Development of Planning Abilities in Normal Aging: Differential Effects of Specific Cognitive Demands

Lena Köstering
University of Freiburg

Christoph Stahl
University of Cologne

Rainer Leonhart, Cornelius Weiller, and Christoph P. Kaller
University of Freiburg

In line with the frontal hypothesis of aging, the ability to plan ahead undergoes substantial change during normal aging. Although impairments on the Tower of London planning task were reported earlier, associations between age-related declines and specific cognitive demands on planning have not been studied. Here we investigated the impact of search depth and goal ambiguity on planning, which impose demands on the depth and breadth of look-ahead processes, respectively. Besides an overall age-related decline in planning accuracy of 106 healthy older adults, differential search depth effects were found: Whereas planning accuracy of subjects in the early 60s was not affected by variations in search depth, between the ages of 65 and 76 years, accuracy was significantly decreased for high versus low levels of search depth. For subjects older than 76, different search depth levels did not further impact on accuracy, which was lowest overall. This nonlinear pattern may reflect differential impairments in fluid abilities and working memory capacity across various stages of older age. As no age-related effects of goal ambiguity were found, normal aging seems to be specifically sensitive to planning demands on the depth but not the breadth of anticipatory search processes. Hence, cognitive functions subserved by the prefrontal cortex experience differential development over the course of normal aging.

Keywords: cognitive aging, prefrontal cortex, planning, Tower of London, structural problem parameters

Supplemental materials: http://dx.doi.org/10.1037/a0032467.supp

Normal aging is known to be accompanied by changes in cognitive functioning across various domains (Craik & Salthouse, 2000; Hedden & Gabrieli, 2004). The frontal hypothesis of cognitive aging assumes this change to predominantly affect functions mediated by the frontal lobes, in particular executive functions underlying higher order conscious control of behavior (for reviews, see Moscovitch & Winocur, 1995; Phillips, MacPherson, & Della Sala, 2002; West, 1996). As a prototypical example of these complex cognitive functions, the ability to plan ahead is especially crucial for successful coping of situations that are beyond everyday routine. That is, planning is necessary whenever automatized schemes and action routines do not suffice to achieve a goal, but rather behavior has to be mentally simulated and its consequences anticipated before action execution (Ward & Morris, 2005).

One of the most commonly used planning instruments is the Tower of London (TOL). The TOL is a disc-transfer task originally developed to assess executive impairments in patients following frontal brain lesions (Shallice, 1982) that has subsequently been validated also as a planning task for healthy subjects (for reviews, see Berg & Byrd, 2002; Kaller, Rahm, Köstering, & Unterrainer, 2011) and is known to critically rely on activity of the prefrontal cortex (see Unterrainer & Owen, 2006, for a review). Advancing age has been consistently found to affect planning performance on the TOL (see Phillips, MacLeod, & Kliegel, 2005; Sullivan, Riccio, & Castillo, 2009, for reviews), thus concurring with the frontal hypothesis. In an early study using a simplified version of the TOL, subjects exhibited an age-related linear decline in performance in terms of number of problems solved and number of moves needed within a set time interval (Allamanno, Della Sala, Laiacona, Pasetti, & Spinnler, 1987). Several other studies similarly found older adults, especially after the age of 60, to generally need more moves to achieve a solution than younger adults (Andrés & van der Linden, 2000; Bugg, Zook, DeLosh, Davalos, & Davis, 2006; Gilhooly, Wynn, Phillips, Logie, & Della Sala, 2002) and to solve fewer problems in the minimum number of moves (De Luca et al., 2003; Phillips, Smith, & Gilhooly, 2002; Robbins et

al., 1998; Zook, Welsh, & Ewing, 2006). Furthermore, older adults have been found to need more time in the planning phase prior to movement execution (Andrés & Van der Linden, 2000; Phillips, Smith, & Gilhooly, 2002; Robbins et al., 1998; also see Phillips et al., 2005).

Extant evidence on age-related planning impairments is based on the most common definition of problem difficulty in terms of the number of moves minimally needed for solution (Owen, Downes, Sahakian, & Polkey, 1990; Shallice, 1982). However, numerous studies have demonstrated that this is too coarse a concept to capture the underlying cognitive processes and that other structural problem parameters also critically affect performance (e.g., Berg, Byrd, McNamara, & Case, 2010; Carder, Handley, & Perfect, 2004; Kaller, Unterrainer, Rahm, & Halsband, 2004; Ward & Allport, 1997; see Kaller et al., 2011, for a review). Structural problem parameters may even differentially affect planning performance in clinical and developmental populations (for examples, see Kaller, Rahm, Spreer, Mader, & Unterrainer, 2008; McKinlay et al., 2008; Rainville, Lepage, Gauthier, Kergoat, & Belleville, 2012). Besides the minimum number of moves to an optimal solution, a problem's difficulty is substantially determined by its *goal hierarchy* and *search depth* (see Figure 1; cf. Kaller et al., 2011). Goal hierarchy refers to the ambiguity with which the sequence of final moves is derivable from the mere configuration of the goal state (e.g., Berg et al., 2010; Kaller et al., 2004; Klahr & Robinson, 1981; Newman & Pittman, 2007; Ward & Allport, 1997): If all balls are to be stacked on one peg, it is unambiguously deducible that the bottommost ball has to be placed into its goal position first, followed by the second-to-last and topmost ball, providing a hierarchy in terms of sequential ordering of the goal moves that have to be executed. In comparison, if balls are to be distributed in a "flat" configuration, that is, each ball on a different peg, the sequence of final moves is not inherently evident from the problem structure, and several alternative move sequences have to be considered and evaluated against each other. The search depth of a problem is another parameter known to affect problem complexity (e.g., Kaller et al., 2004, 2008; McKinlay et al., 2008); it refers to the number of intermediate moves[1] to be considered before execution of the first goal move (Spitz, Webster, & Borys, 1982).

Variations of structural properties thus entail differential task demands in terms of cognitive processing: Greater search depths require that a longer series of intermediate moves leading to the first goal move be identified in consideration of move interdependencies, thereby invoking increased demands on the depth of anticipatory steps along the solution path (Köstering, McKinlay, Stahl, & Kaller, 2012). Furthermore, greater search depths induce higher demands on maintenance processes in working memory because an individual has to hold a particular goal move in mind while devising the intermediate moves necessary to accomplish it (Spitz et al., 1982; also see Ward & Allport, 1997). By contrast, rather than searching ahead a single move sequence and maintaining its steps in working memory, ambiguous goal hierarchies demand a flexible search across the breadth of the problem space among several alternative goal move sequences (Köstering et al., 2012). That is, although sufficient working memory capacity is also needed to mentally represent alternative move sequences that arise from ambiguous goal hierarchies, the critical cognitive demand lies in a flexible changing of working memory representations (cognitive flexibility; cf. Cools, 2006) so as to identify the optimal sequential ordering of goal moves. Cognitive flexibility is supposed to rely on striatal functioning, whereas stable maintenance of working memory representations is subserved by the prefrontal cortex (Cools, 2006), which is affected by normal aging to a much greater extent than the striatum (Hedden & Gabrieli, 2004; Raz & Kennedy, 2009). Hence, age differences can be expected to result primarily from varying demands on cognitive stability. This view is further supported by prior findings revealing that goal ambiguity-related planning impairments are associated with Parkinson's disease (PD), which is characterized by degeneration of the striatum, whereas search depth-related planning impairments are associated with lower dementia screening ratings in PD patients and controls alike (Köstering et al., 2012).

The potential association between age-related decline in planning ability and specific task demands on the breadth versus depth of search processes has not been systematically studied, so that the specific nature of planning deficits in normal aging remains as yet unclear. Therefore, our rationale in the current study was to unravel the influence of these specific task demands on older adults' planning ability by delineating the specific contribution of variations in goal hierarchy and search depth to age-related differences in TOL performance. Prior evidence indicates that preschool children exhibit an age-related increase in the ability to deal with greater search depth of problems (Kaller et al., 2008; also see Spitz et al., 1982), marking variations in the search depth of problems as a sensitive measure of the cognitive development of planning abilities. As normal aging is also known to affect working memory capacity (e.g., Bopp & Verhaeghen, 2005; Norman, Kemper, & Kynette, 1992; Zeintl & Kliegel, 2007), older adults can be expected to be especially prone to increased demands on working memory maintenance through increased depth of search during planning with advancing age. Thus, mirroring the upward slope in children, the downward development of planning abilities in normal aging is expected to be especially associated with a decline in the ability to deal with increased cognitive demands on deeper look-ahead processes induced by greater search depths of problems, but not with a decline in the ability to deal with broader look-ahead processes induced by ambiguous goal hierarchies of problems.

Method

Participants

Participants were recruited as part of a study at the University Medical Center Freiburg comprising the assessment of planning abilities with the Tower of London (TOL) planning task (Shallice, 1982) as well as additional behavioral testing and magnetic resonance (MR) imaging that were not of interest for the current analysis. The study protocol was approved by local ethics authorities. Recruitment was carried out throughout the city of Freiburg with the help of advertisements and information leaflets informing the city residents about the possibility to get feedback on a dementia screening test in addition to a monetary compensation of 25

---

[1] Contrary to a goal move that places a ball into its goal position, an intermediate move does not place a ball into its goal position but is essential to problem solution nonetheless (Kaller et al., 2011).
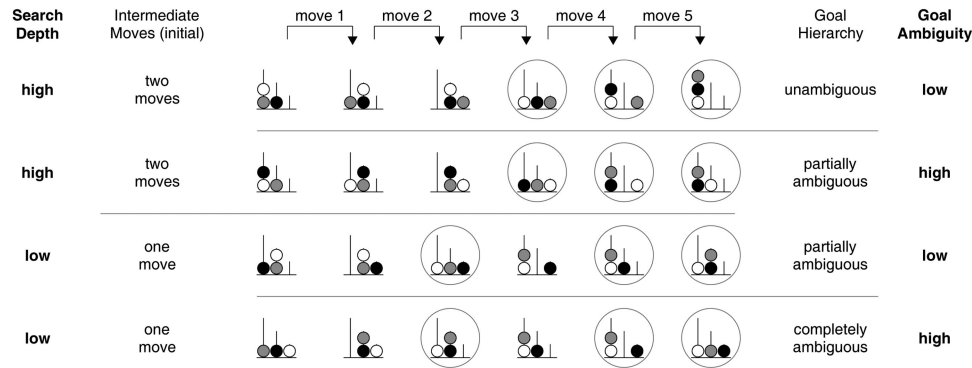
Figure 1. Experimental design of the current study, illustrated for problems with a minimum number of five moves. In the Tower of London problems applied here, two predominant structural patterns are evident: (a) *Search Depth* denotes the number of intermediate moves to be considered before the first goal move. High demands on search depth require consideration of a longer series of intermediate moves and their interdependencies (cf. Kaller et al., 2011). For example, for optimal solution of the problem shown in the top row, the white ball itself has to be moved from its goal peg to the middle peg, thereby enabling that in turn the gray ball can be removed from the goal position of the white ball, which only then can be placed into its goal position. That is, because two intermediate moves and their resulting consequences have to be anticipated before the first goal move, the solution path has to be searched ahead to a greater depth than, by contrast, for the problem shown in the bottom row. This only demands removing the obstacle presented by the gray ball before the white ball can be placed into its goal position, thus requiring mental anticipation of a shorter interdependent move sequence. (b) *Goal Hierarchy* refers to the ambiguity with which the sequential ordering of goal moves is deducible from the mere configuration of the goal state (cf. Kaller et al., 2011). In problems with a tower configuration (see top row), the goal state alone unambiguously implies the goal move sequence in that the white ball has to be placed into its goal position first, followed by the black and the gray ball. By contrast, for problems with flat goal states (see bottom row), no sequential order of goal moves is deducible and several alternatives exist as to whether to first move the gray or the black or the white ball to their goal position. Therefore, rather than being able to readily select a single move sequence, a broad search among the entirety of alternative move sequences has to be conducted to identify the optimal sequential ordering of goal moves. Given that not all possible combinations of three different levels of goal hierarchy are existent for the two levels of search depth in the Tower of London problem space, the levels of goal hierarchy were hierarchically nested under the levels of *Search Depth*, resulting in the factor *Goal Ambiguity*, each featuring two levels with high and low demands (see also Köstering et al., 2012; McKinlay et al., 2008). Circles around states denote problem states resulting from goal moves.

euro. To be eligible for participation, subjects had to be native German speakers and 60 years of age or older. Before subjects were tested, the presence of following exclusion criteria was checked over the phone using a medical history questionnaire: history of or current neurological disorder/cerebral event; current episode of psychiatric disorder; history of major psychiatric disorder (e.g., psychosis); current use of psychotropic medication; alcohol or drug abuse; insulin-dependent diabetes; color blindness; and major impairment of hearing or eyesight. Additionally, during behavioral testing, several control measures were administered: The dementia screening battery devised by the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) at the Duke University Center for the Study of Aging and Human Development (Morris et al., 1989) was used in its authorized German version (Memory Clinic Basel, 1997) to screen for dementia and related mild cognitive impairment (MCI). The CERAD includes an assessment of the Mini-Mental State Exam (MMSE; Folstein, Folstein, & McHugh, 1975) and has demonstrated superior validity compared with the MMSE alone in discriminating older adults with normal cognition from those with MCI (Chandler et al., 2005; Ehrensperger, Berres, Taylor, & Monsch, 2010; Seo et al., 2010). Screening for MCI/dementia was carried out according to the

following procedure: If subjects exceeded the cutoff value of -1.5 standard deviations on more than one of the nine z-transformed CERAD subscales (including the MMSE), their full CERAD profile was evaluated by an experienced neuropsychologist at the Memory Clinic of the University Medical Center Freiburg, and they were included or excluded based on this clinical judgment of MCI. Crystallized verbal intelligence was measured with a German vocabulary test (Mehrfachwahl-Wortschatz-Intelligenztest, MWT–A; Lehrl, Merz, Burkhard, & Fischer, 1991); the Geriatric Depression Scale (GDS; Sheikh & Yesavage, 1986) and the Beck Depression Inventory (BDI–II; Beck, Steer, & Brown, 1995) were used to screen for current depressive symptoms. A short test on visual acuity established that participants' acuity was sufficient for performance of the computerized tasks. Furthermore, fluid-attenuated inversion recovery (FLAIR) and T1-weighted MR images of participants were inspected for incidental findings of neurological abnormalities or events.

Following this procedure, 126 community-dwelling middle-class adults of European ethnic origin were included in the study. Written informed consent (including an option on whether they wanted to be informed about the results of their dementia screening) was obtained before behavioral testing, which lasted 2.5 hr on

average and during which breaks were taken as needed. A subset of 88 participants also completed MR imaging. Individualized written feedback on the CERAD dementia screening test was sent to participants via mail after completion of testing.

Of the 126 participants, two subjects had to be excluded because of an incidental neurological finding. After inspection of control measures, another 16 subjects had to be excluded: Four participants' BDI–II score exceeded the threshold of 20 points. Furthermore, the CERAD screening of 12 participants was clinically indicative of MCI. Last, two participants were excluded after inspection of performance on the TOL during data preprocessing because their overall accuracy rate was below 5%, indicating that no meaningful data had been acquired. In sum, 20 subjects were excluded, leaving data of 106 subjects for statistical analysis.

Participants ranged in age from 60.73 to 89.45 years; 65 of the subjects were women (61.32%; see Table 1). The exact age was computed by dividing the number of days from date of birth to date of the behavioral testing session by 360. Subjects were then assigned to age groups based on the 20th, 40th, 60th, and 80th percentile of the sample age distribution so as to yield five groups of equal size: 60.73–65.83 years (Group 1; $N = 21$), 65.84–69.14 years (Group 2; $N = 22$), 69.15–72.16 years (Group 3; $N = 21$), 72.17–76.48 years (Group 4; $N = 21$), and 76.49 years and older (Group 5; $N = 21$). For comparison of cognitive status between age groups, the total CERAD score was computed as proposed by Chandler et al. (2005): The scores on the subscales Verbal Fluency, Boston Naming Test, Word List Learning, Constructional Praxis, Word List Recall, and Word List Recognition Discriminability were summed and corrected for effects of age, education, and gender (assessed by linear regression with men as the reference group; cf. Chandler et al., 2005) by the following correction formula: total score $+ .449^*$age $- .536^*$years of education $+ 4.536^*$sex. Age groups differed slightly in ratio of women to men; they did not differ from each other on any of the control measures (see Table 1). Analyses of variance (ANOVAs) between age groups on years of education, BDI–II and GDS scores, crystallized intelligence, and mental status assessed by the CERAD total score and the MMSE were all nonsignificant (highest $F = 1.433$, lowest $p = .231$).

Experimental Task and Paradigm

Planning was assessed with the original Shallice Tower of London task (TOL; i.e., three rods of unequal size and three balls; cf. Shallice, 1982) using a structurally balanced set of TOL problems (Kaller et al., 2011) which consisted of 24 trials of three-, four-, and five-move problems (eight trials each) presented in fixed order (for psychometric evaluations, see also Kaller, Unterrainer, & Stahl, 2012). The structural problem parameters search depth and goal hierarchy were manipulated following McKinlay et al. (2008): For each length of optimal solution path, the relative ambiguity of goal hierarchy was nested under the two respective levels of search depth, yielding the two experimental parameters *search depth* (SD) and *goal ambiguity* (GA), with two levels (low and high) each (see Figure 1; also see Kaller et al., 2011; 2012; Köstering et al., 2012). The TOL was presented in a computerized version on a 12.1-in. touch-sensitive screen of an IBM ThinkPad ×41 Tablet laptop computer (IBM; Armonk, NY). At the beginning of the task, subjects were familiarized with the maneuvering of balls (in the primary colors of yellow, blue, and red) and orally instructed to transform the start state of each problem, presented in the bottom half of the screen, into matching the goal state, presented in the upper half of the screen, in the least number of moves possible while following the rules that (a) the tallest peg could accommodate three balls, the second tallest peg two balls, and the shortest peg only one ball; (b) only one ball at a time could be moved; (c) no ball could be placed outside the pegs; (d) only the topmost ball could be moved if several were stacked on one peg; and (e) they should always plan their moves in advance of starting to move any balls. The minimum number of moves to solution for each problem was presented throughout, and the presentation of each trial was limited to 1 min (cf. Shallice, 1982). If the time limit was exceeded on three consecutive trials, the task was automatically aborted.

Data Analysis

Raw data of TOL performance were restructured and aggregated for each participant so as to perform a repeated-measures analysis of variance (RM ANOVA). The main dependent measure was planning *accuracy*, computed as the number of problems solved

Table 1
Demographic Information for Age Groups

Variable	Group 1 (60.73–65.83 years)[a] M (SD)	Group 2 (65.84–69.14 years)[b] M (SD)	Group 3 (69.15–72.16 years)[c] M (SD)	Group 4 (72.17–76.48 years)[d] M (SD)	Group 5 (76.49 years and older)[e] M (SD)	Sample[f] M (SD)
Age	63.39 (1.59)	67.49 (1.11)	71.00 (0.87)	74.30 (1.36)	81.48 (4.34)	71.50 (6.58)
EDU	15.71 (3.32)	15.05 (4.27)	13.86 (3.73)	13.52 (3.88)	14.38 (3.60)	14.51 (3.79)
MWT-A	33.91 (1.41)	32.86 (2.80)	32.81 (2.66)	33.00 (1.84)	32.76 (2.32)	33.07 (2.27)
BDI-II	5.86 (3.64)	6.60 (4.61)	7.81 (4.57)	8.65 (4.45)	6.76 (3.42)	7.11 (4.20)
GDS	1.19 (1.66)	0.96 (1.25)	1.67 (1.49)	1.90 (1.86)	1.10 (1.76)	1.35 (1.62)
CERAD	111.31 (7.01)	110.45 (7.40)	111.89 (8.04)	110.00 (8.83)	111.98 (7.15)	111.12 (7.60)
MMSE	29.14 (1.39)	29.27 (0.83)	29.05 (.80)	29.00 (0.95)	28.57 (1.03)	29.01 (1.03)

Note. EDU = years of education; MWT-A = Mehrfachwahl-Wortschatz-Intelligenztest (assessment of crystallized intelligence; Lehrl et al., 1991); BDI–II = Beck Depression Inventory (Beck, Steer, & Brown, 1995); GDS = Geriatric Depression Scale (Sheikh & Yesavage, 1986); CERAD = Consortium to Establish a Registry for Alzheimer's Disease (Memory Clinic Basel, 1997); MMSE = Mini-Mental State Exam (Folstein, Folstein, & McHugh, 1975).
Gender (female:male): [a] 15:6. [b] 11:11. [c] 14:7. [d] 11:10. [e] 14:7. [f] 65:41.

correctly in the minimum number of moves within the time limit divided by the total number of 24 problems, yielding a percentage value of accuracy. Furthermore, *initial thinking time* (ITT; time from presentation of a problem to movement of the first ball) and *movement execution time* (MET; time from movement of the first ball to problem completion) were recorded for problems completed within the time limit and median-aggregated for each participant across correct and incorrect solutions. In supplementary analyses, the number of *excessive moves* (i.e., number of moves above the minimum number of moves) and the efficiency of problem solution (number of moves divided by minimum number of moves) for problems completed within the time limit were computed to explore effects on other measures of planning performance beyond accuracy (see the online supplementary materials). For all dependent measures, a RM ANOVA was performed using SPSS Version 20 with the within-subject factors *search depth* (SD; low vs. high demands on search depth) and *goal ambiguity* (GA; low vs. high ambiguity of goal hierarchy), and the between-subjects factor *age group* (Groups 1–5). Supplementary analyses also included the within-subject factor *minimum number of moves* (three, four, and five moves).

Results

Accuracy

On average, participants solved 64.15% (standard deviation, ±14.21) of problems. The RM ANOVA on accuracy (percentage of problems solved correctly in the minimal number of moves) yielded a main effect for search depth, $F_{SD}(1, 101) = 26.845$, $p < .001$, partial $\eta^2 = .210$, and for goal ambiguity, $F_{GA}(1, 101) = 66.692$, $p < .001$, partial $\eta^2 = .398$. Participants solved more problems with a low than with a high search depth and also more problems with low than high ambiguity of goal hierarchy. As Figure 2 illustrates, there was also a significant effect for age group, $F_{age}(4, 101) = 4.146$, $p = .004$, partial $\eta^2 = .141$. Tukey's honestly significant difference post hoc pairwise comparisons revealed that subjects in Group 5 solved significantly fewer problems overall than subjects of Groups 1 and 2 (*p*s = .002 and .025, respectively; see also Figure 2).

Furthermore, there was a significant interaction between search depth and age group, $F_{SD*age}(4, 101) = 3.806$, $p = .006$, partial $\eta^2 = .131$. Post hoc Bonferroni-corrected pairwise comparisons of search depth levels for each age group revealed that accuracy for problems with a high search depth was significantly lower compared with problems with a low search depth for Age Groups 2 ($p = .002$), 3 ($p < .001$), and 4 ($p < .001$; see Figure 3, panel A). By contrast, no such difference in accuracy was evident for the youngest (Group 1; $p = .832$) and oldest participants (Group 5; $p = .525$; see Figure 3A). Variance in accuracy for low and high levels of search depth was comparable across age groups (see standard error of means in Figure 3A) and mean accuracy for high search depth ranged from 89.29% (±16.90) to 52.38% (±24.88) in Group 1 and from 76.19% (±23.02) to 36.90% (±24.52) in Group 5 across the minimum number of moves to solution (see Table S02, in the online supplementary materials). Thus, the lack of effect in Groups 1 and 5 was not attributable to a lack of search-depth-related variance per se. It is important to note that there was
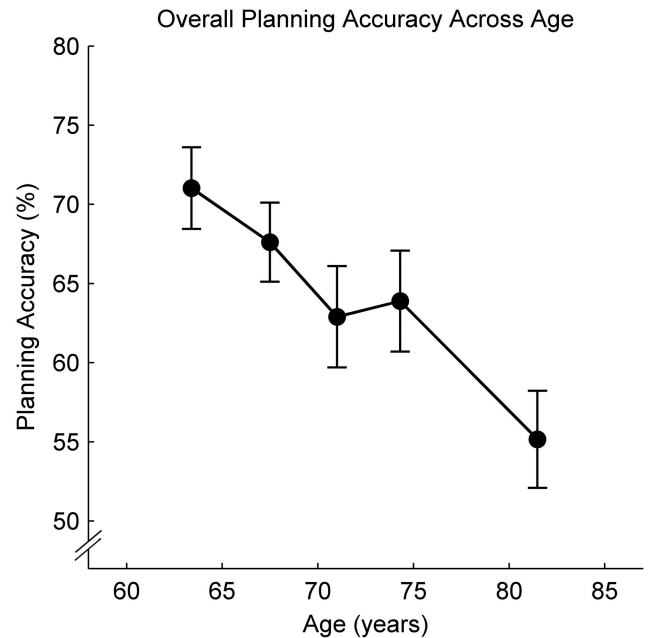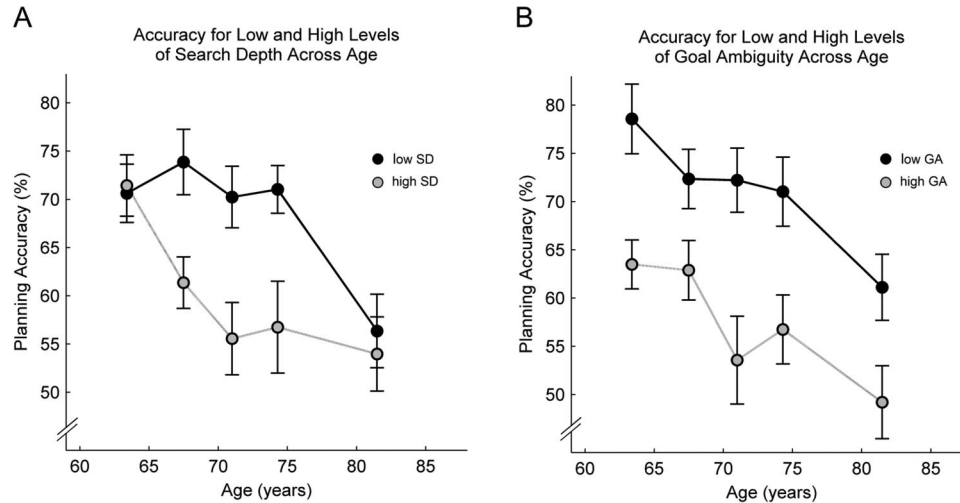

Figure 2. Overall planning accuracy (percentage of problems solved correctly in the minimum number of moves) across age. Error bars indicate standard error of mean. Age groups are marked at their group-specific mean age.

no interaction between goal ambiguity and age group, $F_{GA*age}(4,101) = .834$, $p = .506$ (see Figure 3B).

None of the remaining two- or three-way interactions reached significance. To preclude that results were influenced by age-related differences in the number of time out trials (mean, 2.92 ± 2.03 out of 24 trials), possibly reflecting gross differences in speed of processing, a unifactorial ANOVA on absolute number of time out trials with age group as independent measure was computed. However, age groups did not differ in the number of time out trials, $F_{age}(4, 101) = 1.140$, $p = .342$; Group 1, mean 2.29 (±1.74); Group 2, mean 3.00 (±1.72); Group 3, mean 2.67 (±1.59); Group 4, mean 3.14 (±2.08); Group 5, mean 3.52 (±2.77). Furthermore, overall accuracy for all trials was highly correlated with accuracy for trials completed within the time limit ($r = .867$, $p < .001$).

The results on an overall difference in accuracy between age groups and a differential age-related effect of search depth also held up when we took into account the effect of the minimum number of moves (see the online supplementary materials), which are known to also significantly affect planning accuracy (e.g., Berg et al., 2010; Kaller et al., 2012; Owen et al., 1990; Shallice, 1982). Thus, the interaction of age groups with search depth, but not goal ambiguity, was independent of the minimum number of moves to solution, which is traditionally the standard used to define problem difficulty.

In further supplementary analyses, possible effects on other measures of planning performance—that is, on number of excessive moves and efficiency of problem solutions—were investigated (see the online supplementary materials). For both measures, a significant interaction between search depth and age groups emerged, again following a nonlinear pattern with Group 3 (69.15–72.16 years) being the most affected by variations in search depth (see Figure S01, in the online supplementary materials). There

A

Accuracy for Low and High Levels of Search Depth Across Age

B

Accuracy for Low and High Levels of Goal Ambiguity Across Age

Figure 3. Search-depth- and goal-ambiguity-related effects on planning accuracy (percentage of problems solved correctly in the minimum number of moves) across age. (A) Accuracy for problems with low levels of search depth (low SD; black markers) and high levels of search depth (high SD; gray markers) across age. (B) Accuracy for problems with low levels of goal ambiguity (low GA; black markers) and high levels of goal ambiguity (high GA; gray markers) across age. Error bars indicate standard error of mean. Age groups are marked at their group-specific mean age.

were no interactions between goal ambiguity and age group. In contrast to the main analysis of planning accuracy, however, age groups did not differ overall in number of excessive moves and efficiency of problem solutions. In sum, the differential age-related effect of search depth was consistently evident across several measures of planning accuracy, while the overall effect of age emerged solely for accuracy in terms of percentage of correct solutions (see the online supplementary materials for detailed information).

Initial Thinking Time and Movement Execution Time

Age-related effects on the latency of planning processes were investigated for initial thinking time (ITT) and movement execution time (MET) separately. Mean ITT was 8,840.12 ms ($\pm$2,863.09) overall; 5,907.27 ms ($\pm$1,738.60) for three-move problems; 6,900.21 ms ($\pm$2,483.64) for four-move problems; and 10,785.07 ms ($\pm$5,179.42) for five-move problems. Mean MET was 11,755.73 ms ($\pm$3,032.69) overall; 6,636.28 ms ($\pm$2,308.01) for three-move problems; 9,640.19 ms ($\pm$3,215.36) for four-move problems; and 14,383.77 ms ($\pm$5,529.34) for five-move problems.

The RM ANOVA on ITT with search depth and goal ambiguity as within-subject factors and age group as between-subjects factor revealed main effects for search depth, $F_{SD}(1, 101) = 29.577, p < .001$, partial $\eta^2 = .227$, and goal ambiguity, $F_{GA}(1, 101) = 5.714, p = .019$, partial $\eta^2 = .054$, but no main effect for age group, $F_{age}(4, 101) = 1.705, p = .155$. Also no interaction effects of search depth or goal hierarchy with age group emerged, $F_{SD*age}(4, 101) = .143 \, p = .966$; $F_{GA*age}(4, 101) = .264, p = .901$ (see Figures 4A and B), only an interaction between the two within-subject factors, $F_{SD*GA}(1, 101) = 8.481, p = .004$, partial $\eta^2 = .077$.

The RM ANOVA on MET with search depth and goal ambiguity as within-subject factors and age group as between-subjects factor revealed main effects for search depth, $F_{SD}(1, 101) = 4.818,$

$p = .030$, partial $\eta^2 = .046$, and goal ambiguity, $F_{GA}(1, 101) = 7.787, p = .006$, partial $\eta^2 = .072$, as well as for age group, $F_{age}(4, 101) = 3.806, p = .006$, partial $\eta^2 = .131$. Furthermore, there was a strong trend for an interaction between search depth and age group, $F_{SD*age}(4, 101) = 2.332, p = .061$, partial $\eta^2 = .085$ (see Figure 5A), but not between goal ambiguity and age group, $F_{GA*age}(4, 101) = 1.009, p = .407$ (see Figure 5B). Bonferroni-corrected post hoc pairwise comparisons revealed significantly longer movement execution times in problems with high compared with problems with low levels of search depth for Age Group 3 ($p = .001$), but none of the other groups (Group 4: $p = .088$, all other ps > .600). All remaining interactions did not reach significance.

In sum, results on latency measures revealed that ITTs were unaffected by age and did not show any age-related differences in the impact of search depth and goal ambiguity. METs, by contrast, did show both an overall age effect and a strong trend for an interaction of age with search depth, but not with goal ambiguity, thus mirroring results on planning accuracy.

Discussion

Summary of Results

The objective in the present study was (a) to investigate the effect of normal aging on planning ability as measured with the Tower of London (TOL) and (b) to assess age-related differences in the impact of the structural problem parameters of search depth and goal hierarchy on planning performance. It was assumed that age-related differences in planning accuracy would be especially prominent for problems with greater search depths. In accordance with this assumption, besides an overall age-related decline in accuracy, a significant interaction be-
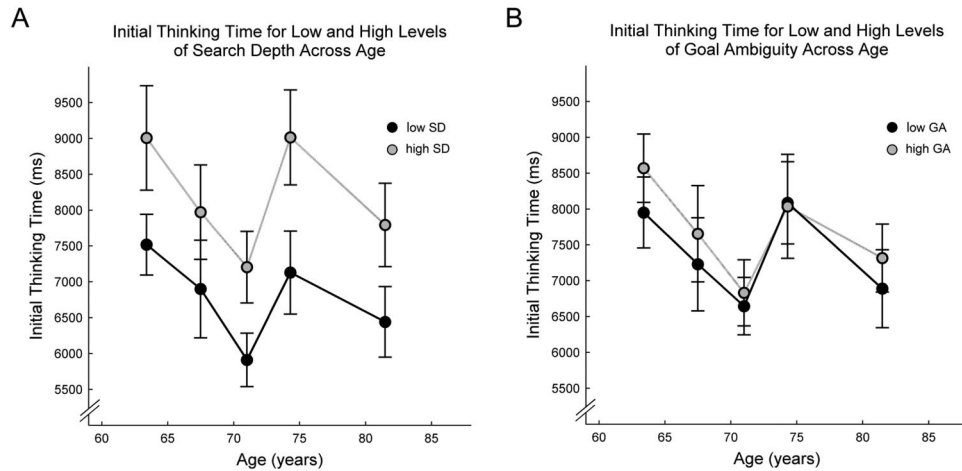
Figure 4. Search-depth- and goal-ambiguity-related effects on initial thinking time (ITT; time from presentation of a problem until movement of the first ball) across age. (A) ITT for problems with low levels of search depth (low SD; black markers) and high levels of search depth (high SD; gray markers) across age. (B) ITT for problems with low levels of goal ambiguity (low GA; black markers) and high levels of goal ambiguity (high GA; gray markers) across age. Error bars indicate standard error of mean. Age groups are marked at their group-specific mean age.

tween search depth and age groups was found: Whereas accuracy of solutions dropped significantly from low to high search depth of problems for the three intermediate age groups (spanning ages 65–76 years), performance of the first (60–65 years) and fifth (76 years and older) age groups was unaffected by the manipulation of search depth (Figure 3A). This interaction effect also held when planning performance was evaluated in terms of the number of excessive moves and in terms of the efficiency of problem solutions, thus demonstrating the robustness of effects. Furthermore, the interaction between age groups and search depth was independent of problem difficulty as

commonly defined by the minimum number of moves to solution. In contrast, the effect of goal hierarchy on planning accuracy was not moderated by participants' age (Figure 3B). Further analyses on the latency of planning processes revealed that ITTs were modulated neither by participants' age nor by its interaction with search depth (Figure 4A). In contrast, METs were significantly affected by participants' age, and an interaction between age groups and search depth, but not goal hierarchy, was evident, demonstrating a similarly nonlinear pattern as the search-depth specific effect on planning accuracy (Figure 5A). Present results hence revealed a specific planning
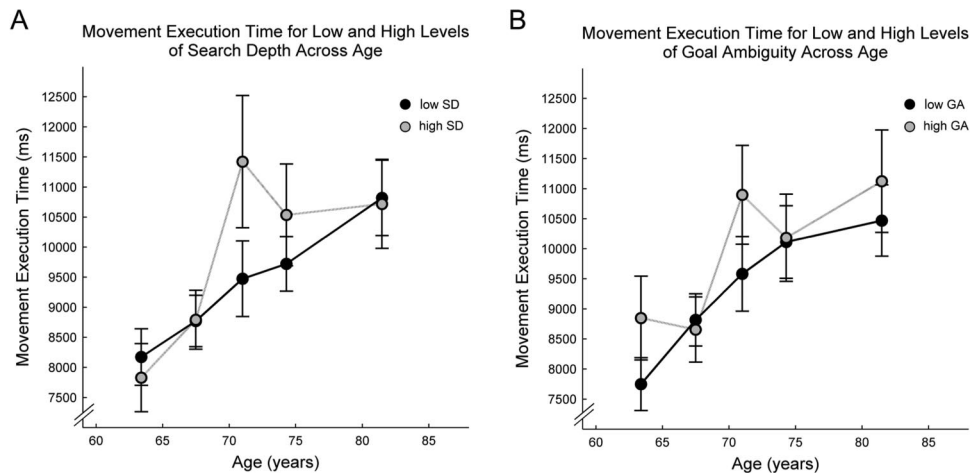
Figure 5. Search-depth- and goal-ambiguity-related effects on movement execution time (MET; time from movement of the first ball until problem solution) across age. (A) MET for problems with low levels of search depth (low SD; black markers) and high levels of search depth (high SD; gray markers) across age. (B) MET for problems with low levels of goal ambiguity (low GA; black markers) and high levels of goal ambiguity (high GA; gray markers) across age. Error bars indicate standard error of mean. Age groups are marked at their group-specific mean age.

deficit in dealing with increased depth, but not breadth, of searching ahead in healthy older adults.

General Age-Related Effects on Planning Performance

Regarding nonspecific age-related differences in planning performance, the significant decline in overall accuracy with advancing age found here is in accordance with former studies that also found age-related effects on the number of TOL problems solved (De Luca et al., 2003; Phillips, Smith, & Gilhooly, 2002; Robbins et al., 1998; Zook et al., 2006; also see Allamanno et al., 1987; Andrés & van der Linden, 2000; Bugg et al., 2006). Yet, the present study is the first to demonstrate this age-related effect on planning accuracy in a sample restricted to older adults. Former studies yielded age-related differences by either comparing a group of older adults with a group of young adults (Andrés & van der Linden, 2000; Zook et al., 2006) or by testing a sample with a very wide age range spanning at least five decades (Allamanno et al., 1987; De Luca et al., 2003; Robbins et al., 1998). In the present study, however, a significant decrease in planning accuracy on the TOL was found despite a more limited age range from 60 to 89 years, which made it more difficult to discover age-related differences than in previous studies.

Given that an effect of age was not evident for the number of excessive moves and the efficiency of problem solutions (cf. the online supplementary materials), accuracy in terms of the percentage of perfect solutions within the minimum number of moves seems to be especially sensitive in detecting age-related differences in planning performance. Despite the time limit of 1 min per trial imposed here (cf. Shallice, 1982), age groups did not differ in the number of time out trials, so that results are not influenced by gross differences in processing speed, which is well known to deteriorate with age and to contribute to executive processes (Salthouse, 1996, 2000). Present results hence indicate that overall changes in planning abilities associated with normal aging are profound enough to be differentiated even between various stages of older age.

However, differing from previous studies (Andrés & van der Linden, 2000; Phillips, Smith, & Gilhooly, 2002; Robbins et al., 1998), here, increasing age was not associated with longer preplanning times. Given that the current study employed the original TOL, which imposes a time limit of 1 min per problem (Shallice, 1982), whereas previous studies did not (Andrés & van der Linden, 2000; Phillips et al., 2002; Robbins et al., 1998), significant differences in preplanning times across older age might only manifest when subjects are not restricted in the time spent on mentally trying to solve a problem.

By contrast, age-related effects emerged here in the time needed for execution of problem solutions. Although movement execution time is influenced by basic processing and motor speed (cf. Berg & Byrd, 2002), it also reflects concurrent planning, that is, planning online during problem solution that helps to reduce working memory load (Davies, 2003, 2005). Thus, the age-related increase in problem execution times but not in preplanning times might, besides a general slowing in speed of processing with advancing age, reflect a stronger engagement in concurrent planning with older age due to limited working memory capacities. Yet, any inferences on the cognitive processes underlying preplanning and movement execution times remain speculative without further investigation of, for instance, eye movement patterns during these phases (e.g., Hodgson, Bajwa, Owen, & Kennard, 2000; Kaller, Rahm, Bolkenius, & Unterrainer, 2009; Nitschke, Ruh, Kappler, Stahl, & Kaller, 2012).

Specific Age-Related Effects of the Depth of Search on Planning Performance

Most important, the present study is the first to demonstrate that the process of normal aging is specifically associated with an impaired ability to deal with increased demands on deeper anticipatory search during planning, while an increased breadth of search does not invoke age-related differences in planning performance. This selective search-depth-related deficit is consistently evident across several measures of planning performance (cf. the online supplementary materials) and further impacts on the time needed for execution and online planning of problem solution. Similarly, Gilhooly, Phillips, Wynn, Logie, and Della Sala (1999) recorded oral protocols of younger and older subjects engaged in solving the TOL and showed that older subjects' mental search demonstrated significantly reduced plan depths (i.e., reduced length of considered move sequences), while no differences emerged in the breadth of search. However, no age-related differences in planning accuracy were found (Gilhooly et al., 1999), whereas here it was possible to delineate search-depth-related differences in accuracy even between several stages of older age.

Present results are also in line with recent evidence indicating that planning impairments related to the breadth of search are specifically associated with PD pathology (McKinlay et al., 2008), while impairments induced by increased depth of search are associated with lower dementia screening ratings in both PD patients and controls (Köstering et al., 2012). Hence, a specifically impaired depth of search during planning possibly marks subclinical antecedents of the conversion from normal to dementing courses of aging (cf. Petersen, 2007). Current findings furthermore accord well with studies on the development of planning processes in children that showed the ability to generate intermediate moves and to deal with greater search depth of problems to emerge during the first decade of life (Kaller et al., 2008; Spitz et al., 1982). It is important to note that no age-related effect for goal hierarchy was found here, which mirrors planning performance of preschool children who do not differ in their ability to deal with goal hierarchies of varying ambiguity (Kaller et al., 2008). Thus, complementary to the search-depth-specific development in planning abilities during childhood, results of the present sample indicate a selective deterioration of in-depth, but not in-breadth, search processes during planning over the course of normal aging.

These findings raise the question of which basic cognitive processes account for this specific age-related deficit. As higher search depths require a goal move to be held in mind during in-depth search for the intermediate moves that will allow it to be accomplished, demands on maintenance processes in working memory are also increased (Spitz et al., 1982; see also Gilhooly et al., 1999). Older adults' performance on disc-transfer tasks is known to draw heavily on working memory functions (Gilhooly et al., 2002; Hull, Martin, Beier, Lane, & Hamilton, 2008; Phillips, Gilhooly, Logie, Della Sala, & Wynn, 2003), and advancing age is well-known to reduce working memory capacity (e.g., Bopp &

Verhaeghen, 2005; Norman, Kemper, & Kynette, 1992; Zeintl & Kliegel, 2007). Thus, it is conceivable that age-related differences in working memory capacity specifically contribute to differences in dealing with greater search depth of problems (cf. Gilhooly et al., 1999).

By contrast, ambiguous goal hierarchies do not tax the stability of information maintained in working memory equally strongly, but rather demand a flexible processing of information so as to deliberate several alternative move sequences. In close relation to this fact, it recently has been shown that age-related decrements in mental set shifting do not result from the switching between different tasks itself. Rather, aging impedes maintaining the competing tasks sets in working memory (Kray & Lindenberger, 2000; Reimers & Maylor, 2005; Wasylyshyn, Verhaeghen, & Sliwinski, 2011), which is suggested to tap into cognitive stability (Cools, 2006). Generating alternative goal-move sequences and selecting from among them the most appropriate one is more akin to the actual set shifting, demanding cognitive flexibility, than to the stable maintenance of tasks sets. Therefore, the lack of age-related effects on dealing with ambiguous goal hierarchies concurs well with these recent findings on task-switching in older adults (Kray & Lindenberger, 2000; Reimers & Maylor, 2005; Wasylyshyn et al., 2011).

The assumption that working memory maintenance processes are particularly taxed by greater search depths is further supported by a search-depth-specific age-related effect of similar pattern in movement execution times, but not preplanning times. That is, whereas greater search depths had a comparable effect on initial thinking times across all age groups, it affected movement execution times specifically at an intermediate-old age. This might indicate a corresponding shift from initial to concurrent planning due to increased working memory demands induced by greater search depth of problems (cf. Davies, 2003, 2005). However, one should bear in mind that this explanation cannot account for the lack of a search depth effect in the oldest group, which should have the lowest working memory capacity. Thus, age-related differences in working memory processes can clearly not be the sole contributor to the search-depth-specific deficit found here.

In this regard, fluid intelligence has been previously found to be the best predictor of TOL performance of older adults, even exceeding the predictive power of chronological age (Zook et al., 2006), which concurs with its contribution to TOL performance of younger adults (Unterrainer et al., 2004; Zook, Davalos, DeLosh, & Davis, 2004). Older adults are also specifically impaired at relational integration in fluid reasoning tasks that demand that information from several relations or dimensions be integrated for problem solution (Viskontas, Holyoak, & Knowlton, 2005; Viskontas, Morrison, Holyoak, Hummel, & Knowlton, 2004). Similarly, TOL problems with greater search depths require increased integration of information. That is, for greater search depths, several interdependencies between different moves have to be identified and integrated so as to avoid suboptimal intermediate moves leading to obstructions of goal moves further along the solution path (cf. Figure 1). Thus, differential decrements in fluid abilities and associated relational integration might be another cognitive process contributing to age-related differences in search-depth-specific planning ability. Given that fluid abilities critically rely on working memory capacity for maintenance of relational information (Viskontas et al., 2004, 2005), these two processes

have possibly interacted with each other in contributing to the nonlinear trajectory of search-depth deficits across older age.

That is, at a young-old age, working memory capacity and fluid abilities are presumably relatively spared and allow for compensation of increased demands on the depth of search processes at least in terms of solution accuracy and movement execution times, solely affecting preplanning times. Conversely, at an advanced old age from the end of the eighth decade on, for which planning accuracy was lowest overall, general age-related deterioration of planning abilities and working memory capacity are present to a much greater extent. Thus, maintenance of relational information on move interdependencies might be generally impaired and insensitive to further specific demands on integration of relational information imposed by higher search depths, so that accuracy of problem solutions as well as movement execution times are not affected any further (cf. Figure 3A).

By contrast, at an intermediate-old age, working memory capacity and planning abilities are not as extensively impaired as in advanced-old age. Therefore, under low demands on in-depth look-ahead processes, planning accuracy is maintained at age-appropriate levels. Increased demands on the depth of search processes and relational integration of relevant information cannot be compensated, however, due to age-related deterioration of underlying cognitive processes. Thus, high search-depth levels unfold to have an especially detrimental impact on performance.

Yet, from the evidence presented here, it cannot be concluded if working memory capacity and fluid abilities are indeed underlying the search-depth effect. Age-related slowing of processing speed might also have differentially contributed to planning performance in subjects of young-, intermediate-, and advanced-old age. Moreover, it remains to be validated whether there really is no search-depth-specific effect on planning accuracy in adults ages 76 and older. Here, this pattern was present across several measures of planning performance, and it was not attributable to a lack of search-depth-related variance per se. Still, it cannot be ruled out that only uncommonly healthy advanced-old adults were inadvertently included in the study, which required sustained attention for neuropsychological testing, thereby obscuring a search-depth effect in the oldest age group. Clearly, further investigations of planning abilities and underlying basic cognitive processes as well as of their developmental trajectories with increasing age are warranted so as to unravel the uncommon nonlinear pattern of specific age-related changes found here.

Conclusion

The results of the present study considerably extend the understanding of the development of planning abilities in normal aging. It was demonstrated that normal aging is associated with deficits in the depth, but not the breadth, of mental search processes during planning. That is, cognitive processes related to the depth of planning ahead and to maintenance and integration of interdependent information are particularly affected by different stages of normal aging. In contrast, flexible processing of planning-relevant information seems to be age-insensitive. Thus, planning abilities, which are prototypical of cognitive functions reliant on the prefrontal cortex of the brain, are not only sensitive to increasing age in general, as posited by the frontal hypothesis of aging, but also to specific cognitive demands. Future studies with healthy older

adults should attempt to delineate the basic cognitive processes and neural underpinnings of these differential deficits in specific cognitive demands during planning so as to provide full understanding of their development in normal aging.

References

Allamanno, N., Della Sala, S., Laiacona, M., Pasetti, C., & Spinnler, H. (1987). Problem solving ability in aging and dementia: Normative data on a non-verbal test. *Italian Journal of Neurological Sciences, 8,* 111–119. doi:10.1007/BF02337583

Andrés, P., & van der Linden, M. (2000). Age-related differences in supervisory attentional system functions. *Journals of Gerontology, Series B Psychological Sciences and Social Sciences, 55,* P373–P380. doi:10.1093/geronb/55.6.P373

Beck, A. T., Steer, R. A., & Brown, G. K. (1995). *Manual for the Beck Depression Inventory–II manual*. San Antonio, TX: Psychological Corporation.

Berg, W. K., & Byrd, D. L. (2002). The Tower of London spatial problem-solving task: Enhancing clinical and research implementation. *Journal of Clinical and Experimental Neuropsychology, 24,* 586–604. doi:10.1076/jcen.24.5.586.1006

Berg, W. K., Byrd, D. L., McNamara, J. P. H., & Case, K. (2010). Deconstructing the tower: Parameters and predictors of problem difficulty on the Tower of London task. *Brain and Cognition, 72,* 472–482. doi:10.1016/j.bandc.2010.01.002

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 60,* P223–P233. doi:10.1093/geronb/60.5.P223

Bugg, J. M., Zook, N. A., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2006). Age differences in fluid intelligence: Contributions of general slowing and frontal decline. *Brain and Cognition, 62,* 9–16. doi:10.1016/j.bandc.2006.02.006

Carder, H. P., Handley, S. J., & Perfect, T. J. (2004). Deconstructing the Tower of London, England: Alternative moves and conflict resolution as predictors of task performance. *The Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology, 57,* 1459–1483. doi:10.1080/02724980343000864

Chandler, M. J., Lacritz, L. H., Hynan, L. S., Barnard, H. D., Allen, G., Deschner, M., . . . Cullum, C. M. (2005). A total score for the CERAD neuropsychological battery. *Neurology, 65,* 102–106. doi:10.1212/01.wnl.0000167607.63000.38

Cools, R. (2006). Dopaminergic modulation of cognitive function - Implications for L-DOPA treatment in Parkinson's disease. *Neuroscience and Biobehavioral Reviews, 30,* 1–23. doi:10.1016/j.neubiorev.2005.03.024

Craik, F. I. M., & Salthouse, T. A. (Eds.). (2000). *The handbook of aging and cognition*. Mahwah, NJ: Erlbaum.

Davies, S. P. (2003). Initial and concurrent planning in solutions to well-structured problems. *The Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology, 56,* 1147–1164.

Davies, S. P. (2005). Planning and problem solving in well-defined situations. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 35–51). Hove, England: Psychology Press.

De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J. A., Proffitt, T. M., Mahony, K., & Pantelis, C. (2003). Normative data from the Cantab: I: Development of executive function over the lifespan. *Journal of Clinical and Experimental Neuropsychology, 25,* 242–254. doi:10.1076/jcen.25.2.242.13639

Ehrensperger, M. M., Berres, M., Taylor, K. I., & Monsch, A. U. (2010). Early detection of Alzheimer's disease with a total score of the German CERAD. *Journal of the International Neuropsychological Society, 16,* 910–920. doi:10.1017/S1355617710000822

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198. doi:10.1016/0022-3956(75)90026-6

Gilhooly, K. J., Phillips, L. H., Wynn, V., Logie, R. H., & Della Sala, S. (1999). Planning processes and age in the five-disc Tower of London task. *Thinking & Reasoning, 5,* 339–361. doi:10.1080/135467899393977

Gilhooly, K. J., Wynn, V., Phillips, L. H., Logie, R. H., & Della Sala, S. (2002). Visuo-spatial and verbal working memory in the five-disc Tower of London task: An individual differences approach. *Thinking & Reasoning, 8,* 165–178. doi:10.1080/13546780244000006

Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: A view from cognitive neuroscience. *Nature Reviews Neuroscience, 5,* 87–96. doi:10.1038/nrn1323

Hodgson, T. L., Bajwa, A., Owen, A. M., & Kennard, C. (2000). The strategic control of gaze direction in the Tower of London task. *Journal of Cognitive Neuroscience, 12,* 894–907.

Hull, R., Martin, R. C., Beier, M. E., Lane, D., & Hamilton, A. C. (2008). Executive function in older adults: A structural equation modeling approach. *Neuropsychology, 22,* 508–522. doi:10.1037/0894-4105.22.4.508

Kaller, C. P., Rahm, B., Bolkenius, K., & Unterrainer, J. M. (2009). Eye movements and visuospatial problem solving: Identifying separable phases of complex cognition. *Psychophysiology, 46,* 818–830. doi:10.1111/j.1469-8986.2009.00821.x

Kaller, C. P., Rahm, B., Köstering, L., & Unterrainer, J. M. (2011). Reviewing the impact of problem structure on planning: A software tool for analyzing tower tasks. *Behavioural Brain Research, 216,* 1–8. doi:10.1016/j.bbr.2010.07.029

Kaller, C. P., Rahm, B., Spreer, J., Mader, I., & Unterrainer, J. M. (2008). Thinking around the corner: The development of planning abilities. *Brain and Cognition, 67,* 360–370. doi:10.1016/j.bandc.2008.02.003

Kaller, C. P., Unterrainer, J. M., Rahm, B., & Halsband, U. (2004). The impact of problem structure on planning: Insights from the Tower of London task. *Cognitive Brain Research, 20,* 462–472. doi:10.1016/j.cogbrainres.2004.04.002

Kaller, C. P., Unterrainer, J. M., & Stahl, C. (2012). Assessing planning ability with the Tower of London task: Psychometric properties of a structurally balanced problem set. *Psychological Assessment, 24,* 46–53. doi:10.1037/a0025174

Klahr, D., & Robinson, M. (1981). Formal assessment of problem-solving and planning processes in preschool children. *Cognitive Psychology, 13,* 113–148. doi:10.1016/0010-0285(81)90006-2

Köstering, L., McKinlay, A., Stahl, C., & Kaller, C. P. (2012). Differential patterns of planning impairments in Parkinson's disease and sub-clinical signs of dementia? A latent-class model-based approach. *Plos ONE, 7,* e38855. doi:10.1371/journal.pone.0038855

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging, 15,* 126–147. doi:10.1037/0882-7974.15.1.126

Lehrl, S., Merz, J., Burkhard, G., & Fischer, S. (1991). *MWT-A: Mehrfachwahl-Wortschatz-Intelligenztest* [Multiple Choice Vocabulary IQ Test]. Göttingen, Germany: Hogrefe.

McKinlay, A., Kaller, C. P., Grace, R. C., Dalrymple-Alford, J. C., Anderson, T. J., Fink, J., & Roger, D. (2008). Planning in Parkinson's disease: A matter of problem structure? *Neuropsychologia, 46,* 384–389. doi:10.1016/j.neuropsychologia.2007.08.018

Memory Clinic Basel. (1997). *CERAD Neuropsychologische Testbatterie (CERAD-NP) [CERAD Neuropsychological Assessment Battery (CERAD-NAB)]*. Basel, Switzerland: Basel University Hospital.

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., Vanbelle, G., Fillenbaum, G., . . . Clark, C. (1998). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD): Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology, 39,* 1159–1165.

Moscovitch, M., & Winocur, G. (1995). Frontal lobes, memory, and aging. In J. Grafman, K. J. Holyoak, & F. Boller (Eds.), *Annals of the New York Academy of Sciences: Vol. 769. Structure and functions of the human*

prefrontal cortex (pp. 119–150). New York, NY: New York Academy of Sciences.

Newman, S. D., & Pittman, G. (2007). The Tower of London: A study of the effect of problem structure on planning. *Journal of Clinical and Experimental Neuropsychology, 29,* 333–342. doi:10.1080/13803390701249051

Nitschke, K., Ruh, N., Kappler, S., Stahl, C., & Kaller, C. P. (2012). Dissociable stages of problem solving (I): Temporal characteristics revealed by eye-movement analyses. *Brain and Cognition, 80,* 160–169. doi:10.1016/j.bandc.2012.05.003

Norman, S., Kemper, S., & Kynette, D. (1992). Adults' reading comprehension: Effects of syntactic complexity and working memory. *Journal of Gerontology, 47,* P258–P265.

Owen, A. M., Downes, J. J., Sahakian, B. J., & Polkey, C. E. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia, 28,* 1021–1034. doi:10.1016/0028-3932(90)90137-D

Petersen, R. C. (2007). Mild cognitive impairment: Current research and clinical implications. *Seminars in Neurology, 27,* 22–31. doi:10.1055/s-2006-956752

Phillips, L. H., Gilhooly, K. J., Logie, R. H., Della Sala, S., & Wynn, V. E. (2003). Age, working memory, and the Tower of London task. *European Journal of Cognitive Psychology, 15,* 291–312. doi:10.1080/09541440244000148

Phillips, L. H., MacLeod, M., & Kliegel, M. (2005). Adult aging and cognitive planning. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 111–139). Hove, England: Psychology Press.

Phillips, L. H., MacPherson, S., & Della Sala, S. (2002). Age, cognition and emotion: The role of anatomical segregation in the frontal lobes. In J. Grafman (Ed.), *Handbook of neuropsychology* (2nd ed., pp. 73–97). Amsterdam, the Netherlands: Elsevier.

Phillips, L. H., Smith, L., & Gilhooly, K. J. (2002). The effects of adult aging and induced positive and negative mood on planning. *Emotion, 2,* 263–272. doi:10.1037/1528-3542.2.3.263

Rainville, C., Lepage, E., Gauthier, S., Kergoat, M-J., & Belleville, S. (2012). Executive function deficits in persons with mild cognitive impairment: A study with a Tower of London task. *Journal of Clinical and Experimental Neuropsychology, 34,* 306–324. doi:10.1080/13803395.2011.639298

Raz, N., & Kennedy, K. M. (2009). A systems approach to the aging brain: Neuroanatomic changes, their modifiers, and cognitive correlates. In W. Jagust & M. D'Esposito (Eds.), *Imaging the aging brain* (pp. 43–70). New York, NY: Oxford University Press. doi:10.1093/acprof:oso/9780195328875.003.0004

Reimers, S., & Maylor, E. A. (2005). Task switching across the life span: Effects of age on general and specific switch costs. *Developmental Psychology, 41,* 661–671. doi:10.1037/0012-1649.41.4.661

Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., Lawrence, A. D., McInnes, L., & Rabbitt, P. M. A. (1998). A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: Implications for theories of executive functioning and cognitive aging. *Journal of the International Neuropsychological Society, 4,* 474–490. doi:10.1017/S1355617798455073

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428. doi:10.1037/0033-295X.103.3.403

Salthouse, T. A. (2000). Aging and measures of processing speed. *Biological Psychology, 543,* 5–54. doi:10.1016/S0301-0511(00)00052-1

Seo, E. H., Lee, D. Y., Lee, J. H., Choo, I. H., Kim, J. W., Kim, S. G., . . . Woo, J. I. (2010). Total scores of the CERAD neuropsychological assessment battery: Validation for mild cognitive impairment and dementia patients with diverse etiologies. *The American Journal of Geriatric Psychiatry, 18,* 801–809. doi:10.1097/JGP.0b013e3181cab764

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 298,* 199–209. doi:10.1098/rstb.1982.0082

Sheikh, J. L., & Yesavage, J. A. (1986). Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. *Clinical Gerontology, 5,* 165–173.

Spitz, H. H., Webster, N. A., & Borys, S. V. (1982). Further studies of the Tower of Hanoi problem-solving performance of retarded young adults and nonretarded children. *Developmental Psychology, 18,* 922–930. doi:10.1037/0012-1649.18.6.922

Sullivan, J. R., Riccio, C. A., & Castillo, C. L. (2009). Concurrent validity of the tower tasks as measures of executive function in adults: A meta-analysis. *Applied Neuropsychology, 16,* 62–75. doi:10.1080/09084280802644243

Unterrainer, J. M., & Owen, A. M. (2006). Planning and problem solving: From neuropsychology to functional neuroimaging. *Journal of Physiology–Paris, 993,* 8–317. doi:10.1016/j.jphysparis.2006.03.014

Unterrainer, J. M., Rahm, B., Kaller, C. P., Leonhart, R., Quiske, K., Hoppe-Seyler, K., . . . Halsband, U. (2004). Planning abilities and the Tower of London: Is this task measuring a discrete cognitive function? *Journal of Clinical and Experimental Neuropsychology, 26,* 846–856. doi:10.1080/13803390490509574

Viskontas, I. V., Holyoak, K. J., & Knowlton, B. J. (2005). Relational integration in older adults. *Thinking & Reasoning, 11,* 390–410. doi:10.1080/13546780542000014

Viskontas, I. V., Morrison, R. G., Holyoak, K. J., Hummel, J. E., & Knowlton, B. J. (2004). Relational integration, inhibition and analogical reasoning in older adults. *Psychology and Aging, 19,* 581–591. doi:10.1037/0882-7974.19.4.581

Ward, G., & Allport, A. (1997). Planning and problem-solving using the five-disc Tower of London task. *Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology, 50,* 49–78. doi:10.1080/027249897392224

Ward, G., & Morris, R. (2005). Introduction to the psychology of planning. In R. Morris & G. Ward (Eds.), *The cognitive psychology of planning* (pp. 1–34). Hove, England: Psychology Press.

Wasylyshyn, C., Verhaeghen, P., & Sliwinski, M. J. (2011). Aging and task switching: A meta-analysis. *Psychology and Aging, 26,* 15–20. doi:10.1037/a0020912

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin, 120,* 272–292. doi:10.1037/0033-2909.120.2.272

Zeintl, M., & Kliegel, M. (2007). How do verbal distractors influence age-related operation span performance? A manipulation of inhibitory control demands. *Experimental Aging Research, 33,* 163–175. doi:10.1080/03610730701192815

Zook, N. A., Davalos, D. B., DeLosh, E. L., & Davis, H. P. (2004). Working memory, inhibition, and fluid intelligence as predictors of performance on Tower of Hanoi and London tasks. *Brain and Cognition, 56,* 286–292. doi:10.1016/j.bandc.2004.07.003

Zook, N. A., Welsh, M. C., & Ewing, V. (2006). Performance of healthy, older adults on the Tower of London Revised: Associations with verbal and nonverbal abilities. *Aging, Neuropsychology, and Cognition, 13,* 1–19. doi:10.1080/13825580490904183