against which

performance by old and older-old adults could be compared. As these experimental materials had not previously been tested on these populations it was important to establish a baseline performance in young adults. Young adults were recruited from the Universities of Swansea and Strathclyde, and old and older-old adults were recruited from the local community in South Wales. Half of each of these groups was randomly assigned to either an early Rp+ condition, or an early Rp- condition.

As in the previous study, the mean NART (National Adult Reading Test) error score was used to calculate the WAIS-R predicted FSIQ (Davis, et al., 1999) which, in turn, was used to estimate premorbid IQ. The Mini Mental State Exam (MMSE) was used to screen for dementia with a score of below 27 as the cut-off. We also applied these tests to our young adults sample for completeness and comparability (see Table 1 for further details). A good level of comparability in intellectual functioning was found across all three age groups in both the early Rp+ and early Rp- conditions. There were no interactions between age and output order on either the MMSE, $F(2, 174) = .88$, $MSe = .74$, *ns*, or the WAIS-R predicted FSIQ, $F(2, 174) = .27$, $MSe = 14.12$, *ns*. As in Study 1, participants were not included if they had reported a previous history of head trauma. No-one was excluded on this basis.

Items in this study were classified in the following way: Rp+ items were practised items from practised categories; Rp- items were unpractised items from practised categories; and NsimRp- items were Nrp items drawn from non-practised categories which shared an implicit category with Rp- items; and NsimC items were unpractised items from unpractised categories drawn from the semantically-similar set but which did not share an implicit category with Rp- items. Finally, due to the potential for list-order effects to occur during the final recall phase, we split the Nrp dissimilar group into two sets: those Nrp items which were reported in the first half of

the recall procedure (NdissC(A)), and those Nrp items which were reported in the latter half of the recall procedure (NdissC(B)). NdissC(A) and (B) were drawn from the semantically dissimilar set which did not share an explicit or implicit category with any of the items that formed either the Rp+ or Rp- sets. See Appendix B for illustration.

Each of the six conditions contained 30 participants. Our study had a 2 (output order: early Rp+, or early Rp-) x 3 (age: young, old, and older-old) x 6 (item type: Rp+, Rp-, NsimRp-, NsimC, NdissC(A) and NdissC(B)) mixed design with repeated measures on the latter factor.

Second-order similarity and implicit categories

The materials used in this study were identical to those previously employed by Anderson and Spellman (1995, Experiments 2, 3a and 4) and comprised 6 categories (cotton, leather, green, soups, loud, and sharp), each containing 6 exemplars (see Appendix B for details). Three of the exemplars in each category were semantically *dissimilar* to exemplars in *other* categories but were semantically *similar* to other members of their *own* category. Appendix B provides a detailed breakdown of the associations between exemplars for different categories. The remaining three exemplars in each category were semantically *similar* to three members of a second implicit category (e.g., pajamas, robe and slacks are originally studied under the cue ‘cotton’ but are also semantically related to belt, boots, and skirt as they are all exemplars of the category ‘clothing’). However, the three remaining items (i.e., curtain, napkin and sheet) are unrelated to any other items outside of their initially studied category.

In the presentation phase, participants were presented with all 36 category-exemplar pairs using the same procedure as in Study 1. During the retrieval practice phase, each participant was presented with three exemplars from two of the available six categories (i.e., 6 Rp+ items in total). Items selected for retrieval practice were *always* derived from the semantically dissimilar set of items (see Appendix B). While these items did not share similarity *between subjects* they did, of course, share similarity *within subjects* (i.e., with the Rp- items). This was to ensure that any second-order inhibition that might be observed was due to the association between Rp- and NsimRp- items and not due to an association with the practised set (i.e., first-order inhibition; see Anderson & Spellman, 1995; Saunders & MacLeod, 2006). For example, participants might practise the exemplars curtain, napkin, sheet, dollar, emerald, and lawn from the categories ‘cotton’ and ‘green’. The three remaining exemplars in each of these two categories would then become the Rp- items (e.g., pajamas, robe, slacks, artichoke, lettuce and pepper) because they shared a semantic similarity with the Rp+ items (*within subjects* similarity). These Rp- items, however, also shared an implicit category with items belonging to another category (i.e., *between subjects* similarity). Rp- items from the category ‘cotton’, for instance, shared an implicit category ‘clothing’ with some of the items from the category ‘leather’ (i.e., belt, boots and skirt). Similarly, Rp- items from the category ‘green’ shared an implicit category ‘vegetables’ with some of the items contained in the category ‘soups’ (i.e., mushroom, onion and tomato). The three exemplars in each of the other two categories formed the NsimRp- items (i.e., six items in total).

There were also six NsimC items which acted as a baseline for measuring second-order inhibition. These six items were the three remaining dissimilar items from the two implicit categories; that is, while they were semantically related to the

NsimRp- items they were not related to the Rp- items (e.g., briefcase, saddle, whip, chicken, clam and turkey).

The remaining 12 exemplars (i.e., 3 exemplars from four categories) comprised the NdissC items (e.g., cannon, grenade, gun, dagger, spear, sword, jackhammer, siren, traffic, needle, tack and thorn). Thus, NdissC items are unpractised items drawn from the semantically dissimilar set. Although NdissC items were semantically related *within* category to one another they were not semantically related *between* category to any other items (i.e., *within subject similarity* but *between subjects dissimilarity*). See Table 2.

Output order

As one of the critical manipulations in our study was to control for list-based output interference (i.e., where items that appear later during recall are less likely to be reported due to the retrieval of earlier items interfering with access to later items), we needed to construct two sets of NdissC items (which we refer to here as set A and set B). In the early Rp+ condition, Rp+ items were cued for recall at final test during the first half of the recall procedure alongside half of the NdissC items (i.e., NdissC(A) items). In contrast, Rp-, NsimRp-, NsimC and the remaining half of the Nrp control items (i.e., NdissC(B) items) were cued for recall in the latter half of the final recall procedure. Conversely, in the early Rp- condition, Rp- items were reported first intermingled with NsimRp-, NsimC, and half the Nrp dissimilar control items (i.e., NdissC(A) items). Rp+ items and the remaining Nrp dissimilar controls (i.e., NdissC(B) items) were then reported in the second half of the recall procedure.

The division of the Nrp dissimilar control group into two sets allowed us to control for list-based output interference. Specifically, in the early Rp+ group, Rp-

recall can be compared to NdisC(B) items as both sets were reported in the second half of the final recall procedure. NsimRp- can also be compared with NsimC items as they were both reported in the second half of the final recall procedure, and Rp+ items could then be compared with NdisC(A) items as both sets of these items were reported in the first half.

In the early Rp- condition, Rp- items were compared with NdisC(A) items as both sets of items were reported during the first half of the recall procedure. NsimRp- items were compared with NsimC items as both sets of items were reported during the first half of recall, and finally Rp+ items were compared with NdisC(B) items as both sets of these items were reported during the latter half of recall.

Procedure

Participants arrived at the laboratory and were greeted by the experimenter. On completing the NART and MMSE, participants were presented with 36 cue-exemplar word pairs (i.e., six categories containing six exemplars as outlined in Appendix B). Each cue-exemplar word pair appeared separately on each page of a study booklet. Participants were given 5 seconds in which to study each word pair. The order of word pairs was randomised for each participant (note that items were not presented in blocked fashion by studied category).

Following completion of the study phase, participants performed the selective retrieval practice procedure. Specifically, participants were presented with cue plus 2-letter-word stems to complete (e.g., cotton-cu_____). Three exemplars from two categories were chosen for each participant (i.e., 6 exemplars in total), and participants were cued to retrieve each item three times each (i.e., 18 prompts in total), and these items were always semantically dissimilar items. Order of

presentation was randomised for each participant during the retrieval practice phase and counterbalanced across the study.

Following retrieval practice, participants worked on unrelated word search puzzles for 60 seconds. Finally, participants completed a cued-recall phase employing the original category cues from the study phase plus the first two letters from the to-be-remembered item. For this phase, there were two conditions: an early Rp+ condition in which Rp+ items were cued in the first half of recall along with the NdissC(A) items. Rp-, NsimRp-, NsimC, and NdissC(B) items were presented for retrieval in the second half of the recall phase (see Table 1 for an example of the assignment of items to this condition). In the early Rp- condition, recall commenced with Rp- items along with NsimRp-, NsimC and NdissC(A) items. Participants were then cued for the retrieval of Rp+ and NdissC(B) items in the second half. Participants were given 5 seconds in which to report each item before moving on to the next item. Within each first half of the final recall phase, order of presentation was randomised but care was taken to ensure that no two items from the same category followed one another. On completion, participants were paid £3 (\$4.73), thanked, and debriefed.

Results and Discussion

Performance during the retrieval practice phase was comparable across conditions and between age groups. Specifically, mean retrieval practice success in the early Rp+ condition was 75.82% ($SD = 23.54$), 78.89% ($SD = 18.53$), and 76.67% ($SD = 18.36$), for young, old, and older-old adults, respectively. Mean retrieval practice success in the early Rp- condition was 76.68% ($SD = 22.11$), 77.78% ($SD = 17.69$), and 74.44% ($SD = 18.94$), for young, old, and older-old adults, respectively.

Note that Holm's sequential Bonferroni approach was employed on all multiple comparisons.

Facilitation Effects

The comparison between Rp+ items and Nrp dissimilar items in the early Rp+, and early Rp-, output groups was based on different Nrp dissimilar sets (A or B). Specifically, in the early Rp+ output condition, Rp+ items were tested in the first half of final recall and so these items were compared with NdissC(A) items because they were also tested in the first half of the final recall session. Conversely, in the early Rp- output condition, Rp+ items were tested in the second half of final recall and so they were compared with NdissC(B) items which were also tested in the second half of recall.

In order to examine whether retrieval practice improved memory for Rp+ items, a 2 (output order: early Rp+ or early Rp-) x 3 (age: young, old, or older-old) x 2 (item type: Rp+ and NdissC A/B) mixed ANOVA was conducted. This revealed an effect of item type, $F(1, 174) = 152.28$, $MSe = .03$, $p < .001$, $\eta^2_p = .47$; an effect of age, $F(2, 174) = 22.61$, $MSe = .88$, $p < .001$, $\eta^2_p = .97$; and an effect of output order, $F(1, 174) = 9.93$, $MSe = .39$, $p < .01$, $\eta^2_p = .05$. There were also significant interactions between age and item type, $F(2, 174) = 4.69$, $MSe = .03$, $p < .01$, $\eta^2_p = .06$; and between output order and item type, $F(1, 174) = 5.49$, $MSe = .03$, $p < .05$, $\eta^2_p = .03$. There was no interaction between age and output order, $F(2, 174) = .48$, $MSe = .02$, *ns*. The interaction between output order, age, and item type was also not significant, $F(2, 174) = .28$, $MSe = .03$, *ns*.

Paired samples t-tests confirmed that in the early Rp+ condition, more Rp+ items were reported than NdissC(A) items in all three age groups. Specifically,

facilitation effects (Rp+ vs. NdissC(A)) emerged for young adults ($M_s = .85$ vs. $.72$, respectively), $t(29) = 2.83$, $p < .01$, $d = .22$; old adults ($M_s = .73$ vs. $.54$, respectively), $t(29) = 4.77$, $p < .001$, $d = .44$; and for older-old adults ($M_s = .73$ vs. $.49$, respectively), $t(29) = 6.28$, $p < .001$, $d = .58$.

In the early Rp- condition, paired samples t-tests similarly revealed that more Rp+ items were reported than NdissC(B) items in all three age groups. Specifically, facilitation effects emerged for young adults ($M_s = .80$ vs. $.59$, respectively), $t(29) = 3.38$, $p < .01$, $d = .28$; old adults ($M_s = .72$ vs. $.47$, respectively), $t(29) = 5.20$, $p < .001$, $d = .48$; and older-old adults ($M_s = .73$ vs. $.36$, respectively), $t(29) = 2.06$, $p < .05$, $d = .13$.

It can also be seen from the pattern of performance that retrieval practice effects increased with age (early Rp+ groups: young M retrieval practice effect = $.13$; old $M = .19$; older-old $M = .24$; early Rp- groups: young $M = .21$; old $M = .25$; older-old $M = .37$), $r(180) = .19$, $p < .01$. It should be noted here, however, that this is most likely to be an artefact in that, for young adults, recall of Rp+ items was already very high and therefore it would be difficult to prime these items much further. In the two older groups, recall of Nrp-control items was much lower, thereby giving greater opportunity for retrieval practice effects to emerge.

Retrieval-induced forgetting

The comparison between Rp- items and Nrp dissimilar items in the early Rp+, and early Rp-, output groups were also based on different Nrp dissimilar sets (A or B). Specifically, in the early Rp+ output condition, Rp- items were tested in the second half of the final recall procedure and so these items were compared with the NdissC(B) items that were also tested in the second half of the final recall session.

Conversely, in the early Rp- output condition, Rp- items were tested in the first half of the final recall and so they were compared with NdissC(A) items which were also tested in the first half of recall.

In order to examine whether retrieval practice had a detrimental impact on retrieval performance for Rp- items, a 2 (output order: early Rp+ or early Rp-) x 3 (age: young, old or older-old) x 2 (item type: Rp- and NdissC A/B) mixed ANOVA was conducted. This revealed an effect of item type, $F(1, 174) = 35.41$, $MSe = .04$, $p < .001$, $\eta^2_p = .17$; an effect of age, $F(2, 174) = 15.78$, $MSe = .62$, $p < .01$, $\eta^2_p = .70$; and an effect of output order, $F(1, 174) = 15.46$, $MSe = .61$, $p < .001$, $\eta^2_p = .08$. There were also significant interactions between age and item type, $F(2, 174) = 9.32$, $MSe = .04$, $p < .001$, $\eta^2_p = .10$; age and output order, $F(2, 174) = 5.25$, $MSe = .21$, $p < .05$, $\eta^2_p = .06$; and between output order, age, and item type, $F(2, 174) = 3.61$, $MSe = .04$, $p < .05$, $\eta^2_p = .04$. The interaction between output order and item type was marginal, $F(1, 174) = 2.93$, $MSe = .04$, $p = .08$, $\eta^2_p = .02$.

Paired samples t-tests confirmed that, in the early Rp+ condition, fewer Rp- items were reported than NdissC(B) items in all three age groups. Specifically, RIF (Rp- vs. NdissC(B)) emerged for young adults ($M_s = .42$ vs. $.64$, respectively), $t(29) = 4.24$, $p < .001$, $d = .38$; old adults ($M_s = .33$ vs. $.50$, respectively), $t(29) = 2.86$, $p < .01$, $d = .22$; and older-old adults ($M_s = .27$ vs. $.39$, respectively), $t(29) = 2.34$, $p < .05$, $d = .16$.

In the early Rp- condition, paired samples t-tests similarly revealed that fewer Rp- items were reported than NdissC(A) items in young adults ($M_s = .45$ vs. $.68$, respectively), $t(29) = 3.36$, $p < .01$, $d = .28$; and in old adults ($M_s = .36$ vs. $.54$, respectively), $t(29) = 3.75$, $p < .001$, $d = .33$. There was, however, no evidence of RIF in the older-old age group in the early Rp- condition. In fact, there was evidence of

some facilitation for Rp- items in comparison to the NdisC(A) baseline ($M_s = .57$ vs. $.44$, respectively), $t(29) = 5.64$, $p < .001$, $d = .52$. It is unlikely that this can be attributed entirely to a drop in performance in the NdisC(A) baseline.

Second-order inhibition effects

As a further test of whether the observed forgetting effects were due to inhibition, we also explored the pattern of second-order effects in both conditions by comparing recall performance for NsimRp- items against NsimC items. A 2 (output order: early Rp+ or early Rp-) \times 3 (age: young, or old or older-old) \times 2 (item type: NsimRp- and NsimC) mixed ANOVA revealed an effect of item type, $F(1, 174) = 62.21$, $MSe = .03$, $p < .001$, $\eta^2_p = .26$; an effect of age, $F(2, 174) = 5.57$, $MSe = .23$, $p < .01$, $\eta^2_p = .06$; but no effect of output order, $F(1, 174) = .48$, $MSe = .02$, $p = .49$. There was no interaction between age and output order, $F(2, 174) = .34$, $MSe = .04$, *ns*, but there were significant interactions between output order and item type, $F(1, 174) = 4.97$, $MSe = .03$, $p < .05$, $\eta^2_p = .03$; between age and item type, $F(2, 174) = 5.78$, $MSe = .03$, $p < .01$, $\eta^2_p = .06$; and between output order, age, and item type, $F(2, 174) = 3.21$, $MSe = .03$, $p < .05$, $\eta^2_p = .04$.

Paired samples t-tests revealed that in the early Rp+ condition, fewer NsimRp- items were reported than NsimC in all three age groups. Specifically, second-order effects emerged for young adults ($M_s = .41$ vs. $.65$, respectively), $t(29) = 2.83$, $p < .01$, $d = .22$; old adults ($M_s = .39$ vs. $.55$, respectively), $t(29) = 3.72$, $p < .001$, $d = .32$; and older-old adults ($M_s = .36$ vs. $.52$, respectively), $t(29) = 3.80$, $p < .001$, $d = .33$.

For the early Rp- condition, paired samples t-tests revealed that fewer NsimRp- items were reported than NsimC in the young and old age groups only. Specifically, second-order effects emerged for young adults ($M_s = .45$ vs. $.64$,

respectively), $t(29) = 3.99, p < .001, d = .35$; and old adults ($M_s = .37$ vs. $.55$, respectively), $t(29) = 3.29, p < .001, d = .28$. Consistent with the absence of first-order effects, there was no evidence of second-order effects in the older-old age group. In fact, there was some facilitation of NsimRp- items relative to NsimC items although this was small, ($M_s = .49$ vs. $.45$, respectively), $t(29) = 3.36, p < .01, d = .28$. See Figure 2.

An interference account would predict that stronger practised Rp+ items in this condition would interfere with the retrieval of related weaker Rp- items. Similarly, we could expect that the early retrieval of Rp+ items would interfere with the subsequent retrieval of NsimRp- items because of their association to Rp- items. In other words, just as first-order forgetting effects can be produced by non-inhibitory means, so can second-order forgetting effects. Where retrieval conditions promote the early retrieval of Rp+ items, we could expect to find first- and second-order forgetting effects to emerge even in the very old as they would be independent of inhibition. Indeed, correlational analyses confirmed that there were significant correlations between size of the inhibitory effect and age (RIF, $r(180) = .24, p < .001$; 2nd order inhibition, $r(180) = .19, p < .01$. Mean inhibition scores were smaller as participant age increased. For young adults: RIF Early Rp+ condition, $M = -.22$; 2nd order, $M = -.22$; RIF Early Rp – condition, $M = -.22$; 2nd order, $M = -.18$. For old adults: RIF Early Rp+ condition, $M = -.19$; 2nd order, $M = -.16$; RIF Early Rp – condition, $M = -.17$; 2nd order, $M = -.18$. Finally, for older-old adults: RIF Early Rp+ condition, $M = -.12$; 2nd order, $M = -.17$; RIF Early Rp – condition, $M = .12$; 2nd order, $M = .04$.

The pattern of first- and second-order forgetting effects observed here indicate that an inhibitory deficit is present but only in the very old. Neither first- nor second-order forgetting effects emerged for older-old adults in the early Rp- condition where

output interference had been minimised. If inhibition had been intact in older-old adults, we could reasonably have expected forgetting effects to have emerged in the early Rp- condition as well as the early Rp+ condition – just as they had with young and old adults. The fact that these effects failed to emerge where output interference had been minimised would suggest that, where this kind of forgetting effect is observed in very old adults, it is unlikely to be a function of inhibition.

General Discussion

Although it has been widely assumed that certain populations such as young children and healthy older adults, and patients with schizophrenia, Alzheimer's disease, ADHD, or frontal lobe damage can suffer from inhibitory deficits, many studies have failed to find evidence of such deficits in episodic memory (e.g., Aslan et al., 2007; Conway & Fthenaki, 2003; Ford, Keating, & Patel 2004; Gómez-Ariza et al., 2009; Moulin et al., 2002; Nestor, Piech, Allen, Niznikiewicz, Shenton & McCarley, 2005; Zellner & Bäuml, 2005).

On first inspection, this would seem to be inconsistent with an inhibitory deficit account. It is important to realise, however, that RIF effects are not always a function of inhibition (see also Anderson & Levy, 2007, for a discussion). In the current article, we demonstrated how the early retrieval of Rp+ items can interfere with the retrieval of weaker Rp- items at test and, thereby, produce RIF effects. It is also worth noting here that many of the failures to find inhibitory deficits in 'at risk' populations have used category-cued recall which leads to the possibility that such output interference may have driven the RIF effects observed. In contrast, when such interference is adequately controlled during the final recall stage of the retrieval practice paradigm, inhibitory deficits become apparent in these populations (e.g.,

Aslan & Bäuml, 2010; Soriano, Jiménez, Román & Bajo, 2009; Storm & White, 2010). The current findings provide further support for this perspective and indicate that, when attempting to measure the effectiveness of inhibitory functioning in episodic memory retrieval in 'at risk' or typically developing populations, particular care needs to be taken to control for these potential confounds.

In our first study, we took particular care to ensure that any observed forgetting effects could only be attributed to inhibition. We accomplished this via the use of independent cues at final test, and by not cueing strong practised Rp+ items at final test. Retrieval-induced forgetting emerged in younger-old and old adults but not in older-old adults. In our second study, we provided an even more stringent test of the age-related inhibitory deficit hypothesis by expressly manipulating whether Rp+ items were cued early, or late, during final test. The rationale here was that, if RIF effects emerged in the early Rp- condition, we could reasonably assume it must be due to inhibition. Similarly, if the prior retrieval of Rp+ items contributed to RIF, we could expect forgetting effects to emerge in the early Rp+ condition for all three age groups. We found that RIF emerged in young adults and old adults, irrespective of output order but, for older-old adults, RIF emerged only when Rp+ items were cued early. Thus, it would seem that where inhibition remains intact, RIF emerges irrespective of the possible contribution of non-inhibitory mechanisms such as interference. Indeed, one possibility is that, where inhibition remains effective such inhibition may have primacy over other potential contributory mechanisms.

More speculatively, perhaps, Study 2 raises the issue of whether the RIF observed in the early Rp+ condition may actually be a function of a different form of inhibition. Bäuml (1998), for instance, raised the possibility that output order effects could be considered a form of inhibition (see also Anderson et al., 1994). In a sense,

both RIF and output interference effects reflect a memory impairment that is caused by the prior retrieval of related material. Thus, there is the possibility that two inhibitory mechanisms may be at play in the retrieval practice paradigm and that these different inhibitory mechanisms may mediate RIF in different age groups.

It is also worth noting that Study 2 showed some evidence of facilitation for Rp- items and NsimRp- items. These effects may be due to increased relational processing in older adults. Previous research using the Deese-Roediger-McDermott (DRM) paradigm has found that item-specific processing decreases false alarms in younger, but not older adults, as well as decreasing veridical recall in older adults (Tun, Wingfield, Rosen & Blanchard, 1998). These findings suggest that older adults rely on shared cues at encoding to improve recall (i.e., relational encoding), while younger adults rely on both shared and unique cues. Tun and colleagues also found that item-specific processing did not reduce the proportion of false alarms relating to weakly related lures in older adults suggesting that activation may spread even further within the associative network than in younger adults - perhaps because of a lack of effective inhibition. Thus, if older adults are more likely to encode information relationally, they may be more likely than younger adults to identify the link between Rp+, Rp- and NsimRp- items. Assuming this to be the case, it is possible that when inhibition is no longer functioning effectively it could actually lead to an increase in the recall of other items because of their relatedness.

In summary, our studies have produced patterns of forgetting in young adults, younger-old, and old adults that are difficult to account for by traditional non-inhibitory explanations, such as interference. We also demonstrated that we could produce or eliminate RIF effects in very old adults simply by manipulating the nature of the retrieval conditions. Our studies provide strong inference that there is an

inhibitory deficit in episodic memory retrieval but that this deficit only becomes apparent once very old age is reached. Our studies also point to the possibility of a decline in inhibitory function as we grow older but this will only be confirmed via longitudinal studies which consider memory performance and inhibitory effectiveness. Clearly, if we are to find ways in which to maintain cognitive performance in elderly adults, we also need to fully understand how retrieval conditions can influence memory performance. Only by doing so, are we likely to meet the various challenges posed by an ageing population. Indeed, only then will we truly be able to determine whether an age-related decline in memory performance is inevitable.

References

- Anderson, M.C. (2003). Rethinking interference theory: Executive control and the mechanisms of forgetting. *Journal of Memory and Language*, 49, 415-445.
- Anderson, M.C., Bjork, R.A., & Bjork, E.L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1063–1087.
- Anderson, M.C., & Levy, B.J. (2007). Theoretical issues in inhibition: insights from research on human memory. In D. Gorfein & C. M. MacLeod (Eds.) *Inhibition in Cognition*. (pp. 81-102). Washington, DC: American Psychological Association.
- Anderson, M. C., & Levy, B. J. (2011). On the relationship between interference and inhibition in cognition. In A. S. Benjamin (Ed.), *Successful Remembering and Successful Forgetting: A Festschrift in Honor of Robert A. Bjork*. (pp. 107-132). Washington, DC: American Psychological Association.
- Anderson, M.C., Reinholz, J., & Kuhl, B., & Mayr, U. (2011). Intentional suppression of unwanted memories grows more difficult as we age. *Psychology and Aging*, 26, 397-405
- Anderson, M.C., & Spellman, B.A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review*, 102, 68–100.
- Aslan, A., & Bäuml, K.-H. (2010). Retrieval-induced forgetting in young children. *Psychonomic Bulletin & Review*, 17, 704–709.
- Aslan, A., & Bäuml, K.-H. (2012). Retrieval-induced forgetting in old and very old age. *Psychology and Aging*, 27, 1027-1032.

- Aslan, A., & Bäuml, K-H (2013). Listwise directed forgetting is present in young-old adults, but is absent in old-old adults. *Psychology and Aging*, 28, 213-218
- Aslan, A., Bäuml, K.-H., & Pastötter, B. (2007). No inhibitory deficit in older adults' episodic memory. *Psychological Science*, 18, 72-78.
- Barber, S.J., & Mather, M. (2012). Forgetting in context: The effects of age, emotion, and social factors on retrieval-induced forgetting. *Memory and Cognition*, 40, 874-888.
- Bäuml, K.-H. (1998). Strong items get suppressed, weak items do not: The role of item strength in output interference. *Psychonomic Bulletin & Review*, 5, 459–463.
- Bjork, R.A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H.L. Roediger & F.I.M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honor of Endel Tulving* (pp. 309-330). Hillsdale, NJ: Erlbaum.
- Camp, G., Pecher, D., & Schmidt, H.G. (2007). No retrieval-induced forgetting using item-specific independent cues: Evidence against a general inhibitory account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 950-958.
- Camp, G, Pecher, D., Schmidt, H.G., & Zeelenberg, R. (2009). Are independent probes truly independent? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 35, 934-942.
- Collette, F., Schmidt, C., Scherrer, S.A. & Salmon, E. (2009). Specificity of inhibitory deficits in normal aging and Alzheimer's disease. *Neurobiology of Aging*, 30, 875-889.

- Conway, M.A., & Fthenaki, A. (2003). Disruption of inhibitory control of memory following lesions to the frontal and temporal lobes. *Cortex*, 39, 667–686.
- Davis, C., Bradshaw, C.M., & Szabadi, E. (1999). The Doors and People Memory Test: Validation of norms and some new correction formulae. *British Journal of Clinical Psychology*, 38, 305-314.
- Earles, J.L., & Salthouse, T.A. (1995). Interrelations of age, health, and speed. *Journal of Gerontology: Psychological Sciences*, 50B, 33-41.
- Ford, R.M., Keating, S., & Patel, R. (2004). Retrieval-induced forgetting: A developmental study. *British Journal of Developmental Psychology*, 22, 585–603.
- Gómez-Ariza, C.J., Pelegrina, S., Lechuga, M.T., Suárez, A., & Bajo, M.T. (2009). Inhibition and retrieval of facts in young and older adults. *Experimental Aging Research*, 35, 83–97.
- Hasher, L. (2007). Inhibition: Attentional regulation of cognition. In H.L. Roediger, Y. Dudai & S.M. Fitzpatrick (Eds.), *Science of memory: Concepts*. (pp. 291-294). New York: Oxford University Press.
- Hasher, L., Stoltzfus, E.R., Zacks, R.T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 163-169.
- Hasher, L., & Zacks, R.T. (1988) Working memory, comprehension, and aging: a review and a new view. In: Bower G.H. (Ed.), *The Psychology of Learning and Motivation*, Vol. 22. (pp. 193-225). San Diego, CA: Academic Press.
- Healey, M.K., Campbell, K.L., & Hasher, L. (2008). Cognitive aging and increased distractibility: costs and potential benefits. In W.S. Sossin, J.C. Lacaille, V.F.

- Castellucci & S. Belleville (Eds.). *Progress in Brain Research, Vol 169*. (pp. 353-363). Amsterdam: Elsevier.
- Hogge, M., Adam, S., & Collette, F. (2008). Retrieval-induced forgetting in normal ageing. *Journal of Neuropsychology*, 2, 463-476.
- Huddleston, E., & Anderson, M.C. (2012). Reassessing critiques of the independent probe method for studying inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1408-1418.
- Hulbert, J.C., Shivde, G., & Anderson, M.C. (2012). Evidence against associative blocking as a cause of cue-independent retrieval-induced forgetting. *Experimental Psychology*, 59, 11-21.
- Jakab, E., & Raaijmakers, J.G.W. (2009). The role of item strength in retrieval induced forgetting. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 35, 607-617.
- Karpicke, J.D., & Roediger, H.L. III. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language*, 57, 151-162.
- Koutstaal, W., Schacter, D.L., Johnson, M .K., & Galluccio, L. (1999). Facilitation and impairment of event memory produced by photographic review. *Memory & Cognition*, 27, 478-493.
- Levy, B.J., & Anderson, M.C. (2008). Individual differences in the suppression of unwanted memories: The executive deficit hypothesis. *Acta Psychologica*, 127, 623-635.
- Lustig, C., Hasher, L., & Tonev, S.T. (2001). Inhibitory control over the present and the past. *European Journal of Cognitive Psychology*, 13, 107-122.
- Lustig, C., Hasher, L., & Zacks, R. (2007). Inhibitory deficit theory: Recent developments in a 'new view'. In D.S. Gorfein & C.M. MacLeod (Eds.), *The*

- place of inhibition in cognition*. Washington, DC: American Psychological Association.
- MacLeod, C.M. (2007). The concept of inhibition in cognition. In D.S. Gorfein & C.M. MacLeod (Eds.), *Inhibition in cognition*. (pp. 3-23). Washington, DC: American Psychological Association.
- MacLeod, C.M., Dodd, M.D., Sheard, E.D., Wilson, D.E., & Bibi, U. (2003). In opposition to inhibition. In B.H. Ross (Ed.), *The psychology of learning and motivation, Vol 43*. (pp. 163-214). San Diego, CA: Academic Press.
- MacLeod, M.D., & Macrae, C.N. (2001) Gone but not forgotten: The transient nature of retrieval-induced forgetting. *Psychological Science, 12*, 148-52.
- MacLeod, M.D., & Saunders, J. (2005). The role of inhibitory control in the production of misinformation effects. *Journal of Experimental Psychology: Learning, Memory & Cognition, 31*, 964-979.
- MacLeod, M.D., & Saunders, J. (2008). Negative consequences of an adaptive process: Retrieval inhibition and memory distortion. *Current Directions in Psychological Science, 17*, 26-30.
- McDowd, J.M. (1997). Inhibition in attention and aging. *Journals of Gerontology: Psychological Science and Social Science, 52B*, 265-273.
- Macrae, C.N., & MacLeod, M.D. (1999). On recollections lost: When practice makes imperfect. *Journal of Personality and Social Psychology, 77*, 463-473.
- Morris, P.E., & Fritz, C.O. (2000). The name game: Using retrieval practice to improve the learning of names. *Journal of Experimental Psychology: Applied, 6*, 124-129.

- Moulin, C.J.A., Perfect, T.J., Conway, M.A., North, A., Jones, R.W., & James, A.N. (2002). Retrieval-induced forgetting in Alzheimer's disease. *Neuropsychologia*, 40, 862–867.
- Nestor, P.G., Piech, R., Allen, C., Niznikiewicz, M., Shenton, M., & McCarley, R.W. (2005). Retrieval-induced forgetting in schizophrenia. *Schizophrenia Research*, 75, 199–209.
- Noreen, S., & MacLeod, M. D. (2015). What do we really know about cognitive inhibition? Task demands and inhibitory effects across a range of memory and behavioural tasks, *PLoS ONE*, 10, 1-21
- Ortega, A., Gómez-Ariza, C.J., Román, P., & Bajo, M.T. (2012). Memory inhibition, aging, and the executive deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 178-186.
- Persad, C.C., Abeles, N., Zacks, R.T., & Denburg, N.L. (2002). Inhibitory changes after age 60 and their relationship to measures of attention and memory. *Journals of Gerontology (Series B): Psychological Sciences and Social Sciences*, 57, 223–232.
- Raaijmakers, J.G.W., & Jakab, E. (2013). Rethinking inhibition theory: On the problematic status of the inhibition theory for forgetting. *Journal of Memory & Language*, 68, 98-122.
- Rundus, D. (1973). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, 89, 63-77.
- Salthouse, T.A., & Babcock, R.L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, 27, 763-776.

- Salthouse, T.A., Kausler, D.H., & Saults, J.S. (1988). Investigation of student status, background variables, and the feasibility of standard tasks in cognitive aging research. *Psychology & Aging, 3*, 29-37.
- Salthouse, T.A., & Mitchell, D.R. (1990). Effects of age and naturally occurring experience on spatial visualization performance. *Developmental Psychology, 26*, 845-854.
- Saunders, J., & MacLeod, M.D. (2006). Can inhibition resolve retrieval competition through the control of spreading activation? *Memory & Cognition, 34*, 307–322.
- Soriano, M.F., Jiménez, J.F., Román, P., & Bajo, M.T. (2009). Inhibitory processes in memory are impaired in schizophrenia: Evidence from retrieval-induced forgetting. *British Journal of Psychology, 100*, 661–673.
- Storm, B.C., Bjork, E.L., Bjork, R.A., & Nestojko, J. (2006). Is retrieval success necessary for retrieval-induced forgetting? *Psychonomic Bulletin & Review, 13*, 1023-1027.
- Storm, B.C., & Levy, B.J. (2012). A progress report on the inhibitory account of retrieval-induced forgetting. *Memory & Cognition, 40*, 827-843.
- Storm, B.C., & White, H.A. (2010). ADHD and retrieval-induced forgetting: Evidence for a deficit in the inhibitory control of memory. *Memory, 18*, 265-171.
- Titz, C., & Verhaeghen, P. (2010). Aging and directed forgetting in episodic memory: A meta-analysis. *Psychology and Aging, 25*, 405-411.
- Tomlinson, T.D., Huber, D.E., Rieth, C.A., & Davelaar, E.J. (2009). An interference account of cue-independent forgetting in the no-think paradigm. *Proceedings of the National Academy of Sciences, 106*, 15588-15593.

- Tun, P.A., Wingfield, A., Rosen, M.J., & Blanchard, L. (1998). Response latencies for false memories: Gist-based processes in normal aging. *Psychology & Aging, 13*, 230-241.
- Veling, H., & van Knippenberg, A. (2004). Remembering can cause inhibition: Retrieval-induced inhibition as cue independent process. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 315-318.
- Verde, M.F., & Perfect, T.J. (2011). Retrieval-induced forgetting in recognition is absent under time pressure. *Psychonomic Bulletin & Review, 18*, 1166-1171.
- Verhaeghen, P., & Salthouse, T.A. (1997). Meta-analyses of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin, 122*, 231–249.
- Wilson, S.P., & Kipp, K. (1988). The development of efficient inhibition: Evidence from directed forgetting task. *Developmental Review, 18*, 86-123.
- Zellner, M., & Bäuml, K.-H. (2005). Intact retrieval inhibition in children's episodic recall. *Memory & Cognition, 33*, 396–404.
- Zellner, M., & Bäuml, K.-H. (2006). Inhibitory deficits in older adults—list-method directed forgetting revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 290–300.

Table 1: Demographic data for participants in Studies 1 and 2.

	Age range	M age (SD)	FSIQ (SD)	MMSE (SD)
STUDY 1				
Younger-old	60-64	62.05 (1.32)	118.34 (6.05)	28 (1.12)
Old	65-69	66.95 (1.28)	115.31 (5.94)	27.75 (.97)
Older-old	70-74	73.05 (1.36)	115.74 (6.44)	27.9 (1.07)
STUDY 2				
<i>Early Rp+ condition</i>				
Young adults	18-34	22.47 (4.80)	105.82 (8.68)	29.57 (.73)
Old	61-69	65.30 (2.44)	105.14 (7.14)	29.43 (.90)
Older-old	70-85	76.93 (4.06)	104.77 (6.16)	29.37 (.76)
<i>Early Rp- condition</i>				
Young adults	18-33	22.20 (3.81)	107.34 (7.71)	29.17 (1.26)
Old	61-69	64.47 (2.47)	104.69 (7.21)	29.40 (.89)
Older-old	70-84	77.23 (4.09)	105.09 (7.63)	29.37 (.89)

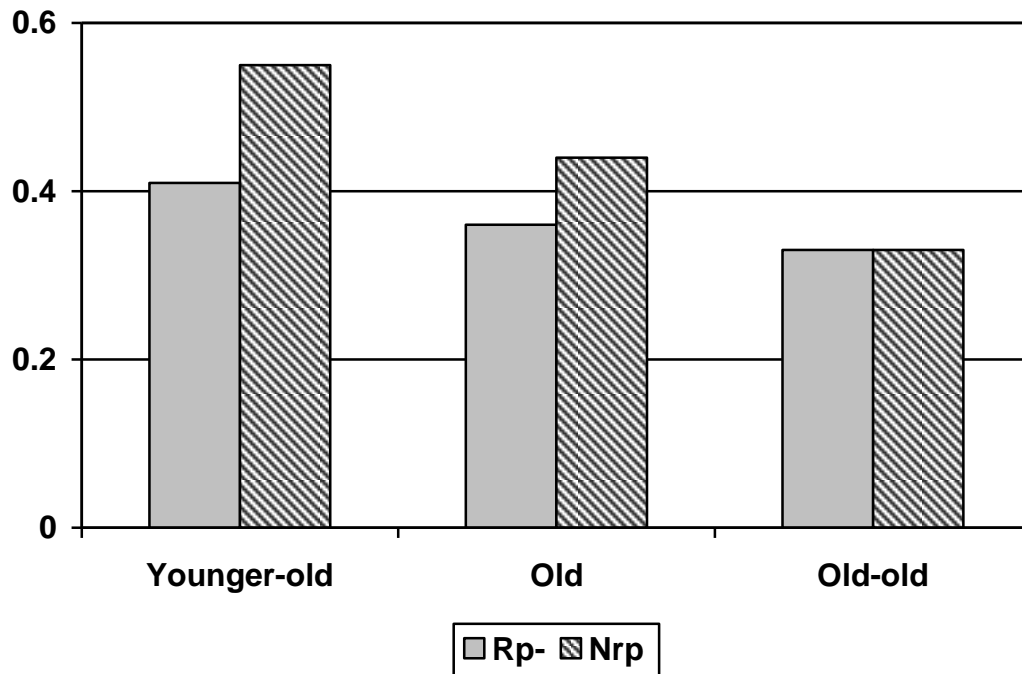
Table 2: Example of materials for early Rp+ condition (Study 2)

Rp+	Rp-
Curtain	Pajamas
Napkin	Robe
Sheet	Slacks
Dollar	Artichoke
Emerald	Lettuce
Lawn	Pepper
NsimRp-	NsimC
Belt	Briefcase
Boots	Saddle
Skirt	Whip
Mushroom	Chicken
Onion	Clam
Tomato	Turkey
NdissC(A)	NdissC(B)
Cannon	Needle
Grenade	Tack
Gun	Thorn
Dagger	Jackhammer
Spear	Siren
Sword	Traffic

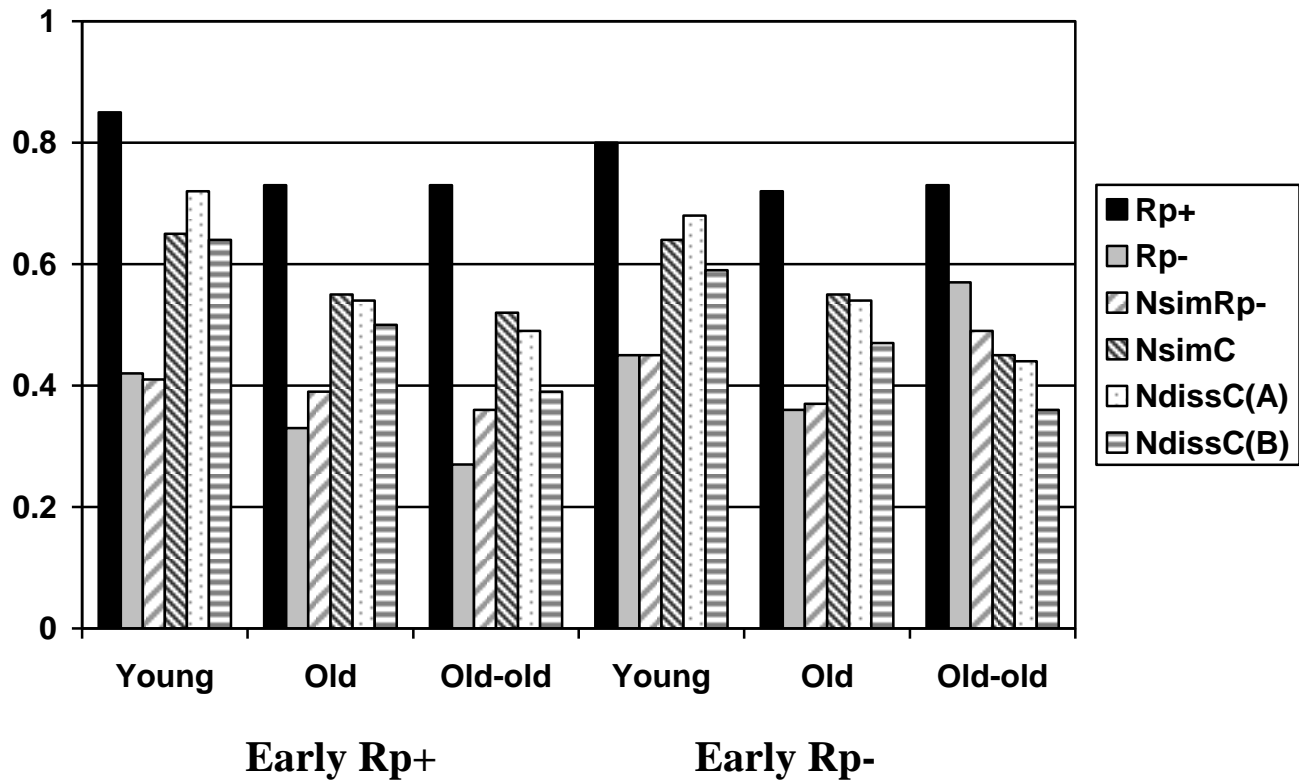
Figure 1: Mean proportion correctly recalled as a function of item type and age (Study 1)

Figure 2: Mean proportion correctly recalled as a function of item type, age and output order (Study 2)

Figure 1



Note. Rp- = unpractised items from practised categories. Nrp = unpractised items from unpractised categories.

Figure 2

Note. Early Rp+ = output of Rp+ items before Rp-, NsimRp-, and NsimC items. Early Rp- = output of Rp-, NsimRp-, and NsimC items before Rp+ items. Rp+ = practised items from practised categories. Rp- = unpractised items from practised categories. NsimRp- = unpractised items from unpractised categories that are semantically related to Rp- items. NsimC = unpractised items from unpractised categories that were drawn from an implicit category but which did not share a relationship with the Rp- items. NdissC(A) = Nrp-dissimilar items tested during first half of recall. NdissC(B) = Nrp-dissimilar items tested during second half of recall.

Appendix A

Cues, independent cues, and item types used (Study 1)

Category	Independent cue	Rp+ items	Rp- and Nrp items
Dishes	Coffee	Bowl, frying pan	Spoon, saucer
Insects	Flying	Ant, caterpillar	Dragonfly, ladybird
Quadruped	Africa	Pig, goat	Elephant, lion
Sports equipment	Round	Mat, bar	Discus, ring
Tools	Mason	Gripper, file	Chisel, trowel
Leather	Clothing	Briefcase, whip	Belt, boots
Green	Vegetable	Emerald, lawn	Artichoke, lettuce
Loud	Weapon	Siren, traffic	Cannon, grenade

Note. Materials taken from Aslan et al (2007)

Appendix B:

Category-exemplar word pairs and their semantic associations (Study 2)

Category						
Exemplar type	COTTON	LEATHER	GREEN	SOUPS	LOUD	SHARP
<i>Semantically dissimilar</i>	Curtain	Briefcase	Dollar	Chicken	Jackhammer	Needle
	Napkin	Saddle	Emerald	Clam	Siren	Tack
	Sheet	Whip	Lawn	Turkey	Traffic	Thorn
<i>Semantically similar</i>	Pajamas*	Belt*	Artichoke^	Mushroom^	Cannon ⁺	Dagger ⁺
	Robe*	Boots*	Lettuce^	Onion^	Grenade ⁺	Spear ⁺
	Slacks*	Skirt*	Pepper^	Tomato^	Gun ⁺	Sword ⁺
<i>Shared implicit category</i>	<i>Clothing*</i>		<i>Vegetable^</i>		<i>Weapon⁺</i>	

Note. Materials based on Anderson & Spellman (1995, Study 2, 3a and 4)

Acknowledgements

We would like to thank Anna Simpson and Marcelle Fernandes for assistance with data collection and to Pauline Hughes' invaluable assistance in recruiting older participants.