Psychology and Aging
1995, Vol. 10, No. 2, 269–283

0882-7974/95/$3.00

Dimensionality of Everyday Problem Solving in Older Adults

Michael Marsiske
Max Planck Institute for Human Development and Education

Sherry L. Willis
Pennsylvania State University

This study investigated individual differences in older adults' everyday problem-solving performance using 3 instruments. Past research, typically using only single measures, has yielded a multitude of findings regarding age effects in everyday problem solving. The present sample consisted of 111 older adults (44 men, 67 women) who ranged in age from 68 to 94 years. Confirmatory factor analyses revealed that, within each of the 3 instruments, subscales representing particular content domains could be reliably identified. There was, however, little relation between the different instruments, and the measures also differed in their relation with chronological age. These results support the view that everyday problem-solving competence is a multidimensional construct, of which previous investigations may only have studied particular dimensions.

The goal of this study was to investigate how older adults' performance on everyday problem-solving tasks varies across methods of assessment. An increasing number of studies have focused on older adults' functioning on cognitive tasks of daily living (Hess, 1990; Poon, Rubin, & Wilson, 1989; Rogoff & Lave, 1984; Sinnott, 1989; Sinnott & Cavanaugh, 1991; Sternberg & Wagner, 1986), but most studies have considered performance only in delimited task domains, using single measures of performance. In addition, there has been little theoretical or empirical integration across studies of adult everyday problem solving. Several reasons exist for this. First, as a clearly identifiable subfield of inquiry, research into the everyday cognition of older adults is relatively new (Woodruff-Pak, 1989). Second, little consensus has been reached on what the defining task properties of everyday cognition might be and what the best methods for assessing these properties are. Indeed, even a standard nomenclature for distinguishing everyday cognition from other kinds of cognition does not yet exist; terms such as *practical problem solving* (Denney, 1989), *everyday problem solving* (Cornelius & Caspi, 1987), *everyday cognition* (Poon et al., 1989), *pragmatics of intelligence* (P. B. Baltes, 1987), and *practical intelligence* (Sternberg & Wagner, 1986), among others, have all been used. This diversity of labels, not surprisingly, reflects the even larger diversity of approaches that have evolved for the measurement of these constructs.

The heterogeneity of measurement means that two important issues remain unresolved in the literature on everyday problem solving. A first issue concerns the nature of everyday problem solving: What tasks best measure everyday problem solving at different points in adulthood, and how is performance on these tasks related to performance on other tasks? Use of single measures without reference to the latent dimensions they represent or an understanding of what aspects of everyday cognition are being assessed provides little information on how the findings of individual studies relate to one another. In the long run, psychological constructs such as everyday problem solving can only be validated through their convergent confirmation by independent measurement procedures (e.g., Campbell & Fiske, 1959) and by their appearance in multiple populations (e.g., Cattell, 1964; Horn & McArdle, 1992; Labouvie, 1980).

Theoretically speaking, should different measures of everyday problem solving be expected to correlate? Research on the development of everyday expertise suggests that everyday prob-

Michael Marsiske, Max Planck Institute for Human Development and Education, Berlin, Germany; Sherry L. Willis, Department of Human Development and Family Studies, Pennsylvania State University.

This article is based on the doctoral research of Michael Marsiske, which was conducted under the direction of Sherry L. Willis.

The research reported in this study was supported by Grant RO1-AG-08082 from the National Institute on Aging and one of the authors was partially supported by National Institute on Aging Training Grant T32-AG-00048. The writing of this article was supported by a postdoctoral research fellowship from the Center for Psychology and Human Development of the Max Planck Institute for Human Development and Education.

Thanks are expressed to Nancy W. Denney, University of Wisconsin—Madison, and Steven W. Cornelius, Cornell University, for permission to include their measures in this study. Thanks also to Ann Horgas, Barbara Klingenspor, Ulman Lindenberger, Heiner Maier, and Ulrich Mayr for their comments on earlier drafts of this article. We gratefully acknowledge the help and cooperation of the students and staff of the Adult Development and Enrichment Project study at many points during this research. We also extend thanks to the other members of the Doctoral Committee, without whom this research could not have been completed: Bonnie J. F. Meyer, K. Warner Schaie, and Steven H. Zarit of Pennsylvania State University. Gratitude is also expressed to John R. Nesselroade, who is now at the University of Virginia, for his advice in the early phases of this research.

Correspondence concerning this article should be addressed Sherry L. Willis, Department of Human Development and Family Studies, S-110 Henderson Building, Pennsylvania State University, University Park, Pennsylvania 16802, or Michael Marsiske, Max Planck Institute for Human Development and Education, Lentzeallee 94, 14195 Berlin, Germany. After July 31, 1995, Michael Marsiske's address will be the Institute for Gerontology, Wayne State University, 87 East Ferry Street, 226 Knapp Building, Detroit, Michigan 48202. Electronic mail should be sent via Internet to slw@psuvm.psu.edu or to marsiske@geroserver.iog.wayne.edu (after July 31, 1995).

lem-solving competence might be highly domain specific (Ceci & Liker, 1986; Charness, 1985; Ericsson & Smith, 1991; Lindenberger, Kliegl, & Baltes, 1992; Salthouse, 1991a). This implies that performance should be correlated across problem-solving measures only if the measures draw on shared knowledge domains. From this perspective, it may be most profitable to view everyday cognition as a more multidimensional construct than the use of global construct labels such as *practical problem solving* suggests. A multidimensional view of cognitive abilities has a long tradition in psychometric conceptions of intelligence (e.g., Cattell, 1971; Guilford, 1959; Thurstone, 1962) and is a key proposition of life span developmental psychology (P. B. Baltes, 1987; P. B. Baltes & Willis, 1982). Extending the multidimensionality conception to everyday problem solving means that increased attention would need to be focused on articulating exactly which aspects of everyday cognition are being studied.

A second unresolved issue in the literature on adult everyday problem solving concerns developmental trajectories. Despite over 7 decades of research on the aging of basic intelligence (e.g., Dixon, Kramer, & Baltes, 1985; Woodruff-Pak, 1989), little is known about how generalizable the findings from this literature are to the aging of everyday cognition. As Rabbitt (1977) and Salthouse (1990) noted, there is an intuitive paradox between older adults' reported difficulties with many laboratory tasks of cognition and the apparent efficacy with which they manage their everyday lives. The first step in resolving this paradox is to more clearly understand the kinds of tasks which do and do not evince performance disadvantages for older adults.

There are two major sets of theoretical predictions about the effects of age on everyday task performance. One theoretical approach is grounded in an expertise-based view of adult cognition (e.g., Charness, 1985; Ericsson & Smith, 1991; Salthouse, 1991b). An articulation of this perspective can be found in the work of P. B. Baltes and his colleagues (e.g., P. B. Baltes, 1987, 1993; Baltes & Baltes, 1990; Dixon & Baltes, 1986; Marsiske, Lang, Baltes, & Baltes, in press; Staudinger, Marsiske, & Baltes, 1993). They hypothesized that older adults may select and be selected into everyday contexts that are largely routine and predictable, and they argued that the pragmatic demands of daily life require individuals to draw largely on their accumulated knowledge systems. These bodies of knowledge may show maintenance or even selective growth throughout the adult life span (e.g., Cattell, 1971; Horn & Hofer, 1992), even in the face of underlying losses in basic information-processing capabilities. Thus, the emergent prediction from some life span contextual models is that the normative developmental trajectory of everyday competence, as suggested by patterns of age-cohort differences, could be one of stability or increment under conditions of appropriate social-cultural support (e.g., P. B. Baltes, 1993). In a related vein, Denney (1989) argued that regularly practiced abilities (i.e., those frequently encountered in daily life) will decline relatively late in the adult life span because exercise or training abates age-related losses, perhaps by automatizing some performance components (see also Berg & Sternberg, 1985).

A second set of theoretical predictions about the age trajectory of everyday cognition may emerge if one views practical problem solving as a complex, "compiled" form of cognition (Salthouse, 1990). In this view, performance on practical tasks is the outcome of one's standing on the particular admixture of abilities needed for a particular task, including domain-specific knowledge acquired through experience and practice (e.g., Berry & Irvine, 1986; Marsiske & Willis, in press; Willis & Schaie, 1986). The implication of such a view is that the same multidirectionality of change that characterizes aging on measures of basic abilities (P. B. Baltes, 1987) should characterize the aging of everyday problem solving.

The empirical findings on older adults' everyday problem-solving capabilities have not yet consistently supported either theoretical perspective. Rather, the research literature contains a diversity of findings regarding age differences (see Smith & Baltes, 1990; Staudinger, Smith, & Baltes, 1992; Sternberg & Wagner, 1986). Early research found that stable or positive cross-sectional gradients characterized the developmental trajectories obtained with adult-relevant tests of everyday information (Demming & Pressey, 1957; Gardner & Monge, 1977). Similarly, and more recently, incremental age trends were also reported by Cornelius and Caspi (1987) for their measure of everyday problem solving in a sample of 20- to 78-year-olds. A positive age trajectory has not been found, however, in many other studies.

Indeed, a close examination of the literature suggests that obtained patterns of age-cohort differences may be relatively measure dependent. For example, a rather different cross-sectional pattern emerged from Denney's laboratory (e.g., Denney & Palmer, 1981; Denney & Pearce, 1989; Denney, Pearce, & Palmer, 1982; Denney, Tozier, & Schlotthauer, 1992). Over several age-comparative studies with participants ranging from 20 to 80 years of age, Denney and her colleagues most commonly reported that peak performances were consistently observed in middle-aged participants. Denney (1989) noted, however, that different developmental functions might be obtained for different types of everyday problems. In one study, for example, when problems were designed to be relevant to younger adults, they performed better than middle-aged or older adults (Denney et al., 1982). A similar "age-match" effect was reported by P. B. Baltes, Smith, and Staudinger (1992) in their research on the everyday cognitive domain of wisdom. The emerging picture, then, is one of substantial heterogeneity in expected age trends for everyday problem solving. Both the specific content of the task and its appropriateness for one's position in the life span seem to regulate obtained age trajectories. Supporting this assertion is one of the few studies that explicitly drew on a multidimensional conception of everyday tasks. In this study, Hartley (1989) found different relationships with age and cohort for each of three tasks investigated in a sample of adults 19 to 84 years of age.

Knowledge about the developmental trajectories of everyday problem solving has also been limited by the relatively little explicit attention paid to the late life span. Few studies have examined the relationship between age and problem solving in participants older than 80 years, despite the potential importance of such research (see also P. B. Baltes, Mayer, Helmchen, & Steinhagen-Thiessen, 1993). Practically, age differences between the young-old and old-old may have important policy implications because the old-old are more likely to be frail and in need of assistance (A. B. Ford et al., 1988; Neugarten, 1984; Palmore, Nowlin, & Wang, 1985) and problem solving may be an important component of everyday competence or the ability to remain independent in the community (Willis, 1991).

Work by Willis and her colleagues provides some insight into the developmental trajectories of everyday problem solving in very late life, as well as the interindividual heterogeneity of such trajectories, by using one kind of everyday task. In one of the few longitudinal investigations of everyday problem-solving performance, Willis and colleagues studied age changes in performance on the Test of Basic Skills (Educational Testing Service, 1977), a measure of problem solving using everyday printed materials, over a 7-year period. Retested participants ranged from 69 to 93 years of age, and most had moved from the young-old to old-old age category during the course of the study. There was significant mean decline at the group level, although the majority (57%) of participants showed no reliable intraindividual decline over the longitudinal interval. Intraindividual losses became normative only for participants who were moving into advanced old age (i.e., participants who were moving from a mean of 77 years of age to 84 years of age; Willis, Jay, Diehl, & Marsiske, 1992; Willis & Marsiske, 1991). Moreover, in another study, individual differences in problem-solving performance with everyday printed materials were found to be related to observed performance on selected critical tasks of daily living (Diehl, Willis, & Schaie, in press).

Thus, by highlighting multiple-measurement strategies and inconsistent age trends, the foregoing discussion suggests that there are a number of unresolved issues in the literature on everyday problem solving in older adults. The present study examined these issues by addressing three questions: (a) What domains or subscales of everyday problems solving are contained in each of the three measures? Do individual measures of everyday cognition represent an internally consistent set of content domains, or are their subscales relatively unrelated to one another? (b) What is the convergence among three methods of everyday task-performance assessment in older adults? Do different everyday problem-solving instruments relate to one another? Can shared second-order constructs be identified? (c) Are similarities and differences among the measures also revealed by their patterns of relationship with age–cohort?

Performance was assessed on three measures of everyday problem solving that are reported in the research literature: the Practical Problems (PP) question list of Denney and Pearce (1989), the Everyday Problem Solving Inventory (EPSI; Cornelius & Caspi, 1987), and Willis and Marsiske's (1993) Everyday Problems Test (EPT). These instruments were selected from the broader array of available tests for two major reasons: (a) They were all designed to be relevant for, and used with, older adult samples, and (b) each measure was thought to yield a relatively broad sample from the possible categories of everyday tasks.

Method

Sample

The sample comprised 111 individuals: 44 men and 67 women. The mean educational level of this sample was 15.22 years (SD = 2.43 years, range = 7–22 years); the mean age was 77.78 years (SD = 5.64 years, range = 68–94 years). The average ratings of general health, vision, and hearing on a six-point Likert-type scale (1 = *very good* to 6 = *very poor*) were health self-rating = 1.80 (SD = 0.88), vision self-rating = 2.22 (SD = 1.09), and hearing self-rating = 2.30 (SD = 1.14). Average annual income was $27,000 (range = $4,000–$50,000). Age was not significantly related to demographic (education and income) or health (health, hearing, and vision) indicators in this sample ($p > .05$); consequently, no mean differences were observed between young-old (68–75 years of age) and old-old (76–94 years of age) participants for these variables ($p > .05$). All participants lived independently in a life-care retirement community in Florida. For their participation, participants were paid $30 plus a $10 completion bonus.

Measures

Practical Problems (PP) Test. This measure was an adaptation of questions administered by Denney and Pearce (1989). Ten everyday problems (e.g., getting the lawn mowed while having a heart condition or maintaining social activities after widowhood), designed to be age relevant, were presented to participants. The participants' task was to generate as many safe and effective solutions as possible for each problem. Two judges rated the effectiveness (1 = *effective*, 0 = *ineffective*) of each solution generated by participants, according to guidelines and feedback provided by N. W. Denney (personal communication, February 1991). Across all 10 problems and all participants, the interrater reliability of effectiveness scores in the present study was comparable to that reported by Denney and Pearce: overall percentage agreement was 99%, and Cohen's kappa was .89. Performance of the present sample on the written form of the test was higher than that reported by Denney and Pearce. In the original study (Denney & Pearce, 1989), participants ages 60–69 had a mean test score of 25.5 (SD = 4.5), and participants ages 70–79 had a mean score of 24.7 (SD = 3.2). When we used the identical scoring method as that used in the Denney and Pearce study, participants in this study had a mean score of 34.8 (SD = 4.7), suggesting that the present sample may have been advantaged on the PP relative to the original Denney and Pearce sample.

N. W. Denney (personal communication, February 1991) recommended that the scoring of the PP used in Denney and Pearce's (1989) study be modified for the present study. In the original study, the number of solutions generated for each problem was rescaled from 1 to 4 (1 = *no safe and effective solutions generated* and 4 = *four or more safe and effective solutions generated*). For the present study, Denney recommended that the absolute number of safe and effective solutions generated be used as participants' scores for each problem. This revised scoring eliminated the problem of ceiling effects (maximum score under the original scoring method was 40), thereby introducing a broader band of individual differences (M = 50.2, SD = 18.9).

In the present study, to permit group testing and self-administration of the PP and in an attempt to reduce method differences between the PP and other measures included in this study, Denney and Pearce's (1989) oral interview format was changed to a paper-and-pencil questionnaire, with questions and instructions kept identical. In addition, because Denney and Pearce's testers also verbally prompted participants during breaks in their solution generation to make sure they had generated as many solutions as they could, these prompts were also built into the top and bottom of every page of the paper-and-pencil version. Thus, every attempt to minimize discrepancies between the oral and written formats of the PP was made.

Everyday Problem Solving Inventory (EPSI): Situational decision making. The EPSI (Cornelius & Caspi, 1987) presented participants with 48 hypothetical problem situations that represented six content domains: (a) consumerism (CONS), (b) complex and technical information management (INFO), (c) home management (HOME), (d) family conflict resolution (FAMILY), (e) resolution of conflicts with friends (FRIEND), and (f) resolution of conflicts with coworkers (WORK). The authors included subsets of problems that were identified as being differentially relevant for different age groups (young, middle-aged, and older adults). For each problem situation, four possible responses were provided to participants. These responses were designed to represent four coping styles: (a) problem-focused action, (b) cognitive problem analysis, (c) passive-dependent behavior, and (d) avoidant thinking and denial. Participants were instructed to imagine that they were in the described situations and were asked to rate the likelihood that they would act in each of the four response modes provided for each situation. Ratings were made on a 5-point scale (1 = "*definitely would not do*" and 5 = "*definitely would do*"). Participants did not simply select one response to describe how they might deal with the problem situation but rated each possible response, thereby producing four ratings for each problem situation.

In this study, the EPSI was scored according to the procedure outlined by Cornelius and Caspi (1987): Participant responses were correlated with judges' ratings (also provided by Cornelius & Caspi) of the effectiveness of each problem solution. Participants' scores, then, represented the degree to which their response patterns approximated optimal response patterns that were identified by judges. Separate scores were obtained for each problem domain as well as for the total measure.

With regard to the psychometric characteristics of the measure, the boosted Spearman-Brown split-half reliabilities of the EPSI in this study were somewhat lower than those reported by Cornelius and Caspi (1987: their coefficients ranged from .53 [CONS] to .77 [INFO], and the total measure coefficient was .92): Reliabilities for the problem-solving domains in the present study ranged from .23 (CONS) to .50 (HOME), and the split-half reliability of the total measure was .78. Unlike the Cornelius and Caspi study, however, the present study included a less age-heterogenous sample of older adults.

Everyday Problems Test (EPT). The EPT (Willis & Marsiske, 1993) was designed to assess older adults' ability to solve problems involving everyday printed materials dealing with seven domains of daily living: food preparation (FOOD), medication use and health behaviors (HLTH), telephone use (PHON), shopping and consumerism (CONS), financial management (FINA), housekeeping and laundry (HOUS), and transportation (TRAN). The domains that were investigated were consistent with the Instrumental Activities of Daily Living Scale (IADL; Lawton & Brody, 1969). In the EPT, the participant was presented with a stimulus (e.g., prescription drug label) and asked to solve two practical problems associated with that stimulus (e.g., calculating how many pills to take over a 2-day period). Stimuli were selected to be age relevant (see Diehl, Willis, & Schaie, 1990, for further details). The EPT had 84 items, and each of the seven scales comprised 12 items. Elsewhere, with a larger sample, Willis and Marsiske (1994) reported that standardized alpha reliabilities of the scales ranged from .62 to .74, and the alpha reliability of the total measure was .94. Furthermore, they found that 1-year, Spearman-Brown test-retest stabilities of the EPT ranged from .74 to .82 for the individual scales, and the 1-year stability of the total measure was .91.

Procedure

Participants took part in two 3-hr testing sessions and also received a homework packet after each testing session. The EPT (Willis & Marsiske, 1993) was administered within the testing sessions, although participants who could not complete all items within the sessions were allowed to finish at home. The EPSI (Cornelius & Caspi, 1987) and the PP (Denney & Pearce, 1989), along with a set of personal and demographic questions, were included as part of a take-home packet. Participants certified that they completed all take-home measures without assistance, and the veracity of this claim was supported empirically: A comparison of mean proportion of correct items for those EPT items done at home versus those done in session revealed no significant differences. All of the measures were administered under untimed conditions.

Results

This study explored the similarities and differences among three measures of everyday problem solving. The results are presented with regard to three major questions: (a) What is the pattern of relationships within measures? What is the dimensionality of individual problem-solving measures? (b) What is the pattern of relationships across measures? Are there latent constructs common to the different measures and scales of everyday problem solving? (c) What is the pattern of age differences in everyday problem-solving performance?

Overview of Model Evaluation

All variables and subscales were standardized into *T*-score metric, with a mean of 50 and a standard deviation of 10. The raw data were checked to examine whether an assumption of multivariate normalcy could be held. Congruent with this assumption, kurtosis estimates fell between 1 and −1 for most of the variables and subscales considered, and only three of the subscales investigated in this study had kurtosis estimates exceeding 2 or −2. Thus, the normalized estimate of Mardia's coefficient of multivariate kurtosis was 1.05. Summary statistics reflecting the skewness, kurtosis, as well as internal consistency reliability of each of the major subscales used in this study are presented in the Appendix. The Appendix also provides the matrix of correlations among variables and subscales used in these analyses.

In this study, all measurement models were tested by the LISREL VIII program and were evaluated on the basis of overall fit indices: goodness of fit index (GFI), adjusted goodness of fit index (AGFI), and standardized root-mean-square residual (RMSR) (Jöreskog & Sörbom, 1993); comparative fit index (CFI) and normed fit index (NFI) (Bentler, 1989; Bentler & Bonett, 1980; Marsh, Balla, & McDonald, 1988); as well as examination of individual parameter estimates provided by the program. Congruent with the work of Akaike (1987) and Carmines and McIver (1981), the ratio of a model's chi-square (χ^2) to its degrees of freedom was also examined; chi-square values less than twice the degrees of freedom were also considered suggestive of acceptable fit. In addition, to compare the fit of nested models, the significance of the difference in likelihood chi-square ratios was tested. Although all models were conducted in covariance metric, results represent standardized solutions.

Dimensionality of Problem-Solving Measures

Practical Problems Test. In previous work with this instrument, Denney and Pearce (1989) reported only a global score for their measure, which implied that all items in the test were equally representative of one underlying construct. To test this assumption, Model PP-1 tested the fit of a general-factor solution to the data (number of effective solutions generated for each of the 10 presented problems) using LISREL VIII

Table 1
Two-Factor Solution for the Practical Problems Test

Problem	Factor loadings: Instrumental	Social/Activity	Unique variance
1	.57		.68
2	.73		.47
3	.76		.42
4		.66	.56
5		.69	.53
6		.93	.14
7	.62		.61
8	.53		.72
9	.62		.62
10		.78	.38

Note. The correlation between these two factors was $r = .82$. Problems 4 and 5 were allowed to have a freely estimated correlated residual of $r = .26$. Practical Problems Test (Denney & Pearce, 1989).

(Jöreskog & Sörbom, 1993). A second model, Model PP-2 was also tested, on the basis of inspection of PP item content. Congruent with other work conducted by Denney and colleagues (e.g., Camp, Doherty, Moody-Thomas, & Denney, 1989), a two-factor representation was tested. One factor contained problems of a social nature (e.g., meeting people in a new town), whereas the second factor was composed of problems of an instrumental nature (e.g., getting groceries in bad weather).

The fit of the general factor Model PP-1 was fairly poor, $\chi^2(35, N = 111) = 79.56, p < .001$ (GFI = .87, AGFI = .80, RMSR = .063, NFI = .85, and CFI = 0.90). Examination of the two-factor Model PP-2 revealed that the fit of this second model was better, $\chi^2(34, N = 111) = 54.81, p < .01$ (GFI = .91, AGFI = .86, RMSR = .048, NFI = .89, and CFI = 0.96), and represented a significant improvement in fit over the general-factor model, $\chi^2(1, N = 111) = 24.75, p < .001$. Although the fit of Model PP-2 was relatively good, both the significant chi-square statistic and inspection of the matrix of standardized residuals suggested room for improvement. Specifically, the results revealed the presence of one, very high residual correlation between two indicators of the Social factor. Therefore, Model PP-3 was tested, in which the residuals of these two indicators were allowed to correlate. The relaxation of this constraint yielded a model with excellent fit, $\chi^2(33, N = 111) = 32.24, p > .05$ (GFI = .95, AGFI = .91, RMSR = .042, NFI = .94, and CFI = 1.00), a significantly better fit than Model PP-2 without correlated residuals, $\chi^2(1, N = 111) = 22.57, p < .001$. The correlated residual between Problems 4 and 5 of the Denney and Pearce (1989) PP was .26 ($z = 3.97, p < .001$). This final two-factor model, Model PP-3, is presented in standardized metric in Table 1 and shows that loadings were generally high (on the instrumental variable they were all above .53, whereas on the social/activity variable, they were all above .66). The correlation among factors was .82, which indicated the presence of a substantial relationship between the two dimensions.[1]

Everyday Problem Solving Inventory. Although Cornelius and Caspi (1987) reported the existence of six problem-solving domains and four coping styles (referred to as "modes of response" in their article) in the EPSI, they presented psychometric information (e.g., descriptive statistics, split-half reliabilities, and correlations across subscales) only for the domain subscales. The present study sought to identify both the domain and coping style subscales of the EPSI with factor analyses. One participant with missing data on the EPSI was eliminated ($N = 110$).

The scoring procedure of the EPSI recommended by Cornelius and Caspi (1987) was not meant to allow factor-analytic confirmation of the domain subscales. The scoring approach, which correlated participants' responses and judges' ratings over all items in a particular domain, yielded only a single indicator per subscale. However, to provide at least two indicators per domain so that factor analysis could be attempted, modified scores that represented correlations between judges and participants over two random split-halves of items in each domain or coping style subscale (also used in the computation of reliabilities discussed earlier) were computed for this study. However, several very low correlations among indicators meant that models based on these scores could not be identified.

The empirical focus was then shifted from confirming the particular domain subscales of the EPSI to exploring the interrelationships among EPSI domain subscales as defined by Cornelius and Caspi (1987). Two hypothesized models were tested. In Model EPSI-1, a general-factor solution was tested with LISREL VIII (Jöreskog & Sörbom, 1993), such that each EPSI domain subscale loaded on a common factor. In Model EPSI-2, the same conceptual classification of tasks that was found with the PP (social vs. instrumental) was tested with the EPSI. Consumer (CONS), Home management (HOME), and Information (INFO) domain subscales were freely estimated as indicators of the instrumental factor, whereas Friend (FRIEND), Family (FAMILY), and Coworker (WORK) domain subscales were specified as loading on the social factor.

The general factor Model EPSI-1 fit well, $\chi^2(9, N = 110) = 15.67, p > .07$ (GFI = .96, AGFI = .90, RMSR = .051, NFI = .91, and CFI = 0.96). Although the two-factor Model EPSI-2 also fit well, $\chi^2(8, N = 110) = 15.20, p > .05$ (GFI = .96, AGFI = .89, RMSR = .049, NFI = .91, and CFI = 0.95), it did not represent a significant improvement over Model EPSI-1, $\chi^2(1, N = 110) = 0.47, p > .10$. Inspecting the two-factor solution, the specified factors were correlated very highly ($r = .95$) and contained little unique variance. Model EPSI-1 was therefore accepted, and the standardized solution for this model is presented in Table 2. As the table shows, the general factor is characterized by relatively high loadings (.57 to .73), with substantial unique variance remaining in each indicator.

[1] Model PP-1 and Model PP-2 used the revised scoring of the PP suggested by Denney (total number of effective solutions generated by participants for each problem). Using Denney and Pearce's (1989) traditional scoring (in which the number of solutions was recoded on a 4-point scale for each problem), an identical pattern of findings was obtained. Although the magnitude of factor loadings, using traditional scoring, was somewhat lower (.36 to .52) the two-factor instrumental/social model, $\chi^2(34, N = 111) = 55.12, p < .02$, fit significantly better than a general-factor model, $\chi^2(35, N = 111) = 64.27, p < .002$, and the two factors were highly correlated ($r = .84$).

Table 2
One-Factor Solution (Factor Loadings) for the Everyday Problem Solving Inventory

Domain subscale	General factor	Unique variance
Consumer	.57	.67
Friend	.62	.61
Home management	.55	.70
Information	.72	.48
Family	.59	.65
Coworker	.68	.53

Note. Everyday Problem Solving Inventory (Cornelius & Caspi, 1987).

Everyday Problems Test. A measurement model for the EPT (Willis & Marsiske, 1993) that was consistent with that reported by Willis and Marsiske (1994) was examined. The EPT had a total of 84 items, with 12 items representing each of seven problem domains. Within each domain subscale, each of its 12 items were scored dichotomously (1 = *correct* and 0 = *incorrect*) and were then summed into three 4-item parcels. Parcels were needed to provide indicators that were more normally distributed than single dichotomous items (Gorsuch, 1983; Little, Das, Carlson, & Yachimowicz, 1993). A covariance matrix among parcels was then produced and examined. The fit of a seven-domain subscale structure (FOOD, HLTH, PHON, CONS, FINA, HOUS, and TRAN) was then tested by the LISREL VIII (Jöreskog & Sörbom, 1993) in Model EPT-1. This model appeared to fit reasonably well, $\chi^2(168, N = 111) = 194.97, p > .07$ (GFI = .86, AGFI = .81, RMSR = .059, NFI = .80, and CFI = 0.96), but inspection of estimated parameters revealed that several estimated subscale intercorrelations approached or slightly exceeded 1.0, which produced a matrix that was not positive-definite or admissible. A second model, Model EPT-2 uniformly rescaled all the correlations estimated in Model EPT-1 by a constant scalar so that the maximum correlation among subscales would be .90. This rescaled subscale intercorrelation matrix was then completely fixed. Factor loadings and unique variances were, as in Model EPT-1, freely estimated. Table 3 presents the resulting factor solution. This model also produced a nonsignificant chi-square statistic, $\chi^2(189, N = 111) = 216.62, p > .05$ (GFI = .85, AGFI = .82, RMSR = .096, NFI = .78, and CFI = 0.96), and did not represent a significant reduction in fit from Model EPT-1, $\chi^2(21, N = 111) = 21.65, p > .10$. As the standardized results in Table 3 show, however, the domain subscales of the EPT, as in the PP, were also highly related, a result that is consistent with that reported by Willis and Marsiske (1994) with a larger sample.

Relationships Among Problem-Solving Measures

The next major question for this study concerned similarities and differences among the problem-solving instruments in terms of the measures' patterns of covariation. A confirmatory factor-analytic approach was used, and factor scores that represented the seven IADL-type subscales of the EPT, the two subscales from the PP (social and instrumental problem solving), and computed scores for the six domain subscales of the EPSI (as reported by Cornelius & Caspi, 1987) were entered into a common data matrix (see Appendix).

Three factor models were tested, again using LISREL VIII (Jöreskog & Sörbom, 1993). Model PS-1 examined the fit of a general-factor solution, in which all subscales from all instruments loaded only on a single common factor; Model PS-2 allowed for the existence of a global common factor (on which all subscales loaded) and three instrument-specific factors (on which only the subscales from a particular measure loaded). In other words, each subscale loaded on both the general factor and an instrument-specific factor. Model PS-3 examined the fit of three instrument-specific factors only (each subscale again had only one loading). All three models were estimated in covariance metric, and results are presented in standardized metric; variable residuals were specified as uncorrelated, and factor correlations were freely estimated.

Model PS-1, the general-factor solution, fit poorly, $\chi^2(90, N = 110) = 285.79, p < .001$ (GFI = .70, AGFI = .60, RMSR = .140, NFI = .61, and CFI = 0.69). Only the EPT subscales had modest to high loadings on this general factor (range = .69–.81); all of the remaining subscales from the other two measures had loadings of less than .27. Unique variance in EPSI and PP subscales therefore exceeded .90.

Model PS-2, which allowed for both a general factor and three instrument-specific factors, fit well, $\chi^2(69, N = 110) = 54.24, p > .90$ (GFI = .94, AGFI = .89, RMSR = .041, NFI = .93, and CFI = 1.00), and the model represented a significant increase in fit over Model 1, $\chi^2(21, N = 110) = 231.55, p < .001$. Inspection of the factor-loading estimates, however, revealed that loadings on the general factor were uniformly low, and only 1 of the 15 loadings on this factor was significantly greater than zero.

The fit of Model PS-3, in which the general factor was eliminated and only three instrument-specific factors were estimated was also very good, $\chi^2(87, N = 110) = 76.96, p > .75$ (GFI = .92, AGFI = .88, RMSR = .049, NFI = .90, and CFI = 1.00) and again represented an improvement in fit over Model 1, $\chi^2(3, N = 110) = 208.83, p < .001$. Furthermore, the three-factor (Model PS-3) model did not fit significantly worse than the less parsimonious four-factor model (Model PS-2), $\chi^2(18)$, $N = 110 = 22.72, p > .10$. The accepted three-factor model, in its standardized solution, is displayed in Figure 1. As the figure shows, no common problem-solving factors underlying the three instruments were identified.

To test the assumption that the confirmatory factor-analytic approach obscured latent relationships among subscales (i.e., by assigning subscales to factors rather than allowing the data to define the factors), a series of exploratory factor analyses were conducted. Four oblique factor structures were tested, ranging from a general-factor to a four-factor solution (and were based on the hypothesized models just discussed). As with the confirmatory factor-analytic findings, these exploratory analyses (in terms of the simple structure pattern of factor loadings and factor correlations) suggested that the best model was one that included three minimally related specific factors, each principally defined by the subscales from only one problem-solving instrument.

Table 3
Seven-Factor Solution for the Everyday Problems Test

Variable	FOOD	HLTH	PHON	CONS	FINA	HOUS	TRAN	Unique variance
	Factor loadings							
FOOD 1[a]	.63							.61
FOOD 2	.52							.73
FOOD 3	.61							.63
HLTH 1		.66						.57
HLTH 2		.39						.85
HLTH 3		.70						.51
PHON 1			.65					.57
PHON 2			.37					.86
PHON 3			.74					.46
CONS 1				.65				.58
CONS 2				.70				.51
CONS 3				.60				.64
FINA 1					.74			.45
FINA 2					.37			.86
FINA 3					.71			.50
HOUS 1						.44		.81
HOUS 2						.79		.38
HOUS 3						.54		.71
TRAN 1							.50	.75
TRAN 2							.68	.54
TRAN 3							.63	.61
	Factor correlations							
FOOD	1.00							
HLTH	.90	1.00						
PHON	.75	.84	1.00					
CONS	.85	.82	.76	1.00				
FINA	.77	.66	.59	.69	1.00			
HOUS	.83	.73	.64	.79	.63	1.00		
TRAN	.81	.84	.90	.90	.80	.86	1.00	

Note. FOOD = food preparation; HLTH = health and medication use; PHON = telephone use; CONS = consumerism/shopping; FINA = financial management; HOUS = housekeeping/laundry; TRAN = transportation use. Everyday Problems Test (Willis & Marsiske, 1993).
[a] Numbers refer to item parcel; three 4-item parcels were created for each domain.

Age–Cohort Differences in Problem Solving

To explore comparability among measures in terms of their cross-sectional age trajectories, individual differences in performance level on each instrument that were attributable to age-cohort were examined. As before, 1 participant with missing data on the EPSI was again eliminated ($N = 110$). Congruent with current research in the analysis of age differences, a structural-equation modeling perspective was used (McArdle & Prescott, 1992; Nesselroade, 1983; Schaie & Hertzog, 1983).

To examine the influence of age on the three measures of everyday problem solving, age was used as a predictor of the three problem-solving factors shown in Figure 1. This model had a satisfactory fit, $\chi^2(99, N = 110) = 80.79, p > .90$ (GFI = .92, AGFI = .89, RMSR = .046, NFI = .89, and CFI = 1.00), and the simplified (omitting reestimated measurement parameters from Figure 1) and standardized results are presented in Figure 2. As shown in Figure 2, age did not account for a large proportion of the variance in any of the problem-solving factors; the only path coefficient significantly greater than zero was the path from age to the EPT ($z = -4.02, p < .001$), which accounted for about 17% of the variance in that factor. This path suggested that performance on the EPT became lower at higher ages. It is important to note that slight but apparent differences in problem-solving factor intercorrelations between Figures 1 and 2 reflect the fact that residual correlations between age-partialed latent constructs are shown in Figure 2.

To test whether the relationship between age and each problem-solving measure was consistent across all subscales (particularly to test whether all subscales in the EPT showed common age effects), a final structural-equation model was tested. In this model, the Age factor was again completely defined by chronological age; that is, the residual variance in age (not accounted for by the factor) was fixed to zero. In addition to loading on their measure-specific factors, however, each problem-solving scale (from each of the three instruments) was allowed to load on the Age factor. In this model, the magnitude of the factor loadings of the subscales on the Age factor is equivalent to the regression relationship when each subscale is predicted by age. The results are displayed in Table 4 and show that no subscale from the EPSI or the PP measures was significantly

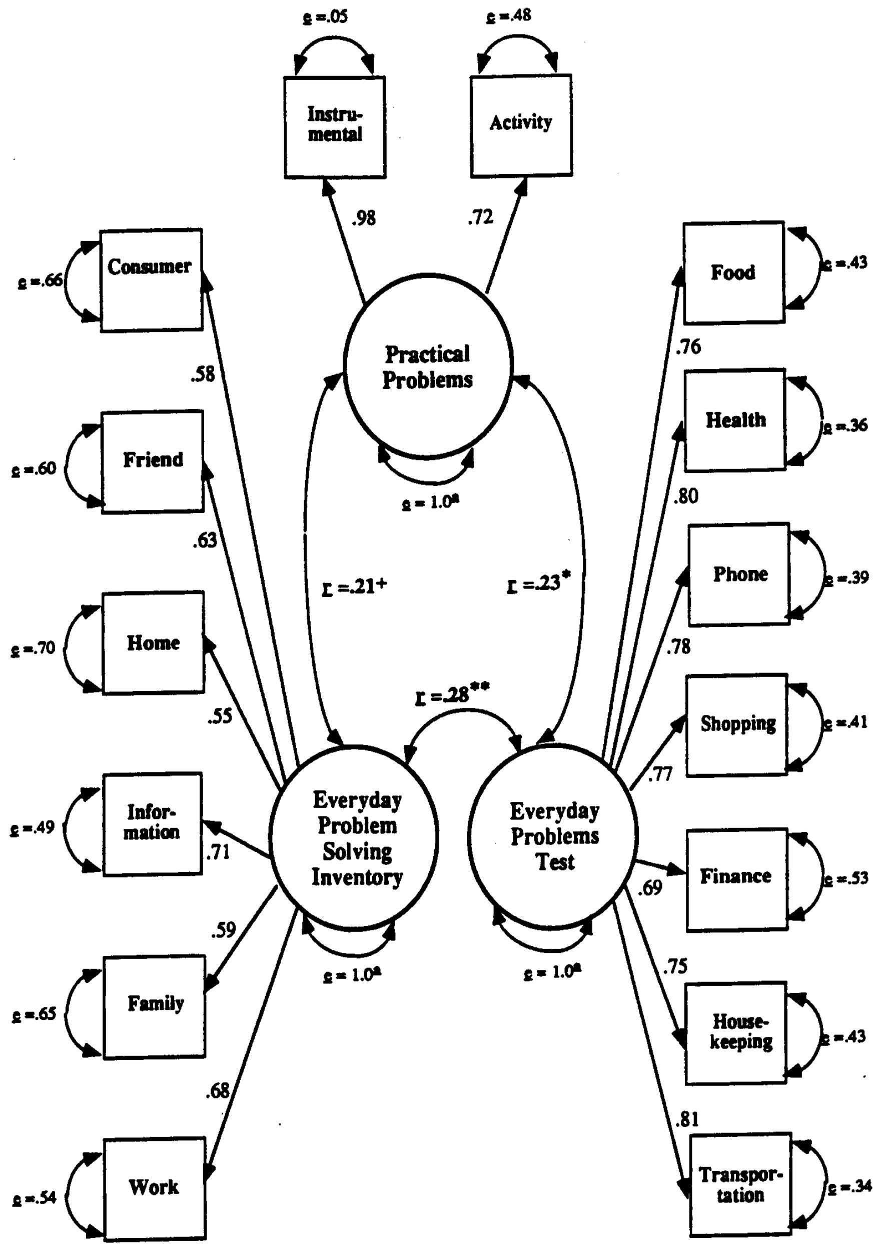

Figure 1. Latent convergence among three measures of everyday problem solving. $^{+}p < .10$, two-tailed. $^{*}p < .05$, two-tailed. $^{**}p < .01$, two-tailed. [a]Fixed parameter.

related to age; conversely, all of the EPT subscales had z values that were lower than -3.01 ($p < .001$). Again, problem-solving factor intercorrelations in Table 4 differ from those in Figure 1 because they are residual correlations, after age has been taken out of all the subscales: The fit of this model was also acceptable, $\chi^2(87, N = 110) = 77.86, p > .70$ (GFI = .92, AGFI = .87, RMSR = .046, NFI = .90, and CFI = 1.00), and the difference in fit from the structural-equation model presented in Figure 2 was not significant, $\chi^2(12, N = 110) = 2.93, p > .10$. As reported earlier, increased age was associated with the production of lower scores on the EPT.

Discussion

The goal of this study was to explore the construct of everyday problem solving from a multivariate perspective, with particular attention to two unresolved issues in the current everyday cognition literature. First, most studies of everyday problem

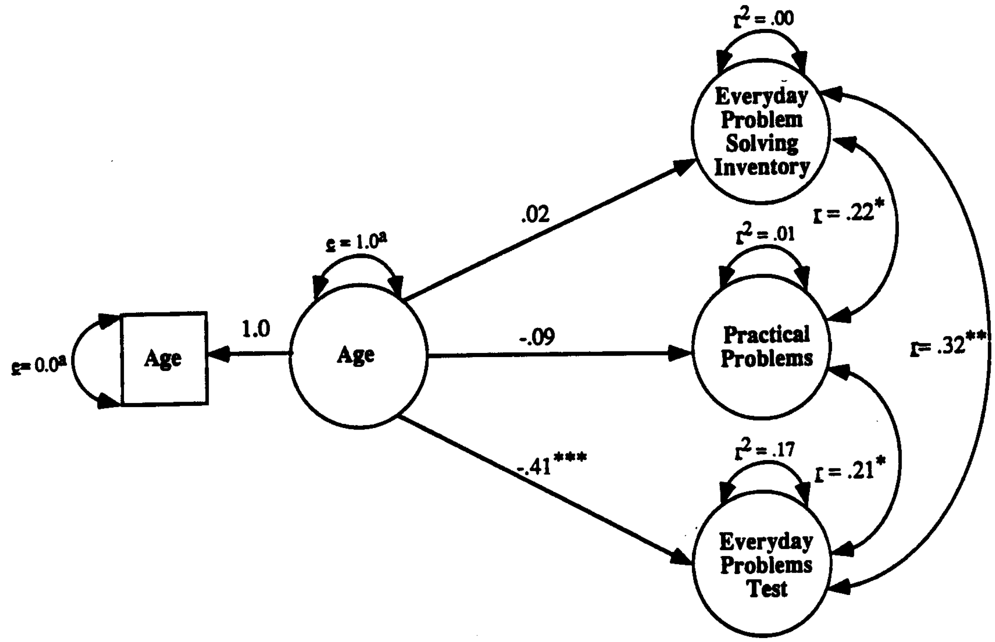

Figure 2. The relationship between age and everyday problem solving factors. *$p < .05$, two-tailed. **$p < .01$, two-tailed. ***$p < .001$, two-tailed. ªFixed parameter.

solving have looked at only performance on single measures. This leaves unaddressed the question of whether there are dimensions of problem solving that can be identified across methods of assessment. A second unresolved issue in research on everyday problem solving concerns its adult developmental trajectory. Although a number of major theories of adult everyday cognition suggest that, at least for some older adults, there may be resilience in everyday functioning, the empirical results have been contradictory. Cross-sectional and longitudinal findings seem to be highly dependent on the measure used, and few studies have examined age trends in participants who are older than 80 years of age.

The data from the present study provide some preliminary information regarding both of these issues. First, the results suggest that within particular instruments, everyday problem solving seems to be a relatively unitary construct. Although distinct subscales can be identified in two of the three measures studied, there is substantial individual consistency of performance across these subscales. Second, the results suggest that there is substantially less interindividual similarity across measures of everyday problem solving. Specifically, three everyday problem-solving measures were relatively unrelated to one another, typically sharing less than 5% of their variance. Third, congruent with the research literature, the findings suggest that age–cohort differences vary according to the instrument used. There was no evidence of a negative age effect for two of the measures studied (the PP and the EPSI). The third measure, the EPT, did show a negative cross-sectional age effect, although age accounted for only 17% of the individual differences in EPT performance. With each additional decade of life, participants performed about two fifths of a standard deviation below younger peers.

What are the reasons for these differences between measures? One possibility emerges from a close examination of their task demands. Although all three measures converge in their inclusion of (a) naturalistic, everyday problems (i.e., problems with high face validity for everyday life), (b) problems designed to be relevant for older adults (although the EPSI contains problems designed to be relevant for younger adults as well), and (c) some problems dealing with instrumental content (e.g., shopping and financial management), there were other potentially important differences. The Denney and Pearce (1989) PP, for example, places its emphasis on how many solutions individuals can devise for particular problems. Consequently, the PP bears a resemblance to measures of ideational fluency used in creativity research (e.g., Guilford, 1950; Marsiske & Willis, in press; Wallach, 1986). In contrast, the EPT (Willis & Marsiske, 1993) appears similar in task demands to measures of document literacy (e.g., Educational Testing Service, 1977; Kirsch & Mosenthal, 1990; Meyer, Marsiske, & Willis, 1993). The EPSI is scored as the correlation between an individual's likelihood of using a given solution for a problem and judges' ratings of that solution's efficacy; thus, high scores on the EPSI may serve as an indication of one's conformity to a particular reference group (i.e., a set of judges) and not only as an indication of cognitive correctness. At the very least, as Cornelius and Caspi (1987) noted, performance on the EPSI may be an indication of both everyday socioemotional and cognitive competence (see also Blanchard-Fields, 1986; Labouvie-Vief, 1982; Staudinger, Lopez, & Baltes, 1995). Taken together, it may be that the three

Table 4
The Relationship Between Age and Everyday Problem-Solving Scales

Variable	Factor loadings: Age	EPSI	PP	EPT	Unique variance
Age	1.00				*0.00*
EPSI: Consumer	0.06	0.59			0.65
EPSI: Friend	−0.01	0.63			0.60
EPSI: Home	0.02	0.55			0.69
EPSI: Information	0.05	0.71			0.49
EPSI: Family	−0.00	0.59			0.65
EPSI: Coworker	−0.03	0.68			0.54
PP: Instrumental	−0.09		1.00		0.00
PP: Social/Activity	−0.12		0.70		0.50
EPT: Food	−0.28			0.70	0.43
EPT: Medication use	−0.34			0.72	0.36
EPT: Telephone use	−0.33			0.71	0.39
EPT: Shopping	−0.30			0.71	0.41
EPT: Financial management	−0.32			0.61	0.53
EPT: Housekeeping	−0.29			0.69	0.43
EPT: Transportation	−0.32			0.75	0.34
Factor intercorrelations					
Age	*1.00*				
EPSI	*0.00*	*1.00*			
PP	*0.00*	0.20	*1.00*		
EPT	*0.00*	0.32	0.20	*1.00*	

Note. Values greater than |0.32| are significantly different from zero, $p < .05$. Italicized values represent fixed parameters. EPSI = Everyday Problem Solving Inventory (Cornelius & Caspi, 1987); PP = Practical Problems Test (Denney & Pearce, 1989); EPT = Everyday Problems Test (Willis & Marsiske, 1993).

measures, although all nominally concerned with measuring the same thing, assess substantively distinct aspects of everyday cognition.

Following from these apparent differences in task demands, the three everyday problem-solving measures may also have differed in the kinds of intellectual abilities needed for their successful performance. Although the relatively small sample size for these analyses precluded consideration of many more constructs, prior research with each of the three measures used in this study provides some clues to possible differences in the intellectual variance assessed by each instrument. With regard to the EPSI, for example, Cornelius and Caspi (1987) reported correlations of .29 with Inductive Reasoning (an indicator of fluid intelligence ability), and .27 with the crystallized intelligence measure of Verbal Facility, in their sample of 20- to 78-year olds. In measures similar in content and format to the PP, Denney and her colleagues (Camp et al., 1989; Denney, 1991) reported similarly modest relationships between everyday problem solving and intelligence ($r = .34$ with the WAIS Information subtest, $r = .29$ with the WAIS Similarities subtest, and $r = .40$ with the Raven Progressive Matrices) in samples of 20- to 80-year olds. Much stronger relationships between intelligence and everyday problem solving were reported by Willis and her colleagues for two measures of document-related everyday cognition. In studies with participants ages 60 to over 90 years, constructs of fluid and crystallized intelligence accounted for between 65% and 80% of variance in everyday problem-solving measures (Willis & Marsiske, 1991; Willis, Marsiske, & Diehl, 1991; Willis & Schaie, 1986). If these correlational findings can be extrapolated to the present study, then one source of the differences between everyday problem-solving tasks may be differential task composition in terms of intellectual demands. This might also explain differences in obtained cross-sectional trajectories. Although most participants in this study were within the age range in which intellectual losses seem to become normative (e.g., P. B. Baltes, 1993; Salthouse, 1991b; Schaie, 1994), only the EPT, which prior research suggests is most related to intelligence, showed negative age effects.

Therefore, the present findings are relevant for life span theory regarding the aging of practical or pragmatic cognition. Several major theorists (e.g., P. B. Baltes, 1993; Denney, 1984) argued that, to the extent that measures of pragmatic or practical problem solving assess those tasks that have actually been selected and practiced within particular person–environment constellations, one might expect to see preserved functioning into relatively late adulthood. The present results do not speak to the relevance of measured everyday tasks to the lives of study participants, but they do seem to support predictions regarding low-to-modest associations between age and everyday problem solving in later adulthood.[2] Even in the measure that shows the largest age differences, the EPT, approximately 83% of the individual differences variance remains unaccounted for after controlling for age. This is consistent with the findings of Poon et al. (1992) who, in their study of participants ages 60 to over 100 years, reported significant negative effects of age on all cognitive measures except for a measure of practical problem solving drawn from Denney's work (Denney, Pearce, & Palmer, 1982).

Before considering some of the further implications of these results, it is important to note several caveats that may limit the generalizability of our findings. First, it must be acknowledged that the sample of old and very old adults considered in this study may have been positively selected and homogeneous with regard to demographic and health characteristics. The relatively restricted age range (68–94 years) in this study may have reduced the magnitude of relationships with age, and the age range included was narrower than in most prior cross-sectional research on this topic (e.g., Cornelius & Caspi, 1987; Denney & Pearce, 1989). Second, our use of a cross-sectional sample limits the developmental conclusions that can be drawn from this study, because it is not possible to separately estimate age effects associated with maturation and senescence from those associated with cohort differences in lifetime opportunity structures (e.g., P. B. Baltes, 1968; Schaie, 1965). Third, because two of

[2] There is some evidence regarding the familiarity and relevance of one measure, the EPT, in older adults' everyday lives. Diehl, Willis, and Schaie (1990) reported that the specific tasks contained within the EPT were rated as highly relevant and familiar to the daily lives of older adults both by health-service professionals and by older adults. In another study, Diehl, Willis, and Schaie (in press) reported that EPT performance substantially predicts observed performance on selected tasks of daily living (food preparation, telephone use, and medication use).

our everyday problem-solving measures were administered at home, without supervision, it is not possible to state conclusively that no assistance was received by any participants. Although participants certified that they had performed take-home tasks independently, it may be that some participants actually did receive help.

Several other limitations to this study center around the structural-equation modeling results. The sample of only 111 older adults is relatively small for structural-equation modeling research. In addition to limiting the number of variables that can be considered in a particular model, the sample size may also limit the statistical power to reject poor models and to discriminate between alternative models and may result in parameter estimates with large confidence intervals. More important, current scholarship on structural-equation models has increasingly emphasized the point that the ordering of variables in any model may be arbitrary and that many other combinations could lead to satisfactory model fit (e.g., MacCallum, Wegener, Uchino, & Fabrigar, 1993). Although the models presented in this article were guided by prior research (using item groupings proposed by the measures' authors), other groupings of items might have led to alternative outcomes, especially for the within-measure factors. Finally, the models discussed in this study were accepted following modifications within a single sample, which causes this study to run the risk of capitalizing on unique features of the sample. Future research that replicates the obtained models in independent samples will be necessary before the findings of measure interrelationships and associations with age can be more confidently accepted.

Despite these possible limitations, we believe that these findings extend our understanding of everyday problem solving in late adulthood in several ways. First, the empirical consequences of the absence of a unifying measurement framework for everyday cognition were revealed. The measures in this study assessed very different constructs, and there is no reason to assume that developmental studies of adult problem solving that use a particular measure will provide results that will lead to useful predictions about other measures, or about the adult developmental trajectory of everyday problem solving in general. Second, by examining practical problem solving at the latent level, we have minimized the likelihood that relationships among variables were attenuated by unreliable measurement variance. That relationships among constructs represented by each measure were of low magnitude suggests that the instruments really did assess relatively unrelated aspects of everyday problem solving.

What might be possible implications of these results for future research and theory about everyday problem solving in late life? Clearly, there seems to be a need for a broader conceptual taxonomy of everyday problem-solving tasks from within which the aspects of problem solving that are assessed by a particular method or instrument can be better understood. Although several researchers have proposed such taxonomies (e.g., M. E. Ford & Nichols, 1988; Schaie, 1978; Willis & Schaie, 1993), little progress has been made in adapting them for use across the life span or for integrating discrepant research findings. From the results of this study alone, and from the literature reviewed, it seems that tasks could be classified within a number of dimensions, including (a) age relevance (e.g., Cornelius, 1984), (b) cognitive demands (e.g., Marsiske & Willis, in press; Salthouse, 1993), (c) noncognitive demands (e.g., Labouvie-Vief, 1982; Staudinger et al., 1995), (d) problem domain (e.g., Ericsson & Smith, 1991; Salthouse, 1990), (e) task criticality (e.g., M. M. Baltes, Mayr, Borchelt, Maas, & Wilms, 1993; Willis, 1991), (f) task novelty (e.g., Berg & Sternberg, 1985; Cornelius, 1984), and (g) ill- versus well-structured tasks (e.g., Wagner, 1986). There has been substantial scholarship about each of these issues, and every problem-solving task can be classified with regard to each dimension. Although this list is not exhaustive, it does suggest that any everyday problem-solving task represents a selection from a broader array.

This list of possible dimensions tries to make explicit a conception of everyday problem solving as multidimensional; unless individual measures focus on similar problem types, it is unlikely that they will be related to one another. The results of low relationships between measures are congruent with, but not directly supportive of, such a multidimensional conception of everyday problem solving. To properly show multidimensionality, multiple measures of those aspects of problem solving that are assessed by a particular instrument would need to be obtained to separate *trait* variance (i.e., variance associated with a particular dimension of problem solving) from *method* or instrument-specific variance (Campbell & Fiske, 1959). In other words, one implication from this study is that future research must focus on issues of convergent and discriminant validity and, more generally, on a multivariate approach to the study of everyday problem solving. The early literature regarding the development of psychometric conceptions of intelligence may serve as a good model for guiding future investigations of practical problem solving (e.g., Ekstrom, French, Harman, & Derman, 1976; Horn & Cattell, 1966; Thurstone, 1962).

In summary, the results from this study show that three measures of everyday problem solving in late life, although themselves measuring subscales of everyday problem solving in an internally consistent way, have little relationship to one another. This research suggests that the inconsistent findings regarding adult age trajectories of everyday cognition may have their roots in an unacknowledged multidimensionality. It may be time to move away from the global construct label of *everyday problem solving* (and its many allied terms) and move toward more precise specifications of those aspects and dimensions that particular investigators are interested in studying.

References

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52,* 317–332.

Baltes, M. M., Mayr, U., Borchelt, M., Maas, I., & Wilms, H.-U. (1993). Everyday competence in old and very old age: An inter-disciplinary perspective. *Ageing and Society, 13,* 657–680.

Baltes, P. B. (1968). Longitudinal and cross-sectional sequences in the study of age and generation effects. *Human Development, 11,* 145–171.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology, 23,* 611–626.

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

Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In M. M. Baltes & P. B. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 1–34). Cambridge, England: Cambridge University Press.

Baltes, P. B., Mayer, K.-U., Helmchen, H., & Steinhagen-Thiessen, E. (1993). The Berlin Aging Study (BASE): Overview and design. *Ageing and Society, 13,* 483–516.

Baltes, P. B., Smith, J., & Staudinger, U. M. (1992). Wisdom and successful aging. In T. Sonderegger (Ed.), *Nebraska symposium on motivation* (Vol. 39, pp. 123–167). Lincoln: University of Nebraska Press.

Baltes, P. B., & Willis, S. L. (1982). Plasticity and enhancement of intellectual functioning in old age: Penn State's Adult Development and Enrichment Project (ADEPT). In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 353–389). New York: Plenum.

Bentler, P. M. (1989). *EQS structural equations manual.* Los Angeles: BMDP Statistical Software.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588–606.

Berg, C. A., & Sternberg, R. J. (1985). A triarchic theory of intellectual development during adulthood. *Developmental Review, 5,* 334–370.

Berry, J. W., & Irvine, S. H. (1986). Bricolage: Savages do it daily. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 361–378). Cambridge, England: Cambridge University Press.

Blanchard-Fields, F. (1986). Reasoning on social dilemmas varying in emotional saliency: An adult developmental perspective. *Psychology and Aging, 1,* 325–333.

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: Theory and application* (pp. 211–228). New York: Praeger.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin, 56,* 124–132.

Carmines, E. G., & McIver, J. P. (1981). Analyzing models with unobserved variables: Analysis of covariance structures. In G. W. Bohrnstedt & E. F. Borgatta (Eds.), *Social measurement: Current issues.* Beverly Hills, CA: Sage.

Cattell, R. B. (1964). Validity and reliability: A proposed more basic set of concepts. *Journal of Educational Psychology, 55,* 1–22.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action.* Boston: Houghton-Mifflin.

Ceci, S. J., & Liker, J. (1986). Academic and nonacademic intelligence: An experimental separation. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 119–142). Cambridge, England: Cambridge University Press.

Charness, N. (1985). Aging and problem-solving performance. In N. Charness (Ed.), *Aging and human performance* (pp. 225–259). New York: Wiley.

Cornelius, S. W. (1984). Classic pattern of intellectual aging: Test familiarity, difficulty, and performance. *Journal of Gerontology, 39,* 201–206.

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

Demming, J. A., & Pressey, S. L. (1957). Tests 'indigenous' to the adult and older years. *Journal of Counseling Psychology, 4,* 450–466.

Denney, N. W. (1984). A model of cognitive development across the life span. *Developmental Review, 4,* 171–191.

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). Cambridge, England: Cambridge University Press.

Denney, N. W. (1991, November). *Intellectual correlates of practical and social problem solving.* Paper presented at the 44th Annual Scientific Meeting of the Gerontological Society of America, San Francisco.

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' 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, 47,* 142–145.

Diehl, M., Willis, S. L., & Schaie, K. W. (1990, November). *Adults' perceptions about the relevance of printed materials for elderly's independent living.* Paper presented at the 43rd Annual Scientific Meeting of the Gerontological Society of America, Boston.

Diehl, M., Willis, S. L., & Schaie, K. W. (in press). Practical problem solving in older adults: Observational assessment and cognitive correlates. *Psychology and Aging.*

Dixon, R. A., & Baltes, P. B. (1986). Toward life-span research on the functions and pragmatics of intelligence. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 203–235). Cambridge, England: Cambridge University Press.

Dixon, R. A., Kramer, D. A., & Baltes, P. B. (1985). Intelligence: Its life-span development. In B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurement and applications* (pp. 301–350). New York: Wiley.

Educational Testing Service. (1977). *Basic Skills Assessment Test: Reading.* Princeton, NJ: Author.

Ekstrom, R. B., French, J. W., Harman, H., & Derman, D. (1976). *Kit of factor-referenced cognitive tests* (Rev. ed.). Princeton, NJ: Educational Testing Service.

Ericsson, K. A., & Smith, J. (Eds.). (1991). *Towards a general theory of expertise: Prospects and limits.* Cambridge, England: Cambridge University Press.

Ford, A. B., Bolmar, S. J., Salmon, R. B., Medalie, J. H., Roy, A. W., & Galazka, S. S. (1988). Health and function in the old and very old. *Journal of the American Geriatrics Society, 36,* 187–197.

Ford, M. E., & Nichols, C. W. (1988). A taxonomy of human goals and some possible applications. In M. E. Ford & D. H. Ford (Eds.), *Humans as self-constructing living systems: Putting the framework to work* (pp. 289–311). Hillsdale, NJ: Erlbaum.

Gardner, E. F., & Monge, R. H. (1977). Adult age differences in cognitive abilities and educational background. *Experimental Aging Research, 3,* 337–383.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Guilford, J. P. (1950). Creativity. *American Psychologist, 5,* 444–454.

Guilford, J. P. (1959). Three faces of intellect. *American Psychologist, 14,* 469–479.

Hartley, A. A. (1989). The cognitive ecology of problem-solving. In L. W. Poon, D. C. Rubin, & B. A. Wilson (Eds.), *Everyday cognition in adulthood and late life* (pp. 300–329). Cambridge, England: Cambridge University Press.

Hess, T. M. (1990). *Aging and cognition: Knowledge organization and utilization.* Amsterdam: Elsevier.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 57*, 253–270.

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. 44–49). Cambridge, England: Cambridge University Press.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18*, 117–144.

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

Kirsch, I. S., & Mosenthal, P. B. (1990). Exploring document literacy: Variables underlying the performance of young adults. *Reading Research Quarterly, 25*, 5–30.

Labouvie, E. W. (1980). Identity versus equivalence of psychological measures and constructs. In L. W. Poon (Ed.), *Aging in the 1980s* (pp. 493–502). Washington, DC: American Psychological Association.

Labouvie-Vief, G. (1982). Dynamic development and mature autonomy: A theoretical prologue. *Human Development, 25*, 161–191.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist, 9*, 179–186.

Lindenberger, U., Kliegl, R., & Baltes, P. B. (1992). Professional expertise does not eliminate negative age differences in imagery-based memory performance during adulthood. *Psychology and Aging, 7*, 585–593.

Little, T. D., Das, J. P., Carlson, J. S., & Yachimowicz, D. J. (1993). The role of higher-order cognitive abilities as mediators of deficits in academic performance. *Learning and Individual Differences, 5*, 219–240.

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin, 114*, 185–199.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103*, 391–410.

Marsiske, M., Lang, F. R., Baltes, P. B., & Baltes, M. M. (in press). Selective optimization with compensation: Life-span perspectives on successful human development. In R. A. Dixon & L. Bäckman (Eds.), *Psychological compensation: Managing losses and promoting gains*. Hillsdale, NJ: Erlbaum.

Marsiske, M., & Willis, S. L. (in press). Practical creativity in older adults' everyday problem solving: Life-span perspectives. In C. E. Adams-Price (Ed.), *Creativity and aging: Theoretical and empirical approaches*. New York: Springer.

McArdle, J. J., & Prescott, C. A. (1992). Age-based construct validation using structural equation modeling. *Experimental Aging Research, 18*, 87–115.

Meyer, B. J. F., Marsiske, M., & Willis, S. L. (1993). Text processing variables predict the readability of everyday documents read by older adults. *Reading Research Quarterly, 28*, 234–249.

Nesselroade, J. R. (1983). Temporal selection and factor invariance in the study of development and change. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life span development and behavior* (Vol. 5, pp. 59–87). New York: Academic Press.

Neugarten, B. L. (1984). Psychological aspects of aging and illness. *Psychosomatics, 25*, 123–125.

Palmore, E. B., Nowlin, J. B., & Wang, H. S. (1985). Predictors of functioning among the old-old: A 10-year follow-up. *Journal of Gerontology, 40*, 244–250.

Poon, L. W., Martin, P., Clayton, G. M., Messner, S., Noble, C. A., & Johnson, M. A. (1992). The influences of cognitive resources on adaptation and old age. *International Journal of Aging and Human Development, 34*, 31–46.

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

Rabbitt, P. (1977). Changes in problem solving ability in old age. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 606–625). New York: Van Nostrand Reinhold.

Rogoff, B., & Lave, J. (Eds.). (1984). *Everyday cognition: Its development in social context*. Cambridge, MA: Harvard University Press.

Salthouse, T. A. (1990). Cognitive competence and expertise in aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed., pp. 310–319). New York: Academic Press.

Salthouse, T. A. (1991a). Expertise as the circumvention of human processing limitations. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise* (pp. 286–300). Cambridge, England: Cambridge University Press.

Salthouse, T. A. (1991b). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.

Salthouse, T. A. (1993). Speed and knowledge as determinants of adult age differences in verbal tasks. *Journal of Gerontology, 48*, 29–36.

Schaie, K. W. (1965). A general model for the study of developmental problems. *Psychological Bulletin, 64*, 92–107.

Schaie, K. W. (1978). External validity in the assessment of intellectual development in adulthood. *Journal of Gerontology, 33*, 695–701.

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

Schaie, K. W., & Hertzog, C. (1983). Fourteen year cohort sequential analyses of adult intellectual development. *Developmental Psychology, 19*, 531–543.

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

Sinnott, J. D., & Cavanaugh, J. C. (Eds.). (1991). *Bridging paradigms: Positive development in adulthood and cognitive aging*. New York: Praeger.

Smith, J., & Baltes, P. B. (1990). A life-span perspective on thinking and problem-solving. In M. Schwebel, C. A. Maher, & N. S. Fagley (Eds.), *Promoting cognitive growth over the life span* (pp. 47–69). Hillsdale, NJ: Erlbaum.

Staudinger, U. M., Lopez, D. F., & Baltes, P. B. (1995). *Wisdom-related knowledge and judgment: An area of psychological functioning located at the interface between personality and intelligence?* Unpublished manuscript, Max Planck Institute for Human Development and Education, Berlin, Germany.

Staudinger, U. M., Marsiske, M., & Baltes, P. B. (1993). Resilience and levels of reserve capacity in later adulthood: Perspectives from life-span theory. *Development and Psychopathology, 5*, 541–566.

Staudinger, U. M., Smith, J., & Baltes, P. B. (1992). Wisdom-related knowledge in a life review task: Age differences and the role of professional specialization. *Psychology and Aging, 7*, 271–281.

Sternberg, R. J., & Wagner, R. K. (Eds.). (1986). *Practical intelligence: Origins of competence in the everyday world*. Cambridge, England: Cambridge University Press.

Thurstone, T. G. (1962). *Primary mental abilities for Grades 9–12* (Rev. ed.). Chicago: Science Research Associates.

Wagner, R. K. (1986). The search for intraterrestrial intelligence. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (pp. 51–83). Cambridge, England: Cambridge University Press.

Wallach, M. A. (1986). Creativity testing and giftedness. In F. D. Horowitz & M. O'Brien (Eds.), *The gifted and the talented: Developmen-*

tal perspectives (pp. 99–123). Washington, DC: American Psychological Association.

Willis, S. L. (1991). Cognition and everyday competence. In K. W. Schaie (Ed.), *Annual review of gerontology and geriatrics* (Vol. 11, pp. 80–109). New York: Springer.

Willis, S. L., Jay, G. M., Diehl, M., & Marsiske, M. (1992). Longitudinal change and the 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. 183–198). Norwell, MA: Kluwer Academic.

Willis, S. L., & Marsiske, M. (1993). *Manual for the Everyday Problems Test.* University Park: Pennsylvania State University.

Willis, S. L., & Marsiske, M. (1994). *Psychometric characteristics of the Everyday Problems Test.* Unpublished manuscript, University Park, Pennsylvania State University.

Willis, S. L., Marsiske, M., & Diehl, M. (1991, November). *Older adults' performance on instrumental activities of daily living and mental abilities.* Paper presented at the 44th Annual Scientific Meeting of the Gerontological Society of America, San Francisco.

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). Cambridge, England: Cambridge University Press.

Willis, S. L., & Schaie, K. W. (1993). Everyday cognition: Taxonomic and methodological considerations. In J. M. Puckett & H. W. Reese (Eds.), *Life-span developmental psychology: Mechanisms of everyday cognition* (pp. 35–53). Hillsdale, NJ: Erlbaum.

Woodruff-Pak, D. S. (1989). Aging and intelligence: Changing perspectives in the twentieth century. *Journal of Aging Studies, 3,* 91–118.