Published in final edited form as:

Neuropsychol Rehabil. 2017 July ; 27(5): 759–776. doi:10.1080/09602011.2015.1107599.

Examining the Impact of Formal Planning on Performance in Older Adults using a Naturalistic Task Paradigm

Chad Sanders and **Maureen Schmitter-Edgecombe**

Department of Psychology, Washington State University, Pullman, WA, USA

Abstract

Thirty-two younger adults (YAs) and 64 older adults (OAs) completed the Amap Task, a naturalistic measure designed to evaluate both the formulation and execution stages of planning. Participants read a map layout of a university apartment and planned an efficient strategy to complete several tasks. To identify whether execution abilities are augmented by formulating a plan, we compared formal (i.e., write out a strategy) and informal (i.e., initiate task when ready without writing out a plan) planning conditions. Participants did not have access to their plan during task execution. Amap scores were also compared with informant-report of everyday functioning. For the formal planning condition, OAs were generally less accurate and less efficient than YAs during the formulation stage, while there were no group differences in total formulation time for the informal planning condition. Across conditions, OAs obtained poorer execution accuracy and efficiency scores, took significantly longer to execute the Amap, and were more likely to omit task parts. Both groups improved from formulation to execution, regardless of formulation performance. Formal planning led to significant improvement in execution efficiency but not accuracy, suggesting that one's approach to the task (i.e., efficiency, finesse) may significantly improve with preplanning even if overall accuracy does not. Furthermore, significant relationships between the Amap formulation and execution scores and informant-report of everyday functioning highlight the ecological validity of the Amap as a measure sensitive to everyday planning abilities.

Keywords

Amap Task; executive functions; aging; formulation; execution

Introduction

Planning ability is an important part of executive functioning, which represents an individual's ability to carry out complex, higher level problems. A variety of day-to-day tasks require individuals to utilize planning abilities, including such activities as preparing for an overnight stay at a conference (e.g., packing toiletries, organizing transport from the airport, packing appropriate clothing) or simply preparing to go out for the evening (e.g., how much money to bring, what to wear). In both instances, an individual must successfully

identify and plan the necessary steps to achieve their goal. Studies have revealed that planning ability declines with age, though findings are less consistent with regards to naturalistic settings where older adults might plausibly be able to compensate for such a decline (see Garden, Phillips, & MacPherson, 2001; Kliegel, Martin, McDaniel, & Phillips, 2007; Jurado & Rosselli, 2007; Allain et al., 2005; Sanders & Schmitter-Edgecombe, 2012). In this study, we evaluate the role of plan formulation in the planning performance of cognitively healthy younger adults (YAs) and older adults (OAs) using an open-ended, naturalistic task in a real world environment.

While there have been numerous standardized neuropsychological instruments designed to assess planning abilities, few clearly separate planning formulation from execution. Theoretically, planning ability has been broken down into a formulation stage and a subsequent execution stage (Grafman, 1989; Shallice, 1982; see also Allain et al., 2005; Sanders & Schmitter-Edgecombe, 2012). The formulation stage involves cognitively designing a plan to solve a problem and incorporating multiple steps that will lead to a goal. The execution stage involves guiding the execution of the plan to a successful conclusion (Allain et al., 2005). Certain neuropsychological tests non-explicitly allow for a pre-execution phase during which it is assumed that participants make use of formulation abilities before beginning the task, such as the Tower of London (TOL) task and the Zoo Map task from the Behavioral Assessment of the Dysexecutive Syndrome test battery (BADS; Wilson, Alderman, Burgess, Emslie, & Evans, 1996). In these cases, the formulation stage of planning is usually represented only by the amount of time that passes before the participant initiates the execution of the task; typically, there is no overt formulation phase built into these tasks. By not explicitly providing a preliminary stage during which participants utilize formulation abilities as they would in daily life, individuals may be forced to rely solely on execution abilities (e.g., online planning, set-shifting), which may hamper overall performance if both formulation and execution abilities are required for successful planning.

Using the TOL task, Phillips and colleagues (1999) provided direct support for the idea that execution abilities are not reliant upon formulation abilities to be successful. The authors implemented distractor tasks during the period of time before participants initiated task execution, which were shown to reduce time spent planning during this supposed formulation phase but had little to no effect on execution performance. In a follow up study with the TOL, Phillips, Wynn, McPherson, and Gilhooly (2001) examined the effect of a plan condition (individuals were given specific instructions to form a mental plan) and a 5-second condition on individuals' ability to mentally carry out the task. Results showed that individuals were more likely to mentally carry out a sequence of moves accurately in the plan versus the 5-second condition. However, there was no effect of the preplanning condition on execution performance or total time when actually carrying out the TOL trials. As there was no formal assessment method to specifically capture what participants were doing during the preplanning phase across these studies, there is no definitive evidence as to the nature of the relationship between the formulation and execution stages of planning.

To separately evaluate the formulation and execution components of planning, in a prior study in our laboratory (Sanders & Schmitter-Edgecombe, 2012), YAs and OAs completed a

naturalistic task called the Amap Task (formerly “the map task”), modeled after the Multiple Errands Test (Alderman, Burgess, Knight & Henman, 2003; Shallice & Burgess, 1991). For the Amap, individuals carried out several tasks in a campus apartment as part of an open-ended task-paradigm that utilizes real-world stimuli. For example, the Amap required participants to retrieve and place several objects at new locations and to answer questions throughout the task (e.g., How many plants are located in the apartment? Is there a broom located in the supply closet?). To be successful, individuals needed to accurately and efficiently interweave the tasks so as to avoid revisiting locations more than once. There were two phases to the Amap Task, a formulation phase and an execution phase. Participants were instructed to write out a plan for completing the task during the formulation phase and subsequently completed the task during the execution phase. We introduced an experimental manipulation by allowing only half of the participants to utilize their written plans during the execution phase in order to identify the potential benefit of utilizing an external aid derived from formulation abilities. Of note, we prioritized ecological validity while designing the Amap as a majority of tests of planning utilize only pen and paper, which may not actually tap the same planning abilities used in daily life.

Overall, we found that OAs performed significantly more poorly than YAs on measures of both formulation and execution accuracy and efficiency. The discrepancies in execution abilities between age groups remained even after taking formulation performance into consideration. We also found that both age groups improved from formulation to execution (i.e., execution scores were better than formulation scores for both groups). Interestingly, having access to their formulated plan during execution had no significant effect on execution accuracy or efficiency for either the YA or OA group. On the one hand, this finding seemed to suggest that the formulation component of planning may have been unrelated to the execution component during task completion, consistent with previously discussed findings from the literature (e.g., Phillips, Wynn, Gilhooly, Della Sala, and Logie, 1999; Phillips, Wynn, McPherson, & Gilhooly, 2001). For example, it is possible that participants’ on-line planning ability may have suggested a different way to carry out the task during execution. It is also possible that the execution phase built upon the formulation phase in some manner, as opposed to being an independent attempt at the same task, resulting in improved performance. Because all participants constructed a plan during the formulation stage, the benefits of planning ahead may have occurred regardless of whether or not the individuals had access to the written plan during task execution.

When evaluating planning ability, an important goal should be to separate and individually assess formulation and execution abilities. However, this becomes increasingly difficult when using stimuli with overly structured instructions for task completion that may suppress some of the abilities that make up execution (e.g., online adjusting, prospective memory). For example, the second phase of the Zoo Map test, theorized to represent plan execution, simply requires participants to complete items one by one on a list in sequential order. This limits the need for more complex plan execution abilities (e.g., online planning, set-shifting). Furthermore, common practice across clinical and research settings has been to focus primarily on the execution stage of planning, with little emphasis on implementing and assessing an overt formulation phase prior to task completion. If execution abilities theoretically build from and expand upon formulation abilities, then this administration

format is problematic. Alternatively, if planning ability predominantly relies on aspects of the execution component and less so on the formulation component, then this mode of assessment may be adequate. These points remain to be clarified in the literature.

Taking these points into consideration, the primary goal of this study was to identify whether having an explicit formulation phase impacts execution performance. Using the previously described Amap Task (Sanders & Schmitter-Edgecombe, 2012), we compared two different administration formats. One format reflected the traditional administration format of laboratory-based planning tasks where there is no overt formulation phase and participants are typically instructed to begin whenever they are ready (i.e., *informal planning* condition). The other administration format required participants to formally plan and write out a strategy prior to task execution (i.e., *formal planning* condition). For both conditions, participants did not have access to their formulated plan during execution in order to directly compare the effect of preplanning on execution abilities.

By investigating the impact of formally planning a strategy prior to execution and comparing it to the more traditional informal planning phase, we can better clarify the nature of the relationship between formulation and execution abilities in planning performance. We were especially interested in the effect of constructing a formal plan prior to task execution and examining whether on-line execution abilities would be augmented in this situation. Based on the theory that execution abilities build upon formulation abilities and jointly constitute planning ability, we hypothesized that individuals in the *formal planning* condition would perform better on the Amap as a result of having an adequate opportunity to utilize formulation abilities before transitioning into execution. Because individuals in the *informal planning* condition were not instructed to formulate a plan, we believed they would be more likely to initiate execution before sufficiently formulating a plan and, thus, not perform as well as individuals in the formal planning condition. However, should execution abilities be more independent and unrelated to formulation abilities, we expect to find no significant differences in execution performance between formulation conditions. Based on our previous study (Sanders & Schmitter-Edgecombe, 2012), we expected OAs to perform more poorly than YAs during both formulation and execution. We further hypothesized that the OAs' Amap performances would benefit more from formal preplanning than the YAs, highlighting a relationship between aging and an increased need to adequately plan ahead.

As a secondary goal, we were interested in comparing Amap outcome measures with informant-report data from the Instrumental Activities of Daily Living: Compensation Scale (IADL-C; Schmitter-Edgecombe, Parsey, & Lamb, 2014), a measure developed to capture early functional difficulties and compensatory strategy use in aging populations. In prior studies, the Amap was found to be related to several standardized neuropsychological tests of executive functioning (i.e., part B from the Trail Making Test [TMT; Reitan & Wolfson, 1985], the Controlled Oral Word Association Test [COWAT; Benton & Hamsher, 1976], the Zoo Map task from the Behavioral Assessment of the Dysexecutive Syndrome test Battery [BADS; Wilson, Alderman, Burgess, Emslie, & Evans, 1996]) as well as the Dysexecutive Questionnaire (DEX) from the BADS battery, supporting the content and convergent validity of the Amap Task (Sanders & Schmitter-Edgecombe, 2012; Sanders, Low, and Schmitter-Edgecombe, in press). To provide support for the ecological validity of the Amap for

capturing planning abilities used in everyday life, we compared IADL-C scores with the Amap.

Methods

Participants

Participants were 64 OAs (42 females, 22 males) ranging in age from 61 to 88 ($M = 71.3$) and 32 YAs (23 females, 8 males; 1 YA case with incomplete demographic information) ranging in age from 18 to 28 ($M = 20.5$). There were no differences in the sex distribution between the OA and YA groups, $\chi^2(1, 95) = .710, p = .400$. Although OAs were more highly educated than YAs, $t(93) = -4.68, p = .000$, the Wechsler Test of Adult Reading (WTAR; Holdnack, 2001) scores did not significantly differ between the older and younger adult groups, $t(77) = -1.89, p = .063$. In addition, none of the Amap Task measures correlated with years of education for either the YA, r s between $-.233$ and $.178$, or OA, r s between $-.089$ and $.189$, groups. Therefore, education was not used as a covariate in the analyses. As seen in Table 1, there were also no significant within-group differences across the formal and informal planning conditions in age and education (YA: t s $< 0.128, p$ s $> .899$; OA: t s $< .801, p$ s $> .426$), or sex distribution [YA: $\chi^2(1, N = 31) = .512, p = .474$; OA: $\chi^2(1, N = 64) = .277, p = .599$].

Older adults were recruited as volunteers responding to newspaper advertisements, brochures distributed in a variety of medical offices, and WSU campus email-newsletters. Younger adults volunteered to participate via a student subject web pool to fulfill class requirements and/or obtain extra credit. The data was collected as part of a larger study investigating the relationship between memory deficits and everyday functioning. Participants completed a battery of standardized and experimental neuropsychological tests in a laboratory setting, followed approximately one-week later by a three-hour battery of testing in a WSU campus apartment. Participants also consented to having study personnel contact a knowledgeable informant via telephone to complete the IADL-C. In exchange for volunteering for the study, OAs were provided with feedback about their cognitive performance in relation to their age peers. Participants were also given travel compensation as well as parking passes to facilitate travel to and parking at WSU for both the laboratory and smart home testing sessions.

All participants were required to be native speakers of English. For the OAs, individuals with a history of significant traumatic brain injury, brain surgery, tumors, stroke, Parkinson's disease, or other medical conditions that may have impacted cognitive functioning were excluded from this study. Screening measures included a telephone-administered clinical interview and the Telephone Interview of Cognitive Status (TICS; Brandt & Folstein, 2003; Brandt, Spencer, & Folstein, M., 1988; Ferrucci et al., 1998). For each participant, review of findings from initial clinical interview, neuropsychological testing, and collateral medical information (e.g., results of laboratory and brain imaging data when available) were used to confirm healthy OAs for inclusion in the final sample.

Materials and Procedures

Participants completed the Amap Task (see Figure 1) in a campus apartment at WSU. The apartment contains a living room, dining room, kitchen, and 3 hallway closets on the ground floor. While there is a bathroom and 3 bedrooms upstairs, including one utilized as the "computer headquarters" of the administrators, the participants only used the ground floor of the apartment for the purposes of testing. The campus apartment contains *smart environment* technologies intended to help individuals maintain independence in their home (e.g., motion detectors, sensor plates); however, no smart environment technologies were incorporated into this study.

Amap Task—For the Amap Task, participants were given a map layout of the WSU smart apartment and a list of activities to be completed, questions to be answered and reported to the administrator, and locations to be visited. Participants were instructed to plan an efficient route that would allow them to complete the requested tasks. The examiner then reviewed the list of tasks with the participant, highlighting objects to be retrieved and placed, and instructing the participant to utilize the key at the top of the map to identify where the items are located. In addition, participants were instructed they would also need to answer several questions throughout the task by saying the answer aloud to the examiner as they accomplish each item. For the *formal planning* condition, participants were told that they had 5 minutes to write out a plan for task completion using the map layout. The participants' written plan was then removed before transitioning to the next phase; none of the participants had access to their written plans during task execution. Participants in the *informal planning* condition were simply instructed to begin whenever they were ready, and they did not have access to pen and paper. After the initial planning stage, all participants were given 10 minutes to execute the task. Participants were prompted for any questions before beginning the task and instructed that they would not be allowed to ask questions once the task began.

IADL-C Questionnaire—For the older adult group, informant-report data was collected for the IADL-C questionnaire. A total of 25 informants in the formal planning condition and 24 informants in the informal planning condition completed the interview. Sixty-nine percent of the informants reported living with or seeing the participant almost daily, while 12%, 4%, and 14% reported seeing the participant 3–5× per week, 1–2× per week, or less than 1× per week. No formal testing of informant cognitive status was conducted. The IADL-C consists of 27 questions pertaining to quality of engagement in several common activities of daily living. Items are distributed between four functional domain subscales, including money and self-management, home daily living, travel and event memory, and social skills. Informant responses to items are based on an 8-point Likert scale, rating levels of independent functioning (1–4) and need for assistance (5–7), with additional options for "unable to complete" or "does not need to complete activity" as well as "no basis for judgment." For tasks completed independently (ratings 1–4), informants report whether the individual uses an aid to carry out the task (e.g., independent, no aid; independent, uses aid to assist; independent but not as well as ever, uses aid to assist; independent but not as well as ever, no aid). An overall total IADL-C score was obtained, with lower scores indicating better ability to carry out everyday activities.

Scoring

Amap Task—As seen in Table 2, Amap scoring involved 19 different task parts and 7 location “hubs,” identical to the Amap Task scoring protocol used in our previous studies (Sanders & Schmitter-Edgecombe, 2012; Sanders, Low, and Schmitter-Edgecombe, 2014). For the *formal planning* condition, outcome measures for the formulation component of planning included formulation accuracy and formulation efficiency. A *formulation accuracy score* was obtained by tallying the number of task parts out of 19 that were incorporated into the participant's plan (e.g., pick up the medication dispenser, retrieve the T.V. remote). A *formulation efficiency score* was determined by the participant's ability to plan ahead to multitask and interweave task parts in a manner that would avoid revisiting locations to complete different tasks. Based on 7 different location “hubs” for task completion, each location the participant included in their formulation was evaluated in terms of one trip or multiple trips needed to carry out the tasks at that particular location (e.g., drop off the medication dispenser AND check for brown sugar in cupboard “A” in one trip). Results were interpreted based out of 7 possible planned hub visits. For the *informal planning* condition, the outcome measure for formulation was simply the total time that elapsed from the end of initial instructions to task initiation in order to parallel the more traditional method of assessing formulation in this manner.

The outcome measures for the execution component of planning included execution accuracy and execution efficiency. Execution accuracy was based on the participant's ability to complete activities, with each task part given a score ranging from 1 to 4. A rating of “1” indicated the task part was completed accurately on the first trip to that particular location, while a rating of “2” indicated the task part was completed accurately but required an additional trip. A rating of “3” was given for task parts that were incomplete due to an omission (e.g., failed to retrieve the dust pan, did not answer microwave question), and “4” indicated an incomplete task part due to an inaccuracy (e.g., incorrect answer given, mustard retrieved instead of relish). Scores of “1” and “2” were considered accurate, while scores of “3” and “4” were considered inaccurate. An overall *execution accuracy score* was obtained by tallying the number of task parts accurately completed (those scored with a “1” or “2”) out of 19. An *execution efficiency score* was determined by the participant's ability to multitask and visit locations only once. Each location “hub” was evaluated for number of times visited during task execution. Multiple visits, single visits with an incomplete task part, and single visits with both task parts completed were tallied individually, with totals interpreted based out of 7 possible single hub visits.

Supplementary Measures—As a supplementary measure for the Amap Task, for those participants in the *formal planning* condition, their formulated plan was broken down into sequenced task parts and directly compared to their subsequent executed sequence to obtain a *route execution adherence score*. This score represented the percentage of formulated task parts adhered to by sequence during the execution phase, providing an estimate of participants' ability to carry out their plan sequence as intended from formulation to execution. In addition, for those participants whose executed sequence was not 100% consistent with their formulated plan, an *adherence efficiency score* was obtained by subtracting the *formulation efficiency score* from the *execution efficiency score*, with a

negative total indicating poorer execution efficiency and a positive total indicating better execution efficiency. Additional supplemental measures for execution performance for both the formal and informal planning conditions included *total time* and *total searching behaviors*. Each instance of searching behavior was tallied regardless of whether the participant was continuously searching for the same item or not (e.g., opening four cupboards in search of the plastic cups was tallied as four searching behaviors).

Analysis

Scoring Reliability

Amap data was scored and double-scored for each participant. Based on total discrepancies in scoring between coders, interrater agreement was >99%, consistent with reliability data from previous studies (Sanders & Schmitter-Edgecombe, 2012; Sanders et al., 2014). This was as expected, given the nature of the Amap Task paradigm and the explicit instructions participants are given to write out their plan, which allows for relatively straightforward scoring with little to no ambiguity.

Statistical tests conducted

Several of the outcome variables were not normally distributed and required transformation. In such cases, SQRT or log (base 10) transformations were utilized to transform variables, based on degree of skewness and with reverse-scoring/reflection for negatively-skewed data. Variables transformed in this manner included Amap formulation and execution accuracy and efficiency, total time, and total searching behaviors, with results falling within normality thresholds (skew and kurtosis statistics $< \pm 1.5$). Transformed variables were then used in all analyses, while means and standard deviations from the raw data were reported to aid interpretation. T-tests were utilized to evaluate group differences in Amap formulation accuracy and efficiency scores for the *formal planning* condition, as well as route execution adherence scores and adherence efficiency scores. For the *informal planning* condition, t-tests were also used to compare the total formulation time between groups (total time elapsed from end of instructions to task initiation). Separate 2 (group: YA, OA) by 2 (condition; formal planning, informal planning) univariate analyses of variance (ANOVAs) were used to examine differences in participants' Amap execution accuracy and efficiency scores, searching behavior, and total time needed for Amap Task completion. Finally, Pearson correlation analyses were utilized to examine the relationship among Amap measures, and Spearman's correlation analyses to examine the relationship between Amap and IADL-C scores for the OA group.

Results

Means and standard deviations for each of the Amap measures for each group are reported in Table 3.

Amap Task Formulation

Formulation Accuracy—For individuals in the formal planning condition, formulation accuracy was measured as the total number of task parts out of 19 incorporated into the

participant's formulated plan. As can be seen in Table 3, t-tests revealed the OA group incorporated fewer task parts into their formulated plans than the YA group, $t(46) = 2.46$, $p < .05$, $d = 0.83$.

Formulation Efficiency—For the formal planning condition, the total number of location hubs incorporated into the participants' plan needing only one visit to complete all relevant task parts was utilized as the measure of formulation efficiency. There was a trend for the YAs to formulate a more efficient plan than the OAs, $t(46) = 1.74$, $p = .088$, $d = 0.50$.

Formulation Time (informal planning)—For the informal planning condition, the outcome measure for formulation was measured as total time (in seconds) that elapsed from the end of initial instructions to task initiation. There was no significant group difference between YAs and OAs in preplanning time, $t(31) = -.381$, $p = .706$, $d = -0.17$.

Amap Task Execution

Execution Accuracy—Total number of task parts completed out of 19 was utilized as the measure of execution accuracy. A group (YA, OA) by condition (formal planning, informal planning) univariate ANOVA revealed that the YAs accurately carried out more task parts than the OA group, $F(1, 92) = 11.8$, $MS = 71.3$, $p = .001$, $\eta_p^2 = .114$. There was no significant effect found for condition, $F(1, 92) = .022$, $MS = .130$, $p = .883$, $\eta_p^2 = .000$, or interaction, $F(1, 92) = .146$, $MS = .880$, $p = .703$, $\eta_p^2 = .002$. Of note, given the mean data for total number of accurately executed task parts for the younger adult group (see Table 3), our ability to determine an interaction effect may have been potentially suppressed by a ceiling effect.

Execution Efficiency—The total number of location hubs visited only once with all relevant task parts completed accurately was used as the measure of execution efficiency. A group (YA, OA) by condition (formal planning, informal planning) univariate ANOVA revealed that the YA group executed the task more efficiently than the OA group, $F(1, 92) = 5.67$, $MS = 14.1$, $p < .05$, $\eta_p^2 = .058$. There was also a significant main effect of condition, with more efficient completion of tasks by participants in the formal planning condition compared to the informal planning condition, $F(1, 92) = 7.55$, $MS = 18.8$, $p = .007$, $\eta_p^2 = .076$. There was no significant group by condition interaction, $F(1, 92) = 1.64$, $MS = 4.08$, $p = .203$, $\eta_p^2 = .018$.

Execution Accuracy and Efficiency (Controlling for Formulation

Performances)—For the formal planning condition, to determine whether the significant group differences found in execution accuracy and efficiency could be accounted for by the OAs poorer formulation performances, we conducted univariate ANOVAs with formulation accuracy and efficiency scores as covariates. Results revealed that group differences in execution scores persisted even after controlling for both formulation outcome measures, with poorer execution accuracy scores, $F(1, 44) = 6.78$, $MS = 5.64$, $p < .05$, $\eta_p^2 = .134$, and efficiency scores, $F(1, 44) = 4.85$, $MS = 1.39$, $p < .05$, $\eta_p^2 = .099$, for OAs in comparison to YAs. This analysis was not carried out for the informal planning condition as there was no difference between age groups in informal preplanning time.

Performance from Formulation to Execution—Paired sample t-tests were used to evaluate performances from formulation to execution. For the YA group, both execution accuracy, $t(15) = 19.1$, $p < .001$, $d = -1.64$, and efficiency, $t(15) = -2.12$, $p = .05$, $d = -0.80$, were significantly higher than their respective formulation scores. For the OA group, execution accuracy scores were higher than their corresponding formulation accuracy scores, $t(31) = 9.60$, $p < .001$, $d = -1.24$, while formulation and execution efficiency scores were not significantly different, $t(31) = -1.05$, $p = .304$, $d = -0.28$.

Supplemental Analyses

Execution Inaccuracy vs. Omission—We were also interested in determining whether the OA group obtained poorer execution scores as a product of attempting task parts inaccurately versus omitting task parts entirely. A group (YA, OA) by condition (formal planning, informal planning) univariate ANOVA revealed the YA group committed fewer omissions than OAs during task execution, $F(1, 92) = 9.43$, $MS = 49.0$, $p = .003$, $\eta_p^2 = .093$. There was no significant main effect for condition, $F(1, 92) = .001$, $MS = .005$, $p = .975$, $\eta_p^2 = .000$, or interaction, $F(1, 92) = .290$, $MS = .151$, $p = .592$, $\eta_p^2 = .003$. Regarding total number of task parts completed inaccurately, a 2 (group) by 2 (condition) univariate ANOVA revealed no significant effect of group, $F(1, 92) = 1.93$, $MS = 1.02$, $p = .168$, $\eta_p^2 = .021$, condition, $F(1, 92) = .039$, $MS = .021$, $p = .843$, $\eta_p^2 = .000$, or group by condition interaction, $F(1, 92) = .354$, $MS = .188$, $p = .553$, $\eta_p^2 = .004$.

Searching—Searching was measured as observed behavior that included visiting non-task related locations (e.g., wrong cupboard) in search of items or answers for questions. As seen in Table 3, a group (YA, OA) by condition (formal planning, informal planning) univariate ANOVA revealed that both groups engaged in a significantly higher number of searching behaviors in the informal compared to the formal planning condition, $F(1, 90) = 4.66$, $MS = 3.14$, $p < .05$, $\eta_p^2 = .049$. There was a trend for OAs to engage in searching behavior more often than YAs, $F(1, 90) = 3.30$, $MS = 2.22$, $p = .073$, $\eta_p^2 = .035$. There was no interaction effect, $F(1, 90) = .060$, $MS = .041$, $p = .806$, $\eta_p^2 = .001$.

Total Time—Total time required for task completion was measured as the time from when the participant began the execution phase to when they indicated they had finished. A group (YA, OA) by condition (formal planning, informal planning) univariate ANOVA revealed the OAs required a significantly larger amount of time for task completion in comparison to YAs, $F(1, 90) = 6.44$, $MS = 41.6$, $p < .05$, $\eta_p^2 = .067$. There was no significant effect of condition, $F(1, 90) = .207$, $MS = 1.34$, $p = .650$, $\eta_p^2 = .002$, or interaction, $F(1, 90) = .262$, $MS = 1.69$, $p = .610$, $\eta_p^2 = .003$.

Route Execution Adherence & Adherence Efficiency—A Route execution adherence score was measured as the percentage of task parts incorporated into the participant's formulated plan that were carried out in similar sequence during task execution. As seen in Table 3, there was no significant group difference in the percentage of total tasks adhered to during task execution, $t(46) = .585$, $p = .562$, $d = 0.18$. There also was no significant group difference in adherence efficiency scores, $t(46) = .980$, $p = .332$, $d = 0.31$.

Correlational Analysis—Correlational analyses were conducted with the OA group to examine the pattern of relations among the Amap formulation and execution scores, as well as the relationship between Amap scores and everyday functioning as measured by the IADL-C. Given the number of correlations completed, we used a more conservative alpha level of .01 when examining for relationships among the Amap measures. For the informal planning condition, Pearson correlations between formulation (preplanning time) and execution accuracy, $r = .238$, $p = .263$, and efficiency, $r = -.107$, $p = .619$, did not reach significance. Similarly, for the formal planning condition, there were no significant correlations between formulation and execution outcome measures (accuracy, efficiency), r s between $-.100$ and $.230$, p s $> .217$. Formulation accuracy and efficiency measures were significantly correlated, $r = .710$, $p < .001$, as were execution accuracy and efficiency measures, $r = .560$, $p < .001$. These findings further support a dissociation between formulation and execution skills in planning ability.

We also conducted exploratory analyses ($p < .05$) to determine the relationship between Amap formulation and execution measures and everyday activities of daily living using informant-report data from the IADL-C. Analyses were conducted separately for both the formal and informal planning conditions for the OA group and based on the four functional domain subscales of the IADL-C, which included money and self-management, home daily living, travel and event memory, and social skills. As can be seen in Table 4, for the informal planning condition, Spearman's correlations revealed no significant relationships between Amap formulation and execution outcome measures and the IADL-C. For the formal planning condition, the correlational analyses revealed that formulation efficiency scores were significantly correlated with informant-report data for the travel and event memory and social skills subscales of the IADL-C (e.g., read and follow a map when traveling) as well as the total IADL-C score. In contrast, execution efficiency was correlated with the money and self-management skills subscale (e.g., ability to prioritize and complete tasks in order of importance).

Discussion

The goal of the current study was to explicitly assess the formulation and execution components of planning ability in healthy younger and older adults using the Amap Task, an ecologically valid, open-ended measure of planning ability (Sanders & Schmitter-Edgecombe, 2012; Sanders, Low, & Schmitter-Edgecombe, 2014). Current standardized tests of planning ability may be limited in this regard, as most are not designed to separately evaluate the formulation and execution stages of planning and do not utilize ecologically valid task stimuli. We also set out to address the issue of methodology in planning assessment, specifically whether execution abilities are differentially impacted by a *formal planning* compared to an *informal planning* stage prior to execution, with the latter representing the traditional approach of focusing more on measuring the execution phase of planning and less on the formulation component.

For the formal planning condition, findings revealed generally poorer formulation accuracy and efficiency scores for the OA group in comparison to YAs. For the informal planning condition, results showed that formulation time (measured as total elapsed time from end of

instructions to task initiation) did not significantly differ by group. Together these findings highlight the limited sensitivity of using total planning time as a measure of quality of plan formulation, consistent with prior research (see Phillips et al., 1999; 2001).

Across conditions, OAs exhibited poorer execution accuracy and efficiency scores compared to YAs. Supplemental Amap measures also revealed that the OA group took significantly longer to execute the Amap and were more likely to omit task parts compared to YAs, with trends suggesting that the OAs also engaged in more searching behavior. Furthermore, the OAs poorer execution scores were shown to persist even after controlling for both formulation accuracy and efficiency scores calculated in the formal planning condition. These findings are consistent with previous research using the Amap Task (Sanders & Schmitter-Edgecombe, 2012; see also Sanders, Low, & Schmitter-Edgecombe, 2014), showcasing the reliability of the Amap as a measure of planning ability with sufficient sensitivity to identify distinct age-related declines in planning ability.

Regarding the planning manipulation, better execution efficiency scores and less searching behavior were found for participants in the formal compared to the informal planning condition. This suggests that participants' ability to efficiently interweave completion of task parts significantly improved following adequate opportunity to construct a plan prior to execution, which also reduced the likelihood they would need to search in order to complete task parts. In contrast to execution efficiency scores, however, execution accuracy scores did not vary by condition. While the execution accuracy findings may be consistent with evidence suggesting that preplanning has no bearing on task execution, regardless of the specificity in planning instructions or even when compared to having no opportunity for preplanning whatsoever (Phillips et al., 1999; 2001), this would not explain the significant effect of formal planning on Amap execution efficiency.

To bridge this gap, we posit that current findings may implicate an inherent hierarchy between accuracy and efficiency, with accurate performance less dependent on a formal versus an informal planning formulation phase while efficiency in performance would benefit more from a structured opportunity to plan ahead. As an example, consider if we were to operationalize grocery shopping. One could conceivably retrieve and purchase necessary goods at a grocery store with 100% accuracy by simply employing an algorithmic strategy to visit every aisle, beginning with those nearest the entrance and concluding with check-out. However, efficiency in this manner would require a more decisive strategy that would avoid visiting unnecessary aisles (e.g., pet food, automobile maintenance, greeting cards) while also taking into account factors affecting the order/sequencing of task parts (e.g., dropping off a prescription at the pharmacy first, collecting frozen items last, coordinating visits to free-food/sample vendors). In the current study, individuals who set out to "visit every aisle" may not have varied in their algorithmic approach based on the format of the initial formulation phase (e.g., those setting out to complete every task and question in the order listed, one after the other).

While efficiency in execution performance improved with formal preplanning, total time required to complete the execution phase did not. This finding may suggest that measures of efficiency are better than time when it comes to assessing planning abilities. However,

additional research is needed to clarify this point for several reasons. First, in the current study the non-significant effect of condition on execution time may have been an artifact of the “balancing” in speed of completion when comparing, for example, the time of a participant who employed an algorithmic approach (e.g. completing tasks one-by-one in order of presentation on the task list) versus the completion time for an individual prioritizing efficiency, here represented as minimal visits to location “hubs” while still able to carry out all relevant tasks. A more thoughtful, efficient approach may have resulted in more time spent using online executive functioning abilities, perhaps balancing the scale between time spent making redundant location visits and time spent self-monitoring, updating, etc. Considering the normative size and layout of the WSU Smart Apartment, it is also feasible the contrast in these approaches ultimately had little effect on speed of completion. Specifically, making multiple visits to the same location to complete task parts may not have unduly slowed overall task completion time, given the maximal distance between locations was typically in the range of 5–15ft., in addition to the relatively brief nature of the Amap task with regards to the overall number of tasks to be completed.

Comparison of formulation and execution accuracy and efficiency scores for the formal planning condition revealed both groups’ scores improved from formulation to execution. This finding suggests that, similar to YAs, OAs can benefit from on-line planning and are able to follow through and improve upon formally planned strategies during execution without an external cue. This would be consistent with evidence in the literature that OAs may be able to compensate for age-related declines in planning ability on ecologically valid tasks of planning ability (Phillips, Kliegel, and Martin, 2006; Kliegel, Martin, McDaniel, & Phillips, 2007; see also Sanders & Schmitter-Edgecombe, 2012). However, both groups may have improved from formulation to execution as a result of preplanning prior to task execution. Because the informal planning approach does not allow direct assessment of what is occurring in the space prior to task execution, this cannot be determined, further highlighting the limitations inherent to the informal planning method.

Highlighting a dissociation between formulation and execution skills in planning ability, our correlational analyses revealed that neither of the Amap formulation outcome measures were associated with execution accuracy or efficiency. In addition, correlations with informant-report data revealed associations between IADL-C subscales that appear more akin to formulation and execution abilities and respective Amap outcome measures. For example, execution efficiency was correlated with the money and self-management skills subscale of the IADL-C. Unlike the other subscales, this subscale is comprised of many items that pertain to complex everyday activities heavily reliant on execution abilities such as on-line planning and set-shifting (e.g., ability to complete tasks in order of importance or find one’s way back to a meeting spot). In combination with previous research where the Amap Task was found to be associated with several standardized neuropsychological tests of executive functioning (Sanders & Schmitter-Edgecombe, 2012; Sanders, Low, & Schmitter-Edgecombe, 2014), the relationship between Amap scores and informant report of everyday functioning highlights the ecological validity of the Amap as a measure sensitive to planning ability in everyday life.

While current findings align with evidence suggesting a need to separately evaluate formulation and execution abilities when assessing planning ability (Phillips, Wynn, McPherson & Gilhooly, 2001; Allain et al., 2005; Basso et al., 2001; Sanders & Schmitter-Edgecombe, 2012; Sanders, Low, & Schmitter-Edgecombe, 2014), this remains a limitation to most current measures across clinical and research settings. Given the nature of the common standardized approach to informally assess pre-execution formulation (e.g., elapsed time from end of initial instructions to task initiation, elapsed time from initial presentation of task stimulus to initiation), it stands to reason most instruments designed in this fashion would be unable to adequately differentiate underlying abilities tapped during formulation and their relationship to execution abilities, if at all. Findings in the literature have indeed been inconsistent in this regard, with some authors finding associations between preplanning time and accuracy and total time in execution (Unterrainer, Rahm, Kaller, Leonhart, Quiske, Hoppe-Seyler et al., 2004) while others have concluded preplanning has no relationship with execution performance (Phillips et al., 1999; 2001). Discrepancy in findings may reflect the nature of the informal preplanning format, which does not allow determination of what is occurring prior to execution and may also be limited by working memory capacity (see Phillips et al., 2001), as well as differences in execution outcome measures and their relationship to formulation abilities.

The current results showed that execution efficiency significantly improved with formal plan formulation while execution accuracy and total time were not significantly impacted. As discussed above, these findings potentially highlight the utility of incorporating a measure of efficiency when assessing planning ability, although future research would be helpful to clarify this point. For example, studies incorporating design considerations for distance versus time could further distinguish the relationship between preplanning and execution time, where redundant or otherwise inefficient task completion would have a more significant impact on not only the length of task execution but also the total distance travelled by the participant and, by extension, other factors such as exertion, stamina, motivation, etc. Consider the Amap paradigm but on a larger scale, such as tasking individuals with completing 19 task parts dispersed across various location "hubs" in a city. With such a paradigm, inefficient performance would undoubtedly have a negative impact on total execution time, perhaps here suppressed or otherwise "balanced" by the scope of the WSU Smart Apartment. In the current study, more efficient performance may have resulted in less wasted trips, though the impact of these excessive trips on overall execution time may have been marginal.

There are several limitations to current findings. Our sample was ethnically homogenous (mostly Caucasian), with discrepancy in educational level between age groups. However, given the constellation of findings illustrating age-related decline in planning ability, this may actually bolster conclusions (i.e. older adult poorer performance across measures despite being generally more highly educated than younger-adult counterparts). Correlational analyses may also have been limited by sample size for certain comparisons with the younger adult group. In addition, the Amap and related modified scoring paradigms are novel in nature, and replication of the Amap task itself may prove difficult across settings. However, the components of the task and its general impetus in design parallels everyday contexts and stimuli, which theoretically would not be difficult to recreate.

Although the real-world apartment where this study took place represents a strength in ecological validity in comparison to laboratory settings, this may yet have impacted performances to some degree. That is, participants were not familiar with the particular apartment layout prior to testing and so there may have been increased need for an overt opportunity to formulate a plan ahead of execution, whereas well-known tasks and/or settings may limit an individual's need in this regard, instead relying on procedural knowledge and expertise to guide behavior.

In conclusion, this study utilized the Amap Task, an ecologically valid, open-ended planning task to separately evaluate the formulation and execution components of planning ability in healthy younger and older adults. While there were no significant group differences in formulation time for the informal planning condition, results revealed age-related declines in formulation accuracy and efficiency scores in the formal planning condition. Our results also showed poorer execution accuracy and efficiency scores for OAs across conditions, with additional findings from supplemental measures collectively delineating an inherent slowing and increased tendency toward omission in planning behavior as part of the aging process. Independent of the quality of the formulated plans, ability of OAs to self-monitor and follow through with intended formulation strategies during execution without an external cue and/or modify and adjust strategies appeared intact. The experimental manipulation comparing formal and informal planning formats revealed that having an overt opportunity to formulate a plan prior to task completion benefitted execution efficiency but did not significantly impact execution accuracy. This suggests that while formal planning may not influence measures that assess for execution accuracy via static correct/incorrect criterion for task completion, execution efficiency (i.e., approach to task completion or "finesse") may significantly improve with formal preplanning, highlighting the utility of implementing a formal planning stage prior to execution when measuring planning abilities. This also has implications for planning ability in everyday life with aging populations. Based on current findings, utilizing formal preplanning strategies would serve as a beneficial compensatory strategy for older adults in everyday life, leading to improved efficiency and "finesse" in everyday activity completion, thereby contributing to improvements in quality of life, self-efficacy, and ability to remain independent in later life. As the adage goes, it is better to work smarter, not harder.

Acknowledgments

This work was supported by the *National Institute of Biomedical Imaging and Bioengineering* under Grant #R01 EB009675.

References

- Alderman N, Burgess PW, Knight C, Henman C. Ecological validity of a simplified version of the multiple errands shopping test. *Journal of the International Neuropsychological Society*. 2003; 9:31–44. [PubMed: 12570356]
- Allain P, Nicoleau S, Pinon K, Etcharry-Bouyx F, Barre J, Berrut G, et al. Executive functioning in normal aging: A study of action planning using the Zoo Map Test. *Brain and Cognition*. 2005; 57:4–7. [PubMed: 15629206]

- Basso D, Bisiacchi PS, Cotelli M, Farinello C. Planning times during traveling salesman's problem: Differences between closed head injury and normal subjects. *Brain and Cognition*. 2001; 46(1–2): 38–42. [PubMed: 11527359]
- Basso D, Lotze M, Vitale L, Ferreri F, Bisiacchi P, Olivetti Belardinelli M, et al. The role of prefrontal cortex in visuo-spatial planning: A repetitive TMS study. *Experimental Brain Research*. 2006; 171(3):411–415. [PubMed: 16710684]
- Benton, AL., Hamsher, K. Manual for Multilingual Aphasia Examination. Iowa City: University of Iowa; 1976.
- Cazzato V, Basso D, Cutini S, Bisiacchi P. Gender differences in visuospatial planning: An eye movements study. *Behavioural Brain Research*. 2010; 206(2):177–183. [PubMed: 19761799]
- Delis, DC., Kaplan, E., Kramer, JH. Delis-Kaplan Executive Function System (D-KEFS). San Antonio, TX: The Psychological Corporation; 2001.
- Franceschi M, Caffarra P, DeVreese L, Pelati O, Pradelli S, Savare R, et al. Visuospatial planning and problem solving in Alzheimer's disease patients: a study with the Tower of London test. *Dementia and Geriatric Cognitive Disorders*. 2007; 24(6):424–428. [PubMed: 17940336]
- Garden SE, Phillips LH, MacPherson SE. Midlife aging, open-ended planning, and laboratory measures of executive function. *Neuropsychology*. 2001; 15(4):472–482. [PubMed: 11761036]
- Grafman, J. Plans, actions and mental sets: Managerial knowledge units in the frontal lobes. In: Perecman, E., editor. Integrating theory and practice in clinical neuropsychology. Hillsdale, NJ: Lawrence Erlbaum Associates; 1989. p. 93–138.
- Jurado MB, Rosselli M. The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*. 2007; 17:213–233. [PubMed: 17786559]
- Kliegel M, Martin M, McDaniel MA, Phillips LH. Adult age differences in errand planning: The role of task familiarity and cognitive resources. *Experimental Aging Research*. 2007; 33:145–161. [PubMed: 17364904]
- Phillips LH, Kliegel M, Martin M. Age and planning tasks: The influence of ecological validity. *International Journal of Aging and Human Development*. 2006; 62(2):175–184. [PubMed: 16541929]
- Phillips LH, Wynn V, Gilhooly KJ, Della Sala S, Logie RH. The Role of Memory in the Tower of London Task. *Memory*. 1999; 7(2):209–231. [PubMed: 10645380]
- Phillips LH, Wynn VE, McPherson S, Gilhooly KJ. Mental planning and the Tower of London task. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*. 2001; 54A(2):579–597.
- Reitan, RM., Wolfson, D. The Halstead-Reitan Neuropsychological Test Battery: Therapy and clinical interpretation. Tucson, AZ: Neuropsychological Press; 1985.
- Sanders C, Schmitter-Edgecombe M. Identifying the nature of impairment in planning ability with normal aging. *Journal of Clinical and Experimental Neuropsychology*. 2012; 34(7):724–737. [PubMed: 22506736]
- Shallice T. Specific impairments of planning. *Philosophical Transactions of the Royal Society of London B*. 1982; 298:199–209.
- Williams, JM. Memory Assessment Scales Professional Manual. Odessa, FL: Psychological Assessment Resources, Inc; 1991.
- Wilson BA, Alderman N, Burgess PW, Emslie H, Evans JJ. Behavioural Assessment of the Dysexecutive Syndrome. Bury St Edmunds: Thames Valley Test Company. 1996

Questions: (Please say the answer to each question out loud)

1. How many plants are located in the apartment?	2	3	4	5
2. What Brand is the Microwave?	Philips	Sharp	Kenmore	Hoover
3. Is there a container of brown sugar located in cupboard “A”?	Yes	No		
4. Is there a broom in the kitchen supply closet?	Yes	No		
5. Is there a snow suit hanging in the hallway closet	Yes	No		
6. Is there a clock sitting on the dining room table?	Yes	No		

Task List

- 7. Retrieve T.V. remote from the kitchen counter and place in front of T.V.
- 8. Retrieve pill holder from dining room table and place in kitchen cupboard labeled “Cupboard A”
- 9. Retrieve water pitcher from hallway “clothes” closet, fill with water from faucet and place in refrigerator
- 10. Bring the following items to the front door: relish, two plastic cups, dustpan (use map to determine location of items)

Figure 1.
Amap Task

Table 1
Demographic Mean Summary Data for the Younger Adult and Older Adult Groups

Variable or Test	Group					
	YA (n = 32)			OA (n = 64)		
	Formal (n = 16)	Informal (n = 16)	SD	Formal (n = 32)	Informal (n = 32)	SD
Age	20.5	1.85	20.4	2.28	71.6	6.12
Education	14.7	1.18	14.7	1.08	17.1	2.59
WTAR	37.1	7.00	39.1	5.99	41.3	6.13

Note. WTAR= Wechsler Test of Adult Reading.

Table 2**Amap Task Scoring Components**

Task Parts (19 total)	Hubs with Respective Task Parts (7 total)
1. Medication dispenser retrieval	Dining room table (1, 17)
2. Medication dispenser placement	Cupboard "A" (2, 14)
3. T.V. remote retrieval	Kitchen counter (3, 19)
4. T.V. remote placement	Supply closet (9, 15)
5. Pitcher retrieval	Hallway closet (5, 16)
6. Pitcher filled with water	Refrigerator (8, 7)
7. Pitcher placement	Exit (11, 12, 13)
8. Relish retrieval	
9. Dust pan retrieval	
10. Plastic cups retrieval	
11. Relish placement	
12. Dust pan placement	
13. Plastic cups placement	
14. Brown sugar question	
15. Supply closet question	
16. Hallway closet question	
17. Dining room table question	
18. Plants question	
19. Microwave question	

Note. For the formulation component of planning, the *formulation accuracy score* is the total number of task parts incorporated into the formulated plan and the *formulation efficiency score* is the total number of hubs incorporated into the plan requiring only one visit to carry out all relevant task parts at that location. For the execution component of planning, the *execution accuracy score* and the *execution efficiency score* are derived in the same manner as the formulation scores, based on actual task performance during the execution phase.

Table 3

Amap Task Scores for Younger Adult and Older Adult Participants

Measures	Group						OA	
	YA							
	Formal (n = 16)	SD	M	Informal (n = 16)	SD	M	Formal (n = 32)	Informal (n = 32)
Formulation Accuracy	14.6	3.18				10.9	5.49	
Formulation Efficiency	2.06	0.98				1.50	1.24	
Formulation (Informal)			14.2	34.9				23.3 67.4
Execution Accuracy	18.4	.814	18.3	1.14	16.4	3.05	16.7	2.76
Execution Efficiency	3.19	1.72	1.81	1.28	1.94	1.88	1.44	1.27
Searching	1.88	2.06	3.36	2.98	3.38	3.75	4.81	3.94
Total Task Time	306	90.0	308	115	346	92.2	365	87.7
Route Execution Adherence	46.9	26.2			42.4	24.9		
Adherence Efficiency	1.13	2.13			.438	2.37		

Notes. Mean scores for informal planning formulation (total elapsed time from end of instruction to task initiation) and total task time are in seconds.

Table 4

Correlations between Amap Task Measures and IADL-C Informant-Report Data for Older Adult Participants

<i>Formal Planning</i>	Amap Task Outcome Variables			
	Formulation Accuracy ^a	Formulation Efficiency ^a	Execution Accuracy ^a	Execution Efficiency ^a
IADL-C, Informant-Report				
Money and Self-management	-.040	-.120	-.242	-.378 *
Home Daily Living Skills	-.077	-.296	.000	-.054
Travel and Event Memory Skills	-.248	-.438 **	-.089	-.177
Social Skills	-.120	-.423 *	.081	-.020
Total Score	-.116	-.372 *	-.177	-.229
<i>Informal Planning</i>	Formulation Time ^b	Execution Accuracy ^c	Execution Efficiency ^c	
IADL-C, Informant-Report				
Money and Self-management	-.265	-.071		-.117
Home Daily Living Skills	-.087	-.005		-.049
Travel and Event Memory Skills	.182	-.180		-.292
Social Skills	.154	-.203		-.006
Total Score	-.207	-.072		-.180

*
p .05;**
p .01.^a = 25,^b = 20,^c = 24.