Age-Related Differences in Adults' Ability to Follow Spoken Instructions

Agnieszka J. Jaroslawska[1], Glen Bartup[1], Alicia Forsberg[2], & Joni Holmes[3]

[1] The University of Edinburgh

[2] The University of Missouri

[3] MRC Cognition and Brain Sciences Unit, The University of Cambridge

Author note

This work was funded by the Centre for Cognitive Ageing and Cognitive Epidemiology at the University of Edinburgh. Materials, data, and analysis code are available on the Open Science Framework (https://osf.io/3pv7e/). Author contributions: Data were collected by GB, and analyzed by AF. Experiments were programmed by AJ and GB, collaboratively designed with the other authors. The manuscript was written by AJ with all the other authors providing feedback. AJ is now at Queen's University Belfast, Belfast, UK.

Correspondence concerning this article should be addressed to Agnieszka J. Jaroslawska, School of Psychology, Queen's University Belfast, Belfast, UK, BT7 1NN. E-mail: a.jaroslawska@qub.ac.uk

Abstract

A growing body of research illustrates that working memory capacity is a crucial limiting factor in our ability to follow spoken instructions. Despite the ubiquitous nature of instruction following throughout the lifespan, how the natural ageing process affects the ability to do so is not yet fully understood. In this study we investigated the consequences of action at encoding and recall on the ability to follow spoken instructions. Younger (< 30 y/o) and older (> 65 y/o) adults recalled sequences of spoken action commands under presentation and recall conditions that either did or did not involve their physical performance. Both groups showed an enacted-recall advantage, with superior recall by physical performance than oral repetition. When both encoding and recall were purely verbal, older adults' recall accuracy was comparable to that of their younger counterparts. When action was involved at either encoding or recall, however, the difference in performance between the two age groups became pronounced: enactment-based encoding significantly improved younger adults' ability to follow spoken instructions; there was no such advantage for older adults. These data show that spatial-motoric representations disproportionately benefit younger adults' memory performance. We discuss the practical implications of these findings in the context of lifelong learning.

Keywords: Working Memory; Following Instructions; Ageing; Enactment; Action Advantage

Age-Related Differences in Adults' Ability to Follow Spoken Instructions

In the course of everyday life people often have to complete novel tasks and use systems or technologies they have not previously encountered. Multi-step instructions are the most common method for communicating how to perform such unfamiliar tasks. The experience of learning through instruction is typically associated with childhood and adolescence, but as the population ages, it is becoming increasingly important to understand age-related changes in the ability to follow instructions. This is because older adults remain engaged in the work force later in life (Helman, Copeland, & VanDerhei, 2010) and learning new skills may serve as a shield against age-related cognitive decline (Park et al., 2014).

Despite the everyday importance of following instructions, relatively little is known about the cognitive processes involved, how our abilities to do so change across the lifespan, or how to improve these abilities. To date, studies of instruction-guided behaviour have recognised a vital role for working memory in maintenaning the content of instructions (e.g., Engle, Carullo, & Collins, 1991; Gathercole, Durling, Evans, Jeffcock, & Stone, 2008; Jaroslawska et al., 2016b; Kim, Bayles, & Beeson, 2008) and demonstrated that memory for instructions can be enhanced by physical movement both during encoding of the instruction sequence and at recall (e.g., Jaroslawska et al., 2016a). Crucially, although cognitive changes in older age have been charted using conventional working memory tasks and other standardised assessments (e.g., Logie, Della Sala, MacPherson, & Cooper, 2007; Deary et al., 2007), there have, to our knowledge, been no systematic evaluations of changes in the ability to follow instructions, or of the possible mechanisms underpinning any such changes as people age.

Understanding the precise mechanisms mediating the translation of verbal instructions into physical actions can provide insights into the constraints on successful instruction-following. This will help us tailor instruction-based information to people of different ages which, in turn, may help facilitate the effectiveness of such information and ultimately accelerate learning. The aim of the work presented here was to establish whether older adults benefit from physical movement when following spoken instructions to the same extent as younger adults. We also sought to clarify whether the ability to follow spoken instructions varies between younger and older participants. In the sections that follow we first present empirical evidence for the link between working memory skills and instruction-following before focussing on the two key phenomena responsible for action benefits in following instructions: the effects of enactment-based encoding and action-based recall.

The Link Between Working Memory and Instruction-Following

Arguably, the most important cognitive skill required for dealing competently with instructions is working memory — the ability to simultaneously process and store information (e.g., Baddeley, 2012; Barrouillet & Camos, 2015; Cowan, 1999; Logie, 2011). Although there is little consensus on the precise definition of working memory (for reviews see Aben, Stapert, & Blokland, 2012; Adams, Nguyen, & Cowan, 2018; Cowan, 2017; Logie, Camos, & Cowan, in press), it is widely agreed that successful working memory performance depends not only on one's ability to maintain information for short periods of time, but also on overall processing efficiency, and the ability to resist the distraction that processing activities create (e.g., Engle, 2002; Engle, Tuholski, Laughlin, & Conway, 1999; Jarrold & Towse, 2006). Working memory capacity varies as a function of age, with older individuals demonstrating performance decrements on tasks requiring the concurrent storage and processing of information (e.g., Bier, Lecavalier, Malenfant, Peretz, & Belleville, 2017; Forsberg, Fellman, Laine, Johnson, & Logie, 2020; Holtzer, Stern, & Rakitin, 2004; Jaroslawska & Rhodes, 2019; Rhodes et al., 2019).

Working memory is a source of individual variation in instruction-following in both children and adults (e.g., Brener, 1940; Engle et al., 1991; Gathercole et al., 2008; Jaroslawska, Gathercole, & Holmes, 2018; Jaroslawska et al., 2016b; Kaplan & White, 1980). Multiple studies have used the dual-task methodology to explore the potential role of different sub-components of the tripartite working memory model (e.g., Baddeley & Hitch, 1974; Baddeley, 1986; Baddeley & Logie, 1999; Logie, 2011) in enacting spoken instructions and verbally repeating them in young adults. Across these studies, memory for instructions was impaired by concurrent activities taxing phonological short-term memory, spatial short-term memory, and executive control, suggesting that the encoding and retention of verbal instructions depends on multiple aspects of working memory (e.g., Jaroslawska et al., 2018; Yang, Gathercole, & Allen, 2014; Yang, Allen, & Gathercole, 2016). An individual differences study conducted with children by Jaroslawska et al. (2016b) further indicated that *verbal* aspects of working memory involved in both simple and complex span tasks played a specific role in performing actions to command.

To our knowledge, the only study directly comparing younger and older adults' ability to follow verbal instructions was conducted by Kim et al. (2008) who designed a task requiring participants to sort coloured pills into containers in response to spoken instructions, mimicking daily scenarios familiar to older adults. The information load of the instructions was manipulated by varying both the number of actions per instruction, and the number of components to be remembered per action (e.g., *Take three white pills on*

Tuesday afternoon and one red pill on Thursday morning). As the information load of the instructions increased, accuracy decreased for both younger and older participants. Crucially, however, this effect was amplified for the older group. Overall, both verbal working memory and age predicted the older adults' abilities to follow spoken instructions. All participants performed more accurately when the instruction contained fewer actions, even if those actions contained more qualifiers (e.g., time of day), suggesting that it was the number of actions that had an effect on the cognitive load and not the complexity of the language. From a practical perspective, these findings suggest that when adults deal with procedural instructions loaded with content, processing will be enhanced when it requires fewer actions.

The Benefits of Enactment-Based Encoding and Action-Based Recall

The recall of instructions can be enhanced by physical movement. Two separate mnemonic effects are of interest here. First is the advantage associated with action-based retrieval, which is observed when sequences of practical instructions are physically performed, rather than verbally repeated, at test (e.g., Allen & Waterman, 2015; Gathercole et al., 2008; Jaroslawska et al., 2016a; Koriat, Ben-Zur, & Nussbaum, 1990; Yang et al., 2014; Yang, Allen, Holmes, & Chan, 2017). Koriat et al. (1990), for example, presented participants with short sequences of action phrases relating to real objects (e.g., *move the eraser*, *lift the cup*, *flip the coin*) and found that action-based recall was significantly more accurate than verbal recall. Crucially, when participants encoded the sequence in anticipation of enacted recall their accuracy on a surprise verbal test was also improved, indicating that the mnemonic advantage reflects the benefits of action planning during encoding. Koriat et al. (1990) hypothesised that planning for actions facilitates the formation of an integrated multimodal representation involving phonological, visual, and motor codes. This multimodal representation (which is not present when verbal repetition is expected) integrates elements from various channels into a coherent representation. Although a robust benefit of action-based recall has been shown in a number of studies with children and young adults (Allen & Waterman, 2015; Gathercole et al., 2008; Jaroslawska et al., 2016a; Koriat et al., 1990; Yang et al., 2014), the advantage is yet to be tested in a healthy older population.

The second effect of interest is the advantage driven by enactment-based encoding, which occurs when the physical performance of to-be-recalled actions at the time of presentation improves their subsequent recall over short delays (e.g., Allen & Waterman, 2015; Charlesworth, Allen, Morson, Burn, & Souchay, 2014; Jaroslawska et al., 2016a; Waterman et al., 2017; Wojcik, Allen, Brown, & Souchay, 2011; Yang et al., 2017). For example, Allen and Waterman (2015) presented sequences of verbal instructions (e.g., *flip the hexagon*, *spin the cross*, *touch the square*) and asked young adult participants to enact each action step using shapes laid out in front of them, as they were presented during encoding. They found that the enactment of actions during the encoding phase significantly facilitated subsequent memory performance. Interestingly, Allen and Waterman (2015) reported that the mnemonic advantage of action-based recall was found to be reduced when combined with enactment during encoding (see also Jaroslawska et al., 2016a; Waterman et al., 2017), indicating that the two phenomena have a common origin. Jaroslawska et al. (2016a) proposed that when performing physical actions during encoding or planning for action recall, participants may actively construct action plans that incorporate spatio-motoric information and representations of intended movements (Koriat et al., 1990; Wolpert & Ghahramani, 2000), which are held in a specialised motor store in working memory (see also Jaroslawska et al., 2018; Smyth & Pendleton, 1989, 1990). In line with Jaroslawska et al.'s (2016a) interpretation, planning to execute action sequences did not enhance recall when the instructions had already been enacted at presentation, because the executed action sequence was already represented in the hypothesised motor store.

Although the beneficial effects of enactment-based encoding have been observed for instruction-following in groups of healthy older adults and patients with mild Alzheimer's disease (Charlesworth et al., 2014), it is not yet known whether this manipulation benefits older and younger adults to a similar extent. Speculatively, physical movement may be less effective at boosting memory performance in older participants given the well established age-related decline in sensorimotor control and functioning. With advanced age come coordination difficulties (e.g., Seidler, Alberts, & Stelmach, 2002), greater spatial and temporal movement variability (e.g., Contreras-Vidal, Teulings, & Stelmach, 1998; Cooke, Brown, & Cunningham, 1989), and slowing of movement (e.g., Diggles-Buckles, 1993). Motor skills also increasingly demand cognitive control with advancing age (e.g., Baltes & Lindenberger, 1997; Li & Lindenberger, 2002). For example, gait and balance require additional higher-level cognitive input as people get older (Yogev-Seligmann et al., 2010). Proposed mechanisms that necessitate the need for greater cognitive control include sensory losses, impaired or less automated sensorimotor performance, and declines in the efficiency of cognitive control (Baltes & Lindenberger, 1997; Li & Lindenberger, 2002; Wingfield, Tun, & McCoy, 2005). Consistent with these findings, the reduced capacity hypothesis suggests that motor tasks exert higher demands on the executive resources of older than younger adults to achieve comparable levels of performance (e.g., Wollesen & Voelcker-Rehage, 2014; Verhaeghen & Cerella, 2002; Verhaeghen, Steitz, Sliwinski, &

Cerella, 2003). A recent study comparing the benefits of self-enactment to visual demonstration at encoding in young adults found that demonstration provided larger boosts to performance than self-enactment (Allen, Hill, Eddy, & Waterman, 2019). Allen and colleagues argued that the benefits of enactment may be smaller than those for demonstration due to the additional cognitive costs of actively generating visuospatial and motoric representations during encoding with enactment (see also Waterman et al., 2017). In the case of demonstration, motoric representations may be automatically recruited by perceptual systems when observing another individual carrying out actions, thereby imposing minimal additional demands on the cognitive system. These findings show that enactment, while enhancing the recall of instructions over spoken repetition, induces cognitive demands in younger adults. Speculatively, increases in cognitive demands elicited by self-enactment might have a particularly detrimental impact on older adults' memory performance, given the already increased demands of motor tasks on the executive system.

The Present Study

The absence of work directly comparing the impact of implementing physical movement in the context of following instructions across the lifespan limits our understanding of the extent to which these manipulations are universally beneficial. The aim of the work presented here was to provide the first direct assessment of instruction-following in groups of younger and older adults. We sought to establish whether the cognitive skills supporting instruction-guided behaviour vary in composition between younger and older adults, and to clarify whether the modes of encoding and recall that can help mitigate immediate memory limits for younger individuals are equally beneficial for older adults. More specifically, we assessed whether older adults' ability to successfully implement sequences of instructions can be enhanced by physical movement in the same way that it can for typically developing children and young adults.

In this study, healthy younger (< 30 years of age) and older (> 65 years of age) participants were required to recall instruction sequences that increased in length in a span-type procedure. The instruction sequences consisted of descriptions of actions to be performed on a set of concrete, three-dimensional props (e.g., *Pick up the red book and then move the yellow pen*). The instructions were either spoken aloud to the participants, or performed one step at a time by the participants during encoding. At recall participants either attempted to repeat the instruction sequence verbally, or to perform it on an array of props laid out in front of them. To maximize standardization of procedures, pre-recorded audio clips were used to present the sequences of instructions. Cognitive tests measuring verbal and visuospatial aspects of working memory (Alloway, 2007) and planning and set-switching (Delis, Kaplan, & Kramer, 2001) were administered alongside the instruction tasks to allow for a partial replication of previous findings from Jaroslawska et al. (2016b) and Kim et al. (2008) and to provide new insights concerning the cognitive mechanisms involved, and how these might vary with age. Unlike Kim et al.'s (2008) study, which used cognitive tests measuring only aspects of verbal memory (i.e., digit span, word span, listening span), visuospatial assessments were included here.

Given the well established age-related deficits in working memory span performance (e.g., Bier et al., 2017; Holtzer et al., 2004; Jaroslawska & Rhodes, 2019), it was predicted that older adults would be less successful at implementing and verbally recalling spoken instructions than their younger counterparts. More specifically, we expected to see significant age effects on the experimental instruction task, as well as on standardised measures of cognitive ability. Based on the findings from Jaroslawska et al. (2016b) and

Kim et al. (2008), it was expected that participants' performance on the verbal complex span task would be a good predictor of instruction task performance. Although previous findings from Charlesworth et al. (2014) suggest that older adults benefit from enactment-based encoding, it could also be the case that the cognitive cost associated with generating physical movements (e.g., Allen et al., 2019) attenuates the benefit of enactment for older participants. Therefore, we made no specific predictions with respect to the impact of enactment-based encoding on older adults' ability to follow spoken instructions. The advantage of action-based recall has not been previously observed in older adults, but the robust evidence for this benefit in other age groups led us to expect that older adult would also benefit from the manipulation and of performing the action sequences at recall (e.g., Gathercole et al., 2008; Jaroslawska et al., 2018; Yang et al., 2014, 2016). However, if participants know they have to enact at recall, they might be planning for actions and forming integrated multimodal representations at encoding (e.g., Koriat et al., 1990), which might come at an additional cognitive cost, and therefore impair performance in older adults.

Method

Participants

Younger and older adults were recruited from the student population of the University of Edinburgh, the Psychology Research volunteer panel, and the wider community of Edinburgh. The final sample consisted of 40 younger adults (80.00% female) with a mean age of 23.33 years (*SD* = 2.38, age range = 18 — 29 years) and 38 older adults (68.42% female) with a mean age of 71.05 (*SD* = 3.00) ranging between 66 and 81 years. Participant characteristics for the final sample are presented in Table 1).

Table 1 Participant characteristics and raw background cognitive test scores (means and standard deviations in parenthesis) split by age group.

	Younger Adults		Older Adults	
	Enactment	No enactment	Enactment	No enactment
N	20	20	19	19
Age	23.65 (1.98)	23.00 (2.73)	71.05 (2.48)	71.05 (3.52)
Years of Education	17.05 (1.82)	16.85 (1.76)	17.32 (4.11)	16.53 (3.58)
MoCA	28.45 (1.47)	28.05 (1.50)	28.22 (1.66)	27.37 (2.06)
AWMA				
Verbal tasks				
Listening recall	19.50 (3.59)	19.40 (3.60)	16.58 (2.97)	17.95 (3.64)
Backward digit recall	25.30 (5.69)	24.10 (5.46)	26.16 (4.73)	26.37 (5.90)
Visuospatial tasks				
Spatial recall	27.10 (5.62)	23.55 (6.10)	20.05 (4.14)	21.37 (4.14)
Odd-one-out	31.45 (4.32)	30.25 (3.95)	28. 21 (3.66)	29.37 (5.45)
D-KEFS				
Card sorting test	13.50 (1.91)	12.80 (1.79)	11.47 (2.37)	11.21 (2.44)
Tower test	20.50 (3.28)	18.40 (2.41)	18.63 (4.18)	19.89 (3.03)

Note: MoCA = Montreal Cognitive Assessment; AWMA = Automated Working Memory Assessment; D-KEFS = Delis-Kaplan Executive Function System.

All participants were fluent speakers of English, with no history of neurological damage, no problems with hearing, and no problems with vision (unless corrected). To ensure that participants over the age of 65 showed no evidence of cognitive dysfunction incommensurate with normal ageing, the Montreal Cognitive Assessment (MoCA, Nasreddine et al., 2005), which is a measure of global cognitive functioning, was administered to all volunteers prior to the main experimental tasks. As there are known issues with the standard cut-off score used to establish cognitive impairment with the MoCA (it appears to have a high false-positive rate as shown by Rossetti, Lacritz, Cullum, and Weiner (2011)), participants were excluded from the study if they scored less than 20 on the MoCA (Waldron-Perrine & Axelrod, 2012). No participants were excluded from the final analysis due to poor performance on the MoCA. There was no clear evidence for a difference in MoCA scores between age groups ($\Delta M = -0.47$, 95% CI $[-1.25, 0.31]$, $t(67.96) = -1.19$, $p = .237$, $\mathrm{BF}_{01} = 2.26$), suggesting that all older adults were cognitively healthy.

Stimuli and Apparatus

Following Instructions. This study used a modified version of the following instructions paradigm developed by Gathercole et al. (2008) and subsequently modified by Jaroslawska et al. (2016a). Participants were required to recall instruction sequences that increase in length in a span-type procedure. Each span consisted of a block of five trials (i.e., five to-be-recalled instruction sequences). Testing began at two actions per sequence (e.g., *Pick up the red book and then flip the yellow pen*), and increased by one action per block (e.g., *Pick up the red book and then flip the yellow pen and then point to the black mug*), until reaching a maximum span of five actions per sequence. The instruction sequences consisted of descriptions of actions to be performed on a set of concrete, three-dimensional props. The objects were a set of four items (a pen, a mug, a book, and a [toy] phone) in each of four colours (red, blue, yellow, and black), with four actions to perform upon them (point to, pick up, flip, move). Action phrases were concatenated using “and then” to produce increasingly longer sequences of instructions that varied in length but not in lexical complexity. The items used in each instruction were selected at random, with the constraint that there was no repetition of colour and object combination in the sequence as a whole. Four alternative lists of 20 instruction sequences (see supplementary materials) were audio-recorded using the Mac OSX system voice ‘Serena’ which simulates a standard British accent. Stimulus delivery was controlled using OpenSesame (version 3.1; Mathôt, Schreij, & Theeuwes, 2012). The audio presentation of each instruction phrase was separated by a 1500ms interval (e.g., *Pick up the red book* — 1500ms — *and then flip the yellow pen*). Trial timings were piloted ahead of data collection to ensure that participants were able to enact each object-action command in the specified amount of time. Performance was scored in terms of the proportion of action phrases recalled correctly (i.e., the proportion of action-colour-object combinations recalled in their correct serial position in the sequence), out of a possible 70 per experimental condition. All elements — actions, objects and colours — recalled in their original position in the instruction sequence were also scored separately. A sample instruction *Pick up the red book and then move the yellow pen* has six features: two actions, two colours, and two objects. Perfect performance over 20 trials (i.e., four blocks of five trials) would result in a score of 210 features correctly recalled (30 features in block one, 45 features in block two, 60 features in block three, and 75 features in block four).

Working Memory. Participants completed four standardised sub-tests from the Automated Working Memory Assessment (AWMA; Alloway, 2007). These included two tests of verbal working memory (i.e., listening recall and backward digit recall) and two tests of visuospatial working memory (i.e., spatial recall and odd-one-out). All measures of working memory were complex span tasks comprising processing components and storage components. Participants were required to recall a series of digits in reverse serial order for the backward digit recall task. In the listening recall test, participants verified whether series of sentences were factually true or false, before recalling the final word of each of the sentences in the order in which they had heard them. In the spatial recall test, pairs of identical shapes were presented on screen. The shape on the right had a dot on it and appeared in one of seven rotated positions. Participants were asked to judge whether the shape on the right was the same as or the mirror-image of the shape on the left for each pair, before recalling the locations of the red dot (from one of the three possible positions) in the order in which they had seen them. In the odd-one-out task participants were presented with sets of three shapes displayed in a row. They were asked to identify the shape that was the odd-one-out in each row, and then recall the location of the odd-one-out shapes in order. Trials were presented in blocks of six. Each task started at a span of one item (except for backward digit recall, which started at two items) and increased in length by one item in each subsequent block. If a participant responded correctly to the first four trials within a block, the program automatically proceeded to the next block (i.e., next span level). If three errors were made within a block, the task was discontinued. Raw scores were reported for all tests (see Table 1). In addition, *Z*-scores were computed for each variable and averaged to provide composite scores for each of the two aspects of working memory (i.e., verbal and visuospatial).

Cognitive Flexibility and Problem-Solving. Two standardised subtests from the Delis-Kaplan Executive Function System (D-KEFS; Delis et al., 2001) were administered: the Sorting Test which is a measure of cognitive flexibility in a problem-solving task, and the Tower Test which measures spatial planning, rule learning, inhibition, and the ability to establish and maintain instructional sets. In the free sorting condition of the Sorting Test, the participants were presented with six mixed-up cards that displayed both perceptual features and printed words. The participants were then asked to sort the cards into two groups, with three cards per group, according to as many different concepts or rules as possible, and to describe the concepts employed to generate each sort. The card set had a maximum of eight target sorts: three based on verbal-semantic information, and five based on visuospatial features of patterns on the cards. The performance was scored in terms of accuracy of the sorting responses and the descriptions of sorting concepts. The materials for the D-KEFS Tower Test included five discs that varied in size and a board with three vertical pegs. Each trial began with the experimenter placing from two to five discs on the pegs in a predetermined starting position and displaying a picture of the tower to be built (i.e., the ending position). The participant's task was to move the discs across the three pegs to build the target tower in the fewest number of moves possible. In constructing the target towers, the participant had to follow two rules: 1) to move only one disc at a time and 2) never to place a larger disc over a smaller disc. The minimum number of moves possible for a correct solution varied from 1 to 26. The performance was scored in terms of: the total number of moves, final achievement (correct or incorrect tower), and number of rule violations. Raw scores were reported for all tests (see Table 1). In addition, *Z*-scores were computed for each variable and averaged to provide a composite score of cogntiive flexibility.

Design and Procedure

The following instructions task involved the manipulation of two variables across two age groups: first, was the way in which instruction were to be encoded; second, the type of recall required. This resulted in a 2 × 2 × 2 mixed design with a within-subject factor of recall type (verbal, action) and between subject-factors of presentation format (enactment, no enactment) and age group (younger adults, older adults).

Participants were invited to the University of Edinburgh and assessed individually in a single testing session which lasted approximately 2 hours. Each participant was seated at a table opposite the test administrator. All 16 items used in the following instructions task were positioned randomly on a large desk within arm's reach of the participant. The object array was in view at all times, but the location of the props varied randomly between conditions. All participants were randomly allocated to one of two presentation conditions: enactment or no enactment. The two versions of the following instructions task (i.e., no enactment/enactment at presentation followed by verbal recall, and no enactment/enactment at presentation followed by action recall) were administered in a counterbalanced order, with different conditions separated by other cognitive tasks. Lists of instruction sequences were randomised across conditions. The instruction sequences were played out loud over speakers. In the no enactment at encoding conditions, participants listened to the instruction sequences without being allowed to manipulate any of the objects. At the end of each trial, participants heard a beep prompting them to recall the sequence by either performing the actions (serial action recall) or repeating them back to the experimenter (serial verbal recall). In the enactment at presentation conditions, participants performed each action phrase immediately after its auditory presentation (i.e., during the 1500ms interval). All instruction sequences were broken down into single action phrases (e.g., *Pick up the red folder* — break — *and then point to the yellow pen*). As with the no-enactment at encoding conditions, recall was requested after presentation of the full instruction sequence, and was either action-based or verbal and signaled with a brief sound. Recall was self-paced in all conditions. Performance was scored manually by the experimenter at the time of testing. All volunteers received a small honorarium (£20) in return for taking part in the study. The experimental protocol was approved by the ethics committee of the University of Edinburgh (271-1617/2).

Data Analysis

Prior to the analysis, proportion correct scores were arcsine-transformed to stabilize the variance. To analyze the data, a model comparison approach based on Bayes factors, implemented with the BayesFactor package in R (Morey & Rouder, 2018; R Core Team, 2018) was used. Bayesian statistics provide a better foundation for probabilistic inference than null hypothesis significance testing (e.g., Kruschke, 2011; Raftery, 1995; Wagenmakers, 2007). In our implementation, Bayes factors (BF) reflect the weight of evidence in favor of omitting a particular component from a model containing all relevant available variables. Bayes factors in favor of a particular main effect (over the null model) or interaction (over main effects only) are reported as BF_{10}, whereas Bayes factors in favor of the null are reported as BF_{01} ($BF_{10} = 1/BF_{01}$).

To obtain Bayes factors we used the default settings of the anovaBF and generalTestBF functions (i.e., ‘medium’ prior scale for fixed effects, and ‘nuisance’ prior scale for the random effect as recommended by Rouder, Morey, Speckman, & Province, 2012), with the modification that ‘whichModels’ was set to ‘top’, to compare linear versions of the full model (M_0), including all main effects and interactions, with each different model in which a given experimental parameter was omitted (M_1). This family of priors was designed to be broadly applicable and invariant with respect to linear transformations of measurement units (Rouder et al., 2012). It was also found to be more conservative than conventional ANOVAs (Rouder, Speckman, Sun, Morey, & Iverson, 2009; Wetzels et al., 2011) and is commonly considered suitable for Bayesian ANOVAs in working memory research (e.g., Oberauer & Eichenberger, 2013; Rhodes, Parra, Cowan, & Logie, 2017). We specified 50,000 Markov chain Monte Carlo (MCMC) iterations and ran an additional 10,000 iterations until the proportional error associated with each Bayes factor was less than 5%, similar to Rhodes, Parra, and Logie (2016). The `anovaBF` function quantifies the strength of evidence B in favor of the full model (M_0) in comparison to the reduced model (M_1 in which the parameter of interest is removed) in light of the data, returning the Bayesian likelihood ratio of M_0 and M_1 (BF_{01}).

The output is interpreted as follows: the observed data is B times more likely under the full model (M_0) than under the reduced model (M_1). A value < 1 indicates evidence that an omitted parameter was important, while $B > 1$ indicates evidence it was not. B can range from 0: (indicating overwhelming support for the full model that includes the parameter M_0), through 1 (indicating equal support for both models), to infinity (providing overwhelming support for the reduced model that omits the parameter M_1; Dienes, 2012). By symmetry, $1/B$ provides evidence against retaining the parameter in the model. A large BF_{10} value indicates strong evidence for including the parameter; i.e., that it was important in predicting the data. Conversely, a large BF_{01} value indicates strong evidence that the parameter was not important. Note, however, that Bayes factors cannot conclusively be interpreted using threshold cut-off points; subjective judgmental interpretation is necessary. Although typically, BF value of 1 is considered as providing no evidence, a BF between 1 and 3 is considered as providing anecdotal evidence, and a BF greater than 3 is considered as providing substantial evidence (Jeffreys, 1961; Wetzels & Wagenmakers, 2012), these labels are subjective, so we apply them only tentatively and urge readers to evaluate the strength of evidence provided by the the Bayes factor values for themselves.

Results

The raw scores obtained on standardized measures of working memory and cognitive flexibility are presented in Table 1. With respect to visuospatial tasks from the AWMA (Alloway, 2007), there was strong evidence indicating that younger adults performed better than older adults on the spatial recall test ($\Delta M = -4.61$, 95% CI [-6.95, -2.28], $t(69.13) = -3.94, p < .001$, $\text{BF}_{10} = 120.53$) but weak evidence for an age difference in scores from the odd-one-out task ($\Delta M = -2.06$, 95% CI [-4.04, -0.08], $t(74.05) = -2.07, p = .042$, $\text{BF}_{10} = 1.49$). For verbal measures of working memory, we found evidence for a difference in listening memory scores ($\Delta M = -2.19$, 95% CI [-3.74, -0.63], $t(76.00) = -2.80, p = .007$, $\text{BF}_{10} = 6.31$), with better performance in the younger group, and inconclusive evidence against a difference in performance on the backward digit recall task by age group ($\Delta M = 1.56$, 95% CI [-0.87, 4.00], $t(76.00) = 1.28, p = .205$, $\text{BF}_{01} = 2.11$). T-tests performed on the scores obtained on the sub-tests of the D-KEFS (Delis et al., 2001), revealed weak evidence against an age difference in spatial planning and inhibition ($\Delta M = -0.19$, 95% CI [-1.71, 1.34], $t(72.02) = -0.24, p = .807$, $\text{BF}_{01} = 4.15$) and strong evidence that younger adults' performance on the Card Sorting test was significantly better than that of their older counterparts ($\Delta M = -1.81$, 95% CI [-2.77, -0.84], $t(70.14) = -3.73, p < .001$, $\text{BF}_{10} = 76.70$).

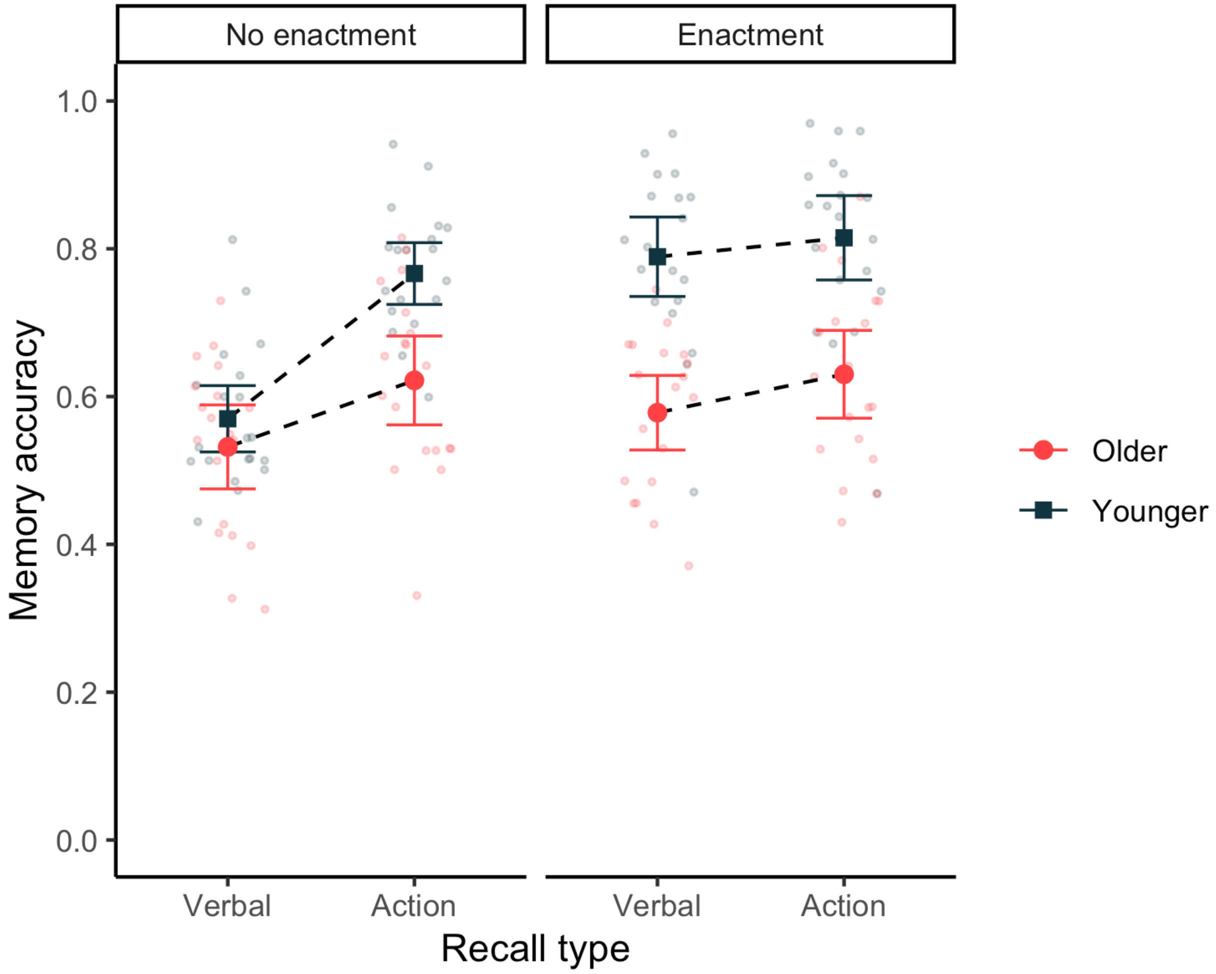

Figure 1 Memory accuracy on the following instructions task, split by presentation format (enactment, no enactment), type of recall (verbal, action), and age group (younger adults, older adults). Points are individual scores (jittered within groups to reduce overlap) with means and 95% confidence intervals overlaid.

Each group's performance on the following instructions task (i.e., the proportion of sequences correctly recalled) is presented on Figure 1. Memory accuracy for each individual feature of the instruction (i.e., action, colour, object) can be found in the supplementary materials. The arcsine-transformed accuracy scores were submitted to a 2 × 2 × 2 mixed ANOVA with a within-subject factor of recall type (verbal, action) and between subject-factors of presentation format (enactment, no enactment) and age group (younger adults, older adults). The analysis revealed strong evidence for an effect of age group ($\mathrm{BF}_{10} = 3.21 \times 10^6$, $\hat{\eta}_p^2 = .312$), presentation format ($\mathrm{BF}_{10} = 93.47$, $\hat{\eta}_p^2 = .127$), and recall condition ($\mathrm{BF}_{10} = 5.87 \times 10^6$, $\hat{\eta}_p^2 = .144$). There was also some evidence for an interaction between the enactment condition and age group ($\mathrm{BF}_{10} = 4.90$, $\hat{\eta}_p^2 = .066$), but not for an interaction between recall condition and age group ($\mathrm{BF}_{01} = 1.36$, $\hat{\eta}_p^2 = .011$). Finally, there was evidence for an interaction between the format of presentation and type of recall ($\mathrm{BF}_{10} = 53.83$, $\hat{\eta}_p^2 = .046$), and weak evidence for a three-way interaction (presentation format × recall type × age group, $\mathrm{BF}_{10} = 2.85$, $\hat{\eta}_p^2 = .020$).

To probe the nature of the three-way interaction, we compared the effect of presentation and recall conditions separately for each age group, and found strong evidence *for* an interaction in young adults ($\mathrm{BF}_{10} = 89.85$, $\hat{\eta}_p^2 = .110$), and weak evidence *against* an interaction in older adults ($\mathrm{BF}_{01} = 2.12$, $\hat{\eta}_p^2 = .006$). The interaction reflects the fact that the mnemonic advantage of action-based recall was reduced following enactment at presentation, indicating that the two effects may be driven by a common mechanism. In line with this interpretation, pairwise comparisons performed on the data from both younger and older participants revealed clear evidence for the advantage of action-based over verbal recall when there was no enactment during encoding (younger group: $\mathrm{BF}_{10} = 6.44 \times 10^4$, $t(19) = -7.79, p < .001$; older group: $\mathrm{BF}_{10} = 76.36$, $t(18) = -4.30, p < .001$) and weak or inconclusive evidence against an action-effect in the enactment condition (younger group: $\mathrm{BF}_{01} = 2.60$, $t(19) = -7.79, p < .001$; older group: $\mathrm{BF}_{01} = 1.05$, $t(18) = -4.30, p < .001$). As demonstrated on Figure 1, when both encoding and recall were purely verbal, older adults' recall accuracy was comparable to that of the younger group. When action was involved at either learning or test, however, the difference in performance between the two age groups became marked. Physical movement disproportionately benefited younger participants, with less clear effects of enactment-based encoding and action-based recall in older adults. Finally, we compared the role of enactment on verbal recall in the two age groups. In younger adults, we observed strong evidence for a difference ($\mathrm{BF}_{10} = 6.49 \times 10^4$, $t(34.26) = -6.40, p < .001$), indicating that enactment benefitted verbal recall. This was not true in older adults ($\mathrm{BF}_{01} = 1.66$, $t(35.52) = -1.29$, $p = .206$), where verball recall accuracy was equivalent irrespective of the mode of presentation.

Next, we assessed whether encoding condition (enactment, no enactment), mode of recall (action, verbal) indices of cognitive ability (verbal working memory composite, visuospatial working memory composite, cognitive flexibility composite) and age group (younger, older) were good predictors of performance on the following instructions tasks. As explained above, *Z*-scores were computed for each sub-test of the AWMA (Alloway, 2007) and D-KEFS (Delis et al., 2001) and averaged to provide composite scores for verbal aspects of working memory, visuospatial aspects of working memory, and cognitive flexibility. We used the generalTestBF function in the BayesFactor R package which allows comparison of models containing both categorical and continuous factors. Recall performance was the dependent variable. Similar to the ANOVA analyses, BF_{10} indicates the strength of the evidence for including the specified parameter, BF_{01} against including it.

For younger adults, there was an effect of enactment condition ($t(74) = 4.36, p < .001$, $\mathrm{BF}_{10} = 893.39$), and recall condition (verbal or action, $t(74) = -4.39, p < .001$, $\mathrm{BF}_{10} = 550.27$) on performance. However, there was no strong evidence for an effect of working memory (for visuospatial composite: $t(74) = 1.02, p = .311$, $\mathrm{BF}_{01} = 1.90$; for verbal composite: $t(74) = 2.11, p = .038$, $\mathrm{BF}_{01} = 2.05$) or cognitive flexibility ($t(74) =$

$0.47, p = .643$, $\mathrm{BF}_{01} = 2.74$). For older adults, recall performance was predicted by enactment condition ($t(70) = 2.64, p = .010$, $\mathrm{BF}_{10} = 3.57$) and recall condition (verbal or action; $t(70) = -3.51, p = .001$, $\mathrm{BF}_{10} = 36.55$), mirroring the pattern observed in the younger cohort. Moreover, performance was affected by visuospatial aspects of working memory ($t(70) = 6.37, p < .001$, $\mathrm{BF}_{10} = 5.25 \times 10^5$) but not by cognitive flexibility ($t(70) = 0.50, p = .622$, $\mathrm{BF}_{01} = 4.51$) or verbal working memory ability ($t(70) = -0.28$, $p = .777$, $\mathrm{BF}_{01} = 4.84$).

Finally, a series of correlations was conducted to explore the links between the composite scores of verbal working memory, visuospatial working memory, and cognitive flexibiity and performance on the following instructions task; these are presented in the supplementary materials (see Table S1 and Table S2). For older adults, performance across all four conditions of the following instructions task was linked with the visuospatial working memory composite. Additionally, older adults' performance in the conditions that *did not* involve enactment at encoding was significantly associated with verbal aspects of working memory. For younger adults, the only significant link emerged between verbal working memory composite and performance in the condition combining enactment at encoding and action at recall.

Discussion

This study set out to investigate, for the first time, whether healthy older adults benefit from enactment-based encoding and action-based recall when following spoken instructions in the same way that children and younger adults do. The key findings can be summarized as follows. First, in line with the extant literature, physical engagement both during presentation and at recall boosted younger adults' accuracy of remembering instructions over short intervals. The novel finding was that enactment during presentation did not improve older adults' instruction following. Second, the mnemonic advantage of action-based recall was diminished following the initial enactment of instruction sequences for both age groups. And finally, the pattern of associations between following instruction abilities and verbal aspects of working memory was inconsistent across age groups. We discuss each of these findings in turn.

In line with a growing literature, the data revealed that planning for or implementing a set of physical actions facilitates working memory performance in younger adults (e.g., Allen & Waterman, 2015; Gathercole et al., 2008; Jaroslawska et al., 2016a; Koriat et al., 1990; Yang et al., 2014, 2017). There were clear beneficial effects of both enactment at encoding and action at recall for this group. With respect to older adults, the inclusion of physical movement during recall reliably improved memory performance *only* when there was no physical enactment during presentation. Crucially, the enactment of instructions during presentation did not benefit recall in older adults. This outcome is particularly interesting because it suggests that older adults are more likely to benefit from the internal reproduction of action representations within working memory without overt output. In other words, older adults seem to benefit from action *planning* but not from physical performance during encoding. It is possible that the cognitive cost associated with generating visuospatial and motoric representtions reduces the benefits of enactment, as proposed by Allen et al. (2019), and because motor imagery and overt motor abilities exhibit different trajectories of age-related decline. Linear decreases in performance as a function of increasing age have been demonstrated with motor tasks such as repetitive finger tapping (e.g., Shimoyama, Ninchoji, & Uemura, 1990), but more complex, non-linear effects are seen in more demanding timed tasks and with visually guided hand movements

(Houx & Jolles, 1993; Kauranen & Vanharanta, 1996). This is particularly relevant in the context of the following instructions task used in the present study which involved the manipulation of three-dimensional props. Furthermore, some estimate that age-related slowing of self-initiated response times begins at 24 years of age (Thompson, Blair, & Henrey, 2014). In contrast, motor imagery ability does not seem to be diminished until much later in life. For example, in a study comparing four age groups, Schott (2012) found that motor imagery skills were better in young adults compared with older adults 70 years and older, but not in older adults 60 to 69 years of age.

Our findings are at odds with those of Charlesworth et al. (2014) who observed beneficial effects of enactment-based encoding for instruction-following in healthy older adults and patients with mild Alzheimer's disease. In the present study, on the other hand, the accuracy of verbal recall in the older group was equivalent irrespective of the mode of presentation. A potential source of this discrepancy may be that serial order recall was not explicitly required in the study conducted by Charlesworth et al. (2014), and pairings of action and objects were scored as correct regardless of the order in which they were recalled in the sequence. Requiring serial order and accurate recall of the positon of the actions and objects in the current design likely increased task demands. Speculatively, this increased task difficulty may have had a moderating effect on the impact of self-enactment on memory accuracy. Notwithstanding, considerably more work needs to be conducted to determine the generalisability and reproducibility of the mnemonic benefits of enactment-based encoding in the context of working memory paradigms.

The present study was designed to further our understanding of whether action-based recall had additive benefits over and above self-enactment at presentation. We found that it did not; the magnitude of the action-recall advantage was diminished when participants acted out the instructions during presentation. This is consistent with Allen and Waterman (2015) and Jaroslawska et al. (2016a) who showed that the effects of enactment during encoding were dependent on the type of recall required. Together, these outcomes suggest that the benefits of action at presentation and recall have a common origin. Both enactment-at-encoding and action-at-recall seem to operate via enhancing the encoding of visuospatial and motoric aspects of a sequence and thus are functionally equivalent. Put differently, both enactment-based encoding and action-based recall recruit additional, and likely highly similar, forms of coding that supplement the verbal input of the instructions (e.g., Jaroslawska et al., 2018). From an applied perspective, this suggests that when instructions are enacted at presentation, there is little scope for further improvement through performing instructions at test. For example, acting out how to use a new medical device, such as an inhaler, during the learning phase at a GP surgery might not benefit later recall. This further supports the idea that additional, or alternative, forms of coding within working memory are useful to supplement performance, in line with Logie's (2011) description of a working memory system as a collection of cognitive functions that can be flexibly deployed in different ways, depending on the task.

Instruction-following in children and older adults has been linked to verbal aspects of working memory both when the instructions were to be performed with physical props and in the more complex context of the virtual school environment (Jaroslawska et al., 2016b; Kim et al., 2008). Jaroslawska et al. (2016b) found this association was entirely restricted to conditions in which the spoken instructions were not enacted at presentation, indicating that recall was supported by verbal storage of the instructions within working memory as well as, in the case of action recall, a proposed motor store (see Jaroslawska et al., 2018). With enactment of the instructions at encoding, however, recall accuracy was related to visuospatial measures of working memory but only when the recall also involved physical movement. In the present study, performance on all versions of the following instructions task was linked with visuospatial aspects of working memory, but that relationship held only for older adults. Moreover, older adults' performance in the conditions that *did not* involve enactment at encoding was associated with verbal aspects of working memory, partially replicating Jaroslawska et al. (2016b). The findings from the younger group were less clear cut: there was a single significant link between verbal working memory composite and performance in the condition combining enactment at encoding and action at recall. Here, baseline differences in spatial ability between younger adults assigned to the two enactment conditions may have confounded the links between task performance and cognitive abilities. Future studies should consider employing a fully repeated measures design to control for individual differences in participants' overall levels of performance. Clearly, more research on this topic needs to be undertaken before the relationship between working memory and instruction-following is clearly understood.

Implications and Future Lines of Enquiry

Instructions are an inherent part of everyday life (e.g., when learning how to use new software or devices, assembling furniture, and following recipes or medication schedules), and impairments in the ability to carry out instructions can lead to difficulty in meeting vocational and family demands, and seriously undermine functional independence in old age. According to data from the World Population Prospects, one in six people in the world will be over 65 years of age in 2050, up from one in 11 in 2019 (Desa, 2019). People over the age of 80 years represent the fastest growing sector in the population of most high-income countries (Hazra & Gulliford, 2017), which increases the need to create a society where people have broad opportunities to learn throughout their lives. And this relies on instruction-following.

The findings of the present study clearly suggest that although incorporating physical engagement within instructions during curricular activities may have the potential to accelerate learning among children and young adults, these means may be less useful for older adults. Older adults benefited from planning for action (i.e., in anticipation of action-based recall), rather than from action per se, indicating that the use of motor imagery, rather than actual physical movement, may be more effective at improving older adults' memory performance. Motor imagery refers to the process of internally reproducing action representations in working memory without overt action output (Decety & Grèzes, 1999). Within the motor imagery literature, it is often assumed that imagery produces an internal forward model that predicts the process of action execution (e.g., Wolpert, 1997), and simulates actual execution by activating similar brain networks (e.g., Jeannerod, 2001, 1995). Motor imagery is used to improve performance in a wide range of disciplines from post-injury rehabilitation to sports (Schuster et al., 2011). Employing motor imagery may reduce the cognitive cost of actually performing actions during encoding, but nevertheless provide additional visuospatial and motoric codes necessary to improve recall. Future research should endavour to directly compare the effect of self-enactment and motor imagery and further explore the potential of using motor imagery practice to improve instruction-following, and potentially accelerate life-long learning, among the elderly. Another potential line of enquiry could address the issue of trial timing by systematically manipulating the amount of time spent encoding/enacting each step of the instruction. Although all participants taking part in the present experiment had ample time to enact each object-action phrase, a longer encoding period could possibly better enable positive effects to emerge in the older group (cf. Rhodes et al., under review).

Conclusion

The work reported here addressed an important question of how the natural ageing process affects the ability to follow spoken multi-step instructions. Our study illustrates how enactment-based encoding and action-based recall can boost younger adults' ability to recall sequence of instructions. The current findings also indicate that there is potential for the motor skills to be compromised in the older age ranges, which may limit the effectiveness of overt physical movement aimed at improving working memory performance. Crucially, the results reported here point to the possibility of using motor imagery as a mnemonic strategy for older individuals.

References

Aben, B., Stapert, S., & Blokland, A. (2012). About the distinction between working memory and short-term memory. *Frontiers in Psychology*, *3*, 301.

Adams, E. J., Nguyen, A. T., & Cowan, N. (2018). Theories of working memory: Differences in definition, degree of modularity, role of attention, and purpose. *Language, Speech, and Hearing Services in Schools*, *49*(3), 340–355.

Allen, R. J., Hill, L. J., Eddy, L. H., & Waterman, A. H. (2019). Exploring the effects of demonstration and enactment in facilitating recall of instructions in working memory. *Memory & Cognition*, 1–11.

Allen, R. J., & Waterman, A. H. (2015). How does enactment affect the ability to follow instructions in working memory? *Memory & Cognition*, *43*(3), 555–561.

Alloway, T. P. (2007). *Automated working: Memory assessment: Manual*. Pearson.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47–89). Elsevier.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active* *maintenance and executive control* (pp. 28–61). New York, NY: Cambridge University Press.

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology and Aging*, *12*(1), 12.

Barrouillet, P., & Camos, V. (2015). *Working memory: Loss and reconstruction*. Psychology Press.

Bier, B., Lecavalier, N. C., Malenfant, D., Peretz, I., & Belleville, S. (2017). Effect of age on attentional control in dual-tasking. *Experimental Aging Research*, *43*(2), 161–177.

Brener, R. (1940). An experimental investigation of memory span. *Journal of Experimental Psychology*, *26*(5), 467.

Charlesworth, L. A., Allen, R. J., Morson, S., Burn, W. K., & Souchay, C. (2014). Working memory and the enactment effect in early alzheimer's disease. *ISRN Neurology*, *2014*.

Contreras-Vidal, J. L., Teulings, H., & Stelmach, G. (1998). Elderly subjects are impaired in spatial coordination in fine motor control. *Acta Psychologica*, *100*(1-2), 25–35.

Cooke, J., Brown, S., & Cunningham, D. (1989). Kinematics of arm movements in elderly humans. *Neurobiology of Aging*, *10*(2), 159–165.

Cowan, N. (1999). An embedded-processes model of working memory. *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*, *20*, 62–101.

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*(4), 1158–1170.

Deary, I. J., Gow, A. J., Taylor, M. D., Corley, J., Brett, C., Wilson, V., … others. (2007). The lothian birth cohort 1936: A study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatrics*, *7*(1), 28.

Decety, J., & Grèzes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Sciences*, *3*(5), 172–178.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-kaplan executive function system(D-kefs): Examiner's manual: Flexibility of thinking, concept formation, problem solving, planning, creativity, impluse control, inhibition*. Pearson.

Desa, U. (2019). World population prospects 2019: Highlights. *New York (US): United Nations Department for Economic and Social Affairs*.

Dienes, Z. (2012). Using bayes to interpret non-significant results. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 34).

Diggles-Buckles, V. (1993). Age-related slowing. In *Sensorimotor impairment in the elderly* (pp. 73–87). Springer.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19–23.

Engle, R. W., Carullo, J. J., & Collins, K. W. (1991). Individual differences in working memory for comprehension and following directions. *The Journal of Educational Research*, *84*(5), 253–262.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309.

Forsberg, A., Fellman, D., Laine, M., Johnson, W., & Logie, R. H. (2020). Strategy mediation in working memory training in younger and older adults. *Quarterly Journal of Experimental Psychology*, 1747021820915107.

Gathercole, S. E., Durling, E., Evans, M., Jeffcock, S., & Stone, S. (2008). Working memory abilities and children's performance in laboratory analogues of classroom activities. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *22*(8), 1019–1037.

Hazra, N. C., & Gulliford, M. (2017). Evolution of the "fourth stage" of epidemiologic transition in people aged 80 years and over: Population-based cohort study using electronic health records. *Population Health Metrics*, *15*(1), 18.

Helman, R., Copeland, C., & VanDerhei, J. (2010). The 2010 retirement confidence survey: Confidence stabilizing, but preparations continue to erode. *EBRI Issue Brief*, (340).

Holtzer, R., Stern, Y., & Rakitin, B. C. (2004). Age-related differences in executive control of working memory. *Memory & Cognition*, *32*(8), 1333–1345.

Houx, P. J., & Jolles, J. (1993). Age-related decline of psychomotor speed: Effects of age, brain health, sex, and education. *Perceptual and Motor Skills*, *76*(1), 195–211.

Jaroslawska, A. J., Gathercole, S. E., Allen, R. J., & Holmes, J. (2016a). Following instructions from working memory: Why does action at encoding and recall help? *Memory & Cognition*, *44*(8), 1183–1191.

Jaroslawska, A. J., Gathercole, S. E., & Holmes, J. (2018). Following instructions in a dual-task paradigm: Evidence for a temporary motor store in working memory. *Quarterly Journal of Experimental Psychology*, *71*(11), 2439–2449.

Jaroslawska, A. J., Gathercole, S. E., Logie, M. R., & Holmes, J. (2016b). Following instructions in a virtual school: Does working memory play a role? *Memory & Cognition*, *44*(4), 580–589.

Jaroslawska, A. J., & Rhodes, S. (2019). Adult age differences in the effects of processing on storage in working memory: A meta-analysis. *Psychology and Aging*, *34*(4), 512–530.

Jarrold, C., & Towse, J. N. (2006). Individual differences in working memory. *Neuroscience*, *139*(1), 39–50.

Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia*, *33*(11), 1419–1432.

Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *NeuroImage*, *14*, S103–S109.

Jeffreys, H. (1961). Theory of probability, 3rd edn oxford: Oxford university press.

Kaplan, C. H., & White, M. A. (1980). Children's direction-following behavior in grades ks. *The Journal of Educational Research*, *74*(1), 43–48.

Kauranen, K., & Vanharanta, H. (1996). Influences of aging, gender, and handedness on motor performance of upper and lower extremities. *Perceptual and Motor Skills*, *82*(2), 515–525.

Kim, E. S., Bayles, K. A., & Beeson, P. M. (2008). Instruction processing in young and older adults: Contributions of memory span. *Aphasiology*, *22*(7-8), 753–762.

Koriat, A., Ben-Zur, H., & Nussbaum, A. (1990). Encoding information for future action: Memory for to-be-performed tasks versus memory for to-be-recalled tasks. *Memory & Cognition*, *18*(6), 568–578.

Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, *6*(3), 299–312.

Li, K. Z., & Lindenberger, U. (2002). Relations between aging sensory/sensorimotor and cognitive functions. *Neuroscience & Biobehavioral Reviews*, *26*(7), 777–783.

Logie, R. H. (2011). The functional organization and capacity limits of working memory. *Current Directions in Psychological Science*, *20*(4), 240–245.

Logie, R. H., Camos, V., & Cowan, N. (n.d.). *Working memory: State of the science*. Oxford University Press.

Logie, R. H., Della Sala, S., MacPherson, S. E., & Cooper, J. (2007). Dual task demands on encoding and retrieval processes: Evidence from healthy adult ageing. *Cortex*, *43*(1), 159–169.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324.

Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of bayes factors for common designs*. Retrieved from https://CRAN.R-project.org/package=BayesFactor

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The montreal cognitive assessment, moca: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699.

Oberauer, K., & Eichenberger, S. (2013). Visual working memory declines when more features must be remembered for each object. *Memory & Cognition*, *41*(8), 1212–1227.

Park, D. C., Lodi-Smith, J., Drew, L., Haber, S., Hebrank, A., Bischof, G. N., & Aamodt, W. (2014). The impact of sustained engagement on cognitive function in older adults: The synapse project. *Psychological Science*, *25*(1), 103–112.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–164.

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rhodes, S., Doherty, J. M., Jaroslawska, A. J., Forsberg, A., Belletier, C., Naveh-Benjamin, M.,… Logie, R. H. (in press). Exploring the influence of temporal factors on age differences in working memory dual task costs. *Psychology & Aging*.

Rhodes, S., Jaroslawska, A. J., Doherty, J. M., Belletier, C., Naveh-Benjamin, M., Cowan, N., … Logie, R. H. (2019). Storage and processing in working memory: Assessing dual task performance and task prioritization across the adult lifespan. *Journal of Experimental Psychology: General*, *148*(7), 1204–1227.

Rhodes, S., Parra, M. A., Cowan, N., & Logie, R. H. (2017). Healthy aging and visual working memory: The effect of mixing feature and conjunction changes. *Psychology and Aging*, *32*(4), 354.

Rhodes, S., Parra, M. A., & Logie, R. H. (2016). Ageing and feature binding in visual working memory: The role of presentation time. *The Quarterly Journal of Experimental Psychology*, *69*(4), 654–668.

Rossetti, H. C., Lacritz, L. H., Cullum, C. M., & Weiner, M. F. (2011). Normative data for the montreal cognitive assessment (moca) in a population-based sample. *Neurology*, *77*(13), 1272–1275.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default bayes factors for anova designs. *Journal of Mathematical Psychology*, *56*(5), 356–374.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237.

Schott, N. (2012). Age-related differences in motor imagery: Working memory as a mediator. *Experimental Aging Research*, *38*(5), 559–583.

Schuster, C., Hilfiker, R., Amft, O., Scheidhauer, A., Andrews, B., Butler, J., … Ettlin, T. (2011). Best practice for motor imagery: A systematic literature review on motor imagery training elements in five different disciplines. *BMC Medicine*, *9*(1), 75.

Seidler, R. D., Alberts, J. L., & Stelmach, G. E. (2002). Changes in multi-joint performance with age. *Motor Control*, *6*(1), 19–31.

Shimoyama, I., Ninchoji, T., & Uemura, K. (1990). The finger-tapping test: A quantitative analysis. *Archives of Neurology*, *47*(6), 681–684.

Smyth, M. M., & Pendleton, L. R. (1989). Working memory for movements. *The Quarterly Journal of Experimental Psychology*, *41*(2), 235–250.

Smyth, M. M., & Pendleton, L. R. (1990). Space and movement in working memory. *The Quarterly Journal of Experimental Psychology Section A*, *42*(2), 291–304.

Thompson, J. J., Blair, M. R., & Henrey, A. J. (2014). Over the hill at 24: Persistent age-related cognitive-motor decline in reaction times in an ecologically valid video game task begins in early adulthood. *PloS One*, *9*(4), e94215.

Verhaeghen, P., & Cerella, J. (2002). Aging, executive control, and attention: A review of meta-analyses. *Neuroscience & Biobehavioral Reviews*, *26*(7), 849–857.

Verhaeghen, P., Steitz, D. W., Sliwinski, M. J., & Cerella, J. (2003). Aging and dual-task performance: A meta-analysis. *Psychology and Aging*, *18*(3), 443.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems ofp values. *Psychonomic Bulletin & Review*, *14*(5), 779–804.

Waldron-Perrine, B., & Axelrod, B. N. (2012). Determining an appropriate cutting score for indication of impairment on the montreal cognitive assessment. *International Journal of Geriatric Psychiatry*, *27*(11), 1189–1194.

Waterman, A. H., Atkinson, A. L., Aslam, S. S., Holmes, J., Jaroslawska, A. J., & Allen, R. J. (2017). Do actions speak louder than words? Examining children's ability to follow instructions. *Memory & Cognition*, *45*(6), 877–890.

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspectives on Psychological Science*, *6*(3), 291–298.

Wetzels, R., & Wagenmakers, E.-J. (2012). A default bayesian hypothesis test for correlations and partial correlations. *Psychonomic Bulletin & Review*, *19*(6), 1057–1064.

Wingfield, A., Tun, P. A., & McCoy, S. L. (2005). Hearing loss in older adulthood: What it is and how it interacts with cognitive performance. *Current Directions in Psychological Science*, *14*(3), 144–148.

Wojcik, D., Allen, R., Brown, C., & Souchay, C. (2011). Memory for actions in autism spectrum disorder. *Memory*, *19*(6), 549–558.

Wollesen, B., & Voelcker-Rehage, C. (2014). Training effects on motor–cognitive dual-task performance in older adults. *European Review of Aging and Physical Activity*, *11*(1), 5–24.

Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, *1*(6), 209–216.

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, *3*(11), 1212–1217.

Yang, T., Gathercole, S. E., & Allen, R. J. (2014). Benefit of enactment over oral repetition of verbal instruction does not require additional working memory during encoding. *Psychonomic Bulletin & Review*, *21*(1), 186–192.

Yang, T.-x., Allen, R. J., & Gathercole, S. E. (2016). Examining the role of working memory resources in following spoken instructions. *Journal of Cognitive Psychology*, *28*(2), 186–198.

Yang, T.-x., Allen, R. J., Holmes, J., & Chan, R. C. (2017). Impaired memory for instructions in children with attention-deficit hyperactivity disorder is improved by action at presentation and recall. *Frontiers in Psychology*, *8*, 39.

Yogev-Seligmann, G., Rotem-Galili, Y., Mirelman, A., Dickstein, R., Giladi, N., & Hausdorff, J. M. (2010). How does explicit prioritization alter walking during dual-task performance? Effects of age and sex on gait speed and variability. *Physical Therapy*, *90*(2), 177–186.