Journal of Adult Development, Vol. 6, No. 2, 1999

Age Differences on a Procedurally Oriented Test of Practical Problem Solving[1]

Douglas A. Hershey[2,3] and Alicia H. Farrell[4]

Numerous studies have focused on developmental differences in practical problem-solving abilities. However, researchers have largely ignored the relationship among age, procedural knowledge, and problem-solving performance. The theoretical model that guided the present investigation suggests that as individuals age, they are exposed to a variety of real-world problems, which in turn should lead to the acquisition of problem-solving scripts. In the present study, a procedurally oriented, practical problem-solving measure was administered to 200 individuals aged 20 to 69 years. For each problem, subjects were required to order a set of discrete behaviors into an optimal solution sequence. These solution sequences were then compared with optimal solutions that domain area experts had established in advance. As expected, age was found to be positively related to prior problem-solving experience. However, contrary to expectations, the quality of subjects' solutions showed slight, negative age differences.

KEY WORDS: Aging; problem solving; everyday; practical; script; procedural; schema.

INTRODUCTION

Many practical problems require solutions that involve planning and coordinating multiple actions or procedures (Rebok, 1989). From simple familiar tasks, such as doing the grocery shopping, to less familiar situations, such as determining what to do if you are involved in an automobile accident, individuals rely on domain-specific procedural knowledge to solve problems. Despite a phenomenal growth in the literature on everyday cognition (Poon, Rubin, & Wilson, 1989; Puckett & Reese, 1993; Sinnott, 1989), relatively little research has focused on two important issues: (1) the relationship between procedural knowledge and practical (or everyday) problem-solving abilities and (2) how the quality of individuals' procedurally based solutions change across the life span as a function of problem-solving experience.

There are good reasons to assume that the ability to solve real-world, procedurally oriented problems should improve as a function of age. As developmental researchers have pointed out, advancing age is accompanied by increases in life experience in numerous real-world problem-solving domains (Baltes, 1993; Camp, Doherty, Moody-Thomas, & Denney, 1989; Charness, 1989; Denney, 1989; Ericsson & Charness, 1994; Hoyer, Rybash, & Roodin, 1989; Puckett, Reese, & Pollina, 1993; Walsh & Hershey, 1993; Willis & Schaie, 1993). Following from this premise, one would expect to find age to be positively correlated with problem-solving performance on measures that encompass multiple problem-solving domains. However, studies of age differences in ev-

[1]Portions of the findings contained in this article were presented at the annual meeting of the Eastern Psychological Association (April 1995) in Boston.

[2]Oklahoma State University, Stillwater, Oklahoma 74078.
[3]Correspondence should be directed to Douglas A. Hershey, Department of Psychology, Oklahoma State University, 215 North Murray Hall, Stillwater, Oklahoma 74078.
[4]George Mason University, Fairfax, Virginia 22030.

1068-0667/99/0400-0087$16.00/0

eryday problem solving have revealed that, in general, increases in age are not necessarily associated with increases in the ability to solve practical problems.

To date, the most comprehensive series of studies on practical problem solving were conducted by Denney and her colleagues, (Denney & Palmer, 1981; Denney & Pearce, 1989; Denney, Pearce, & Palmer, 1982; Denney, Tozier, & Schlotthauer, 1992; Heidrich & Denney, 1994). In these studies, individuals (typically ranging in age from 20 to 80 years) were asked to solve a set of between 6 and 10 real-world problems. Examples of the types of problems used in the Denney studies are as follows: What should you do if you are stranded in a blizzard? What should you do if you receive threatening telephone calls? What should you do if you buy a vacuum cleaner from a door-to-door sales person and it stops working 3 weeks later? The quality of subjects' performance on these multiproblem measures was based on the number of safe and effective solutions they generated for each of the various problems. In all five of the studies just cited, Denney found that older adults were less effective at solving everyday problems than were middle-aged individuals. Furthermore, in two of these studies (Denney et. al., 1982; Heidrich & Denney, 1994), she found that older adults also performed more poorly than younger individuals. This general pattern of age-related decline was found even when the problems were specifically designed to be highly familiar to older individuals (Denney et al., 1982). The consistency of findings across these studies led Denney to conclude that "it appears to be . . . difficult, if not impossible, to develop a set of [practical problems] . . . on which older individuals will do better than middle-aged or younger individuals" (1989, p. 341).

In contrast to the work of Denney and her colleagues which showed a general pattern of age-related decline, the study by Cornelius and Caspi (1987) revealed a small but significant positive linear relationship between age and everyday problem-solving performance. These investigators used a different methodology to examine the relationship between age and everyday problem-solving performance. They presented subjects (age range, 20–78 years) with brief descriptions of 48 practical problems (e.g., What would you do if your landlord refused to make costly repairs?). For each problem, subjects were presented with four solution options that were designed to correspond to four problem-solving modalities: (1) problem-focused actions (e.g., make the repairs yourself), (2) cognitive problem analysis (e.g., try to understand your landlord's point of view), (3) passive-dependent behaviors (e.g., try to get someone else to settle the dispute), and (4) avoidant thinking and denial (e.g., simply accept the situation). For each of the 48 problems, subjects were asked to indicate how likely they would be to select the actions associated with each of the four modalities. The quality of each subject's response pattern was assessed by correlating his or her likelihood ratings of problem solutions with those of a panel of independent judges whose averaged likelihood responses were deemed to be optimal. The investigators found that the quality of individuals' solutions increased significantly as a function of age ($\beta_{age} = .19$), which led them to further examine whether age differences in problem familiarity could have been responsible for the observed developmental improvement. Contrary to expectations, older adults reported being less familiar with the problems than did younger and middle-aged adults, and there was no difference in the self-reported familiarity levels of the latter two groups. Therefore, the theoretically based relationship among age, problem-solving experience, and problem-solving ability suggested earlier (which was also of interest to Cornelius & Caspi) was not tested in this study because age was not found to be significantly correlated with problem familiarity.

Willis and her colleagues carried out yet another set of studies that focused on aging and everyday problem solving (Marsiske & Willis, 1995; Willis, Jay, Diehl, & Marsiske, 1992; Willis & Marsiske, 1991). The goal of this line of work was to identify age-related changes in *late adulthood* (i.e., approximately the sixth through the ninth decades of life). Two of the studies (Willis et al., 1992; Willis & Marsiske, 1991) used a longitudinal approach, whereas the third (Marsiske & Willis, 1995) used a cross-sectional design. Although each of the three studies used somewhat different tasks, which focused on different problem-solving domains, a consistent pattern of age-related change was identified. Across all three studies, everyday task performance was found to decline with increasing age. Moreover, the magnitude of the observed decline was found to increase as a function of increases in age. That is, in the Willis and Marsiske (1991) longitudinal study, the average decline seen within individuals tested at 63 years of age and then again at 70 years of age was somewhat less than the decline seen in the same individuals when they were tested at age 70 and then again at age 77. The greatest drop in performance was seen within individuals between the ages of 77 and 84. In sum, the findings from all three Willis et al. studies clearly indicate that everyday problem-solving performance declines precipitously in adults older than 60 years.

One possible reason for the equivocal findings across the Denney et al. studies, the Cornelius and Caspi (1987) study, and the Willis et al. studies is that age was not necessarily associated with increasing levels of domain-specific experience. That is, the problems used across all these studies were "real-world" problems. However, that criterion did not ensure that these problems were encountered by subjects with equal frequency at all stages of the life span. Consider, as an example, Denney's vacuum cleaner and blizzard problems (described previously). It is not clear that knowledge about problems such as these would increase during a person's life span, given their relatively rare rates of occurrence. This being the case, one might not expect older individuals to earn higher scores than those of younger and middle-aged individuals when they are asked to solve laboratory simulations of these problems. Similarly, in the Cornelius and Caspi (1987) study, older adults were *less* familiar with the set of problems than were middle-aged and younger subjects, which would lead one to surmise that the older subjects possessed less domain-specific knowledge about those problems. In this article, we take the position that one's domain-specific problem-solving experiences should be directly related to the quality and amount of one's knowledge of a task. To that extent, we believe it is critical to demonstrate that age is related to problem familiarity if one starts with the premise that problem-solving abilities will increase during a lifetime.

The theoretical model that guided the present investigation suggests that increases in age should be accompanied by increases in experience across multiple practical problem-solving domains. This experience, in turn, should lead to general age-related improvements in problem-solving performance. Furthermore, for most real-world problems, what is learned through experience are *problem-solving scripts* (Hershey, Walsh, Brougham, Carter, & Farrell, 1998; Hershey, Walsh, Read, & Chulef, 1990; Hershey, Wilson, & Mitchell-Copeland, 1996; Walsh & Hershey, 1993), which are task-specific solution procedures that specify a course of action that will resolve a problem situation. On the basis of this theoretical perspective, we believe that the solution-generation task used by Denney and the likelihood-rating task used by Cornelius and Caspi might have undervalued older adults' procedurally based experiential knowledge and thus underestimated their ability to effectively solve everyday problems. Both the solution-generation task and the likelihood-rating task require subjects to be familiar with the relative effectiveness of a variety of solutions for a particular problem. We contend that most people are unlikely to have this form of knowledge readily associated with various problems. Rather, we think that when individuals are confronted with a specific problem situation, they rely on detailed knowledge of one or perhaps two procedurally based solutions that have proven reliable in the past.

The present investigation was based on the premise that individuals' problem-solving scripts become increasingly veridical and well defined during adulthood as a function of increasing life experience. Furthermore, age-related improvements in the quality of individuals' knowledge structures should lead to improvements in the ability to solve a variety of everyday problems. To test this hypothesis, we developed a seven-item test of procedurally based problems and administered it to 200 adults. For each problem, subjects were shown a list of different actions that might be taken to generate a solution and then asked to order those actions into what they believed to be the best possible sequence. To assess the quality of subjects' performance, we scored their solution sequences by comparing them with a set of optimal solutions. (These optimal solution sequences were determined through consultation with experts when we initially designed the problems.) As suggested previously, we expected to find that older individuals would report having experienced the various problems contained on the test more often than younger persons would. Furthermore, on the basis of this anticipated relationship between age and prior problem-solving experience, we hypothesized that older adults would provide better solutions to the problems than those provided by younger individuals.

In developing the problem-solving measure, we recognized that subjects might know how to solve a particular problem but may be unwilling (or unable) to carry out the steps necessary to effectively resolve the situation. For example, older adults might know how to change a flat tire but may be unable to carry out all the required steps because of age-related declines in physical strength. Therefore, in addition to a measure that assessed the quality of subjects' scripts, we thought that it would be informative to have a second dependent measure that assessed the subjects' likelihood of completing the course of action that they had specified. Toward that end, subjects were asked to identify any behaviors that were included as part of their solution sequences that they would be unwilling or unable to carry out. We then examined the relationship between age and the willingness of subjects to carry out sequenced actions.

METHOD

Participants

Participants were 200 men and women who ranged from 20 to 69 years of age. To ensure that a sufficient number of individuals of different ages were represented across the adult life span, we adopted a sampling design that specified the testing of 20 men and 20 women from each of the five decades included in the targeted age range (i.e., 20–29, 30–39, 40–49, 50–59 and 60–69). Thus, the final sample included 100 men and 100 women equally distributed across the age range. The overall mean age of the sample group was 43.97 years (SD = 14.55). Subjects were found to be relatively well educated, having completed an average of 14.6 years of formal education (SD = 3.52). Educational level was not found to be systematically related to either age or gender. Participants were residents of the greater Washington, DC, area who were sampled at various locations throughout the community (e.g., service organization meetings, laundromats, libraries, etc.). All subjects voluntarily participated in the study, without remuneration. The attrition rate was low; only three individuals discontinued testing once the measure had been administered.

Materials and Procedure

The test consisted of a series of vignettes that described seven hypothetical real-world problem situations. The problem situations involved (1) an automobile accident, (2) a child bitten by a dog, (3) a gas leak in the kitchen, (4) a dead telephone, (5) an unconscious woman in the park, (6) a flat tire, and (7) gold coins stolen during a move. For each problem, subjects were presented with a list of actions that might be taken to resolve the situation. The subjects' task was to specify, in each of the seven cases, the sequence of actions that would most effectively and efficiently lead to a resolution of the problem. All subjects were tested individually or in groups of two or three. Each subject was told at the outset of the test that he or she would have an unlimited amount of time to complete the measure. Descriptions of each of the seven problem scenarios are contained in Appendix A (most of the vignettes that subjects responded to were appreciably longer than these descriptions, containing additional contextual information and limiting circumstances). Furthermore, a complete example of one of the problems can be found in Appendix B, including all options and foils.

Test booklets were constructed that contained one problem per page. Below each problem, a number of *action steps* were listed. Among these actions were a subset of valid *solution steps,* which, if properly sequenced, would lead to a safe and effective resolution of the problem. Interspersed among the solution steps were a number of *foils,* which were steps that (a) could not be carried out given the constraints of the problem (e.g., trying to check the fuse on a dead telephone, given that telephones do not have fuses) or (b) would potentially interfere with a safe and effective solution (e.g., turning on the top burners of a stove if you smell a gas leak, which could lead to an explosion). Each problem contained eight or nine action steps. Of these steps, five or six were valid solution steps and three or four were foils. Problems were counterbalanced within test booklets in an effort to guard against order effects.

Subjects were asked to complete each problem by following six clearly defined steps. First, they were to read the problem description and each of the possible action steps. Second, they were to eliminate any action steps that they thought should not be carried out, by drawing lines through those steps. Third, they were to number the remaining steps in the order in which the actions should ideally be carried out. Fourth, they were asked to circle any actions that they had sequenced that they personally would not, or could not carry out if they were actually in that situation. Fifth, they were asked to indicate approximately how many times they had faced this problem (or a very similar problem) in the past. Finally, they were asked to answer the following question after completing each problem: *How confident are you that the numbering sequence you have specified is the best course of action?* Responses to this question were made on a 7-point Likert-type scale (1 = *not at all* *confident;* 4 = *somewhat confident;* 7 = *extremely confident*). Each of the six steps contained in the instructions were outlined in brief on each page of the test booklet. Before beginning the test, subjects were shown a model of a completed sample problem. The time required to complete the task, from the beginning of the first problem to the conclusion of the final problem, was covertly recorded by the experimenter. Upon completion of the measure, subjects were asked to make a global confidence rating of the quality of their solutions by using the 7-point Likert-type scale described previously.

One of our primary objectives in developing scenarios for the measure was to choose a set of realistic, nontrivial problems that would be perceived by individuals of all ages as problems worth solving. We believed that the use of significant, real-world scenarios (as opposed to novel or trivial problems) would help to ensure that subjects would be motivated to solve the problems in a thoughtful and conscientious fashion. Much of the literature on adult everyday cognition has emphasized the routine, overlearned nature of everyday problems (Willis & Schaie, 1986); however, we attempted to select a range of problems that varied in terms of their level of familiarity. As Meacham and Emont (1989) pointed out, studies of practical problem solving need not be based on tasks that are familiar, ordinary, or routine, but rather on tasks that possess ecological face validity. Finally, we constructed each problem to have a single, efficient, procedurally based solution.[5] For certain problems, more than one possible solution could be adopted. However, for each problem, one solution was deemed by our domain area experts to be maximally efficient. These seven optimal solutions were used as "gold standards" against which the quality of subjects' solutions was judged.

Scoring Procedure

A subject's total score on the measure was based on the contribution of two independent scoring components: (1) *positive points,* which were earned by correctly ordering *pairs* of valid solution steps, and (2) *penalty points,* which were accumulated by including foils as part of the solution sequence. Therefore, for problems that contained six valid solution steps, five valid paired sequences were possible (i.e., step 1 followed by step 2, step 2 followed by step 3, and so on). One-hundred positive points were awarded for a problem if all pairs of valid solution steps were ordered in the correct sequence.[6] When subjects made an error in specifying the correct sequence, *partial credit* was awarded (e.g., if four of five possible paired sequences were correctly specified, 80 of the 100 possible points were awarded). Partial credit was assigned to ensure that subjects' solutions were not simply judged to be right or wrong, but were instead, defined according to varying degrees of solution quality.

As mentioned previously, subjects not only scored positive points by correctly ordering pairs of valid solution steps, but also lost points by including foils as part of the solution sequence. A maximum of 100 *penalty points* could be incurred in this fashion. Therefore, if a problem contained three foils, each foil included as part of the solution sequence would result in the loss of 33.3 points. If a problem contained four foils, each sequenced foil would result in the loss of 25 points.

Individuals' raw scores for each problem were calculated by subtracting the number of penalty points from the number of positive points earned. Thus, the raw score could theoretically range from −100 points to +100 points. For ease of interpreta-

[5]The advice of two or three domain area experts was obtained when we were developing the correct sequence of actions for each problem. Fortunately, the problems were fairly well structured (Simon, 1973). That is, they had a clear-cut solution sequence that would lead to a successful outcome. In all but one problem, experts agreed on a series of steps that would most effectively and efficiently lead to a solution. In the one problem for which there was disagreement (the missing coins problem), one expert's opinion differed only slightly from those offered by the other two consultants. Upon subsequent contact, the discordant expert agreed that the solution offered by the other two would lead to an acceptable resolution of the situation. The consensus script was then adopted as the optimal sequence for that problem. In recruiting our expert consultants, we made every effort to ensure that these individuals qualified as experts, by virtue of both their training *and* their experience. For example, two directors of training for the American Red Cross (who are responsible for developing, reviewing, and teaching CPR procedures) and an emergency medical technician were consulted for the unconscious-woman-in-the-park scenario, and two emergency response operators from two gas companies were consulted for advice regarding how to resolve the gas leak scenario.

[6]In scoring the solutions, we considered only the "pairwise correctness" of *adjacent* solution steps. We did this to ensure that a sequence error early in a solution would not lead to subsequent multiple errors, which would thereby make it a more serious error than one that occurred later in the sequence.

tion, raw scores were converted to percentage scores by rescaling them from 0 to 200, then dividing by a denominator of 200. Finally, a total score for the entire test was derived by calculating the mean of the individual item percentages across all seven problems.[7]

RESULTS

The primary goal of the study was to examine the nature of the relationships among age, experience, and the quality of test performance on the procedurally based measure of practical problem solving. In line with this goal, this section reports separate analyses that focus on pairwise combinations of these three variables,[8] followed by a set of multivariate LISREL (linear structural relations) analyses. Additional analyses are reported that help provide further insight into the relationship between aging and problem-solving performance.

Throughout this section, age-related hypotheses are tested by using regression analyses in which age is treated as a continuous predictor variable. However, to facilitate the interpretation of age-related trends, we occasionally present tables in which age is treated as a discontinuous variable. In such cases, the age range is always decomposed into five discrete, nonoverlapping groups (i.e., groups based on the five decades contained in the age range—20s, 30s, 40s, 50s, and 60s).

Experience and Age

An important assumption underlying the present investigation was that, on the whole, older adults would be more likely than younger individuals to have experienced the set of specific problems contained on the practical problem-solving measure. To test this assumption, we examined the relationship between age and self-reported prior-experience scores. Table I shows the number of times that subjects reported having experienced each of the seven problems, as a function of age. Although the experience scores for each of the seven problems do not always show monotonic increases as a function of age, the mean experience scores across all seven items reveal that older adults were somewhat more likely than younger individuals to have experienced the set of problems.

A total experience score was computed (based on a set of *T*-score transformations, $M = 50$, $SD = 10$) to assess the subjects' overall level of experience with the entire set of problems. This set of scores forms the bottom row of Table I. As can be seen by reading scores from left to right across the bottom row, experience levels increase linearly from the 20s age group through the 50s age group and then level off. A regression analysis using age as a predictor of total experience revealed that age and experience were significantly related, $F(1, 198) = 12.16, p < .01$, $R^2 = .058$. This finding provides important evidence for one of our fundamental assumptions: that older adults possessed significantly more experience at solving the set of practical problems contained on the measure than did younger individuals.

Age and Problem-Solving Performance

A critical goal of the present study was to examine the relationship between age and quality of performance on the practical problem-solving measure. Mean solution-quality scores and standard deviations for each of the seven problems and the full scale are shown in Table II. These scores, which for presentation purposes are arranged into five discontinuous age groups, reveal a general pattern of negative age differences. The total score values presented from left to right across the bottom row of the table reveal that individuals in their 20s and 30s produced the highest mean scores on the measure (55.1% and 55.0%, respectively), individuals in their 40s and 50s produced somewhat lower scores (53.1% and 52.5%, respectively), and individuals in their 60s produced the lowest scores overall (50.6%). To test the significance of this apparent age-related trend, we regressed the subjects' ages on their total scores on the measure. This analysis revealed age to be a significant predictor of test performance, $F(1, 198) = 4.28, p < .05, R^2 =$

[7] A detailed summary of the scoring system and a complete copy of the test can be obtained upon request from the first author.

[8] A number of pairwise computations reported throughout the results section take the form of bivariate regressions, whereas others take the form of Pearson correlations. We acknowledge the conceptual equivalence of these two analyses. The relationships that were considered to be causal in nature on the basis of theoretical grounds are reported as bivariate regressions (i.e., *X* predicts *Y*). In the absence of an a priori causal hypothesis, bivariate relations are reported as (nondirectional) Pearson correlations (i.e., *X* is associated with *Y*).

Table I. Mean Self-Reported Experience Scores[a] and Standard Deviations[b] as a Function of Age

Problem	Age group: 20s	30s	40s	50s	60s	All subjects
Auto accident	1.00	1.15	1.05	1.32	1.58	1.22
	(1.01)	(1.39)	(1.13)	(1.69)	(1.88)	(1.46)
Dog bite	0.03	0.10	0.15	0.13	0.30	0.14
	(0.16)	(0.30)	(0.58)	(0.40)	(0.91)	(0.54)
Gas leak	0.48	1.03	0.83	1.52	0.87	0.95
	(1.09)	(1.94)	(1.57)	(2.93)	(1.91)	(1.99)
Dead phone	3.72	2.75	3.92	5.97	3.02	3.88
	(6.27)	(2.53)	(7.98)	(9.68)	(4.26)	(6.68)
Woman in park	0.08	0.13	0.49	0.40	0.32	0.28
	(0.35)	(0.40)	(2.14)	(1.52)	(2.06)	(1.50)
Flat tire	1.82	3.22	2.82	3.55	2.98	2.88
	(2.55)	(4.34)	(4.00)	(4.95)	(3.94)	(4.04)
Missing coins	0.08	0.00	0.33	0.62	0.60	0.33
	(0.35)	(0.00)	(0.94)	(1.21)	(1.19)	(0.91)
Total score[c]	46.07	48.11	49.94	53.68	52.20	50.00
	(5.34)	(7.88)	(10.47)	(12.47)	(10.77)	(10.00)

[a]All means except those in the far right row (All subjects) are based on an N of 40 individuals per age group. Means in the body of the table for each problem were calculated by using unadjusted (raw score) values.
[b]In parentheses.
[c]To derive total experience scores, we first converted scores for each problem to T-scores (i.e., with a distributional mean of 50 and standard deviation of 10).

.021. The small but statistically significant negative age difference (standardized $\beta_{age} = -.15$) is shown in the scatterplot of the data (see Fig. 1). A comparison between the predicted mean score for a 20-year-old and the predicted mean score for a 69-year-old (representing the two endpoints of the age range) reveals a five-percentage-point drop in performance. It is important to note, however, that there is substantial overlap in the distribution of scores for individuals of various ages. Therefore, although the regression

Table II. Mean Solution-Quality Scores[a] and Standard Deviations[b] as a Function of Age

Problem	Age group: 20s	30s	40s	50s	60s	All subjects
Auto accident	65.8	63.0	63.6	62.0	58.2	62.5
	(24.8)	(22.4)	(22.9)	(21.4)	(16.9)	(21.8)
Dog bite	49.5	51.6	51.9	48.8	53.6	51.1
	(17.1)	(21.8)	(15.2)	(17.4)	(19.3)	(18.2)
Gas leak	44.1	39.2	38.6	40.8	36.0	39.7
	(17.6)	(18.1)	(17.6)	(15.0)	(18.3)	(17.4)
Dead phone	52.5	57.5	50.4	50.8	45.0	51.2
	(26.0)	(21.1)	(22.4)	(23.3)	(21.9)	(23.5)
Woman in park	48.1	49.0	46.0	43.0	43.9	46.0
	(22.7)	(21.7)	(20.0)	(21.2)	(19.0)	(20.9)
Flat tire	64.7	71.2	77.5	70.6	68.6	70.5
	(26.4)	(30.6)	(18.6)	(26.2)	(25.1)	(25.7)
Missing coins	61.0	53.4	43.4	51.5	48.9	51.7
	(22.8)	(24.2)	(23.6)	(19.9)	(22.0)	(23.1)
Total score[c]	55.1	55.0	53.1	52.5	50.6	53.3
	(9.97)	(12.61)	(9.11)	(9.21)	(11.61)	(10.6)

[a]Percentages. All means except those in the far right row (All subjects) are based on an N of 40 individuals per age group.
[b]In parentheses.

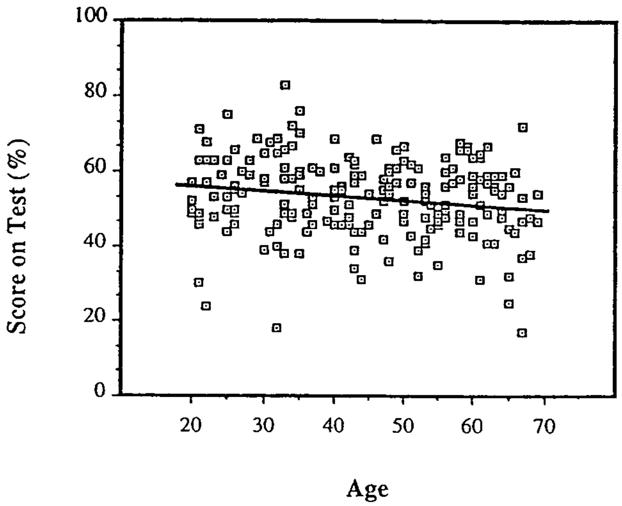

Fig. 1. Scatterplot of age versus total score on measure (in percentage points).

line in Fig. 1 indicates the presence of a small negative age difference, one can look to the variability in performance across the age range to see that increases in age are not *necessarily* associated with low levels of problem-solving performance.

Further examination of the test score data revealed that problem-solving performance was also related to the subjects' level of formal education and their gender. Educational level and test performance were found to be significantly correlated, $r(198) = .20, p < .01$, with more-educated individuals scoring higher on the test than less-educated persons. Furthermore, a *t*-test revealed that males ($M = 54.9\%$, $SD = 9.28$) scored significantly higher on the measure than females did ($M = 51.6\%$, $SD = 11.61$), $t(198) = 2.27, p < .05$. These findings suggest that a more accurate estimate of the relationship between age and problem-solving ability could perhaps be obtained by statistically controlling for the effects of gender and education.

A hierarchical multiple-regression analysis was conducted to test whether age was still a significant predictor of problem-solving performance once the variance associated with gender and education had been partialled from the test scores. Education and gender were simultaneously entered into the regression equation as the first block of predictors, after which age was entered as the sole predictor in the second block. As expected, the first block of variables was found to account for a significant amount of variance on the test, $F(2, 197) = 5.98, p < .01, R^2 = .057$. When age was entered into the equation in the second block, the R^2 increased from .057 to .075, which reflected a 1.8% increase in accounted-for variance. This change in R^2 was found to be statistically significant, $F(3, 196) = 3.78, p < .05$, which indicated that age still accounted for a significant portion of unexplained variance once the variance associated with gender and education had been partialled from the test scores.

A set of seven bivariate regressions were computed to examine the relationship between age and solution quality for each of the test items. Of the seven tests, only the missing coins item was found to be statistically significant, $F(1, 198) = 4.77, p < .05$, with increases in age associated with increases in solution quality.

Finally, a set of nonlinear analyses were calculated to further explore the relationship between age and solution quality. Seven power polynomial regression analyses explored the linear and quadratic relationship between age and performance for each problem. An eighth polynomial regression used linear and quadratic instances of age as predictors, controlling for education and gender, and the *aggregate* solution-quality score (an average over all seven problems) as the criterion. Only one of these analyses—the missing coins problem—indicated a significant nonlinear trend, $F(2, 197) = 4.60, p < .05, R^2 = .044$. Both the linear ($\beta = -1.16$) and the quadratic ($\beta = 1.02$) age functions for this problem were significant at the .05 level. However, the larger pattern of findings suggests that age was not systematically related to performance in a nonlinear fashion.

Experience and Problem-Solving Performance

As indicated previously, the conceptual model that guided this research suggests that individuals' prior experience with the problems on the measure should be positively associated with the quality of their problem-solving efforts. To test this hypothesis, we conducted a regression analysis in which subjects' total experience scores were used to predict individuals' total scores on the measure. This analysis revealed an unexpected finding, namely, that experience was not significantly related to problem-solving performance, $F(1, 198) = 1.19$, *ns*, $\beta = .08$. One possible explanation for this curious finding is that perhaps too much information was lost in the process of aggregating subjects' problem-specific experience scores into a single, global measure of experience.

Next, we sought to determine whether the quality of subjects' solutions varied as a function of differing levels of familiarity with each of the various problems. An inspection of the mean experience scores across all subjects (the far right column in Table I) revealed that the seven problems could be categorized into three familiarity levels. Two of the seven problems (the dead telephone and the flat tire) could be characterized as familiar, with subjects having experienced these situations three or four times during their lifetime. Two additional problems could be characterized as moderately familiar (the automobile accident and the gas leak), with subjects having experienced these situations once before, on average. Finally, three of the problems (the dog bite, the woman in park, and the missing coins) could be characterized as relatively unfamiliar problems, inasmuch as only a few subjects had ever experienced these situations. On the basis of our theoretical model, we expected to find solution-quality scores to be highest for the most familiar problems and lowest for the least familiar problems. An inspection of the mean solution-quality scores revealed monotonic increases as a function of increases in prior problem-solving experience. The mean solution-quality score was 49.6% (SD = 13.7) for the unfamiliar problems, 51.1% (SD = 14.5) for the moderately familiar problems, and 60.9% (SD = 18.0) for the familiar problems. Paired Bonferroni t-tests revealed that performance scores for the familiar problems were significantly larger than those for the moderately familiar problems, $t(199) = 6.82$, $p < .001$, and unfamiliar problems, $t(199) = 8.13$, $p < .001$. Solution-quality scores for the moderately familiar and unfamiliar problems were not found to differ, $t(199) = 1.21$, *ns*.

A set of analyses were also carried out that examined the relationship between subjects' experience levels for each problem and their solution-quality scores. For each of the familiar and moderately familiar problems (dead telephone, flat tire, automobile accident, gas leak), bivariate regressions were computed in which the experience score was regressed on solution quality. These tests failed to reveal significant outcomes for all but the flat tire problem, $F(1, 198) = 10.09$, $p < .01$.

For the three unfamiliar problems, there was a restricted range of scores, and positive skew was associated with the prior-experience distributions. Given that these conditions violate the assumptions of linear regression, we conducted group difference analyses to assess the strength of the experience–performance relationship. Specifically, for each of the three unfamiliar problems, the experience variable was dichotomized (i.e., forming "experienced" and "no prior experience" groups), then independent sample t-tests were used to compare the quality of subjects' solutions. These analyses, like the regression analyses reported previously, failed to indicate that experience was systematically related to solution quality: dog bite, $t(198) = .25$, *ns*; woman in the park, $t(198) = 1.16$, *ns*; and missing coins, $t(198) = .56$, *ns*.

Finally, a set of nonlinear analyses were calculated to further explore the relationship between experience and solution quality. Specifically, four power polynomial regression analyses examined the linear and quadratic effects of the experience–performance relationship for individual problems.[9] A fifth power polynomial regression used aggregate experience scores (both linear and quadratic) as predictors, and the *aggregate* solution-quality score (an average over all seven problems) as the criterion. Only one of these analyses—the flat tire problem—indicated a significant nonlinear trend, $F(2, 197) = 13.27$, $p < .01$, $R^2 = .119$. Both the linear ($\beta = .69$) and the quadratic ($\beta = -.43$) components of this test were found to be significant at the .01 level. However, the broader pattern of findings suggests that experience was not systematically related to performance in a nonlinear manner.

Confidence Ratings

Relationships were then examined between subjects' confidence ratings and (a) their prior-experience scores, (b) their age, and (c) their mean solution quality. These analyses involved the confidence ratings for each problem, in addition to a single, global post-task confidence rating in which participants assessed the overall quality of their performance.

Four separate bivariate regression analyses were computed to determine the extent to which prior-experience scores determined confidence levels (cf. footnote 9). The experience–confidence rela-

[9]The three experience–performance relationships for the unfamiliar problems (dog bite, missing coins, woman in the park) were not tested for nonlinear effects as a result of unsuitable distributional characteristics for the experience measure. For the same reason, these three unfamiliar problems were not included as part of the experience–confidence analyses that follow.

tionships were statistically significant for each of the four individual problems tested: automobile accident, $F(1, 198) = 31.88$, $p < .01$; gas leak, $F(1, 198) = 33.94$, $p < .01$; dead telephone, $F(1, 196) = 22.33$, $p < .01$; flat tire, $F(1, 198) = 42.81$, $p < .01$. Moreover, subjects' mean experience scores were significant predictors of their global confidence rating, $F(1, 198) = 31.19$, $p < .01$. Age was not found to be systematically related to the quality of subjects' individual or global confidence ratings, except for the gas leak problem ($r[198] = .16$, $p < .05$), with older subjects making higher confidence ratings. Finally, subjects' global confidence ratings were found to be significantly associated with their mean solution quality, $r(198) = .16$, $p < .05$, with higher confidence ratings associated with higher levels of solution quality.

Multivariate Analyses

An examination of the multivariate relationships among age, experience, and problem-solving performance was carried out using the LISREL VIII program (Jöreskog & Sörbom, 1993). Of the multiple models considered, four were selected for presentation (see Table III) on the basis of either theoretical grounds (i.e., our originally hypothesized model) or goodness of fit. Table III displays parameter estimates for each of the four models (i.e., standardized beta-weight values) and three statistical fit indices: (1) the chi-square statistic, (2) the goodness-of-fit index (GFI), and (3) the root-mean-square residual (RMSR).

Our originally hypothesized model is labeled *Model One* in Table III. It represents subjects' ages regressed on the composite experience variable, which, in turn, is regressed on the solution-quality marker. The statistical fit indices suggest, however, that this model is a poor fit to the data, $\chi^2(1, N = 200) = 5.80$, $p < .02$; $GFI = .98$; $RMSR = .07$. The modification index revealed that an increase in goodness of fit could be obtained by adding a direct link between age and solution quality. This link, when added to Model One, forms the basis of Model Two. The second model, which is a fully saturated model, is a good fit given the data, $\chi^2(0, N = 200) = 0.00$, $p = 1.00$. In addition to the hypothesized age $\rightarrow$ experience $\rightarrow$ quality relations, Model Two reveals a significant negative relationship between age and solution quality ($\beta = -.17$). Thus, relative to the first model, the second model captures unaccounted-for variance between age and performance quality that is not mediated by the subjects' prior level of problem-solving experience.

We had expected to find a relatively strong relationship between experience and solution quality, which was not the case. Therefore, we considered alternative model configurations in an effort to identify a better predictor of the performance variable. Model Three is an extension of Model One in which subjects' mean confidence level across the seven problems is added as a mediator between experience and performance. In this instantiation of the model, age predicts experience, experience predicts confidence, and confidence predicts solution quality (see Hershey & Wilson, 1997, for a discussion of aging, confidence ratings, and problem-solving perfor-

Table III. Parameter Estimates and Statistical Fit Indices for the Four Structural Equation Models

	Model[a]			
Regression link	One	Two	Three	Four
Age → Experience	.24*	.24*	.24*	.24*
Experience → Quality	.08	.12	—	—
Age → Quality	—	−.17*	—	−.15*
Experience → Confidence	—	—	.37*	.37*
Confidence → Quality	—	—	.16*	.16*
Statistical fit indices				
df	1	0	3	2
χ^2	5.80	0.00	6.38	1.79
p	.02	1.00	.09	.41
GFI	.98	—	.98	1.00
RMSR	.07	—	.05	.03

[a]Asterisks indicate significant beta weights (i.e., t values > 2.0).

mance). The statistical fit indices indicate that this configuration also provides a good fit to the data, $\chi^2(3, N = 200) = 6.38, p < .09$; $GFI = .98$; $RMSR = .05$. Moreover, the model has intuitive appeal. That is, one would expect extensive experience to lead to high levels of confidence, and high levels of confidence (to the extent that it is associated with perceived self-efficacy) should have a positive impact on the quality of the solutions. However, modification indices for the third model suggested that a better overall fit could be obtained by adding a direct link between age and problem-solving performance (comparable to the link added in Model Two).

The fourth model in Table III is identical to Model Three, with the addition of a regression link between age and solution quality (see Fig. 2 for a graphic representation). This model not only improves the overall goodness of fit relative to the third model ($\chi^2[2, N = 200] = 1.79, p < .41$; $GFI = 1.00$; $RMSR = .03$), but also provides for a richer theoretical representation of elements important to the problem-solving process. Of the various models tested, this one appears to best represent the data, while also maintaining the essential structure of the hypothesized relations among core constructs.

Positive Points and Penalty Points

Recall that a subject's total score on the test was based on the contribution of two independent scoring components: (1) positive points, which were earned by correctly ordering pairs of valid solution steps, and (2) penalty points, which were accumulated by including foils (i.e., steps that would impede the problem-solving process) as part of the solution sequence. Although the total score variable is a reasonable measure of the *overall* quality of subjects' solution sequences, it does not, as a summary score, provide any information regarding the independent contributions of positive points and penalty points. Given that one of the goals of the present study was to examine the determinants of age effects in practical problem solving, we believed that it was important to establish whether there were age differences in the relative number of positive and penalty points scored. To address this issue, we carried out two separate regression analyses, one that used positive points as the criterion variable and a second that used penalty points as the criterion. In both analyses, age was used as the sole predictor variable. Subjects' ages were not found to be significantly related to either positive points, $F(1, 198) = 3.50$, *ns*, or penalty points, $F(1, 198) = 1.50$, *ns*, which suggests that older and younger subjects were equally proficient at sequencing valid solution steps, and at identifying and eliminating foils.

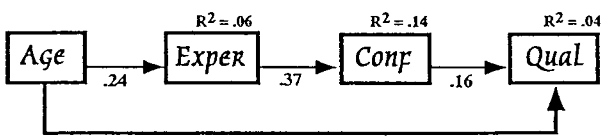

Fig. 2. Structural equation model (Model Four) showing standardized beta weights among age of subject (Age), mean prior-experience level (Exper), global confidence rating (Conf), and solution quality (Qual). Also shown are R^2 values for each endogenous variable in the model.

Time to Completion

A number of studies have shown that older adults take longer than younger adults to complete a variety of cognitive tasks (see Cerella, 1990, and Salthouse, 1985, for reviews). To determine whether age and problem-solving performance were related in the present experiment, we examined the amount of time that each subject took to complete the measure.[10] Consistent with the findings of other developmental studies, the results of this study showed age to be positively correlated with time on task, $r(178) = .25, p < .01$, with older subjects taking longer than younger individuals to complete the measure. Fig. 3 shows the bivariate scatterplot of age versus time to completion. As can be seen in the figure, the subject who required the least time completed the measure in 10 min, whereas the subject who took the most time finished in 75 min. The mean time to completion across all subjects was 27 min ($SD = 9.5$). The regression line in Fig. 3 indicates an age-related increase in the amount of time required

[10]The amount of time required to complete the test was inadvertently not recorded for 20 of the 200 subjects. Therefore, all analyses regarding the amount of time it took to complete the task are based on a reduced sample of 180 individuals.

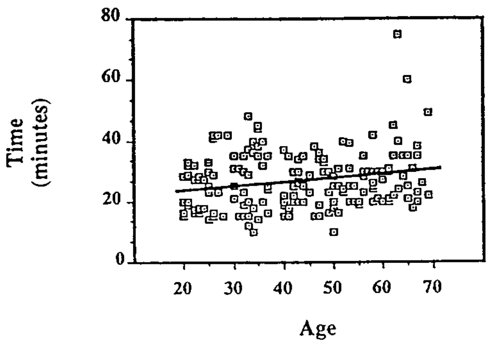

Fig. 3. Scatterplot of age versus amount of time required to complete the measure (in min).

to finish the measure. The mean predicted time to completion for a 20-year-old was approximately 23 min, whereas the mean predicted time to completion for a 69-year-old was 31 min. This 8-min difference represents a 34% increase in time on task. However, the time required to complete the measure was not found to be correlated with the quality of subjects' solutions, $r(178) = -.04$, *ns*. Additional analyses revealed that neither gender nor educational level was related to the amount of time taken to complete the test.

Willingness to Carry Out Sequenced Actions

A final set of analyses were conducted to determine whether subjects would be willing to carry out the actions that they had sequenced (recall that subjects circled any actions that they believed to be part of the solution sequence that they would not or could not personally carry out). Of the 200 individuals tested, 160 (80%) had circled at least one solution step from among the set of seven solution sequences. For each of the problems, the mean proportion of circled items to the total number of sequenced items is shown in Table IV. For ease of interpretation, these proportions are displayed in a format similar to that used in Table I and II, with the age range divided into five discrete, nonoverlapping groups. These data reveal that, on the whole, subjects were unwilling to carry out an average of 10% of the solution steps they had sequenced. A regression analysis revealed that the overall proportion of circled steps was not significantly predicted by subjects' age, $F(1, 198) = 1.28$, *ns*. The proportion of circled steps, however, was found to be negatively correlated with subjects' level of education, $r(198) = -.17$, $p < .05$, which indicates that more-educated individuals were more likely to carry out the steps that they had sequenced. It was also found that females ($M = 0.12$, $SEM =$

Table IV. Mean Proportion of Circled Items to Sequenced Items, as a Function of Age[a]

Problem	Age group[b] 20s	30s	40s	50s	60s	All subjects[b]
Auto accident	.05 (.09)	.07 (.13)	.07 (.11)	.05 (.11)	.06 (.10)	.06 (.11)
Dog bite	.11 (.15)	.10 (.18)	.14 (.15)	.13 (.18)	.09 (.14)	.11 (.16)
Gas leak	.20 (.24)	.11 (.15)	.15 (.14)	.08 (.16)	.21 (.23)	.15 (.19)
Dead phone	.13 (.20)	.11 (.13)	.13 (.15)	.10 (.16)	.19 (.17)	.13 (.17)
Woman in park	.16 (.24)	.05 (.10)	.09 (.16)	.10 (.20)	.13 (.21)	.11 (.19)
Flat tire	.07 (.21)	.05 (.14)	.07 (.21)	.02 (.06)	.20 (.30)	.08 (.21)
Missing coins	.07 (.12)	.03 (.07)	.07 (.12)	.07 (.20)	.08 (.14)	.07 (.14)
Total score	.11 (.10)	.07 (.09)	.10 (.09)	.08 (.10)	.14 (.12)	.10 (.10)

[a]All means except those in the far right row (All subjects) are based on an N of 40 individuals per age group. The mean portions reported in the body of the table were calculated by dividing the number of circled steps in a particular solution sequence by the total number of steps (i.e., valid solution steps plus foils) included in that sequence.
[b]Standard deviations are in parentheses.

.12) circled significantly more steps per problem than males did (M = 0.08, SEM = .08), $t(198) = 2.87$, $p < .01$.

DISCUSSION

Taken together, the results of the present study reveal an interesting paradox. As anticipated, older subjects were found to have had more experience than younger individuals at solving the seven practical problems. This alone was not a surprising finding, given the 30 plus years of life experiences the older subjects had over their younger counterparts. What we did find puzzling, in light of this positive relationship between age and problem-solving experience, was that younger adults were able to generate better solutions to the problems than those generated by older adults. Also puzzling was the finding that subjects' global experience scores were uncorrelated with the solution-quality index. On the basis of our theoretical model, we had anticipated that the level of prior problem-solving experience would be positively correlated with both age and solution quality, and, in turn, age would be positively related to the quality of the solutions.

A number of possible explanations exist for why the quality of subjects' solutions revealed a pattern of negative age differences. However, on the basis of the available data, it is difficult to favor one explanation over the others. First, developmental declines in basic processing resources such as working memory (Craik & Jennings, 1992; Craik, Morris, & Gick, 1990) and attention (Hartley, 1992; McDowd & Birren, 1990) could have contributed to the observed age effect. Relative to other everyday problem-solving tasks, the task used in the present study was fairly demanding from a processing resource standpoint. Each problem required subjects to encode and hold in working memory as many as nine action steps prior to determining the appropriate sequence of behaviors. This basic requirement of the task could have presented difficulties for some of the older participants in the sample.

A second possible explanation for the slight negative relationship between age and performance is that general age-related cognitive declines may be stronger than any buffering offered by domain-specific knowledge. Hypothetically, had levels of problem-solving experience been equivalent across age groups, older adults, as a result of generalized losses in basic cognitive abilities, might have performed *much more poorly* on the task than younger subjects (i.e., more poorly than they performed in the present investigation). A recent study by Clancy and Hoyer (1994) found support for the notion that knowledge can attenuate age-related performance losses on domain-relevant tasks. According to this explanation, the high levels of domain-specific experience possessed by older adults should allow them to maintain their problem-solving abilities until late in life, despite losses in processing resources. This notion of a trade-off between domain-specific crystallized abilities and fluid abilities has been discussed at length elsewhere (cf. Hershey, 1995).

Third, age-related declines in higher-order reasoning abilities could have contributed to the negative relationship between age and the quality of subjects' solutions. Our procedurally oriented task involved at least two substantial, independent reasoning components: (1) the ability to eliminate foils and (2) the ability to integrate the remaining valid action steps into an efficient and effective solution sequence. Older adults in the sample may have been disadvantaged relative to younger individuals at carrying out these tasks, given a substantial body of literature that shows age-related declines in reasoning skills (see Salthouse, 1992, for a review).

A fourth possible reason for the lack of a significant positive relationship between age and solution quality may have had to do with limited age differences in experience levels. Across the set of problems, age *was* found to be significantly positively related to experience levels. However, perhaps the relatively small magnitude of these age-related prior-experience differences limited the strength of the relationship between age and solution quality. Had the older adults in the sample been more "expert" on the set of problems, age and solution quality may have been positively related.

Fifth, it is also possible that older adults scored more poorly on the measure than younger individuals because they were less practiced at taking paper-and-pencil tests. Presumably, many of the older subjects had not taken a test of this kind in decades, whereas younger subjects would have been exposed to measures like this in the not too distant past in either high school or college.

Finally, older adults' problem-solving scripts might have interfered with their ability to do well on the task. That is, they may have found it difficult to specify optimal solution procedures if the action steps provided did not match their mental models (Gentner & Stevens, 1983) of the actions for solving the various problems. In contrast, younger adults, who would be less likely to possess rigid scripts, may have relied more on their deductive reasoning abilities than on their preexisting procedural knowledge to solve the problems.[11]

Regardless of the reason for the unanticipated observed age/performance relationship, it is well worth emphasizing that the magnitude of the association between age and solution quality was relatively small. Although the negative slope of the regression line was statistically significant, there was only a five-percentage-point difference in the average solution-quality scores for subjects at the two ends of the age range. Moreover, many of the oldest subjects' solutions were better than those of younger individuals, which suggests that increases in age are not necessarily associated with poor performance on the task. In other words, it is important to note that older subjects' solutions were not poor (in absolute terms); they were just poorer, on average (in relative terms), than those produced by younger subjects.

The theoretical model we initially proposed suggested that age should be positively related to problem-solving experience, which in turn should be positively related to the quality of individuals' solutions. Although this model was rejected as a poor fit given the data, the multivariate (LISREL) analyses did provide general support for a somewhat modified version of this model. We found a reasonable fit between the data and a model that included subjects' overall confidence level in the analysis as a mediator of the relationship between experience and solution quality. This simple addition to the model has intuitive appeal in that (a) subjects' confidence levels could be expected to increase incrementally as a function of the number of times they had experienced the problems, and (b) subjects who are more confident in their problem-solving abilities are likely to have higher levels of task-specific self-efficacy, which in turn should have a positive influence on problem-solving performance. Although self-efficacy was not specifically measured in the present study, the link between it and problem-solving performance has been well documented elsewhere (Lachman & Jelian, 1984; Pajares & Miller, 1994; Wege & Möller, 1995). The other important difference between our initially hypothesized model and Model Four (shown in Fig. 2) was the addition of an unmediated direct relation between age and solution quality. This statistically significant negative relationship suggests that age accounts for appreciable variance in problem-solving performance over and above the positive impact age has on solution quality as mediated by experience and confidence. It is this negative-age-to-solution quality relationship that we believe reflects age-linked deficits in working memory capacity and deductive reasoning abilities, difficulties associated with taking paper-and-pencil tests, and/or overly rigid solution scripts.

[11]Unfortunately, we did not collect data that would allow us to evaluate the rigidity of subjects' scripts. However, we acknowledge the need to obtain this form of data in future studies of this type.

The relationship between age and problem-solving abilities identified in the present study differed from the pattern of results found by Cornelius and Caspi (1987) and Denney and her colleagues (Denney & Palmer, 1981; Denney & Pearce, 1989; Denney, et al., 1982; Denney, et al., 1992; Heidrich & Denney, 1994). Cornelius and Caspi found modest age-related improvements in everyday problem-solving performance, using a task that required subjects to indicate how likely they would be to adopt various solution strategies. In the Denney studies—which required subjects to generate a variety of effective solutions to practical problems—the opposite effect was found. Namely, her data revealed a pattern of substantial age-related decline, particularly following middle age. In the present study—which required subjects to recall procedural knowledge subroutines to generate solutions—we found a slight negative relationship between age and problem-solving performance. Interestingly, the developmental function found in this study was intermediate to increasing and decreasing age-related functions shown in the Cornelius and Caspi (1987) and Denney studies, respectively. Taken together, the results of these studies suggest the likelihood of an Age × Task interaction, in which the developmental profile one might expect to find would depend on the precise requirements of the task at hand. The stated goal of the Denney studies, the Cornelius and Caspi study, and the present study was to determine whether there were age-related differences in everyday problem-solving abilities. However, at the level of the basic requirements of the three tasks, it is apparent that we were asking our subjects to engage in different types of processing. Therefore, it is not particularly surprising to find markedly different developmental outcomes when we look across tasks.

The lack of a significant positive relationship between subjects' self-reported experience scores and the quality of their solutions was a second unexpected finding. Perhaps we were overly optimistic to assume that repeated experiences with a problem should have led to the development of efficient and effective problem-solving scripts. Maybe the set of problems included on the measure had not been experienced frequently enough to facilitate the development of high-quality scripts. Or, as pointed out earlier, perhaps older adults' preexisting scripts interfered with their ability to generate an optimal solution. Cornelius and Caspi (1987) also found prior-experience levels to be uncorrelated with everyday problem-solving performance. In their discussion of this counterintuitive finding, they concluded that ". . . some people may experience these problems more frequently than do others, and some people may be more proficient than others in solving everyday problems. Familiarity with these problems, however, does not appear to be related to people's skill or preference in solving them" (p. 151). On the basis of our data, we have no choice but to concur with this conclusion. Perhaps one could empirically demonstrate a link between experience and practical problem-solving performance by designing a test that contains problems that vary more widely in terms of their level of familiarity (i.e., from problems that are extremely rare to those that which occur *every day*). At least two other alternative explanations could account for the lack of a significant relationship between experience and solution quality. Subjects' self-reports of prior experience may have been unreliable, which would make this a flawed marker of their true experience levels. Or, the specific set of problems that were used may not have been representative of problems that exhibit a relationship between experience and problem-solving performance. Alternatively, all three of the preceding explanations may have contributed to the observed nonsignificant effect.

Finally, it is worth noting that many of our subjects could generate reasonable solutions for the unfamiliar problems even though they had never experienced them. Most subjects had never encountered an unconscious individual, lost gold coins during a move, or had their child bitten by a stray dog. However, most were able to organize the action steps provided for these problems into fairly accurate solution sequences. This point is significant because it highlights the notion that much of one's acquired procedural knowledge can come through nonexperiential learning opportunities (e.g., through modeling, education, and second-hand stories). Or, it could also be the case that some subjects scored well on the measure, in the absence of experience, because of their strong deductive reasoning abilities.

Other analyses revealed that older subjects took significantly longer to solve the problems than did younger subjects; however, older subjects reported being no less likely than younger individuals to carry out the action steps they had sequenced. Admittedly, when we were designing the study, we assumed that we might find older adults less willing than younger persons to carry out sequenced actions, particularly for the more physically oriented problems such as changing a flat tire or turning off the natural gas in the event of a leak. This assumption, however, appears to be unfounded, at least to the extent that we can take subjects' willingness reports at face value. We were also intrigued to find that more-educated individuals reported being more likely than less-educated individuals to carry out sequenced actions, and that educational levels were positively correlated with solution quality. This latter finding can perhaps be explained by the fact that a person's level of educational attainment is typically found to be positively correlated with intelligence (Horn & Hofer, 1992; Matarazzo, 1972), and intelligence levels, in turn, have been found to be correlated with problem-solving abilities (Cornelius & Caspi, 1987; Heidrich & Denney, 1994). In addition, more-educated individuals might be more test savvy than less-educated individuals, which also might help to explain the positive relationship between educational level and solution quality.

The developmental conclusions that can be drawn from the present study are limited by the fact that cross-sectional samples were selected for analysis (Baltes, Cornelius, & Nesselroade, 1979; Schaie, 1983). In general, the negative age differences seen in cross-sectional studies typically exceed the declines found in longitudinal studies, presumably because of the influence of cohort effects (Schaie, 1994). Therefore, we speculate that the small negative correlation between age and problem solving performance might have been nonsignificant, or even positive, had a single group of subjects been studied over time. In addition, the generalizability of our findings is limited by the fact that our measure was based on a relatively small set of problems. Although other investigators have also relied on measures that contain fewer than a dozen problems (cf. the Denney studies cited previously), a stronger instrument would contain multiple problems that represent of a variety of independent problem-solving domains (e.g., health, transportation, finances, family life). Although it would have been desirable to have used a more comprehensive measure, it would have been much more difficult to administer such a test given the relatively time-consuming nature of our task. In fact, a number of our subjects complained of fatigue at having to solve what we believed to be a minimal set of seven problems.[12] Perhaps future studies of everyday problem-solving could be designed that use a more time-efficient methodology, while also capturing the essence of the relatively detailed procedural knowledge that subjects bring to the task. Finally, we acknowledge that our conclusions regarding adult age differences in everyday problem-solving are limited by the relatively young age range of our sample. A maximum subject age of 69 years might be considered "young-old" by some cognitive aging researchers; it is unclear how "old-old" individuals (i.e., those older than 70 years) would have faired on our measure.

We suggested in the introduction that previous studies of everyday cognition might have underestimated the competence of older adults by ignoring the relationship between procedural knowledge and problem-solving abilities. Our goal, in designing the task used in the present study, was to develop a problem-solving instrument that would allow us to empirically establish a link between life-span experiences and problem-solving competency. In pursuing this goal, we employed a methodology that focused on the more molecular aspects of the problem-solving process by analyzing the specific actions that subjects would engage in to reach a solution.[13] We consider this approach to be a unique contribution to the literature in that it represents a significant departure from the more molar methodologies used by other researchers.

Although it has been suggested that psychologists should discontinue the search for a general nomothetic function that represents everyday problem-solving abilities (Denney, 1989), we believe that it still may be possible to identify such functions. In attempting to do so, investigators will need to pay close attention to the specific types of abilities required by particular tasks. That is, the possible existence of an Age $\times$ Task interaction, suggested by different developmental functions across different tasks, indicates that the nature of the developmental effect one can expect to find will be determined by the types of cognitive operations subjects are required to carry out. This being the case, we believe it would be fruitful to identify the developmental trends associated with a variety of everyday problem-solving tasks. In particular, multitask studies (cf. Marsiske & Willis, 1995) could provide valuable insights into the types of situations in which individuals of different ages would be prone to failure or success.

[12]In pilot work, we tested a version of the instrument that contained nearly twice as many problems; however we found that an unacceptably large number of subjects were unwilling to complete such a lengthy measure.

[13]One issue we were unable to examine using the methodology we adopted, however, was that of individual differences in problem-solving style. In a previous study, Blanchard-Fields, Jahnke, and Camp (1995) identified age differences in the style individuals adopted to solve ill-defined interpersonal problems. Future studies might profitably explore individual differences in problem-solving style by using ill-defined procedurally oriented problems (as opposed to the well-defined problems used in the present investigation).

APPENDIX A: DESCRIPTIONS OF THE PROBLEM SCENARIOS

1. You are running errands downtown one afternoon. When entering an intersection your car is hit by another car. Both vehicles sustain damage, and your car can no longer be driven. What should you do?
2. You are baby-sitting a friend's child when the child is severely bitten by a neighborhood dog. What should you do?
3. You return home from an evening out and you smell the distinctive odor of natural gas upon entering your home. The odor is strongest near the stove. What should you do?
4. You pick up your telephone to call a friend, but you do not hear a dial tone. What should you do?
5. You are walking in the park when you notice a woman sitting on a nearby bench fall to the ground, unconscious. What should you do?
6. You are taking a drive in the country one morning when you hit a broken bottle and get a flat tire. Help is unavailable. What should you do?
7. You hired a moving company to pack and transport your belongings to a city 300 miles away. Upon unpacking, you realize that a collection of gold coins is missing. What should you do?

APPENDIX B: SAMPLE COMPLETED PROBLEM (THE GAS LEAK PROBLEM)

[Foils have been lined out, and the remaining valid solution steps have been ordered in the proper sequence.]

Instructions: (1) Read entire problem. (2) Cross out things you should not do. (3) Number remaining steps in correct order. (4) Circle the number for steps you would not carry out. (5) Answer the experience-level and confidence-level questions at bottom of page.

Upon returning home from an evening out, you smell the distinctive odor of natural gas. The odor seems to be strongest in the kitchen, near the gas stove. What should you do?

___ ~~Press and hold down the pilot light reset button.~~
___ ~~Turn off the incoming gas supply at the meter located outside the house.~~
_3_ Locate the gas shut-off valve behind or under the stove.
_2_ Check to see if the pilot lights have gone out.
_4_ Turn the valve stem so that it is perpendicular to the gas line.
_5_ Contact a plumber or the gas company in order to arrange for repairs.
_1_ Ventilate the area by opening a door or a window.
___ ~~Turn the top burners on and off to determine if they are all working.~~

Have you personally ever had to deal with this problem (or a very similar problem) in the past? Check yes or no. If yes, indicate approximately how many times.
NO ___
YES ___ ⟶ Approximately how many times? ___

How confident are you that the numbering sequence you have specified above is the best course of action? (circle one number)

1	2	3	4	5	6	7
Not at All Confident			Somewhat Confident			Extremely Confident

ACKNOWLEDGMENTS

The authors are grateful to Jo Ann Wilson and Greg Keyser for their help in designing the test and to Rene Hines, Vicki Shaffer, and Marti Anderson for their untiring assistance in collecting the data. We also acknowledge the immense amount of time and effort Vicki Shaffer contributed to the project in scoring the raw data, and Jason Allaire's assistance with the data analysis.

REFERENCES

Baltes, P. B. (1993). The aging mind: Potential and limits. *The Gerontologist, 33,* 580–594.

Baltes, P. B., Cornelius, S. W., & Nesselroade, J. R. (1979). Cohort effects in developmental psychology. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 61–87). New York: Academic Press.

Blanchard-Fields, F., Jahnke, H. C., & Camp, C. (1995). Age differences in problem-solving style: The role of emotional salience. *Psychology and Aging, 10,* 173–180.

Camp, C. J., Doherty, K., Moody-Thomas, S., & Denney, N. W. (1989). Practical problem solving in adults: A comparison of problem types and scoring methods. In J. D. Sinnott (Ed.), *Everyday problem solving* (pp. 211–228). New York: Praeger.

Cerella, J. (1990). Aging and information-processing rate. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 201–221). New York: Academic Press.

Charness, N. (1989). Age and expertise: Responding to Talland's challenge. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 437–456). New York: Cambridge University Press.

Clancy, S. M., & Hoyer, W. J. (1994). Age and skill in visual search. *Developmental Psychology, 30,* 545–552.

Cornelius, S. W., & Caspi, A. (1987). Everyday problem solving in adulthood and old age. *Psychology and Aging, 2,* 144–153.

Craik, F. I. M., & Jennings, J. M. (1992). Human memory. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 51–110). Hillsdale, NJ: Erlbaum.

Craik, F. I. M., Morris, R. G., & Gick, M. L. (1990). Age differences in working memory. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 247–267). Cambridge, England: Cambridge University Press.

Denney, N. W. (1989). Everyday problem-solving: Methodological issues, research findings, and a model. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 330–351). New York: Cambridge University Press.

Denney, N. W., & Palmer, A. M. (1981). Adult age differences on traditional and practical problem solving measures. *Journal of Gerontology, 36,* 323–328.

Denney, N. W., & Pearce, K. A. (1989). A developmental study of practical problem solving in adults. *Psychology and Aging, 4,* 438–442.

Denney, N. W., Pearce, K. A., & Palmer, A. M. (1982). A developmental study of adults'x performance on traditional and practical problem solving tasks. *Experimental Aging Research, 8,* 115–118.

Denney, N. W., Tozier, T. L., & Schlotthauer, C. A. (1992). The effect of instructions on age differences in practical problem solving. *Journal of Gerontology: Psychological Sciences, 47,* 142–145.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49,* 725–747.

Gentner, D. R., & Stevens, A. L. (1983). *Mental models.* Hillsdale, NJ: Erlbaum.

Hartley, A. A. (1992). Attention. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 3–49). Hillsdale, NJ: Erlbaum.

Heidrich, S. M., & Denney, N. W. (1994). Does social problem solving differ from other types of problem solving during the adult years? *Experimental Aging Research, 20,* 105–126.

Hershey, D. A. (1995). Influence of age and gender on estimates of long-term financial growth functions. *Aging and Cognition, 2,* 231–250.

Hershey, D. A., Walsh, D. A., Brougham, R., Carter, S., & Farrell, A. (1998). Challenges of training pre-retirees to make sound financial planning decisions. *Educational Gerontology, 24,* 447–470.

Hershey, D. A., Walsh, D. A., Read, S. J., & Chulef, A. S. (1990). The effects of expertise on financial problem solving: Evidence for goal directed problem solving scripts. *Organizational Behavior and Human Decision Processes, 46,* 77–101.

Hershey, D. A., & Wilson, J. A. (1997). Age differences in performance awareness on a complex financial decision making task. *Experimental Aging Research, 23,* 257–273.

Hershey, D. A., Wilson, T. L., & Mitchell-Copeland, J. (1996). Conceptions of the psychological research process: Script variation as a function of training and experience. *Current Psychology, 14,* 293–312.

Horn, J. L., & Hofer, S. M. (1992). Major abilities and development in the adult period. In R. J. Sternberg & C. A. Berg (Eds.), Intellectual development (pp. 44–99). Cambridge, England: Cambridge University Press.

Hoyer, W. J., Rybash, J. M., & Roodin, P. A. (1989). Cognitive change as a function of knowledge access. In M. L. Commons, J. D. Sinnott, R. A. Richards, & C. Armon (Eds.), *Adult development: Vol. 1. Comparisons and applications of developmental models* (pp. 293–305). New York: Praeger.

Jöreskog, K. & Sörbom, D. (1993). *LISREL VIII user's reference guide.* Mooresville, IN: Scientific Software.

Lachman, M. E., & Jelian, E. (1984). Self-efficacy and attributions for intellectual performance in young and elderly adults. *Journal of Gerontology, 39,* 577–582.

Marsiske, M., & Willis, S. L. (1995). Dimensionality of everyday problem solving in older adults. *Psychology and Aging, 10,* 269–283.

Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). New York: Oxford University Press.

McDowd, J. M., & Birren, J. E. (1990). Aging and attentional processes. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 222–234). New York: Academic Press.

Meacham, J. A., & Emont, N. C. (1989). The interpersonal basis of everyday problem solving. In J. D. Sinnott (Ed.), *Everyday problem solving* (pp. 7–23). New York: Praeger.

Pajares, F., & Miller, M. D. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem solving: A path analysis. *Journal of Educational Psychology, 86,* 193–203.

Poon, L. W., Rubin, D. C., & Wilson, B. A. (1989). *Everyday cognition in adulthood and late life.* New York: Cambridge University Press.

Puckett, J. M., & Reese, H. W. (1993). *Mechanisms of everyday cognition.* Hillsdale, NJ: Erlbaum.

Puckett, J. M., Reese, H. W., & Pollina, L. K. (1993). An integration of life-span research in everyday cognition: Four issues. In J. M. Puckett & H. W. Reese (Eds.), *Mechanisms of everyday cognition* (pp. 3–16). Hillsdale, NJ: Erlbaum.

Rebok, G. W. (1989). Plans, actions, and transactions in solving everyday problems. In J. D. Sinnott (Ed.), *Everyday problem solving* (pp. 100–122). New York: Praeger.

Salthouse, T. A. (1985). Speed of behavior and its implications for cognition. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 400–426). New York: Academic Press.

Salthouse, T. A. (1992). Reasoning and spatial abilities. In F. I. M. Craik & T. A. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 167–211). Hillsdale, NJ: Erlbaum.

Schaie, K. W. (1983). What can we learn from the longitudinal study of adult psychological development? In K. W. Schaie (Ed.), *Longitudinal studies of adult psychological development* (pp. 1–19). New York: Guilford Press.

Schaie, K. W. (1994). The course of adult intellectual development. *American Psychologist, 49,* 304–313.

Simon, H. A. (1973). The structure of ill-structured problems. *Artificial Intelligence, 4,* 181–201.

Sinnott, J. D. (1989). *Everyday problem solving: Theory and applications.* New York: Praeger.

Walsh, D. A., & Hershey, D. A. (1993). Mental models and the maintenance of complex problemsolving skills into old age. In J. Cerella, J. Rybash, W. Hoyer, & M. L. Commons (Eds.), *Adult information processing: Limits on loss* (pp. 553–584). New York: Academic Press.

Wege, J. W., & Möller, A. T. (1995). Effectiveness of a problemsolving training program. *Psychological Reports, 76,* 507–514.

Willis, S. L., Jay, G. M., Diehl, M., & Marsiske, M. (1992). Longitudinal change and prediction of everyday task competence in the elderly. *Research on Aging, 14,* 68–91.

Willis, S. L., & Marsiske, M. (1991). Life span perspective on practical intelligence. In D. E. Tupper & K. D. Cicerone (Eds.), *The neuropsychology of everyday life: Issues in development and rehabilitation,* (pp. 183–197). Boston: Kluwer Academic Publishers.

Willis, S. L., & Schaie, K. W. (1986). Practical intelligence in later adulthood. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 236–268). New York: Cambridge University Press.

Willis, S. L., & Schaie, K. W. (1993). Everyday cognition: Taxonomic and methodological considerations. In J. M. Puckett & H. W. Reese (Eds.), *Mechanisms of everyday cognition* (pp. 33–53). Hillsdale, NJ: Erlbaum.