

Planning and task management in older adults: Cooking breakfast

FERGUS I. M. CRAIK

Rotman Research Institute, Toronto, Ontario, Canada

and

ELLEN BIALYSTOK

York University, Toronto, Ontario, Canada

The article describes a simulated “cooking breakfast” task in which participants must remember to start and stop cooking five foods so that all the foods are “ready” at the same time. In between starting and stopping operations, the participants also carried out a “table-setting” task as a filler activity. The breakfast task yields various measures of multitasking and executive control. Groups of younger and older adults performed the task; half of the participants in each group were bilinguals and the other half were monolinguals. The results showed substantial age-related decrements in most measures of executive control. Additionally, older bilinguals showed some advantages in task management over their monolingual peers.

The ability to formulate an efficient plan is a crucial aspect of adaptive living. It is also necessary to retrieve the plan at appropriate times, and keep the overall goal in mind while carrying out actions required by the plan’s subgoals. The complex cognitive activities comprising planning are thought to be mediated largely by the frontal lobes (Luria, 1966; Stuss & Benson, 1986), and this assumption is strongly supported by clinical studies showing impairments of planning in patients with frontal lobe lesions (Shallice, 1982; Shallice & Burgess, 1991). There is general agreement that planning is not a unitary construct (Morris & Ward, 2005; Stuss & Alexander, 2000). Instead, the ability to formulate and execute an efficient plan appears to depend on a variety of cognitive functions whose exact contributions vary substantially, depending on the subject matter of the planned activity and on the planner’s experience with similar activities. In this vein, Burgess and colleagues recently proposed that “planning is an activity that can be supported by very many different cognitive processes—perhaps hundreds—and that different subsets of these are tapped by different planning tasks” (Burgess, Simons, Coates, & Channon, 2005). As

a consequence, different tests of planning may not correlate highly with each other, and the way to understand performance is to analyze a task in terms of the cognitive components necessary for its successful completion and to assess task participants for their strength or weakness on these cognitive components.

Good planning clearly involves many high-level cognitive skills, including the ability to look ahead and mentally envisage the relevant situation and the ability to anticipate potential problems and think of ways to overcome them. When the plan is implemented, it is necessary to keep the big picture in mind, carry out actions at appropriate times, evaluate progress and adjust the plan if necessary, and avoid continuing with one set of actions when it is more efficient to switch to another set. In terms of the cognitive constructs studied in laboratory settings, good planning requires abstract reasoning and problem solving: especially, perhaps, analogical reasoning (in order to bring relevant expertise to bear on the current situation), working memory, prospective memory, and the ability to switch sets flexibly. Planning can be assessed in a variety of ways, many of which are discussed in a recent collection on the topic (Morris & Ward, 2005).

An important decision in constructing such assessments is how much the test situation should resemble real-life planning activities. At one extreme, researchers have studied logical problems such as the Tower of Hanoi or its occidental counterparts (see, e.g., Davies, 2005; Owen, 2005), but this approach may generalize only weakly to real-life tasks. Alternatively, one can study planning “in the field,” an approach that has been used successfully in the shopping plan test (Burgess et al., 2005); the problem here is in isolating the specific cognitive components that may be impaired. An intermediate position has been taken

This research was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to the first author, a grant from NSERC to the second author, and a grant from the Canadian Institutes of Health Research to both authors. We are grateful to Jane Logan for her contributions to this study and her work in collecting the data. Correspondence concerning this article should be addressed to F. I. M. Craik, Rotman Research Institute, 3560 Bathurst Street, Toronto, ON, M6A 2E1 Canada (e-mail: fcraik@rotman-baycrest.on.ca).

Note—This article was accepted by the previous editorial team, when Colin M. MacLeod was Editor.

by others who have constructed board games or virtual reality situations to mimic real-life settings (see, e.g., Morris, Kotitsa, & Bramham, 2005; Rendell & Craik, 2000). This approach has the benefits of allowing substantial experimental control while also allowing participants to bring their personal experience to bear on task problems and solutions. The present article describes such a task.

The main purpose of the present study was to assess the effects of normal aging on complex planning performance. In order to keep the task manageable, quantifiable, and controlled, we used a laboratory-based task, but one in which the context and goals were very familiar to participants of all ages—namely, cooking breakfast. Several writers have mentioned meal preparation as being a good example of a real-life planning task (see, e.g., Shallice, 1982; Ward, 2005), in that various subgoals have to be monitored and accomplished, while also bearing in mind the overall goal of having the various foods cooked and ready to eat at some specified time. In her master's thesis, Sheila Kerr (1991) developed a "cooking breakfast" task programmed on a Macintosh computer. The participant's task was to start "cooking" five breakfast foods by clicking on relevant icons, to monitor progress of the cooking, and to stop each food at its specified time in such a way that the five foods were "ready" at the same time. In Kerr's task, the five foods were bacon, oatmeal, coffee, eggs, and toast, and their cooking times varied from 2.5 min (toast) to 10 min (bacon). Between starting and stopping cooking each food, the participant also carried out a table-setting task as a background activity. Table setting was also simulated on the computer; the participants dragged knives, forks, plates, glasses, and so on to their correct places by means of the computer mouse. To start and stop cooking each food, the participant clicked on the food's icon on the home screen. This brought up a screen for that food, containing START and STOP buttons and also a time indicator, in the form of a filled vertical bar whose height dropped in real time, once it was started, from the total cooking time for that food to zero. Table setting was also carried out on a separate screen. The task was to set as many places as possible, but also to remember to check the progress of each food, and to start and stop cooking at appropriate times.

The major finding from Kerr's experiment was that older adults "burned their breakfasts." Although all of the participants were late in stopping the foods, the inaccuracies were far greater for the older adults. For the young adults, the average discrepancy was about 30 sec, but the participants in the group of older adults (mean age = 72 years) were 87 sec late, on average. The participants were allowed to check the cooking progress of each food as often as they wanted to, and the results showed that the younger participants checked more often than the older adults. Finally, older adults started cooking each food later than their younger counterparts, and this difference increased as the task progressed. Kerr concluded that the profiles of performance on the breakfast task confirmed laboratory findings of age-related drops in the ability to carry out prospective memory and working memory

tasks, and age-related drops in the efficiency of planning and related executive functions.

The present experimental task was based on Kerr's (1991) work but was reprogrammed for the PC and extended in various ways, to be described. The essence of the task remained the same: The participants had to start and stop cooking five foods so that they would be ready at the same time, while also completing as many table settings as possible. The measures recorded included the difference between required and actual cooking times for each food, the range of stop times for the foods, the number of times each food was checked while cooking, the discrepancy between the actual and ideal time each food was started, and the proportion of task time occupied by carrying out the filler task—namely, setting the table.

Our assumption was that the whole task reflected the ability to formulate and execute a plan: to maintain the overall goal in mind (preparing the meal so that all foods are properly cooked and ready at the same time) while also monitoring and executing the subgoals (setting the table, remembering to start foods cooking at appropriate times, checking the progress of foods that were cooking, and stopping each food at its specified cooking time). These overall goals depend on specific cognitive processes: *working memory*—to hold both the general plan and the progress of each food in mind; *prospective memory*—to remember to start and stop foods appropriately; and *perseveration*—to suspend table setting when some other activity was required. In the Method section, we propose specific measures for each of these cognitive functions, but it should be noted that the choice of a real-life task means that the specificity of measurement of individual cognitive functions is limited, although the applicability of past experience is enhanced. Moreover, just as planning is not a unitary construct, neither should working memory and prospective memory be regarded as unitary constructs. Rather, these terms connote related sets of cognitive operations whose form and properties will vary substantially, depending on materials, participants, degree of contextual support, and other variables.

On the basis of Kerr's (1991) results and general findings that prospective memory, working memory, and executive functions typically become less efficient with age, we expected that older adults would perform the task less well than their younger counterparts. We also expected that the different task measures would further illuminate the strengths and weaknesses associated with cognitive aging by showing different degrees of decline for the various processes being measured. Previous work has found substantial age-related decrements in such processes as inhibition, working memory, prospective memory, and set switching (Craik & Jennings, 1992; Mayr, 2001; Meiran & Gotler, 2001; Zacks, Hasher, & Li, 2000). Also, many researchers have pointed out the similarities between behavioral effects found in normal aging and in mild frontal dysfunction (Craik, Morris, Morris, & Loewen, 1990; Moscovitch & Winocur, 1992; West, 1996), and this parallel is strengthened by reports that the frontal lobes are among the parts of the brain showing the greatest age-

related losses, both in volume (Raz, 2000) and in blood flow (Gur, Gur, Obrist, Skolnick, & Reivich, 1987).

Given these demonstrated age-related deficits in the components of planning, older adults should be poor planners. Yet interestingly, Garden, Phillips, and MacPherson (2001) found no midlife decline on two open-ended planning tasks, despite finding age-related decrements on laboratory tests of executive functioning. Similarly, Martin and Ewert (1997) concluded that age does not necessarily predict performance in complex cognitive tasks such as planning. The key to this apparent mismatch between poor performance on cognitive components and sustained performance on global planning tasks may lie in the highly practiced experience of older adults on real-life planning tasks resembling those given in experimental studies. That is, successful experience in many similar situations may compensate for declining performance in such fundamental abilities as working memory, inhibitory control, and set switching. In this regard, Ormerod (2005) has stressed the role played by retrieval of stored knowledge from long-term memory in successful planning and execution of complex tasks. On the other hand, Kliegel, McDaniel, and Einstein (2000) investigated the performance of younger and older adults on an open-ended laboratory planning task (a modified version of the six-element task devised by Shallice & Burgess, 1991) and did find age-related decrements in planning, initiation, and execution of the set of tasks. The existing literature on age-related differences in cognitive planning is critically examined in a useful chapter by Phillips, MacLeod, and Kliegel (2005). The present study provides further evidence on this question.

The study also had a second purpose—to explore the effects of bilingualism on planning behavior. In a series of studies of bilingual children, Bialystok (2001) found that bilingualism conferred advantages in tasks requiring inhibitory control. The suggested explanation was that bilingual children constantly switch from one language to another, and that the inhibitory “dampening” of one language system while using the other is therefore highly practiced. The benefits of such practice then generalize to other situations involving cognitive control. These ideas were extended to the study of aging by Bialystok, Craik, Klein, and Viswanathan (2004) using the Simon task, in which responses to visual stimuli were either congruent or incongruent with their displayed position. The results showed that the “Simon effect” (longer RTs to incongruent stimuli) was somewhat smaller in middle-aged bilinguals

than in monolinguals of the same age, showing that the bilinguals were more able to override the prepotent same-side response in the incongruent condition. This small bilingual advantage in middle-aged adults was greatly amplified in a group of older adults (60–80 years), leading to the conclusion that bilingualism protects older adults against some aspects of age-related cognitive decline. We therefore wanted to see if this age-related bilingual advantage also applied to richer and more complex situations requiring cognitive control. Accordingly, half of the participants in the present experiment were fluent bilingual speakers and the other half were monolinguals.

To summarize, we describe a novel “cooking breakfast” task that assesses planning behavior and its cognitive components. Given the well-documented drops in performance shown by older adults on working memory, prospective memory, and set-switching tasks, we expected to find age-related deficits on the breakfast task. Additionally, if bilingualism confers an advantage on aspects of executive processing, and this advantage increases with age, we expected to see superior performance by the bilinguals on some measures—at least in the older participants.

METHOD

Participants

Sixty participants were tested, of whom half were younger adults between 18 and 30 years old (mean age = 20.2 years) and half were older adults between 60 and 80 years old (mean age = 69.6 years). In each age group, half of the participants were monolingual and half were bilingual, as defined in the next section. The four groups were matched approximately on their years of education, as shown in Table 1.

Instruments and Tasks

Background and language questionnaires. A general background questionnaire was used to establish the age, education level, and health status of each participant. The health survey inquired about major and minor health problems as well as medications that each participant was taking at the time of the study.

The language background questionnaire contained questions about the number of languages spoken, where each of those languages was learned, the means by which they were learned (e.g., in the home, at school, etc.), how often each language had been used throughout life, and how often each language was used at present. The questionnaire also included a self-rating scale to determine proficiency in understanding, speaking, reading, and writing in each language. Bilingual participants were those who demonstrated that they had spoken two languages every day from an early age (young adults, 6 years or younger; older adults, 10 years or younger) and had continued to use two languages up to the present, with excellent

Table 1
Mean Scores and Standard Deviations of
Background Measures by Age and Language Group

Group	Age (Years)		Education (Years)		Mill Hill		Alpha Span		Digit Span	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Younger monolinguals	20.9	2.8	13.8	1.9	19.5	3.1	34.8	7.6	66.3	23.3
Younger bilinguals	19.5	1.2	13.3	1.1	16.4	3.6	36.5	8.5	74.1	22.9
Older monolinguals	70.3	4.3	16.1	3.5	24.1	3.4	33.5	12.7	71.8	27.5
Older bilinguals	68.8	6.1	15.3	3.7	17.9	6.1	21.4	10.7	52.5	23.0

competence in both. The languages spoken by the younger bilingual participants, in addition to English, were Szechuan, Hindi, Spanish, Slovak, Dari, Gujarati, Cantonese, Vietnamese, Bengali, Russian, and Polish. The languages spoken by the older bilingual participants, in addition to English, were French, Gujarati, Urdu, Hindi, German, Punjabi, Czech, and Bengali.

Digit span and alpha span. The digit span task is a measure of short-term verbal memory. The experimenter read a list of single-digit numbers (for example, 5–4–8), at the rate of one digit per sec, and the participant recited back the digits in the same order. The task began with strings of three digits and included two trials at each digit length until the participant failed to recall at least one list of a particular length. Credit was given for partially correct recall responses; 1 point was awarded for each digit in a correctly ordered consecutive pair, with the restriction that each digit could only be counted for 1 point altogether. For example, if the list was 8–7–4–3–9–1–5 and recall was 8–7–4–9–3–5, the participant would receive 1 point each for 8, 7, and 4 only, because neither 9, 3, nor 5 occurs in a correctly ordered pair. The maximum score for each string length was the number of digits in that string. In the alpha span test, the experimenter read a list of words in a fixed random order (for example, “chair–zebra–apple”) and the participant attempted to recite the list back in alphabetical order (“apple–chair–zebra”). The task was administered and scored in the same way as the digit span test. The alpha span task (Craik, 1986) is taken to be a measure of working memory, on the grounds that items must be stored briefly and then reordered mentally before being given as responses.

Mill Hill Vocabulary Scale. The Mill Hill Vocabulary Scale (Raven, 1960) is a multiple choice test of receptive vocabulary. The test contains 34 words in total, but the first item is for illustration, leaving 33 test items. Each word is presented with six response choices, and the participants choose the response that most

closely resembles the target word in meaning. The participants are encouraged to take as long as they need to answer the questions, and to guess if they are unsure of the correct response. One point is awarded for each correct choice, yielding a maximum score of 33.

The breakfast task. The breakfast task was conducted on a PC using a 34 cm × 27 cm touch screen monitor to facilitate responding, especially for the older participants. The task was to prepare a virtual breakfast consisting of five foods and to set a table while the foods were cooking. Each food had a different cooking time, ranging from 1 min to 4.5 min. The participant began cooking each food by touching a START button located above the name of each food, and stopped cooking by touching the corresponding STOP button. Once the food started cooking, the name of the food was highlighted in green, and when it was stopped, the name turned red. An analog timer counted down the number of minutes until the food was finished cooking in real time. Each timer consisted of a vertical column beneath its food icon; time was registered by the height of a filled portion whose level dropped continuously throughout cooking from the start time to zero (Figure 1). If the participant overcooked the food, the timer continued to measure time until the STOP button was touched, but there was no visible indication of the amount of overcooking. The participants were required to cook each food for its specified time and to have all the foods ready at the same time. There was no signal from the timer when the food was ready, so the participant was responsible for monitoring the time in order to stop each food at the appropriate time and for calculating when to start each food to maximize the foods’ simultaneous completion.

To increase the demands on planning and prospective memory, the cooking was conducted in conjunction with a distractor task—namely, setting the table. Forks, knives, spoons, and plates were arranged at the bottom of the screen and had to be placed in the correct positions in four place settings around a square table (see

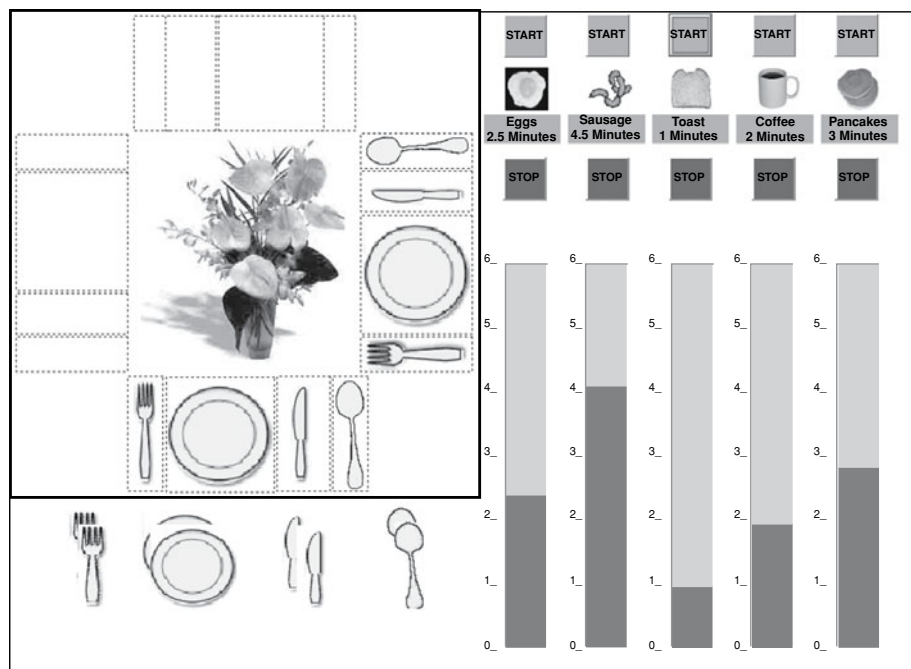


Figure 1. A sample screen for the 1-screen condition of the breakfast task, showing foods in the process of being cooked and the table partly set. Neither the toast nor the coffee has been started, so their cooking indicators remain at 1 and 2 min, respectively. The eggs and pancakes have each been cooking for a few seconds, and the sausages have been cooking for almost 30 sec.

Figure 1). The places were set by touching an item and then touching its destination on the table. When all four place settings were complete, the items returned to their original places at the bottom of the screen and the participant continued with four new place settings. The participants were told to set as many places as they could and that part of their success would be determined by the number of places set. Nonetheless, it was made clear that successful cooking of the meal was the primary task, and that the main criteria of success were to avoid burning the food and to have the foods ready at the same time.

The task was presented in three conditions of varying complexity; we varied the complexity by changing the number of screens that needed to be active to cook the foods. The first condition, called the 1-screen condition, showed the five foods and the table on the same screen. An illustration of this condition is shown in Figure 1. The participants needed only to direct their attention to different parts of the display to manage the progress of each of the foods and the table setting.

In the second, 2-screen condition, the table appeared on one screen and the five foods and their time bars appeared on a different screen. In this condition, therefore, the participants needed to switch between the two screens when they were performing either the cooking or the table setting, and needed to keep in mind the state of progress in the other. The main screen with the table setting contained a set of icons down the left side representing each of the foods, with a symbol that indicated its state of cooking. In order to start, stop, and check the foods, the participant switched from the main screen to the food screen by touching the icon for any one of the foods. Similarly, the food screen contained a button called RETURN TO TABLE in the top left corner.

In the third and most complex condition, there were 6 screens. The main table screen was similar to that in the 2-screen condition, but touching any of the food icons down the left side produced a screen specific to that food. Therefore, in addition to the main table screen, there were 5 separate food screens, 1 for each food. To switch from one food to another, the participant first returned to the table screen and then touched the icon for the other food.

Procedures

Testing began with an explanation of the breakfast task and an initial practice condition using two foods and a single screen. This was followed by the general background questionnaire. The participants then returned to the breakfast task to repeat the practice with two foods in the 1-screen condition, followed by the 1-screen experimental condition. Following this, the participants were given the language background questionnaire and the alpha span task. These activities were followed by practice for the 2-screen condition, and the participant had the opportunity to try the practice alone and ask questions before proceeding to the experimental condition. In the next break, the experimenter administered the digit span task and the Mill Hill Vocabulary Scale. The practice for the 6-screen breakfast task was then introduced, and again the participants took as long as they needed to become familiar with the task and to ask questions. The 6-screen experimental condition was run, and the experimental session ended. The total time for the participants to complete all these tasks was approximately 1 h for younger adults and 1.5 h for older adults.

In the pilot work, we discovered that older adults had a great deal of trouble understanding the task if they were introduced to the 6-screen version first. For that reason, we deliberately tested all the participants in the same order: 1-screen, 2-screen, and then 6-screen. This aspect of the procedure had the effect of confounding order and complexity, but in our judgment it was preferable to make sure that all of the participants understood the task fully and were familiar with the procedures rather than counterbalancing the order and running the risk of obtaining spurious results as a function of the participants' imperfect grasp of methods and goals.

RESULTS AND DISCUSSION

Background Measures

The mean ages and the mean scores for each of the alpha span, digit span, and Mill Hill vocabulary tests by age and language group are presented in Table 1. For each measure, a two-way ANOVA for age group (younger, older) and language group (monolingual, bilingual) compared scores across groups. With regard to education, the older adults had received more years of formal schooling than the younger adults [$F(1,56) = 3.2$, $MS_e = 7.8$, $p < .03$], with no difference between the language groups ($F < 1$). The analysis of the Mill Hill vocabulary scores indicated higher performance by the older adults [$F(1,56) = 7.7$, $MS_e = 17.9$, $p < .007$] and monolinguals [$F(1,56) = 17.9$, $MS_e = 17.9$, $p < .0001$], with no interaction.

In the alpha span task, there were main effects of age group [$F(1,56) = 10.0$, $MS_e = 101.2$, $p < .002$] and language [$F(1,56) = 4.0$, $MS_e = 101.2$, $p < .05$], as well as an interaction between them [$F(1,56) = 7.1$, $MS_e = 101.2$, $p < .01$]. A Scheffé analysis of contrasts ($p < .05$) indicated that the only significant difference was that the older bilinguals scored lower than the other three groups. There were no main effects of age or language in the digit span scores, but the interaction of age and language was significant [$F(1,56) = 4.65$, $MS_e = 588$, $p < .05$]. Table 1 shows that the interaction again reflects the lower score obtained by the older bilinguals, although a Scheffé analysis revealed no reliable differences among the four means. The poorer performance of the older bilinguals on the span tasks, especially the alpha span, is likely a reflection of their weaker vocabulary knowledge. For the whole sample, there was a correlation between Mill Hill vocabulary scores and alpha span [$r(58) = .27$, $p < .03$], which was stronger for the older participants [$r(28) = .52$, $p < .003$], and between alpha span and digit span scores [$r(58) = .47$, $p < .0002$; for the older adults, $r(28) = .47$, $p < .008$]. Therefore, it is reasonable to suggest that the lower span scores for the older bilinguals are attributable to their weaker vocabulary. It should also be acknowledged that the absence of significant differences between older monolinguals and bilinguals in the Mill Hill and digit span tests is likely attributable to low power, given that $n = 15$ per subgroup.

Cooking Measures

As the participant performed the task, the program recorded all the keypresses (i.e., screen touches) in a detailed timeline. These sequences were used to calculate several measures. The first variable was the disparity between the actual time the food cooked and the time required to cook that food. This disparity is referred to as the *discrepancy* for each food, calculated as the absolute value of the difference in time between the actual and required cooking times (either undercooking or overcooking). The formula was

$$\text{Discrepancy} = |\text{actual time} - \text{ideal time}|.$$

This variable represents the ability to remember to stop the food at the right time, and so relies heavily on prospec-

tive memory. If the discrepancy score is low (i.e., close to zero), the participant was able to remember that the food needed attention and to disengage from the table-setting distractor task to stop the food.

The mean discrepancy scores for each food are shown in Figure 2. A three-way ANOVA for age group, language group, and condition indicated that the discrepancies were greater for the older adults than the younger ones [$F(1,56) = 18.6, MS_e = 196.8, p < .0001$], with no differences between language groups ($F < 1$) and no interaction. There was a main effect of condition [$F(2,112) = 38.4, MS_e = 84.2, p < .0001$], in which the discrepancies for all conditions were significantly different from each other. However, there was also an interaction of condition and age group [$F(2,112) = 12.5, MS_e = 84.2, p < .0001$]. There were two reasons for this interaction. The first was that the difference in discrepancy scores between age groups was not significant for the 1-screen condition ($F = 1.8, n.s.$). The second was that there was a different pattern of effects for each of the age groups: For the younger adults, only the contrast between the 1-screen and 6-screen conditions was significant [$F(1,28) = 27.9, p < .001$], but for the older adults, all three conditions were reliably different from each other (all $ps < .01$). Although the discrepancy scores indicate deviations from the specified cooking time, in almost all cases the deviations indicate overcooking; very few of the participants stopped the clock before the specified time. It is also noteworthy that there were no age differences in the 1-screen condition, in which no switching across screens was required.

We interpret the age \times condition interaction as showing an age-related decline in prospective memory ability. There are no prospective memory demands in the 1-screen

condition, because all of the foods and their time indicators are always present on the screen, and there are no age differences in this condition (Figure 2). The 2-screen condition requires the participant to remember to switch from the table-setting screen to the food screen, and an age-related decrement in performance was found in this case. The prospective memory load is even greater in the 6-screen condition, and age differences were largest in that condition (Figure 2). The conclusion that these results reflect an age-related decline in prospective memory is consistent with recent literature (Einstein, McDaniel, Manzi, Cochran, & Baker, 2000; Henry, MacLeod, Phillips, & Crawford, 2004) but is based on a task that is more realistic and may provide more “environmental support” than most laboratory tasks.

The second measure was the *range of stop times*—the difference in seconds between stopping the first and the last foods. Ideally, this number should approach zero, because the participants were told that one of the main goals was to have all the foods finish cooking at the same time. The range was calculated as the difference in time between these two events; that is,

$$\text{Range} = \text{time last food stopped} - \text{time first food stopped.}$$

This measure reflects planning ability and, to some extent, working memory. The lower the range, the more the participant was able to anticipate the outcome (by accounting for the different cooking times for each food) and ensure that the foods would finish at the same time. It also indicates the extent to which the participant understood the order inherent in the different cooking times and followed that order throughout the task.

The mean scores indicating the range of stop times produced by each group for each screen condition are shown

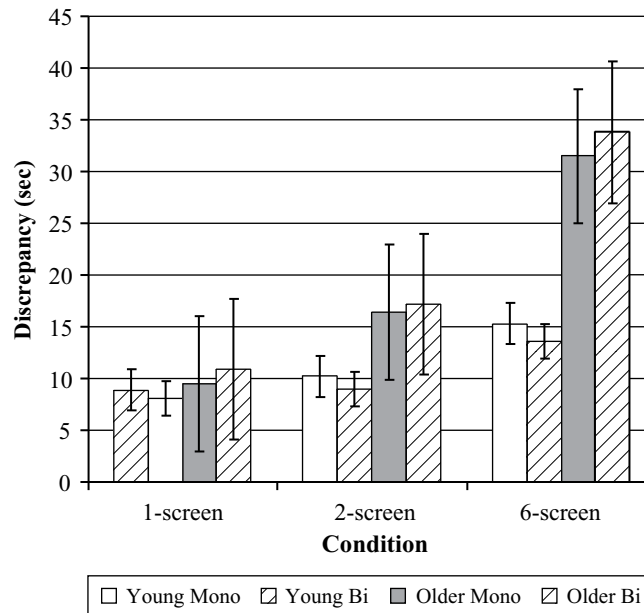


Figure 2. Mean discrepancy scores (and standard errors) between specified and actual cooking times as a function of age, language, and condition.

in Figure 3. A three-way ANOVA for age, language, and condition showed that the older adults produced a larger range than the younger adults [$F(1,56) = 10.4$, $MS_e = 1,306$, $p < .002$], with no difference between language groups ($F < 1$) and no interaction. As before, there were differences among the three conditions [$F(2,112) = 12.3$, $MS_e = 541.4$, $p < .0001$] and an interaction of age group and condition [$F(2,112) = 5.2$, $MS_e = 541.4$, $p < .007$]. Unlike the discrepancy scores, the difference between the age groups was significant only for the 6-screen condition. In fact, all three conditions were equivalent for the younger adults [$F(2,56) = 1.8$, n.s.], but the older adults performed more poorly on the 6-screen condition than on the other two [$F(2,56) = 10.6$, $MS_e = 858$, $p < .0001$].

Range scores indicate the ability to prepare a meal in which all of the foods are ready at the same time—possibly the central criterion of good meal preparation in real life. It is thus a measure of global planning and the ability to keep this plan in mind while performing the more local tasks of starting and stopping specific foods and setting places at the table. Range scores thus measure overall planning ability and a form of working memory involving retention of the overall plan in mind throughout the task. Older adults showed deficits on the measure, and the age-related decline in performance was amplified by task complexity. Language yielded no statistically significant effects, but it is interesting to note that whereas the older monolinguals' range scores increased regularly from the 1-screen to the 6-screen condition, the older bilingual group maintained relatively low range scores in the first two conditions and appeared to suffer a collapse of performance only on the 6-screen version. This pattern is seen again in the deviation from ideal start time measure de-

scribed next. The nonsignificant difference between older monolinguals and bilinguals in the 2-screen condition is again likely attributable to the small number of participants in each subgroup.

The third measure was the *average deviation of start times*. In order to achieve a low range of stop times score, the participants had to calculate ideal start times for each food. These ideal times changed with each new food. For example, the second food should have been started 90 sec after the first food was started, but if it was not started until 110 sec had elapsed, there would be no perfect time to start the third food. The third food should have been started 120 sec after the first food (for them to finish together), but 30 sec after the second food (for *them* to finish together). The best that could be done in this case is the average between 120 sec (for Food 1) and $110 + 30 = 140$ sec (for Food 2); that is, Food 3 should ideally have been started 130 sec after Food 1. The ideal start times for Foods 4 and 5 can be calculated following the same general principle—by determining the best start time, given the actual start times of all previous foods. The measure calculated for each participant was the average deviation between the ideal and actual start times for Foods 2–5.

This measure provided a further indicator of planning ability—in this case, weighted also with a strong prospective memory component. The lower the average deviation, the more successful the participant was in starting the five foods within an acceptable range so that they all finished cooking at close to the same time. If the average deviation over the five foods was high, then the participant either failed to start the series of foods in their correct order or failed to remember to start one or more foods at the appropriate time. The participants needed to hold the various

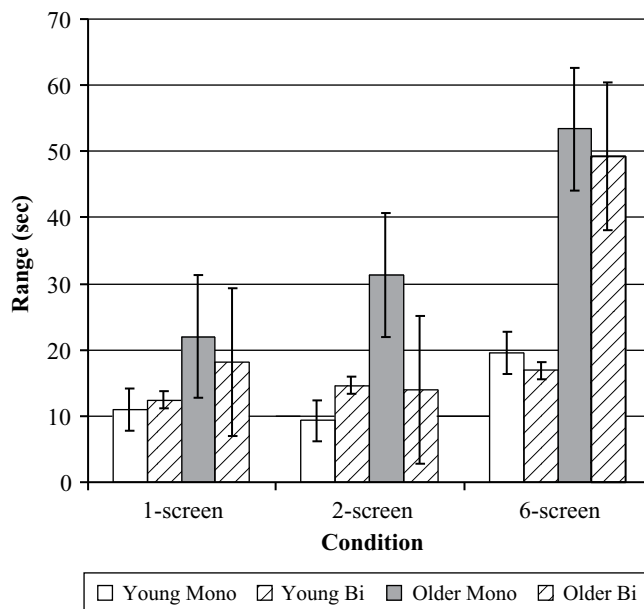


Figure 3. Mean range of stopping times between first and last food stopped (and standard errors) as a function of age, language, and condition.

cooking times in mind and estimate the best time to start each new food.

The average start time deviations are shown in Figure 4. There were main effects of age group [$F(1,56) = 13.5, MS_e = 167.3, p < .0006$] and condition [$F(2,110) = 23.8, MS_e = 57.9, p < .0001$], indicating longer times for older adults and more complex conditions, but also an interaction of age and condition [$F(2,110) = 6.2, MS_e = 57.9, p < .003$]. The interaction occurs because the difference between age groups is not quite significant for the simple 1-screen condition [$F(1,55) = 3.3, p = .08$], but it is significant and progressively greater in the 2-screen and 6-screen conditions. Further, the differences between all of the conditions are significant for the older adults (all $ps < .01$), but the difference between the 2-screen and 6-screen conditions is only marginal for the younger adults [$F(1,28) = 3.4, p = .07$]. The increased difficulty of the 2-screen condition was not further elevated in the 6-screen version for the younger adults. There were no reliable effects or interactions involving language but, as shown in Figure 3, the older bilinguals have smaller deviations from the ideal start times than the older monolinguals in the 1-screen and 2-screen conditions, yet again appear to be overwhelmed by the complexity of the 6-screen condition.

Starting each food at its optimal time is perhaps the most difficult component of the breakfast task. It is straightforward in the 1-screen condition, in which all indicators are visible (Figure 1), because each successive food should be started as the cooking indicators of the previous foods reach the specified cooking time for that food. In the 2-screen condition, the indicators are again visible, provided the participant checks the cooking screen at the appropriate times. In the 6-screen condition, how-

ever, the participants must remember the progress of all of the previously started foods when deciding that the next food should start cooking. Thus, the start time measure reflects overall planning with both prospective memory and working memory components that increase their involvement in performance as the task conditions increase in complexity.

Fourth, the number of times that the participants checked the progress of the foods when they switched screens was calculated for each of the 2- and 6-screen conditions. Because all of the information was on the same screen for the 1-screen condition, it was not possible to assess the degree to which the participants monitored the two tasks. In the other two conditions, each instance in which the participant touched a food icon (to switch to either the whole set of foods in the 2-screen condition or any one of the foods in the 6-screen condition) was added to a cumulative total. The number of checks indicates the prospective memory and monitoring ability of the participants.

The average numbers of food checks carried out by each group for the 2- and 6-screen conditions are shown in Figure 5. The younger adults checked cooking progress more frequently than did the older adults [$F(1,56) = 11.2, MS_e = 27.7, p < .002$], with no effect of language group and no interaction (both F s < 1). Both age groups checked more frequently for the 6-screen condition [$F(1,56) = 112.4, MS_e = 8.5, p < .0001$], but there was also an interaction of condition and age group [$F(1,56) = 5.5, MS_e = 8.5, p < .02$]. Although the younger adults checked more often than the older adults in both conditions, the difference between the two age groups was greater in the 6-screen condition [$F(1,56) = 10.7, p < .001$] than in the 2-screen condition [$F(1,56) = 6.9, p < .01$].

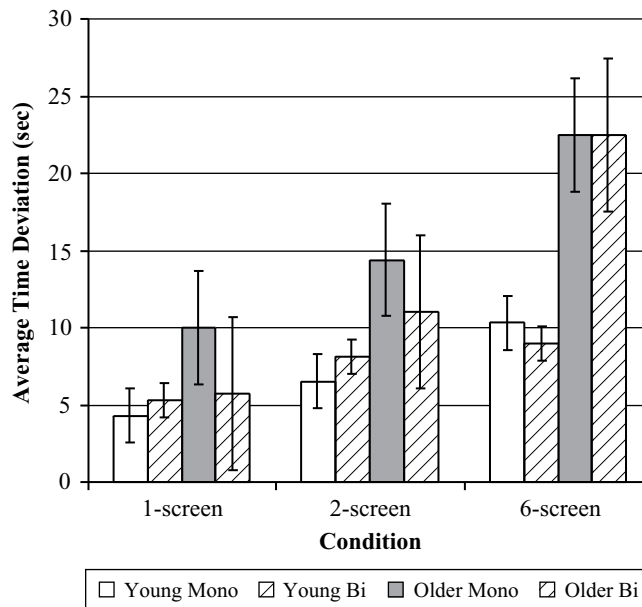


Figure 4. Mean deviation of start times from ideal start times (and standard errors) as a function of age, language, and condition.

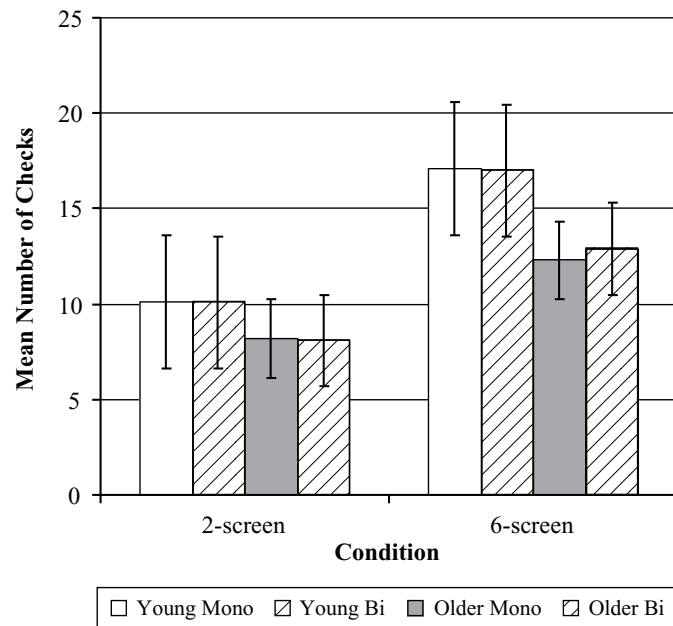


Figure 5. Mean number of food checks (and standard errors) as a function of age, language, and condition.

No penalties were associated with checking the progress of cooking, so it might have been expected that older adults would check more often. That was not the case, however; this finding was also reported by Kerr (1991). Checking allows the participant to ascertain the time remaining before returning to the cooking screen to stop a food. It therefore seems likely that greater numbers of checks would be associated with smaller discrepancy scores. Surprisingly, the correlations between number of checks and cooking measures were essentially zero in the younger group (see Table 3 and related discussion) but were significant in the older group (Table 3).

Because there was a moderately significant correlation between the number of checks and the discrepancy scores for the older adults, an ANCOVA on discrepancy scores was conducted using number of checks as a covariate. The purpose was to determine whether the discrepancy scores for the two age groups would be more comparable if the participants in each group engaged in the same amount of checking behavior. The ANCOVA showed that age continued to be significant after controlling for number of checks [$F(1,57) = 22.3, p < .0001$], with least squared means (and standard error) for discrepancies of 16.0 (2.8) and 31.1 (2.8) for the younger and older groups, respectively. Therefore, the difference in cooking accuracy is not accounted for by differences in the amount of checking undertaken by the participants in each group.

The importance to performance of checking the foods is surprisingly small, implying that success in the breakfast task is more attributable to such "internal" executive functions as working memory and prospective memory. It is also worth noting that in an experiment reported by Harris and Wilkins (1982), in which participants had to signal

a number of prespecified clock times while watching an interesting film, the participants also made inefficient use of checking. They could check the time on a clock behind them, but it was found that on 25% of the occasions on which people looked at the clock in the 10 sec prior to the target time, they still forgot to perform the target task. Checking updates the relevant information in working memory, but retaining that information and translating it into appropriate action is presumably undermined by switching attention to the other activity—in the present case, setting the table.

Table-Setting Times

The final group of measures comprised a set of variables for table-setting performance. The first of these is termed *table setting*, and was expressed as the proportion of total task time spent on table setting. The second was *inappropriate table setting*—namely, the amount of time spent setting the table when the cooking required attention, specifically because an ideal start time or stop time had passed. This measure was expressed as the proportion of table-setting time that should have been used to attend to cooking. The proportion of task time spent in setting the table is obviously the complement of *cooking time*—the proportion of overall task time devoted to the cooking operations. Our assumption was that higher proportions of cooking time reflected concern and care for the main purposes of the task—namely, to avoid burning the foods and to ensure a small range of stopping times. Thus, high values of table-setting proportions should reflect somewhat inefficient performance. Inappropriate table setting is a measure of perseverance: If the participant's percent of time on inappropriate table setting is high, it implies

that the participant made inefficient use of prospective memory, was unable to disengage from the current activity, and was inefficient in switching between tasks.

The total amounts of time spent performing the breakfast task are shown in Table 2. These means are the averages over the three conditions. Older adults took longer (mean = 328 sec) than the younger adults (mean = 301 sec), and this difference was significant [$F(1,56) = 7.5$, $MS_e = 1,505$, $p < .008$], with no effect of language ($F < 1.0$) and no interaction. Table 2 also shows the lengths of time spent in table setting and the proportions of total time taken up in table setting. These proportions are similar for the first three groups but lower for the older bilinguals. An ANOVA on the proportions of total time spent setting the table revealed a significant effect of language [$F(1,56) = 14.1$, $MS_e = 0.01$, $p < .0004$] and a marginally reliable effect of age [$F(1,56) = 3.6$, $MS_e = 0.01$, $p = .06$]. The age \times language interaction was significant, however [$F(1,56) = 12.9$, $MS_e = 0.01$, $p < .0007$], reflecting the fact that the language effect is carried entirely by the older group. Table 2 also reports the mean numbers of places set, showing that younger adults (mean = 48.0) set more places than older adults (mean = 30.2), a difference that was reliable [$F(1,56) = 51.0$, $MS_e = 273.7$, $p < .0001$], with no effects of language or of the age \times language interaction.

The finding that older adults took longer to perform the task yet set fewer places is a reflection of their lesser dexterity and reduced processing speed, and perhaps of their lesser familiarity with computer games. The finding that older bilinguals devoted a smaller proportion of task time to table setting means that they spent proportionally more time attending to cooking operations (129 sec, on average, compared with 73 sec, for the older monolinguals, and 81 sec, on average, for the two younger groups). This greater attention to cooking on the part of older bilinguals resulted in better performance on some but not all measures of task efficiency, as discussed later.

We also measured the amounts of time spent setting the table when it was inappropriate to do so because some food should be started or stopped. These times are expressed as proportions of table-setting time spent inappropriately, and are given in Table 2. Older adults were more likely to continue setting the table when they should have been cooking [$F(1,56) = 10.9$, $MS_e = 0.01$, $p < .002$], but bilinguals in both groups were less likely to continue table setting inappropriately [$F(1,56) = 4.2$, $MS_e = 0.01$,

$p < .05$]. The age \times language interaction was not significant ($F = 1.3$, n.s.). We interpret this tendency to continue table setting when other operations should be performed as a form of perseverative behavior—a reluctance to switch to a different activity—in line with the notion that older adults exhibit mild signs of frontal lobe dysfunction (Grady & Craik, 2000; West, 1996). The finding that bilingual participants of both age groups showed less inappropriate table setting may be construed as a sign of their greater degrees of cognitive control (Bialystok et al., 2004), and speculatively, as enhanced frontal efficiency for their respective age groups.

A more detailed examination of the way that the task time was allocated between cooking and table setting was examined by means of regression analyses. For each of the four age \times language groups, a regression model was calculated to identify the factors that determined how participants divided their time between these tasks. The independent variables entered into the model were the three background measures (digit span, alpha span, and Mill Hill) and the number of checks carried out in the 6-screen condition. The dependent variable was the proportion of task time spent in table setting as a function of the total amount of time each participant spent on the task in the 6-screen condition. In each model, only one independent variable contributed significantly to the results. For the younger monolinguals ($R^2 = .69$), there was a significant effect of alpha span [$t(1) = -3.7$, $p < .004$]; for younger bilinguals ($R^2 = .83$), there was a significant effect of number of food checks [$t(1) = -5.6$, $p < .0002$]; for older monolinguals ($R^2 = .43$), there was a significant effect of alpha span [$t(1) = -2.5$, $p < .03$]; and for older bilinguals ($R^2 = .39$), there was a significant effect of number of food checks [$t(1) = -2.3$, $p < .04$]. In both age groups, therefore, the division of attention between cooking and table setting was related to different factors for the monolinguals and bilinguals.

These results are only suggestive, because they are based on only 15 participants in each group. Nonetheless, the patterns are clear in showing that the allocation of time between table setting and cooking was determined more by differences in bilingualism than by differences in age. Specifically, the proportion of time spent table setting was determined by alpha span for monolinguals and by the number of food checks for bilinguals, with no systematic differences attributable to age group. The implication is that within the monolingual groups, those participants

Table 2
Mean Scores and Standard Deviations for Times Spent on Aspects of the Task and Mean Numbers of Places Set

Group	Total (sec)		Setting (sec)		Cooking (sec)		Prop. Set.		Inappropriate		Places Set	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Younger monolinguals	306	19.7	225	23.2	81	12.2	0.73	.04	0.15	.09	48	8.4
Younger bilinguals	296	7.3	216	24.9	80	24.8	0.73	.08	0.12	.07	48	9.5
Older monolinguals	318	44.2	245	44.0	73	24.2	0.77	.08	0.31	.13	32	8.3
Older bilinguals	338	60.2	209	56.3	129	40.3	0.62	.11	0.23	.24	29	11.6

Note—Prop. Set., proportion of total time spent setting table; Inappropriate, proportion of setting time in which cooking operations should have been performed.

with good working memory abilities keep track of the various operations that must be carried out, and so perform the task efficiently. In the bilingual groups, on the other hand, there appears to be greater reliance on active control measures, and in these groups, the individuals who carry out more frequent checks are the ones who allocate their time more efficiently between table setting and cooking.

Correlations Among Measures

To investigate relations among the cooking measures and between the background measures and cooking measures, these variables were entered into a correlation matrix separately for the two age groups. These correlations (Pearson's r) were calculated only for the results from the most complex 6-screen condition. Table 3 displays correlation coefficients for both age groups; the top right quadrant shows values relating background measures to cooking time measures for the younger group; the bottom left quadrant shows the corresponding correlation coefficients for the older group. The top left quadrant shows relations among the background measures, and the bottom right quadrant shows relations among the cooking measures. In both cases, correlations for younger adults are shown above the diagonal, and those for older adults are below the diagonal.

Alpha span and digit span are related significantly, as one might expect, and alpha span is also related to vocabulary level in the older group. This result suggests that verbal ability is a major determinant of this working memory measure in older adults, whereas alpha span is more determined by "domain-free" executive control processes in younger adults. Relations among cooking measures (Table 3, bottom right quadrant) are also somewhat higher for the older group. It might be expected that discrepancy, range, and start times should all be positively related to each other (smaller discrepancies and a smaller range represent efficient cooking) and negatively related to number of checks (more checks = smaller discrepancies), a pattern that holds for the older group but holds only weakly for the younger group. In both groups, the

deviation between ideal and actual start times is related to the range of stop times; those participants who closely monitored the foods were also able to have the foods finish cooking at close to the same time.

Differences in the background variables did not predict cooking performance particularly well. In fact, correlations for the older group (Table 3, bottom left) are all small and nonsignificant. For the younger group (Table 3, top right), alpha span and digit span are both related negatively to discrepancy, range, and start times (larger spans = smaller range and discrepancies), with two of the correlations reaching significance. Thus, span measures predicted the main cooking measures weakly, but only in the younger group. Again, this result may indicate that span measures themselves may reflect different underlying abilities in different groups; for example, that alpha span and even digit span may reflect verbal ability in the older group. Similarly, the cooking measures may reflect some degree of expertise in real-life meal preparation (as discussed in a later section). The general point is that while executive functions undoubtedly play a major role in performing the breakfast task efficiently and in performing span tasks effectively, these tasks also draw on specific areas of expertise. Whether two executive tasks correlate will, therefore, depend both on the balance (within each task) between some abstract executive ability and domain-specific expertise and on the similarity between the domains of expertise tapped by the two tasks.

Validity of the Breakfast Task

The participants found the task interesting and challenging. One common criticism of experimental studies of cognitive aging is that older adults may judge laboratory tasks to be artificial and pointless; in turn, this reduces their motivation and results in their performance being underestimated. In contrast, the participants in the present study enjoyed performing the breakfast task, and showed their involvement by such comments as, "I don't much like sausages, but I'll cook them anyway" and "actually, I usually drink tea at breakfast time." We therefore conclude

Table 3
Correlations ($df = 28$) Among Background Measures, Among Cooking Measures, and Between Background and Cooking Measures for Both Age Groups

	Alpha	Digit	MHV	Discrep.	Range	Start	Check
Alpha	–	+0.41	+0.18	–0.45	–0.30	–0.26	+0.04
Digit	+0.47	–	–0.05	–0.19	–0.36	–0.22	–0.25
MHV	+0.52	+0.31	–	–0.03	+0.18	+0.20	+0.25
Discrep.	+0.03	–0.10	–0.07	–	–0.22	–0.21	–0.03
Range	+0.05	–0.12	–0.14	+0.20	–	+0.81	–0.04
Start	–0.14	–0.18	–0.17	+0.52	+0.58	–	–0.02
Check	–0.12	–0.12	+0.01	–0.38	–0.22	–0.16	–

Notes—Correlations for the 30 younger participants are in the top right quadrant (between background and cooking measures) and above the diagonal of the top left quadrant (correlations among background measures) and bottom right quadrant (correlations among cooking measures). Correlations for the 30 older participants are in the bottom left quadrant and below the diagonal of the top left and bottom right quadrants. Correlations significant at $p < .05$ are printed in bold type. Alpha = alpha span; Digit = digit span; MHV = Mill Hill Vocabulary Scale; Discrep. = discrepancy; Start = deviation from ideal start times; Check = mean number of checks per condition.

that the task is realistic and involving and that participants are motivated to succeed. Another positive feature is that good performance does not depend on either speed or fine motor control. In theory, older adults could perform as well as their younger counterparts; the finding that they do not reflect an age-related decline in various executive and memory functions.

Some preliminary evidence for the validity of the breakfast task comes from a study of older adults reported by Edwards and Ryan (2004). Part of the study involved cooking a simple meal in the participant's own home while setting two places for the meal. Directions for cooking and place setting were provided. Additionally, the meal was prepared while a talk radio station was playing and a telephone call was received, to provide real-life interference. Of the various measures recorded, the most relevant for the present comparison is the number of strategies employed to facilitate (1) completing steps within the task in the required order and (2) switching attention from one step to the next. In a second session, the same participants performed the present computerized breakfast task, so it was possible to correlate the breakfast task measures with the real-life strategies. The results showed that the number of sequencing strategies in the real-life task correlated significantly with deviation from ideal start times in the breakfast task, $r = -.47$ for the 6-screen condition, $r = -.50$ for the 2-screen condition, and $r = -.51$ for the 1-screen condition. Sequencing strategies also correlated with range in the 2-screen condition ($r = -.45$) and in the 1-screen condition ($r = -.54$). Finally, sequencing strategies also correlated with the number of checks ($r = +.47$) and with the number of place settings ($r = +.53$) in the 6-screen condition. These correlations provide some reassurance that the breakfast task has real-life relevance and measurement validity.

Another important aspect of task validity is the effect on processing complexity from changing the number of screens that needed to be simultaneously active. All of the participants were tested in the order 1-screen, 2-screen, and 6-screen, to give participants the opportunity to practice the various operations and ensure that they were clear about procedures and task requirements when they tackled the most complex 6-screen condition. This increase in complexity had systematic effects on each of the planning processes captured by the various measures. The demands on working memory and prospective memory clearly increase from the 1-screen to the 6-screen condition, and these increasing demands are reflected in the declining performance of older adults, in particular (see Figures 2, 3, and 4). The absence of age-related effects on the 1-screen condition confirms that the older participants understood the task and could perform it as well as young people, under the simplest conditions. Therefore, the age-related decline in measures as the conditions moved eventually to the 6-screen version supports the argument that decrements in performance reflect problems of executive control. Equally important, the decrements across the three conditions were not the same for the various measures, indicating a somewhat different impact of aging on the processes measured by these variables.

Because the task required planning in real time and performing actions at specific points in the future, performance is probably based in some measure on the accuracy of time perception. Time perception over these kinds of intervals is known to be less efficient in older people (Craik & Hay, 1999), so this may also contribute to the age-related decrements on the discrepancy and accuracy scores. Against this possibility, there was no penalty for checking the progress of cooking, so an efficient participant could presumably have compensated for poorer time judgments by checking more often.

Effects of Aging

The measures shown in Figures 2, 3, and 4 for discrepancy, range of stop times, and deviations from ideal start times reveal a consistent pattern. Age-group differences are small or nonexistent on the 1-screen condition, somewhat larger on the 2-screen condition, and substantial on the 6-screen condition. This progression indicates that the last condition is most sensitive to the effects of normal aging, a result that we attribute to the greatly increased working memory load and greater need for prospective remembering in this condition. The present results thus add to established findings of age-related decrements in working memory (Craik & Jennings, 1992; Park et al., 2002; Zacks, Hasher, & Li, 2000) and in prospective remembering (Einstein et al., 2002; Henry et al., 2004), using a task that is more realistic than most laboratory tasks. They also indicate that these decrements in working memory and prospective memory are confined to situations that are more challenging, involving higher memory loads and greater processing demands.

Younger adults set more places (Table 2), an unsurprising finding given the general slowing of manual and mental dexterity associated with aging. More interestingly, older adults spent a greater proportion of their table-setting time inappropriately, during times when foods should have been started or stopped (Table 2). Thus, older adults demonstrated greater perseveration by being more reluctant to switch to a different activity, in line with the notion that older adults exhibit some mild signs of frontal lobe dysfunction (Grady & Craik, 2000; West, 1996).

Effects of Bilingualism

The effects of bilingualism on measures of cognitive control are smaller and more subtle than the effects of normal aging. Given the results of Bialystok and colleagues (Bialystok, 2006; Bialystok et al., 2004), we expected little or no difference between the performance levels of monolinguals and bilinguals in the young adult groups, but predicted that older bilingual participants would perform better than their monolingual counterparts on at least some measures of cognitive control.

In both age groups, the monolinguals and bilinguals performed equivalently on the main breakfast task measures of discrepancy, range, and start times (Figures 2–4), measures that we take to reflect prospective memory and working memory. However, the results showed a bilingual advantage on two table-setting measures of planning and

perseveration: Bilinguals spent a smaller proportion of the overall task time in table setting, and a smaller proportion of table-setting time in inappropriate table setting (Table 2). Both results suggest that bilingual participants made more effective use of their time in the task, and were more efficient in switching to food-related operations when it was appropriate. These effects were larger for the older participants. Inspection of Figures 3 and 4 reveals a similar advantage for older bilinguals over older monolinguals in the 2-screen condition, although neither effect reached statistical significance. In contrast, there is no hint of a bilingual advantage in the discrepancy scores (Figure 2).

Another way of considering how the participants divided their attention between the two tasks is to examine the factors that predict the proportion of time individuals spent setting the table. For both age groups, the ability of monolinguals to divide their time between the two tasks was determined by alpha span, a measure of working memory, but the ability of bilinguals to allocate time to table setting was determined by the number of food checks they made. In other words, monolinguals were guided by individual differences in working memory that allowed them to perform the task to a certain level of expertise, but the bilinguals controlled their performance more actively. This difference, revealed by the series of regression analyses, indicates better planning and monitoring by the bilinguals at both ages.

Speculatively, planning and executive control in the breakfast task may occur at both global and local levels (see also Ormerod, 2005). That is, participants may concentrate on stopping each food at its appropriate time and thereby obtain good discrepancy scores (efficient local planning), yet fail to coordinate the starting and stopping of all five foods in concert, thereby obtaining poor scores on the range of stop times and start times measures (poor global planning). In those terms, older bilinguals show no advantage on the local planning measures, yet exhibit good global planning abilities relative to their monolingual peers.

The breakfast task provides a laboratory-based method to assess planning and other aspects of higher-level cognitive control. The present results demonstrate that the task is sensitive to the effects of aging and also to differences in task management between older monolinguals and bilinguals. These individual differences, in turn, presumably reflect underlying differences in working memory, prospective memory, and task switching, although correlations between specific measures from the breakfast task and memory span measures were generally small (Table 3). We suggest that planning is not a unitary construct and that relations to its presumed components will depend both on the particular planning task under consideration and on the specific measures of working memory, prospective memory, and task switching selected. Further work is under way to document these relations more extensively, and also to explore the relations between performance on the breakfast task and planning performance in real-life settings.

REFERENCES

- BIALYSTOK, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. New York: Cambridge University Press.
- BIALYSTOK, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology*, **60**, 68-79.
- BIALYSTOK, E., CRAIK, F. I. M., KLEIN, R., & VISWANATHAN, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology & Aging*, **19**, 290-303.
- BURGESS, P., SIMONS, J., COATES, L., & CHANNON, S. (2005). The search for specific planning processes. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 199-227). New York: Psychology Press.
- CRAIK, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities: Mechanisms and performances* (pp. 409-422). Amsterdam: Elsevier.
- CRAIK, F. I. M., & HAY, J. F. (1999). Aging and judgments of duration: Effects of task complexity and method of estimation. *Perception & Psychophysics*, **61**, 549-560.
- CRAIK, F. I. M., & JENNINGS, J. M. (1992). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of aging and cognition* (pp. 51-110). Hillsdale, NJ: Erlbaum.
- CRAIK, F. I. M., MORRIS, L. W., MORRIS, R. G., & LOEWEN, E. R. (1990). Relations between source amnesia and frontal lobe functioning in older adults. *Psychology & Aging*, **5**, 148-151.
- DAVIES, S. P. (2005). Planning and problem solving in well-defined domains. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 35-52). New York: Psychology Press.
- EDWARDS, D., & RYAN, J. D. (2004, September). *Identifying the real life implication of mild stroke*. Paper presented at the American Congress of Rehabilitation Medicine—American Society of Neurorehabilitation Joint Annual Conference, Ponte Vedra, FL.
- EINSTEIN, G. O., MCDANIEL, M. A., MANZI, M., COCHRAN, B., & BAKER, M. (2000). Prospective memory and aging: Forgetting intentions over short delays. *Psychology & Aging*, **15**, 671-683.
- GARDEN, S. E., PHILLIPS, L. H., & MACPHERSON, S. E. (2001). Midlife aging, open-ended planning, and laboratory measures of executive function. *Neuropsychology*, **15**, 472-482.
- GRADY, C. L., & CRAIK, F. I. M. (2000). Changes in memory processing with age. *Current Opinion in Neurobiology*, **10**, 224-231.
- GUR, R. C., GUR, R. E., OBRIST, W. D., SKOLNICK, B. E., & REIVICH, M. (1987). Age and regional cerebral blood flow at rest and during cognitive activity. *Archives of General Psychiatry*, **44**, 617-621.
- HARRIS, J. E., & WILKINS, A. J. (1982). Remembering to do things: A theoretical framework and an illustrative experiment. *Human Learning*, **1**, 123-136.
- HENRY, J. D., MACLEOD, M. S., PHILLIPS, L. H., & CRAWFORD, J. R. (2004). A meta-analytic review of prospective memory and aging. *Psychology & Aging*, **19**, 27-39.
- KERR, S. (1991). *Prospective memory, aging, and lapses of intention*. Unpublished MA thesis, University of Toronto.
- KLIEGEL, M., MCDANIEL, M. A., & EINSTEIN, G. O. (2000). Plan formation, retention, and execution in prospective memory: A new approach and age-related effects. *Memory & Cognition*, **28**, 1041-1049.
- LURIA, A. R. (1966). *Human brain and psychological processes*. New York: Harper & Row.
- MARTIN, M., & EWERT, O. (1997). Attention and planning in older adults. *International Journal of Behavioral Development*, **20**, 577-594.
- MAYR, U. (2001). Age differences in the selection of mental sets: The role of inhibition, stimulus ambiguity, and response-set overlap. *Psychology & Aging*, **16**, 96-109.
- MEIRAN, N., & GOTLER, A. (2001). Modelling cognitive control in task switching and ageing. *European Journal of Cognitive Psychology*, **13**, 165-186.
- MORRIS, R. G., KOTITSA, M., & BRAMHAM, J. (2005). Planning in patients with focal brain damage: From simple to complex task performance. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 153-180). New York: Psychology Press.
- MORRIS, R. [G.], & WARD, G. (Eds.) (2005). *The cognitive psychology of planning*. New York: Psychology Press.

- MOSCOVITCH, M., & WINOCUR, G. (1992). The neuropsychology of memory and aging. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 315-372). Hillsdale, NJ: Erlbaum.
- ORMEROD, T. C. (2005). Planning and ill-defined problems. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 53-70). New York: Psychology Press.
- OWEN, A. M. (2005). Cognitive planning in humans: New insights from the Tower of London (TOL) task. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 135-151). New York: Psychology Press.
- PARK, D. C., LAUTENSCHLAGER, G., HEDDEN, T., DAVIDSON, N. S., SMITH, A. D., & SMITH, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology & Aging*, *17*, 299-320.
- PHILLIPS, L. H., MACLEOD, M. S., & KLIEGEL, M. (2005). Adult aging and cognitive planning. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 111-134). New York: Psychology Press.
- RAVEN, J. C. (1960). *Guide to the standard progressive matrices*. London: H. K. Lewis.
- RAZ, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 1-90). Mahwah, NJ: Erlbaum.
- RENDELL, P. G., & CRAIK, F. I. M. (2000). Virtual week and actual week: Age-related differences in prospective memory. *Applied Cognitive Psychology*, *14*, S43-S62.
- SHALLICE, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London, Series B*, *298*, 199-209.
- SHALLICE, T., & BURGESS, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114*, 727-741.
- STUSS, D. T., & ALEXANDER, M. P. (2000). Executive functions and the frontal lobes: A conceptual view. *Psychological Research*, *63*, 289-298.
- STUSS, D. T., & BENSON, D. F. (1986). *The frontal lobes*. New York: Raven.
- WARD, G. (2005). Planning and the executive control of thought and action. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 89-110). New York: Psychology Press.
- WEST, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*, 272-292.
- ZACKS, R. T., HASHER, L., & LI, K. Z. H. (2000). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 293-358). Mahwah, NJ: Erlbaum.

(Manuscript received August 19, 2004;
revision accepted for publication June 9, 2005.)