Monitoring Memory in Old Age: Impaired, Spared, and Aware

Alan D. Castel, Catherine D. Middlebrooks, and Shannon McGillivray
The Oxford Handbook of Metamemory
Edited by John Dunlosky and Sarah (Uma) K. Tauber

Subject: Psychology, Cognitive Psychology
DOI: 10.1093/oxfordhb/9780199336746.013.3

Abstract and Keywords

Although a variety of memory changes accompany old age, an important question is the degree to which older adults are aware of these changes, and how older adults may or may not accurately monitor their own memory. Monitoring refers to the ability to assess how well one will remember certain information at a later time. In some cases, older adults may be overconfident about memory performance, whereas, in other situations, older adults may be highly aware of their memory abilities and demonstrate appropriate predictions and insight. This chapter will provide an overview of current research regarding this topic to shed light on the degree to which metacognitive monitoring may be intact in older adults. We discuss why this has implications for how older adults can use strategies to selectively remember important information, as well as future directions for metacognitive aging research.

Keywords: metamemory, aging, monitoring, judgments of learning, confidence, memory, older adults

Aging can lead to a variety of changes in memory across the adult lifespan (for a review, see Hess, 2005; McDaniel, Einstein, & Jacoby, 2008; Zacks & Hasher, 2006). Given how memory abilities can change as we get older, the ability to accurately monitor what we will later remember, or forget, has important practical and theoretical implications. In many cases, healthy older adults are highly aware of potential declines in memory (Horhota, Lineweaver, Ositelu, Summers, & Hertzog, 2012), and this can be a source of concern and anxiety (Hess, 2005). Although many older adults may be concerned about changes in their memory abilities, it is critical to consider whether these assumptions and insights about memory reflect older adult’s actual memory capacity and abilities. Monitoring refers to the ability to assess how well one will remember certain information at a later time. There may be instances where older adults feel that they can remember recently encountered information, only to forget this information at a later time, leading to initial overconfidence. However, given lifelong experience with factors that lead to successful remembering, as well as memory challenges and failures, older adults may also be highly aware of the fallibility of memory. While memory impairments can be very concerning for older adults, the inability to effectively monitor, predict, and assess one’s memory can have obvious negative consequences. It can result in disappointment or embarrassment in situations where individuals are not able to remember what they think or say they will be able to, such as names or appointments. However, despite declines in explicit memory abilities, recent work suggests that efficient monitoring of memory may be relatively intact in healthy older adults (Hertzog & Dunlosky, 2011), allowing for accurate and strategic monitoring of what information will later be remembered.

What does “older adult” mean and who is a healthy older adult? As is customary in the cognitive aging literature, “older adults” are defined as individuals over the age of 65. We acknowledge that this age range criteria is more a matter of convention, as certain adults at the age of 80 may behave, in terms of memory performance, like that of a far younger person, and vice versa. Indeed, there is considerable variability in performance when using chronological age as a grouping variable (see Hultsch, MacDonald, Hunter, Levy-Bencheton, & Strauss, 2000;
Salthouse, 2013). We refer to healthy older adults as those who are living independently in the community, report good health, and are free of any diagnosed neurological disorder. The typical comparison group of younger adults is usually university students who volunteer for course credit, or are paid, and are typically in the age range of 18–25 years. Admittedly, one limitation of most of the studies discussed in this review is the use of an extreme age groups design, generally involving comparisons of college students (often from highly selective universities) and healthy/active older adult volunteers (who are often interested and concerned about changes in memory). A greater range of age sampling throughout the lifespan can provide important insight regarding the trajectory of changes that accompany aging (see Castel et al., 2011; Hertzog, Sinclair, & Dunlosky, 2010, for more inclusive lifespan samples). Recently, collecting data via the Internet may allow for a broader, diverse, sample and age range, but this method may also have some selection issues. In addition, longitudinal designs can provide important insight regarding change with age, as opposed to age-related differences observed from the cross-sectional designs more typically used in most of the studies discussed in this review.

The present chapter reviews methods used to assess potential age-related differences in the ability to monitor memory. These methods provide insight into the mechanisms used by younger and older adults to regulate memory. Further, older adults may use these mechanisms to selectively remember important information. While the review focuses on metacognitive monitoring of memory, this monitoring ability has direct implications regarding which behaviors younger and older adults engage in to improve memory (see Hertzog, this volume). For example, if someone believes that an item is not well learned, he or she might then choose to study that item for a longer period of time or use an alternate memory-enhancement strategy (e.g., Dunlosky, Kubiak-Silman, & Hertzog, 2003). In addition, both younger and older adults choose to study high-value information, at the expense of lower-value items, when they have control over which, and how many, items to study (Castel, Murayama, Friedman, McGilivray, & Link, 2013; Price, Hertzog, & Dunlosky, 2010). Thus, there are clear and important links between the monitoring operations and the control operations, which can lead to improvements in memory. We also outline how older adults monitor retrieval operations, such as feelings of knowing and tip-of-the-tongue (TOT) experiences, and assess confidence at test. We conclude by highlighting future directions regarding novel ways to assess and improve memory monitoring across the adult lifespan.

Potential Declines in Monitoring Memory in Older Adults

One of the most common methods to assess people’s prediction regarding how likely they will be to later remember certain information is the judgment of learning (JOL) (see Rhodes, this volume, for a description and review of JOL methods and recent research). For example, participants may be told that they are going to be shown a list of words; for each word, they are asked to assess how well they have learned the word and how likely they would be to recall this word on a later memory test. Alternatively, individuals may be asked to predict the percentage or number of items they believe they will be able to recall (i.e., to provide global JOLs). Early research on this topic typically examined absolute accuracy, such that the average JOL (e.g., 60%) is compared with the percentage of items actually remembered (e.g., 40%). Thus, absolute accuracy allows for the investigation of whether individuals’ JOLs reflect a pattern of underconfidence or overconfidence (such is the case in the example provided), or are well matched with actual memory performance, leading to a measure of calibration between JOLs and memory. Alternatively, relative accuracy examines whether the JOLs assigned by an individual can distinguish between which information is later remembered versus forgotten (see Nelson, 1984, 1996), and this approach is discussed later in the chapter. In the early work that first examined absolute (but not relative) accuracy, Bruce, Coyne, and Botwinick (1982) asked younger and older adults to predict the number of items they would be able to recall from a list of words. They were allowed as much time as they needed to study the words before completing a free recall test. Although JOLs were similar for younger and older adults, older adults recalled significantly fewer words. Thus, younger adults’ JOLs were more accurate relative to their actual recall performance, whereas older adults displayed a pattern of overconfidence in their judgments. Similar findings of global JOL inaccuracies among older adults have been demonstrated in other studies (e.g., Connor, Dunlosky, & Hertzog, 1997), thus providing initial support for the notion that older adults may suffer from memory and metamemory impairments.

It has also been suggested that age differences in metamemory abilities may contribute to poorer working memory performance (Bunnell, Baken, & Richards-Ward, 1999). Bunnell and colleagues investigated this hypothesis by assessing JOLs during a working memory task. They found that older adults’ JOLs prior to list presentation and after recall reflected patterns of overconfidence, compared with actual memory ability, whereas younger adults
demonstrated greater metamemory accuracy. The authors conclude that the differences observed in metamemory accuracy seem to reflect age-associated differences in awareness of working memory functioning.

More Recent Evidence of Preserved Monitoring in Older Adults

More recent research has suggested that some elements of metacognitive functioning remain intact with advancing age. For example, not all studies have found evidence of impairments in metacognitive judgments of learning (e.g., Dunlosky & Connor, 1997; Murphy, Sanders, Gabrieleski, & Schmitt, 1981). Although Connor et al. (1997) reported some overconfidence in global JOL accuracy among older adults (this was reduced in later experiments), their study also suggests that, on an item-by-item basis, older and younger adults may display similarly accurate JOLs. In a series of experiments, older and younger adults studied word pairs and gave either immediate or delayed JOLs (see Rhodes, this volume, for more details about this methodology). Immediate JOLs were made directly after study of the word pair, whereas delayed JOLs were made after a 10-item delay. Furthermore, prior to the starting the study phase of the experiment, participants were asked to give global JOLs (i.e., how many items they would be able to remember out of 60). As noted, older adults gave slightly more inaccurate global JOL predictions than did younger adults, although both age groups demonstrated greater accuracy when asked to give delayed rather than immediate JOLs. Similarly, on an item-by-item basis, younger and older adults continued to show increased calibration (the degree to which JOLs matched later performance) when JOLs were delayed versus immediate. Importantly, older adults did not demonstrate lower accuracy compared with younger adults on either the immediate or delayed JOLs when assessed on an item-by-item basis. These results indicate that, under some conditions, older adults are capable of monitoring their memory performance as effectively and as accurately as younger adults.

Devolder, Brigham, and Pressley (1990) investigated both predictions and global postdictions of memory performance across three samples of participants utilizing nine learning/memory tasks, including more ecologically valid tasks. These tasks included word learning, prose passages, face-name learning, and an appointment-keeping task. Regarding prediction accuracy (judgments made prior to learning), older adults were less accurate than younger adults on three of the nine tasks and more accurate on the appointment-keeping task; no age-related differences were found on the remaining five tasks. Older adults displayed comparable postdiction accuracy overall compared with younger adults, although individual analyses revealed that older adults were less accurate on three of the nine tasks and more accurate than younger adults on the remaining six tasks. Furthermore, improvements in accuracy were observed from predictions to postdictions on all nine tasks among older adults and on eight of the nine tasks among younger adults, thus indicating greater improvement in accuracy for the older participants. Overall, these data suggest that metacognitive performance awareness in older adulthood may remain relatively stable and accurately reflect changes in memory that occur with age.

Hertzog and colleagues have also examined JOL accuracy and monitoring abilities among younger and older adults (Hertzog, Dunlosky, Powell-Moman, & Kidder, 2002). Participants were presented with either related or unrelated word pairs (e.g., pants-belt or target-seed, respectively) for 10 seconds each and asked to give immediate JOLs. Both age groups exhibited better overall recall performance for related pairs compared with unrelated pairs, although older adults recalled fewer items. However, older adults’ average JOLs were closer (and nearly accurate) to their actual memory performance, whereas younger adults were slightly less accurate in the direction of underconfidence. Both groups gave higher JOLs for related compared with unrelated pairs, indicating that their judgments were sensitive to associative relatedness, although older adults were slightly more responsive to this type of information. The authors concluded that older adults do not, in fact, display age-related impairments in monitoring accuracy and, at times, may even be able to better calibrate memory expectations to actual performance. More recent work has shown that older adults account for associative memory deficits when monitoring their memory for words and proper names (Berry, Williams, Usualieva, & Kilb, 2013), thus suggesting some awareness of the distinction between item and associative memory impairments (see Naveh-Benjamin, 2000) and the challenges that accompany remembering proper names.

Absolute and Relative Accuracy of Judgments of Learning and Implications for Aging

Based on the prior review of research, an important distinction regarding JOLs and memory performance is the absolute versus relative accuracy of these judgments (see Rhodes, this volume). Early research typically
examined absolute accuracy, such that the average JOL (e.g., 60%) is compared with the percentage of items actually remembered (e.g., 40%). Thus, absolute accuracy allows for the investigation of whether individuals’ JOLs reflect a pattern of underconfidence or overconfidence or are well matched with actual memory performance, thereby leading to a measure of calibration between JOLs and memory. Alternatively, relative accuracy examines whether the JOLs assigned by an individual can distinguish between which information is later remembered versus forgotten, providing a measure of resolution, and is typically assessed using gamma correlations (Nelson, 1984, 1996). Simply put, higher relative accuracy (i.e., a positive correlation) occurs when higher JOLs are given to information later recalled and lower JOLs are given to information forgotten at test (see Higham et al., this volume, for important discussions on this topic). This is a crucial issue for cognitive aging, as older adults typically recall fewer items than do younger adults but may, for whatever reason, be using a JOL–scale range similar to that of younger adults. Thus, older and younger adults may use a similar range in which most JOLs are distributed, for example, between predictions of between 40 and 60%. Although younger adults’ memory performance may fall within this range for a given task (for example, recalling about 50% of the material), older adults’ performance is frequently lower, thus leading to apparently poorer calibration.

Although early research found evidence for age-related impairments in absolute accuracy (e.g., Bruce et al., 1982; Murphy et al., 1981), more recent investigations into the effects of aging on the relative accuracy of JOLs have often found little to no age-related differences (Connor et al., 1997; Hertzog et al., 2010; Hines, Touron, & Hertzog, 2009) and have even found slightly more accurate performance by older adults (Hertzog et al., 2002; see Hertzog & Dunlosky, 2011, for a recent review). Hertzog et al. (2010) conducted one of the most thorough and informative studies to date, using a large cross-sectional sample of people ranging from 18 to 81 years of age (including middle aged adults). Participants studied related and unrelated word pairs and made immediate JOLs; after this study–judgment phase, they received a test of paired-associate recall. Both age groups successfully predicted that they would remember the related pairs better than the unrelated pairs, although older adults remembered fewer items than the younger adults. The cross-sectional regression lines for JOL resolution (as measured by gamma correlations between JOLs and recall) as a function of age for all paired-associate items are presented in Figure 1. Older adults recalled fewer items and assigned lower JOLs, compared to younger adults, but no age-related declines in resolution across the adult lifespan were present in the sample. In fact, significant age-related increases were found in the aggregate gamma correlations.


Although some studies report a decline in absolute accuracy with age, others show a more complex pattern of findings. For example, Bruce et al. (1982) found that older adults' JOLs were lower than those of younger adults, whereas Murphy et al. (1981) found no age differences. Hertzog et al. (2010) also observed that older adults' JOLs were lower than those of younger adults, but they noted that this pattern was not consistent across all age groups. In addition, they found that older adults' JOLs were more accurate than those of younger adults, suggesting that older adults are better at generating accurate JOLs.

One important issue is to determine whether younger and older adults use similar cues when making JOLs, as the specific type of cue one relies on when forming a JOL could potentially lead to either an increase or decrease in metacognitive accuracy, depending on the validity of that particular cue (Koriat, 1997). Both age groups appear to be highly sensitive to item-relatedness, a factor that has a strong effect on later cued recall. As noted in a recent review by Hertzog and Dunlosky (2011), other studies have also shown that both younger and older adults effectively use similar encoding cues when making JOLs, which influence later memory. For example, the fluency or ease with which people generate mediators during associative learning seems to influence their JOLs, with faster generation leading to higher JOLs (Hertzog, Dunlosky, Robinson, & Kidder, 2003). Importantly, the magnitude of this
fluency effect is similar for older and younger adults’ JOLs (Robinson, Hertzog, & Dunlosky, 2006). Overall, the similar influences of differential cues on JOLs for people of different ages, along with equivalent JOL resolution, argues for relatively spared ability to monitor encoding operations across the adult life span.

This sparing suggests that older adults may be able to use metacognitive strategies or awareness to help overcome or compensate for age-related declines in memory performance. The ability to use metacognitive insight in a strategic manner is consistent with the theory of selective optimization with compensation (Baltes & Baltes, 1990), which suggests that successful aging is linked to older adults’ ability to selectively invest limited cognitive resources into areas that yield optimal returns. Thus, accurate metacognitive insight might have a more direct impact on memory performance in the ability to modify attention and goal-directed processing in a strategic manner (Castel, McGillivray, & Friedman, 2012; Hertzog & Dunlosky, 2011). Delaying JOLs improves resolution, as it encourages people to rely on long-term memory and retrieval factors when making JOLs, and not merely encoding fluency (for a recent metanalytic review, see Rhodes & Tauber, 2011). Importantly, delayed JOLs have equivalently high resolution for older and younger adults (e.g., Connor et al., 1997), thus suggesting that older adults may benefit from this form of retrieval to enhance monitoring.

Beyond JOL Measures of Monitoring in Older Adults

Monitoring with Consequences:

Although JOLs provide useful measures of people’s ability to assess their memory, there are no specific consequences associated with making either accurate or inaccurate JOLs. Thus, participants may have low motivation to exert the necessary effort to accurately assess whether they might remember or forget certain information. However, in the real world, forgetting can have consequences, especially when you did not anticipate forgetting certain information (e.g., you think you would remember a spouse’s birthday, but if you forget, there will be consequences).

To determine how consequences (in terms of gains and losses) may influence metacognitive judgments, McGillivray and Castel (2011) employed a novel paradigm in which participants were presented with lists of words paired with varying point values that indicated how much that word was worth (e.g., book 5, train 15, drill 1). Thus, some words were more important to remember. As participants were shown each word, they had to “bet” (yes or no) which items they would be able to remember. Participants received points for those items on which they originally bet “yes” and then subsequently recalled, but lost points for those items on which they had bet “yes” and then failed to recall. Thus, participants could choose which items, and how many items, to bet on. If a
participant bet on an item, she would receive whatever points were associated with that item if she was later able to recall it, but would lose those points if she failed to recall it. Participants did not lose points if they recalled items that they did not bet on; accordingly, they could recall any items at test, but would only receive points for items on which they had initially bet. The novel aspect of this study, compared to the more standard JOL paradigm, was that there were rewards (i.e., points) associated with accurately monitoring and predicting which items would be recalled and penalties if one failed to do so. Furthermore, individuals engaged in six study-test trials, with different words on each list, in order to assess the effects of task experience. Both younger and older adults strategically bet on and recalled more of the high-point than low-point value items, and there were no age differences in memory performance for the highest-valued items, a finding that is consistent with previous literature (Castel, Benjamin, Craik, & Watkins, 2002; Castel, Farb, & Craik, 2007). As is indicated in Figure 2, regardless of age group, participants were highly overconfident on initial lists (i.e., they bet on more items than they were actually able to recall), but this was significantly reduced with task experience. Furthermore, as shown in Figure 2, overall point scores on each list improved with task experience, and older and younger adults’ scores were comparable on later lists as a function of metacognitively accurate betting practices, despite the fact that older adults recalled less information overall. This suggests that older adults implemented strategies that actually led to improved calibration on the later lists in order to achieve goal-relevant outcomes (see also Hertzog, this volume).

Using Subjective Memory States to Predict Later Memory:

Recently, Soderstrom, McCabe, and Rhodes (2012) investigated whether older adults can successfully monitor the age-related declines in recollection (the detailed recall of specific events) typically evident in episodic memory tests. Using a novel procedure, younger and older adults were asked, while studying a series of word pairs, to indicate whether they expected to have a subjective feeling of recollection or familiarity during a later recognition memory test. These predictions of later subjective memory experiences, based on the remember–know procedure, are also known as judgments of remembering and knowing (JORKs). Specifically, participants predicted whether word pairs would be remembered (i.e., accompanied by recollective details, such as remembering some specific thoughts or events associated with the encoding of the pair) or known (i.e., accompanied by a sense of familiarity devoid of these specific and vivid recollective details) or forgotten on a later test. Compared with actual test performance, older adults were highly overconfident in predicting remembering, whereas younger adults’ predictions more closely corresponded with actual remembering. These data suggest that older adults may have difficulties monitoring age-related declines in recollection. Older adults may be overconfident when imagining retrieval states, and this may be tied to confidence judgments made during recognition memory tests (see later section on confidence judgments at test and overconfidence/hyperconfidence in older adults).

Task Experience and Underconfidence with Practice when Monitoring Learning

Multiple study-test blocks, particularly those consisting of unique items per block, serve as informative designs when determining whether or not younger and older adults are similarly able to update their monitoring after some degree of task experience and, if so, how. Some studies have found that older adults’ ability to accurately update metacognitive predictions are impaired compared to younger adults (Matvey, Dunlosky, Shaw, Parks, & Hertzog, 2002; Price, Hertzog, & Dunlosky, 2008), while other studies have found comparable benefits of task experience (Dunlosky & Hertzog, 2000; Hertzog & Dunlosky, 2011; McGillivray & Castel, 2011; Tullis & Benjamin, 2012). In general, research suggests that, when learning new items on each successive list, both older and younger adults lower their predictions and correct their initial overconfidence with task experience.

The use of multiple trials in the investigation of strategic metacognitive monitoring is often necessary, and evidence suggests that selectivity (the ability to effectively recall high-value items relative to lower values) may only emerge with task experience (Castel, Balota, & McCabe, 2009; McGillivray & Castel, 2011). For example, in Figure 2, both age groups enhanced their score after several study-test blocks consisting of different words and point values across lists. However, the typical design usually assesses the impact of task experience on metacognitive accuracy by presenting participants with the same set of information at least twice and examining the degree of improvement in predictions and strategy usage across lists (e.g., Rast & Zimprich, 2009; Tauber & Rhodes, 2012). Task experience and feedback may be particularly important for older adults (e.g., Jacoby, Wahlheim, Rhodes, Daniels, & Rogers, 2010) to learn to calibrate their predictions and confidence judgments with actual memory performance. Online monitoring needed for accurate predictions may tax attentional and working memory systems.
that can become compromised in old age (Bieman-Copland & Charness, 1994; Craik, 2002; Craik & Byrd, 1982; Hasher & Zacks, 1988), and older individuals may require more time and experience to adopt appropriate strategies and reach levels of performance on par with younger adults (McGillivray & Castel, 2011; Rogers, Hertzog, & Fisk, 2000; Touron, Hoyer, & Cerella, 2004).

When people study and are tested on the same set of material multiple times, a phenomenon often observed among both younger adults and older adults is the underconfidence-with-practice (UWP) effect. After repeated trials or testing, younger adults may give JOLs that underestimate memory performance (Koriat, Sheffer, & Ma’ayan, 2002). Rast and Zimprich (2009) extended this research to examine potential age differences. As in similar studies, participants were presented with word pairs and asked to provide either immediate JOLs or delayed JOLs (the delay averaged ~45 seconds). The design consisted of either two (Experiment 1) or five (Experiment 2) trials, with the same words featured during each trial. Within the first experiment, both younger and older adults were significantly overconfident in their judgments compared to accuracy on the first trial (older adults exhibited this trend to a greater extent). By the second trial, however, older adults exhibited improved calibration of their JOLs, while younger adults showed the UWP effect. Experiment 2 largely replicated these findings, in that only younger adults exhibited an UWP effect. Furthermore, as observed in other studies, delayed JOLs were more accurate than immediate JOLs for both age groups, indicating that more realistic expectations of memory performance occur if an adequate delay is provided.

One way to explain the UWP effect is that, after an initial study-test trial, participants rely on their memory for previous study-test experiences rather than monitoring the current learning of the information (memory-for-past-tests; see Finn & Metcalfe, 2007). Tauber and Rhodes (2012) had young and older adults learn word pairs and make JOLs before completing a memory test in each of three study-test trials on the same material. The results confirmed that both young and older adults relied on the memory-for-past test heuristic as a basis for JOLs. Both age groups showed enhanced recall on later lists, and although JOLs were higher than recall on List 1, recall was higher than JOLs on later lists (see Figure 3). In terms of measures of resolution, changes in memory for prior tests across trials were nominal for both groups. Furthermore, only the most recent past test influenced JOLs; earlier tests were unrelated to later judgments. In general, JOLs were influenced by prior-trial JOLs and were related to subsequent memory performance on the same trial. These findings suggest that, when making JOLs during repeated study-test cycles of the same material, memory for prior tests can guide both younger and older adults’ monitoring.

**Expertise, Emotional Material, and the Role of Social Context**

Levels of expertise, or prior knowledge, within a specific domain may facilitate memory performance, but the potential for expertise to either help or hurt one’s ability to monitor memory may become more pronounced with age. There is substantial work regarding the “curse of knowledge” (Hogarth, 1981), in which certain individuals who have extensive expertise in a particular domain find it extremely difficult to think about problems from the perspective of people who do not have this same or similar knowledge base. Recent work has shown that older
adults may exhibit similar biases when assessing information for which they have high degrees of familiarity, such as names of actors who were famous in 1950 versus 1990. Toth, Daniels, and Solinger (2011) found that older adults had lower JOL accuracy for 1950s actors despite having higher recollection for, and knowledge about, them. Thus, prior knowledge may serve as a “double-edged sword,” increasing the availability of details that facilitate memory, but also increasing feelings of familiarity that can reduce the accuracy of memory predictions.

Prior knowledge, social communication needs, and/or context can also influence what older adults recall. While younger and older adults can accurately monitor how well emotionally negative material will be remembered, older adults may not appreciate the increased memorability of emotionally positive information (Tauber & Dunlosky, 2012). Adams and colleagues have also investigated contextual manipulations with respect to participants’ memories for stories by varying to whom participants (both younger and older women) were asked to retell a story, either an experimenter or a young child (Adams, Smith, Pasupathi, & Vitolo, 2002). When the listener was an experimenter, the younger women recalled more propositional content than did the older women, but this age difference disappeared when the listener was a young child. Furthermore, when the listener was a child, participants engaged in more elaborations and repetitions while retelling the story, but the older women were more adaptive in appropriately adjusting the complexity levels, given the age of the listener. These findings underscore the importance of the context (in particular, social contexts) in which one is asked to recall information and the degree to which given contexts differentially motivate younger and older adults. In some cases, older adults may be especially astute at monitoring the knowledge base of another, such as a child, in order to effectively communicate information.

Given apparently intact monitoring abilities, healthy older adults may also use awareness of limited memory resources to selectively remember information that is relevant, consistent with prior knowledge, or that may be pertinent for future use. For example, older adults remember the prices of market-value grocery items (eggs $2.99), but not prices that are outlandish or inappropriate (milk $17.99; Castel, 2005), or when remembering the price of common items from the past versus the items’ price in the more abstract future (Castel, McGillivray, & Worden, 2013). When asked to imagine packing for a trip, older adults focus on remembering items that are essential and worry less about forgetting less important items (McGillivray & Castel, submitted), or they may focus on remembering important health-related information, such as critical medication side effects (Friedman McGillivray, Murayama, & Castel, in press). In addition, older adults will remember trivial facts that they found interesting and forget the less interesting ones (McGillivray, Murayama, & Castel, submitted). In general, older adults may use prior knowledge and interest to guide attention and encoding, and this may be based on a qualitatively different form of monitoring relative that used by younger adults.

Monitoring Recall Output and Forgetting in Episodic Memory

Although most of the work reviewed suggests that older adults can effectively monitor memory at encoding, very little work has examined how older adults monitor which, and how much, information will be forgotten at testing. Some studies have examined this form of monitoring via postdictive judgments, which, following a recall test, generally involves asking participants to indicate how many items they think they correctly remembered, given the total number of items that were presented during encoding (e.g., Bunnell et al., 1999; Hertzog, Saylor, Fleece, & Dixon, 1994). The consistent finding from such studies is that there are little to no age-related differences in the ability to accurately monitor performance that has already occurred (Baker, Dunlosky, & Hertzog, 2010; Brigham & Pressley, 1988; Bunnell et al., 1999; Devolder et al., 1990; Hertzog et al., 1994, 2010). Thus, older and younger adults are aware of how much they were able to remember under these conditions.

Other studies examining output monitoring have found age-related impairments in monitoring one’s own performance on a memory test. For example, older adults may be unaware of having told a particular story to someone multiple times (Jacoby, 1999). Within an experimental context, older adults were more likely than younger adults to classify previously recalled words as unrecalled (e.g., Einstein, McDaniel, Smith, & Shaw, 1998; Koriat, Ben-Zur, & Sheffer, 1988; Marsh, Hicks, Cook, & Mayhorn, 2007). Similarly, studies that examined memory-monitoring accuracy at the time of retrieval (i.e., indicating whether an answer that came to mind has a high or low probability of being correct) have shown age-related impairments (e.g., Kelley & Sahakyan, 2003; Pansky, Goldsmith, Koriat, & Pearlman-Avnon, 2009; Rhodes & Kelley, 2005). Despite some evidence of age-related deficits, recent work suggests that older adults may be just as aware as younger adults of how much information
they have forgotten when learning and recalling lists of categorized items (Halamish, McGillivray, & Castel, 2011), suggesting that certain types of monitoring of forgetting may remain relatively intact in old age.

Monitoring Retrieval and Feelings of Knowing in Older Adults

Many people, and perhaps more often older adults, may remark that “I can’t remember it, but I’d know it if I saw it.” The subjective experience of being unable to explicitly recall something, like an acquaintance’s name, but being confident in one’s ability to recognize it were it to be presented, is referred to as a “feeling-of-knowing” (FOK; see Thomas, Lee, & Hughes, this volume). Expressing a FOK reflects a prospective, metacognitive monitoring of retrieval processes, in that a judgment is being made about the plausibility of successfully retrieving and/or recognizing the sought-after information from one’s memory in the future (e.g., Hart, 1965; Hertzog & Dunlosky, 2011). Evidence suggests that people base their judgments of retrievability on contextual factors associated with the sought-after target information, particularly cue-familiarity and target-accessibility (Dunlosky & Metcalfe, 2009; Koriat & Levy-Sadot, 2001; Thomas, Bulevich, & Dubois, 2011). FOK judgments have been investigated in terms of their magnitude (high or low FOKs) and, predominantly, their resolution (i.e., the correlation between reported FOKs and subsequent recognition performance).

A number of differences have been noted regarding the magnitude and resolution of FOKs across the adult lifespan. Older adults tend to provide lower FOK ratings than do younger adults, even when the item is subsequently recognized, indicating lower levels of confidence in expected recognition performance (Hertzog & Touron, 2011; MacLaverty & Hertzog, 2009; Sacher, Isingrini, & Tacconat, 2013). These lower ratings may be a consequence of associative binding deficits and difficulties with context retrieval (Naveh-Benjamin, 2000); the greater the quantity of retrievable contextual details for an unrecalled item, the higher FOK judgment ratings tend to be (Koriat, 1993; Schwartz & Metcalfe, 1992; Souchay, Moulin, Clarys, Tacconat, & Isingrini, 2007). While there is some evidence for an apparent deficit in FOKs in older adults, findings are sometimes inconsistent regarding the degree, if any, to which older adults display specific deficits in episodic FOKs, relative to younger adults.

With respect to FOK resolution, there is no current evidence of age-related differences for semantic memories (Allen-Burge & Storandt, 2000; Anooshian, Mammarella, & Hertel, 1989; Butterfield, Nelson, & Peck, 1988). This may reflect the fact that access to, and successful retrieval of, semantic memories is generally well maintained across the lifespan (Nilson, 2003; Zacks & Hasher, 2006), factors that are critical to FOK judgment formation and accuracy (Bacon, Danion, Kauffmann-Muller, & Bruant, 2001; Koriat, 1993). Conversely, FOK resolution for episodic memories generally shows age effects, with older adults displaying lower levels of resolution than younger adults (e.g., Perrotin, Isingrini, Souchay, Clarys, & Tacconat, 2006; Souchay & Isingrini, 2012).

Two of the more prominent theories to have been put forth to explain this age-related discrepancy in impaired episodic versus accurate semantic FOKs are the inferential-deficit hypothesis and the memory-constraint hypothesis. The inferential-deficit hypothesis posits that older adults’ inferences about the likelihood of their successfully recognizing an unrecalled item are based on faulty evaluations of accessible contextual information that might facilitate recognition (Hertzog, Dunlosky, & Sinclair, 2010; Sacher et al., 2013; Thomas et al., 2011). Effectively and accurately utilizing this information in a manner that leads to FOK resolution depends on executive activities associated with frontal lobe function (Janowsky, Shimamura, & Squire, 1989; Perrotin, Tournelle, & Isingrini, 2008; Souchay, Isingrini, & Espagnet, 2000), which has been shown to decline with age (e.g., Braver & Barch, 2002). The memory-constraint hypothesis, on the other hand, argues that older adult FOK resolution is not a result of inferential deficits, but, rather, of original encoding strength (Hertzog et al., 2010). Older adults’ FOK resolution increases with increasing quality of encoding (e.g., strategy use), in some cases alleviating age differences entirely (MacLaverty & Hertzog, 2009; Sacher et al., 2013; Souchay et al., 2007). In fact, recent work by Eakin, Hertzog, and Harris (in press) using name-face associations has found no age-related differences in FOK resolution for either semantic memory (famous faces) or episodic memory (nonfamous novel faces), despite older adults showing poorer episodic memory, in terms of recollection, for the nonfamous faces compared to younger adults.

Tip-of-the-Tongue Experiences and Aging

A TOT experience, within the domain of metacognitive study, refers to the very strong, subjective feeling of
knowing a target item or piece of information but being unable to retrieve it at that particular moment (Brown & McNeill, 1966; Salthouse & Mandell, 2013; Schwartz, 2008; see also Schwartz & Cleary, this volume). According to the heuristic-metacognitive account, TOT experiences are believed to result from inferences about item retrievability, based on cue availability (of a semantic, syntactic, or phonological nature) or item fluency (Kornell & Metcalfe, 2006; Metcalfe, Schwartz, & Joaquin, 1993; Schwartz & Metcalfe, 2011). From a psycholinguistic perspective, theories of blocking (e.g., Burke, MacKay, Worthley, & Wade, 1991; Jones, 1989) and partial activation (also known as transmission-deficit hypothesis; e.g., MacKay & Burke, 1990; James & Burke, 2000) have also been proposed to explain TOT occurrence.

Regardless of whether one approaches TOTs from a psycholinguistic or metacognitive standpoint, diary- and laboratory-based research provides consistent evidence that older adults experience more frequent TOTs than do their younger counterparts, particularly with respect to proper nouns (Evrand, 2002; Gollan & Brown, 2006; Heine, Ober, & Shenaout, 1999; Schwartz & Frazier, 2005; Shafto, Stamatakis, Tam, & Tyler, 2010; but see White & Abrams, 2002). No consistent age differences appear evident, however, in terms of TOT resolution (Brown, 2012), and when younger adults do show better resolution than older adults, it may be a consequence of the time constraints of laboratory tasks (Brown, 2012; Schwartz, 2002). Diary studies suggest that older adults take more time to resolve TOTs than do younger adults (Burke et al., 1991; Heine et al., 1999), potentially resulting from age-related declines in cognitive processing speeds (Salthouse, 1994); the limited time in laboratory tasks may lead to underestimations of older adult TOT resolution rates.

To explain the age-related increases in TOT frequency, research has predominantly adopted either a decrement view or an incremental-knowledge view, though the two are not necessarily mutually exclusive. The decrement view holds that, as a consequence of general cognitive declines, older adults experience weakened connections within their lexical networks, making successful retrieval of target items increasingly difficult (Schwartz & Frazier, 2005). The incremental-knowledge theory postulates that older adults experience more TOTs than do younger adults because they have more semantic knowledge and, thus, more opportunities for information to become temporarily inaccessible (Dahlgren, 1998; Gollan & Brown, 2006), although recent work suggests that controlling for differences in knowledge does not eliminate these age differences (Salthouse & Mandell, 2013).

From a more metacognitive perspective, age-related increases in TOT frequency may be indicative of active metacognitive monitoring processes reflecting age-related retrieval deficits (Schwartz & Frazier, 2005). Knowing that one does, indeed, know the target item, but just cannot recall it at the moment, reflects more precise monitoring than just knowing whether it is presently recallable or not, potentially leading to longer search efforts for the target than if it were to be completely dismissed as not recallable. The fact that these processes are engaged is not directly related to their accuracy, as some work has shown that controlling for memory performance (i.e., retrieval success) does not mitigate age differences, meaning that these age-related differences in TOT frequency may stem from potential age-related deficits in metacognitive monitoring (Salthouse & Mandell, 2013).

**Familiarity and Heightened Confidence at Test in Older Adults**

Retrieving an item from memory does not speak to the fact that one actually encountered that specific item. Anyone who has ever argued, “I could be wrong” or “don’t quote me on that” has provided a retrospective confidence judgment, a metacognitive assessment of the probable accuracy of a retrieved memory (Dunlosky & Metcalfe, 2009; Hertzog & Dunlosky, 2011). Given what is well known about the fallibility of memory (e.g., Schacter, 1999, 2013), how good are people, particularly older adults, at recognizing when their memory is likely to be accurate versus inaccurate? This process of monitoring knowing and retrieval has been investigated predominantly in terms of calibration, the degree to which overall confidence ratings match overall performance. In general, people tend to be systematically underconfident or overconfident in their performance within a particular task (Dunlosky & Metcalfe, 2009), although this can depend on a number of variables.

As proposed by the metamemory hypothesis (Brewer & Sampaio, 2006; Brewer, Sampaio, & Barlow, 2005), confidence judgments may be formed on the basis of multiple memory processes, predominantly utilizing source information, and one’s beliefs about what the products of these operations mean with respect to the probable accuracy of a retrieved target memory. Such processes, which are also implicated in the formation of metacognitive predictions and JOLs, can include feelings of familiarity, fluency, self-consistency, and other
heuristics, such as accessibility, distinctiveness, and representativeness (Brewer & Sampaio, 2012; Dodson & Schacter, 2002; Gilovich, Griffin, & Kahneman, 2002; Koriat, 2012; Tversky & Kahneman, 1974).

With respect to memory for general or frequently encountered knowledge, confidence judgment accuracy and calibration does not appear to significantly differ by age (Dahl, Allwood, & Hagberg, 2009; Dodson, Bawa, & Slotnik, 2007; Pliske & Mutter, 1996). For recent, episodic events, however, evidence suggests that older adults, compared to younger adults, tend to express greater levels of confidence for incorrectly remembered items, even when memory accuracy is equated across age groups (Chua, Schacter, & Sperling, 2009; Dodson & Krueger, 2006; Fandakova, Shing, & Lindenberger, 2013; Norman & Schacter, 1997). For example, Figure 4 (data and figure from Dodson, Bawa, & Krueger, 2007, Experiment 2) displays both cued-recall accuracy and confidence judgment calibration for older adults with 10 seconds study time per word pair (Young-10s), younger adults with 2 seconds study time per word pair (Young-2s), and older adults with 10 seconds study time per word pair (Older). Even when young adults were equated with older adults in terms of cued-recall accuracy and, presumably, encoding quality (when comparing the Young-2s and Older), older adults showed poorer calibration (leading to a higher calibration error score), meaning that the older adults are not as well-calibrated as younger adults when making confidence judgments at test. This pattern of miscalibrated confidence judgments also extends to tasks specifically designed to elicit false memories, such as the misinformation paradigm (Loftus, Miller, & Burns, 1978). Older adults appear to be more likely than younger adults to express high confidence in their false memories for task-provided misinformation, indicating a mismatch between confidence ratings and memory performance (Dodson & Krueger, 2006; Karpel, Hoyer, & Toglia, 2001; Mather & Johnson, 2003; Mitchell, Johnson, & Mather, 2003).

The high-confidence errors made by older adults may be partly owing to flawed associative binding and miscombinations of item-source memories (Kroll, Knight, Metcalfe, Wolf, & Tulving, 1996; Lyle, Bloise, & Johnson, 2006; Naveh-Benjamin, Hussain, Guez, & Bar-On, 2003). In addition, older adults may tend toward an overreliance on familiarity at test in various settings (e.g., “Your face is familiar, but where do I know you from?”), as a result of recollection impairments (see Jacoby, 1999). The use of these processes when retrieving information may then lead to documented increases in episodic misrecollections and false recognition rates (Jacoby, Bishara, Hessels, & Toth, 2005; Kensinger & Schacter, 1999; Rhodes, Castel, & Jacoby, 2008; Tun, Wingfield, Rosen, & Blanchard, 1998). Miscombining source features would not affect the application of the aforementioned metacognitive processes, but the concluding confidence judgment would be unknowingly based on false premises. This also has some implications for more real-world situations, such as eyewitness memory, in which older adults may display poorer memory compared to younger adults but may report strong confidence.

Related to a miscombination of source features, older adults may also have difficulty inhibiting irrelevant source information (Jacoby & Rhodes, 2006), the retrieval of which may inappropriately influence metacognitive assessments and, therefore, confidence judgments (Brewer, Marsh, Clark-Foos, & Meeks, 2010). Older adults also tend to falsely recall and recognize semantically related information (Norman & Schacter, 1997), suggesting the use of gist-based retrieval processes, which are often associated with high levels of confidence for related, but not originally presented, information. Recollection quality has also been advanced as an explanation for high-
confident information, as have general, overarching monitoring impairments, potentially resulting from variables ranging from anxiety (e.g., age-related stereotype threat; see Hess, 2005) to age-related working memory declines (Wong, Cramer, & Gallo, 2012).

Summary

The present review outlines the mechanisms that appear to be spared (e.g., relative monitoring of encoding, aspects of monitoring forgetting) in older adults, as well as potential age-related impairments (e.g., output monitoring, overconfidence at test). Older adults also benefit from task experience, such that they can update judgments and use strategic metacognitive monitoring to selectively remember important information. Older adults may report some changes in FOK states and more frequent TOT experiences, but, under some conditions, show strong resolution between these states and later memory retrieval. Overall, as this review illustrates, despite certain changes and impairments in memory, older adults may be well aware of how (their) memory works and which aspects decline with age. However, older adults may exhibit certain biases in metacognition, especially when relying on gist- or familiarity-based processing during recognition memory tests, thus leading to heightened (but misaligned) confidence judgments at retrieval.

Future Directions

The next wave of research regarding how older adults monitor their memory may consider several interrelated issues. The history of metacognitive monitoring and cognitive aging teaches us that the methodology used to examine accuracy can impact the conclusions. Early work suggested that monitoring, in terms of absolute accuracy, was impaired in older adults, while more recent work suggests that relative accuracy remains intact. The measures that are developed and used to assess these constructs are critical, and novel ways of assessing monitoring and control will provide useful insight regarding how aging impacts metamemory. Future research will benefit from continuing to develop measures that can assess older adults’ beliefs about monitoring, as well as decisions regarding control operations (see also Hertzog, this volume). In addition, these measures need to be consequential (e.g., McGilivray & Castel, 2011), such that people will make these metacognitive predictions with the knowledge that being overconfident or underconfident will lead to particular consequences.

There is a tendency in most experimental tasks to focus on one, or sometimes several, distinct and experimentally manipulated variables (e.g., item difficulty, presentation time) that can influence memory and monitoring. Future research could consider how older adults take into account interference, retrieval conditions, and anxiety/stereotype threat, for example, when making judgments regarding memory performance, as well as how these variables may lead to biases in metacognitive accuracy. It may also be important to assess how beliefs about memory change can influence older adults’ predictions (e.g., Horhota et al., 2012; Plaks & Chasteen, 2013), with the hope that new research can shed light on how older adults’ perspectives and beliefs about memory can shape how they engage in online monitoring of memory accuracy (see also Mueller, Tauber, & Dunlosky, 2013 for relevant research with younger adults) and the degree to which memory is reconstructive and modified by retrieval efforts.

More practical, applied research may help in terms of understanding how monitoring may or may not work effectively in everyday settings, especially those concerning older adults. It is important to know whether research findings with word pairs translate to more complex settings. For example, older adults may feel that they will remember information that a doctor recently told them, or take certain medications, but this potential use of prospective memory may not always be reliable (see also R. Smith, this volume). Thus, older adults may benefit from knowing how to effectively remember critical information when presented with a large amount to remember, especially in settings that involve medical or financial decision making.

Finally, as we get older, we may need to learn to be more judicious about what we try to remember. Older adults may, in fact, be good at prioritizing what is important to remember and may use prior knowledge to supplement this ability (see Castel, 2008; Castel et al., 2012; Hess, Rosenberg, & Waters, 2001). In addition, older adults may also be especially astute at monitoring what other people (such as a spouse) might remember, or need/want to know, such that effective communication can be accomplished to enhance memory in various applied settings. New research that examines how younger and older adults can selectively monitor which, and how much, information
should be remembered will shed light on important links between monitoring and control operations, thus potentially leading to new methods of enhancing memory efficiency in both younger and older adults.

Acknowledgments

We thank Matt Rhodes and Chris Hertzog for helpful comments on an earlier version of the chapter. Preparation of the chapter was supported in part by a grant from the National Institutes of Health (National Institute on Aging), Award Number R01AG044335.

References


Monitoring Memory in Old Age


Monitoring Memory in Old Age


---

**Alan D. Castel**  
Alan D. Castel, Department of Psychology, University of California, Los Angeles, Los Angeles, CA

**Catherine D. Middlebrooks**  
Catherine D. Middlebrooks, Department of Psychology, University of California, Los Angeles

**Shannon McGillivray**  
Shannon McGillivray, Department of Psychology, Weber State University, Utah