



Published in final edited form as:

Psychol Aging. 2010 June ; 25(2): 271–288. doi:10.1037/a0019106.

Developing a Tool for Measuring the Decision-Making Competence of Older Adults

Melissa L. Finucane, PhD¹ and Christina M. Gullion, PhD²

¹East-West Center, Honolulu, HI

²The Center for Health Research, Kaiser Permanente Northwest, Portland, OR

Abstract

The authors evaluated the reliability and validity of a tool for measuring older adults' decision-making competence (DMC). Two-hundred-five younger adults (25-45 years), 208 young-older adults (65-74 years), and 198 old-older adults (75-97 years) made judgments and decisions related to health, finance, and nutrition. Reliable indices of comprehension, dimension weighting, and cognitive reflection were developed. Unlike previous research, the authors were able to compare old-older with young-older adults' performance. As hypothesized, old-older adults performed more poorly than young-older adults; both groups of older adults performed more poorly than younger adults. Hierarchical regression analyses showed that a large amount of variance in decision performance across age groups (including mean trends) could be accounted for by social variables, health measures, basic cognitive skills, attitudinal measures, and numeracy. Structural equation modeling revealed significant pathways from three exogenous latent factors (crystallized intelligence, other cognitive abilities, and age) to the endogenous DMC latent factor. Further research is needed to validate the meaning of performance on these tasks for real-life decision making.

Keywords

Decision-Making Competence; Aging; Person-Task Fit Framework; Measurement Methods

Introduction

Measuring older adults' decision-making competence (DMC) reliably and validly is critical for (1) identifying individuals whose decision-making abilities are impaired; (2) avoiding inappropriate removal of decision-making power from competent individuals, and (3) advancing theories of when and how decision makers across the adult lifespan are able to meet the demands of various decision environments. Even though robust assessment tools are essential for modeling and assisting successful aging, few tools provide a performance-based analysis of older adults' DMC. In this article, we report progress in an ongoing research program that aims to measure older adults' decision skills using behavioral decision tasks.

Approaches to Measuring DMC

A common approach to measuring DMC is via intuitive clinical judgment. Typically, (expert) physicians are the "gold standard," judging individuals' DMC based on their

clinical experience and their impression of the patient using clinical and contextual information (Grisso, Appelbaum, & Hill-Fotouhi, 1997; Janofsky, McCarthy, & Folstein, 1992; Kim, Caine, Currier, Leibovici, & Ryan, 2001). Studies have shown, however, that judgments of the same individual may vary across physicians (Kitamura & Kitamura, 2000; Marson, 1994; Marson, McInturff, Hawkins, Bartolucci, & Harrell, 1997). In addition, clinicians often assess individuals in isolation, with relatively little behaviorally based information. Consequently, the reliability and external validity of their judgments may be limited.

In an attempt to clarify methods for measuring decision-making abilities, several instruments have been developed (e.g., Fitten, Lusky, & Hamann, 1990; Grisso et al., 1997; Kitamura et al., 1998; Schmand, Gouwenberg, Smit, & Jonker, 1999; Stanley, Guido, Stanley, & Shortell, 1984). Most instruments assess medical competencies (e.g., competence to consent to treatment) and many have been tested on impaired, ill, or hospitalized persons. Few researchers have developed instruments to evaluate day-to-day decision making. Furthermore, most existing psychogeriatric instruments sum a mixture of abilities (Vellinga, Smit, van Leeuwen, van Tilburg, & Jonker, 2004), even though some authors argue convincingly that the separate scores of different abilities should not be added (Grisso et al., 1997; Palmer, Nayak, Dunn, Appelbaum, & Jeste, 2002). Overall, questions of reliability and validity remain for many existing tools (Betz, 1996; Clark & Watson, 1995; Ryan, Lopez, & Sumerall, 2001).

Performance-based measures of DMC are rare. Recent exceptions include the Adult Decision-Making Competence index (A-DMC) (Bruine De Bruin, Parker, & Fischhoff, 2007) and the Youth Decision-Making Competence index (Y-DMC) (Parker & Fischhoff, 2005). These indices use a variety of experimental tasks developed originally to describe general cognitive processes. The authors report correlated responses across tasks comprising the indices, suggesting a “positive manifold” of decision-making performance (Stanovich & West, 2000). However, the composite nature of the tasks obscures decision-maker strengths and weaknesses at the level of specific decision skills. Furthermore, some items in these indices may not be appropriate for very old populations (e.g., asking the probability of getting in a car accident while driving is not relevant for many old-older adults who do not drive).

An alternative approach to developing performance-based measures of DMC by Finucane and colleagues (Finucane, Mertz, Slovic, & Schmidt, 2005; Finucane et al., 2002) has emphasized the need for standardized measurement of the multiple specific skills that comprise competent decision processes. Developing items applicable across the adult lifespan, these researchers have attempted to capture functionally different capacities that may be served by different cognitive abilities and information-processing modes. However, initial research suggested that some items had low discriminability and indices of some skills (e.g., consistency) had low reliability. In addition, limited measurements of decision-maker characteristics in earlier studies restricted examination of the construct validity of the tasks. Importantly, samples did not include sufficient individuals representing very old adults—arguably the group of greatest interest and concern given demographic trends toward an increasing proportion of old-older adults in the population.

Understanding the Multidimensional Nature of Decision-Making Competence

Competent decision making requires several key skills including the ability to understand information, integrate information in an internally consistent manner, identify the relevance of information in a decision process, and inhibit impulsive responding. Performance on these skills is expected to reflect the degree of congruence between characteristics of the decision maker and the demands of the task and context (Finucane et al., 2005).

DMC as Compiled Cognition—One way of looking at DMC is as a complex, “compiled” form of cognition (Appelbaum & Grisso, 1988; Park, 1992; Salthouse, 1990; Willis & Schaie, 1986) dependent on basic cognitive abilities, such as crystallized and fluid intelligence, memory capacity, and speed of processing (Kim, Karlawish, & Caine, 2002; Schaie & Willis, 1999) which change with increasing age (MacPherson, Phillips, & Sala, 2002; Salthouse & Ferrer-Caja, 2003). According to this view, age-related changes in DMC are likely to have similar characteristics to age-related changes in these basic cognitive abilities.

For instance, comprehension may be affected by an individual’s inductive and arithmetic reasoning abilities, such as problem identification, decomposition, step sequencing, information manipulation, and perceptions of consequences (Allaire & Marsiske, 1999). Comprehension will suffer as text complexity increases or readability decreases (Meyer, Marsiske, & Willis, 1993) and outstrips the reader’s reasoning skills and memory capacity (Allaire & Marsiske, 1999; Finucane et al., 2002; Meyer, Russo, & Talbot, 1995; Park, Willis, Morrow, Diehl, & Gaines, 1994; Tanius, Wood, Hanoch, & Rice, 2009).

Similarly, decline in memory and processing speed and greater reliance on simpler strategies among older adults may make them more prone than younger adults to inconsistency in decision making across alternative framings of statistically equivalent information (Finucane et al., 2005; Finucane et al., 2002; Weber, Goldstein, & Barlas, 1995). Holding constant the relative importance of decision dimensions (attributes) across situations that differ only superficially is important because it both generates and reflects reliable preferences. Some research suggests that while age-related cognitive deficits may increase susceptibility to framing effects, age-related strengths (e.g., added experience) may counteract the effects when applicable (Rönnlund, Karlsson, Laggrens, Larsson, & Lindström, 2005).

Compared with comprehension and consistency, researchers have paid less attention to how the skills of identifying the relevance of information and inhibiting impulsive responding are associated with age-related decline in cognitive abilities. In assessments of competence to consent to research, Appelbaum, Griss, Frank, O’Donnell, and Kupfer (1999) found that better appreciation of the relevance of information to one’s personal situation was related to lower age (for women) and having prior research experience. However, the small number of participants in these studies and the small range of responses limit the interpretability of the results. Also, their measures were designed only for patient populations and did not directly examine patients’ appreciation of the importance of specific dimensions of information in their decision processes.

The ability to inhibit impulsive responding has been examined recently by Frederick (2005), who found that younger adults’ performance on the *Cognitive Reflection Test* was correlated positively with measures of their general cognitive ability and capacity to make “patient” (or risk seeking) decisions. Cognitive reflection among older adults has not been studied, but analytic oversight of impulsive tendencies might be expected to decline with age because self-monitoring depends on basic cognitive abilities.

Some authors argue that good decision making depends also on numeracy, the ability to work with numbers and to understand basic probability and mathematical concepts (Schwartz, Woloshin, Black, & Welch, 1997). Hibbard, Peters, Dixon, and Tusler (2007) showed via path analysis of data collected from a sample of adults 18-64 years that numeracy contributes to comprehension of hospital performance information, which in turn contributes to choosing a higher quality hospital. Peters and colleagues (Peters, Dieckmann, Västfjäll, Mertz, & Slovic, 2009; Peters et al., 2006) reported that compared with less numerate individuals, more numerate individuals are more likely to retrieve and use

appropriate numerical principles and are less susceptible to framing effects in judgments about student work quality, mental patient violence, and the attractiveness of gambles. However, with the exception of one study, the samples recruited by Peters et al. were restricted mostly to younger adults. Some reports suggest that numeracy tends to be lower in older than younger adults (Kutner, Greenberg, & Baer, 2005), but the role of numeracy in decision-making competence remains an open question.

Experience and Automaticity—A second way of conceptualizing DMC is as a function of experience and automaticity. Experiencing the demands of daily life helps individuals to develop expert knowledge and automatic processes (which become increasingly selective and domain-specific with age, Baltes & Baltes, 1990; Salthouse, 1991). Older adults have been shown to perform better than younger adults on age-appropriate problems with which they are experienced (Artistico, Cervone, & Pezzuti, 2003; Brown & Park, 2003). Thus, DMC may be preserved with age to the extent that decision tasks are routine and predictable (cf. Baltes, 1993).

Associative and automatic processes may become more salient with age because of a strategic avoidance of more deliberative processing (Salthouse, 1996; Smith, 1996), shifts in motivation (Carstensen, 1995, 1998), or increased knowledge permitting reliance on gist (Adams, Smith, Nyquist, & Perlmutter, 1997; Hess, 1990; Reyna & Brainerd, 1995; Reyna, Lloyd, & Brainerd, 2003; Shanteau, 1992). These changes may result in more misunderstanding of information or greater vulnerability to cognitive biases with increasing age (Johnson, 1990) when decision tasks demand a thorough deliberation of information. However, studies show that both analytic and experiential or automatic processes are needed for making sound but efficient decisions. “Dual-process” models (Damasio, 1994; Epstein, 2003; Kahneman, 2003) imply that decision makers given the same information can use qualitatively different processing, potentially resulting in better or worse decisions depending on the match between the mode and task. For instance, a person who relies primarily on a feeling about the gist of information may have difficulty with literal (factual) comprehension problems, but be entirely consistent across similar problems.

Evidence suggests that more analytic information processing seems to be related to better decision making, at least on the types of non-prose behavioral decision tasks used in this study. For instance, an information processing style that reflects the tendency to think hard about problems, such as “need for cognition” (Cacioppo & Petty, 1982) or “analytic” (Epstein, Pacini, Denes-Raj, & Heier, 1996), has been correlated with fewer framing errors (Smith & Levin, 1996). In-depth processing of information may help decision makers see through irrelevant differences in framings that are normatively equivalent (Bruine De Bruin et al., 2007; although see LeBoeuf & Shafir, 2003; Parker, Bruine De Bruin, & Fischhoff, 2007).

Addressing Limitations of Previous Research

Developing a tool for measuring the DMC of older adults requires the identification and rigorous assessment of performance-based tasks that reliably capture distinct decision skills and correlate meaningfully with other variables known to influence decision processes. The present work addresses the limitations of previous research on aging and DMC in several ways. Tasks measuring comprehension and consistency are refined; new tasks measuring the ability to identify the weight placed on specific dimensions of information and the ability to inhibit impulsive responding in decision processes are developed and evaluated. A large, age-balanced sample has been recruited to permit comparison of the performance of younger, young-older, and old-older adults. Finally, measures of individual differences have

been expanded to include at least two marker tests of each of several basic cognitive abilities, to ensure robust analysis of the construct validity of the DMC indices.

Aims of the Current Study

The current study has three main aims.

Aim 1: Assemble items ranging in difficulty to measure specific decision skills (comprehension, consistency, dimension weighting, cognitive reflection) across the adult lifespan.

Aim 2: Evaluate the reliability of indices of comprehension, consistency, dimension weighting, and cognitive reflection in terms of internal consistency.

Aim 3: Evaluate the convergent and discriminant construct validity of the DMC indices in terms of the degree to which they correlate as expected with each other and with other relevant constructs (including age, social variables, health, cognitive abilities, attitudes and self-perceptions, decision style, numeracy) via (1) analyses of variance, (2) hierarchical regression, and (3) structural equation modeling.

Since one of our main interests is to provide a way of measuring differences between age groups, we tested the hypothesis of a negative association between age and performance on DMC indices. We also tested the hypothesis of a positive association between cognitive abilities and performance on DMC indices.

Methods

Participants

We recruited 611 participants from the Kaiser Permanente Hawai'i membership and the general community in Honolulu, Hawai'i. Each participant was paid \$40; for individuals tested outside our research center (e.g., at a retirement community), an additional \$5 per participant was donated to the participating organization.

Eligibility criteria were evaluated in a telephone-screening interview. The requirements were: at least 8th grade education, able to read English well or very well, not depressed (scoring less than 3 on the PHQ-2, Kroenke, Spitzer, & Williams, 2003), not anxious (scoring less than 3 on a 2-item anxiety questionnaire developed for this study), no physical ailment that would prevent participation (e.g., blindness, severe arthritis), and a score of 18 or higher on the ALFI-MMSE, a mini-mental status examination (Roccaforte, Burke, Bayer, & Wengel, 1992).

Materials and Procedure

Overview—Each participant completed a questionnaire booklet (average completion time was about 90 minutes for older adults and 45 minutes for younger adults). The full set of items in the tool for assessing DMC is provided in an online supplement (referred to hereafter as the Appendix), which can be found at <http://www.kpchr.org/research/public/OurResearchContent.aspx?pageid=58>. The booklet contained multiple pen-and-paper tasks designed to measure comprehension, consistency, dimension weighting, and cognitive reflection. (Additional filler tasks were presented but are not relevant to the present paper and thus not reported.) The content of the tasks was relevant to making choices in one of three domains that older adults are likely to encounter: (1) health (e.g., selecting a health care plan); (2) finance (e.g., selecting a mutual fund), and (3) nutrition (e.g., choosing among food products). Other sections of the booklet collected

information about decision style, self-perceptions, and demographics. The booklet was printed in 14-point font wherever possible to accommodate age-related vision deficits. Finally, the participants completed a battery of cognitive abilities measures administered in-person. The battery took about 30 minutes to complete.

Comprehension Measures (COM)—The comprehension problems had one of two formats. One format presented a table of options with their values on a set of dimensions (choice array), followed by literal and/or inferential comprehension questions. The second format involved a similar tabular presentation, but with instructions to choose an option that satisfies given criteria. In both formats, the correct answer was unambiguous and these items were scored using an item key. There were seven “simple” problems (COM-1 to COM-7) and seven “complex” problems (COM-8 to COM-14). Simple problems involved three to five options with up to seven dimensions; each complex problem described at least eight options on at least five dimensions. The score range (0-19) for the Comprehension Index is greater than the number of problems because some problems were accompanied by two questions. A sample comprehension problem is shown in Figure 1; the full set of problems is provided in the Appendix. Eight problems were developed de novo; six problems were based on problems used by Finucane et al. (2005), revised to increase their difficulty (by adding options or changing values in the choice array). Problems COM-1, COM-4, COM-6, COM-9, COM-12, and COM-13 in the present study were based on problems C-2, C-6, C-10, C-4, C-7, C-11, respectively, in Finucane et al. (2005).

Consistency Measures (CON)—The consistency measures were designed to measure the ability to hold constant the relative importance of decision dimensions across superficially different contexts. Russo et al.’s (1975) “ordering” paradigm was used for all consistency problems because Finucane et al. (2005) reported that a task based on this paradigm in the health domain showed moderate difficulty and good discriminability. Participants were presented with a total of nine problems to measure consistency: five simple problems and four complex problems. An example consistency problem is shown in Figure 2; the full set of problems is provided in the Appendix. Problems CON-1 and CON-6 were based on adaptations of problems I-1 and I-2 about health used by Finucane et al. (2005). Other consistency problems were developed de novo.

In the present study, each simple consistency problem (CON-1 to CON-5) presented individuals with a choice array of four options described on three or four dimensions, whereas each choice array in the complex problems (CON-6 to CON-9) presented 10 options. Each problem was presented twice. In one presentation, the options were shown in an invariant random order; in the second presentation, the options were shown ordered highest to lowest on one of the dimensions (the two presentations were separated by several other tasks). For each presentation, participants indicated their first choice. A pair of responses was scored as consistent when a participant selected the same option in both presentations.

Dimension Weighting (DW)—Developed de novo, the dimension-weighting (DW) measures again presented a choice array in each item, but assessed whether respondents reported accurately on which dimension they weighted most heavily in making a choice. When respondents do not do what they say they want to do, we can interpret that as an inability to maximize utility according to their preferences. This is not just inconsistency, but a failure or mistake indicating irrational behavior from a utility maximization perspective. Participants were presented with a total of 12 problems to measure their insight into their own dimension weighting: six simple problems [DW-1 to DW-6] and six complex problems [DW-7 to DW-12]. An example DW problem is shown in Figure 3; the full set of problems is provided in the Appendix.

In this series, each choice array presented values on several dimensions of a set of options. The values were designed so that the best value on the dimension considered most important corresponded with a unique option as the best choice. (Pilot testing identified the value differences necessary to create dominant options.) Participants were asked to indicate (a) which plan was his/her first choice and (b) which item of information (dimension) was most important to him/her when making that choice. Each simple problem described five options on five dimensions and each complex problem described eight options on eight dimensions.

Cognitive Reflection (CR)—Frederick's (2005) *Cognitive Reflection Test (CRT)* measures ability to suppress erroneous answers that spring to mind "impulsively." We interpret errors as indicating an inability to favor analytic over intuitive processes when needed in judgment and decision making. Participants were presented with a total of six problems (see Appendix) to measure cognitive reflection (CR): three problems from Frederick's original test and three new problems designed to parallel the original items. An example of one of the new problems is: "Soup and salad cost \$5.50 in total. The soup costs a dollar more than the salad. How much does the salad cost?"

Cognitive Ability Marker Tests—Each participant completed a battery that included at least two marker tests of each of four broad, second-order dimensions of intelligence: fluid intelligence (Gf), crystallized intelligence (Gc), memory span (Ms), and perceptual speed (Ps). This four-factor structure has been confirmed in older adults via confirmatory factor analyses (Baltes, Cornelius, Spiro, Nesselroade, & Willis, 1980). The measures were selected from a large body of factor-referenced tests that have been developed and validated over decades of research on the structure of mental abilities (Carroll, 1974; Cattell, 1971; Ekstrom, French, Harman, & Derman, 1976; Horn & Hofer, 1992; Salthouse, 1993; Thurstone, 1928; Wechsler, 1987). Tests required reading at 8th-grade level; we used enlarged print to assure legibility for older adults. Given the large number of tests in this battery, we used a fractional design to minimize participant burden, while ensuring that all tests were administered to some respondents and all possible pairs of constructs were represented. Table 1 shows that the design resulted in a total of 10 groups of participants, each providing five to seven measures taking between 27 and 35 minutes total to complete. Each test was taken by between four and eight groups. Group assignment was random and balanced over age strata. Table 2 shows the number of participants who actually were measured on each test assigned to their group

Crystallized intelligence: (Gc) was represented by three tests of verbal comprehension with varying difficulty. The first, drawn from the Primary Mental Abilities (PMA) collection of tests (Ekstrom et al., 1976), was *Recognition Vocabulary (V-1)*. We used the mean of the two V-1 part scores in analyses. The other verbal ability tests were Salthouse's (1993) *Synonyms* and *Antonyms* tests.

Fluid intelligence: (Gf) was represented by two measures of induction. The first measure was drawn from *Raven's Advanced Progressive Matrices* (Raven, Raven, & Court, 1998) and included three sample problems followed by the 18 odd-number items. The second measure of Gf was the PMA *Locations* test (Ekstrom et al., 1976); we used the mean of scores from the two sections of this test in analyses.

Memory span: (Ms) was represented first by *Backward Digit Span* from the Wechsler Adult Intelligence Scale III (WAIS-III) (Wechsler, 1987). The second measure was the *Paired Associates* test by Salthouse and Babcock (1991); we used the mean of scores from the two parts of this test in analyses.

Perceptual speed: (Ps) was represented by the primary ability of speed of processing. One measure was the WAIS-III *Digit Symbol Substitution* test. The second measure was the *Letter Comparison* test (Salthouse & Babcock, 1991); we used the mean of scores from the two parts of this test in analysis.

Other Measures

Decision style: The first two measures of decision style were presented during the in-person testing session. The *Rational-Experiential Inventory-Short Form (REI-S)* was used to capture individual differences in two factorially independent processing modes (Rational and Experiential), as proposed by Cognitive-Experiential Self-Theory (Denes-Raj & Epstein, 1994; Epstein, 1994; Epstein et al., 1996). Our second in-person measure of decision style was obtained from the *Risk-Benefit Rating* task (Finucane, Alhakami, Slovic, & Johnson, 2000). The score from this task used in analyses was the correlation between risk and benefit ratings on the 23 paired items within individual; a higher negative correlation indicated greater reliance on an experiential decision style. The third measure, the *Jellybeans* task (Denes-Raj & Epstein, 1994), was included in the questionnaire booklet. For analysis, the score was recentered at 0 and consequently ranged from -3 (preferred Bowl A, a more experiential decision process) to +3 (preferred Bowl B, a more analytic decision process).

All other measures (described below) were included in the questionnaire booklet.

Attitudes and self-perceptions: To measure experience and motivation, respondents were asked to indicate their agreement on a 4-point scale (1=strongly disagree, 2=disagree, 3=agree, 4=strongly agree) with two statements: (1) "I am very experienced with the kinds of decisions I made during this study;" (2) "When completing the tasks in this study, I was motivated to take each task seriously." Respondents were also asked to rate their skill in using tables and charts to make decisions on a 4-point scale (1=poor, 2=fair, 3=good, 4=excellent). Self-perceived memory (compared with people of the same age) was rated on a 5-point scale (1=much worse, 2=slightly worse, 3=about the same, 4=slightly better, 5=much better). Self-reported driving frequency during the past year was measured using a 5-point scale (1=never, 2=rarely, 3=off and on, 4=frequently, 5=almost every day).

To measure perceived need for decision support, respondents were asked to rate on a four-point scale (1=none, 2=a small amount, 3=moderate amount, 4=a lot) how much assistance they would seek if they had to choose a (1) health plan, (2) nutritional plan, or (3) checking account, in the next year. For each individual, we calculated an index of decision support by averaging responses across the 3 items.

Perceived decision self-efficacy was measured using 5 items adapted from Löckenhoff and Carstensen (2007). A sample item is: "Imagine you need medical care. How confident are you that you can make decisions that would help you get the best care?" Each item was rated on a 7-point scale from "cannot do at all" (1) to "certainly can do" (7). For each individual, we calculated an index of self-efficacy by summing responses across the 5 items.

Self-rated health: Respondents were asked to rate their "physical health" in comparison with others of the same age and sex on a 5-point scale (1= excellent, 2=good, 3=average, 4=fair, 5=poor). Respondents used the same scale to rate their "mental health." Respondents were asked to indicate whether they take prescription medication for each of 24 conditions (e.g., asthma, arthritis, cancer); the sum of positive responses was recorded as "number of medications taken." Finally, respondents were asked to indicate the number of times they visited a doctor or nurse in the last month (excluding overnight stays and dental visits) and the number of nights they stayed overnight in a hospital in the past year.

Numeracy: Our measure of numeracy is the total number of correct responses to 3 items testing comprehension of probabilistic information. The items are from a scale developed by Schwartz, Woloshin, Black, and Welch (1997). An index score was created by summing the number of correct responses for each respondent. We chose this three-item measure over a longer eight-item measure (Lipkus, Samsa, & Rimer, 2001) because the two measures are moderately correlated ($r_s = .48, p < .001$) but the shorter measure minimized participant burden (Donnelle, Hoffman-Goetz, & Arocha, 2007).

Demographics: We asked participants to report their birth date, gender, education, income category, and ethnicity.

Statistical Analysis

We used multiple imputation (SAS® PROC MI, data augmentation with Markov Chain Monte Carlo data generation) to fill in missing test scores. Three old-older participants did not complete any of the in-person tests and were excluded from the imputed dataset. Since most missing data were missing by design in the in-person test battery, these data can be considered missing completely at random. There were no missing data on the numeracy items or DMC index scores; skips were scored as an incorrect response. Eight imputations were created, in order to ensure efficiency of at least 95% (relative to asymptotic results) in all variables included in the planned analyses. In addition to the in-person tests, the imputation model included demographic features, responses to questions about other individual differences, and DMC index scores and other scores from the booklet.

All analyses were done using SAS® (Release 9.1). Analyses were carried out identically over all imputations, and results were combined using Rubin's rules (Rubin, 1987). This process involves adjusting the degrees of freedom, in order to obtain unbiased p-values when sample size depends in part on imputed data and the standard error estimates include between-imputation variability. Because data are complete by imputation, the sample size does not vary between analyses.

The analysis first examined the psychometric performance of the DMC tasks, including difficulty and discrimination. Next, we verified the internal consistency of the various multi-item scores. Then we discarded poorly performing tasks and adjusted scores to maximize psychometric quality. This work was followed by multi-step validation strategy of the DMC index scores, including evaluation of both convergent and discriminant validity and whether numeracy might be a mediator on the association between cognitive skills and DMC, using methods given by Baron and Kenny (1986).

Finally, structural equation modeling (SEM) of the covariance matrix, using SAS PROC CALIS, was used to test hypotheses about the association of participant age, numeracy, and cognitive abilities with DMC performance. Modeling proceeded in two steps: first verifying the fit of the hypothesized measurement model in the endogenous and exogenous variables and then adding and testing the fit of the hypothesized structural model linking the endogenous and exogenous latent factors (Mueller & Hancock, 2008). Observed variables were centered on their mean. For overall models, the mean was overall participants, and for analyses carried out by age group, the data were centered on the age group mean. The analyses were conducted using SAS® PROC CALIS, Rel 9.1.

The objective of the measurement model is to verify that the observed variables belong to their respective hypothesized latent factor, as indicated by the *t*-test on the estimate of the λ (slope) coefficient. The variable was dropped from the model if this test was not significant, constraining the parameter to 0 had a non-significant effect on model fit (χ^2 test on change in $-2*\log(\text{likelihood})$ ($p < .05$ for both tests), and modification tests did not suggest the

variable had a stronger association with a different latent factor. If modification tests suggested a variable belonged on another factor, we moved it to that factor and determined whether this improved the fit. Modifications were made one at a time.

The final measurement model was changed into the structural model by adding an equation in which the DMC latent factor is a function of the exogenous latent factors. The new parameters added at this stage were also examined for significance, and the model was simplified (by eliminating unneeded paths) where indicated.

Results

Participant Characteristics

As Table 3 shows, we were successful in enrolling approximately equal percentages of females in the three age groups. The three groups were similar in household income and education, but differed in race/ethnicity distribution and MMSE ($p < .0001$).

Internal Consistency of Predictor Measures

Internal consistency (Cronbach's α) was moderate for perceived need for decision support (0.71 overall, and 0.71, 0.65, 0.73, for younger, young-older, and old-older adults, respectively) and high for perceived decision self-efficacy (0.86 overall, and 0.83, 0.87, 0.87, for younger, young-older, and old-older adults, respectively). Cronbach's α for numeracy was 0.53 overall (0.54, 0.50, 0.48 for younger, young-older, and old-older adults, respectively).

Aim 1: Assemble Items Ranging in Difficulty

Overall, we were able to assemble items for measuring each of the decision skills with a considerable range of difficulty levels. Item difficulties within age subgroups ranged from 0.25 to 0.93 for the comprehension items, 0.30 to 0.85 for the consistency items, 0.55 to 0.88 for the dimension-weighting items, and 0.16 to 0.65 for the cognitive-reflection items (difficulties for each item are provided in the Appendix). Generally, to enable discrimination across a full range of abilities, a set of test items should include items with a range of difficulties, varying from around .10 to .90 (Betz, 1996) This range is covered by the items assembled, suggesting that these items would produce a range of scores. This selection of test items also ensured that even the least able respondents would be able to answer a few items correctly.

Aim 2: Evaluate the Reliability of the Indices of Decision-Making Competence

Aim 2(a): Internal Consistency—We created an index score for each respondent by summing the number of correct responses across the sets of problems in each of the four decision skills: comprehension, consistency, dimension weighting, and cognitive reflection. We obtained Cronbach's α on each of these four indices, then investigated whether we could improve coefficient alpha by removing items less related to the total score, dropping these items in a stepwise fashion. We found that the most reliable index was created by retaining all problems. Table 4 shows the final internal consistency coefficients for each index for each age group. Cronbach's α is adequate for all but the consistency index, which was dropped from further analysis.

Aim 2(b): Correlations of Items with Each Other—The correlations between the comprehension index and the dimension-weighting and cognitive-reflection indices are 0.46 and 0.60, respectively. The correlation between dimension weighting and cognitive

reflection is 0.30. All correlations are significant ($p < .0001$) and positive, indicating a positive manifold of tasks, with relative consistency in performance across indices.

A principal-components analysis of the items comprising the overall comprehension, dimension-weighting, and cognitive-reflection indices resulted in three factors accounting for 27.5% of the variance (see Table 5). With the exception of COM-1, COM-2, COM-8b, and COM-11, at least one factor loading for each variable is at least .30, suggesting internal consistency, in the sense that the problems capture an underlying construct for each of the main skills needed for competent decision-making. Generally, the items that comprise the first factor capture individuals' ability to use information (i.e., to combine or transform dimension values) and include mainly comprehension items. The items that comprise the second factor capture the ability to identify the importance of information (i.e., to recognize salient information or to understand one's own weighting scheme) and includes mainly dimension weighting. A third factor, composed of only two cognitive-reflection items may represent some sub-component of problem solving.

Aim 3: Evaluate the Construct Validity of the Indices of Decision-Making Competence

Aim 3(a): Analyses of Variance—Figure 4 shows the mean difficulty (proportion correct) of the comprehension, dimension-weighting, and cognitive-reflection items for each age group. As expected, one-way analysis of variance, with age as a between-groups factor, revealed a significant main effect of age for the comprehension index ($F[2,605] = 37.6, p < .0001$), the dimension-weighting index ($F[2,605] = 10.2, p < .0001$), and the cognitive-reflection index ($F[2,605] = 9.2, p < .0001$). Contrasts between age groups confirmed the differences were significant, with the sole exception of the young-old versus old-old contrast on cognitive reflection, $p = .0594$. Compared with younger adults, young-older and old-older adults generally experienced more difficulty on every type of task. Mean index scores (and standard deviations) for each age group for each index (Table 6) confirm that the age groups are ordered consistently on age with respect to each DMC index.

Aim 3(b): Correlations and Regressions—Covariation between age and other variables suggests that the differing performances of younger, young-older, and old-older adults on the DMC indices might be accounted for by differences in these variables across age groups, including mean age trends. As shown in Table 7, age groups differed on measures of health, cognitive abilities, and numeracy and on some measures of attitudes, self perceptions, and decision style. Significant differences were not always in the direction of decline with increasing mean age. For instance, means on physical health and performance on crystallized intelligence measures were positively associated with mean age across age groups.

Table 8 shows correlations of the three DMC indices with the measures of individual differences. Most of the variables in Table 8 correlate significantly with the index scores, indicating that several factors predict decision skills. However, significant intercorrelations exist among the individual difference variables as well (see Appendix). Given that the contributions of the various predictors are not necessarily independent of each other, we conducted stepped regression modeling to estimate the relative influences of age and individual difference variables on the DMC indices.

We conducted a hierarchical sequence of multiple regressions on data from the combined sample for each of the three most reliable index scores: comprehension, dimension weighting, and cognitive reflection. Primarily, we were interested in (1) how much age-related variance exists in each index; (2) excluding age, how much variance could be explained by social variables, health measures, basic cognitive abilities, attitudinal and self-

perception measures, and decision style; and (3) how much age-related variance is independent of the other variables in the model.

Results of the regressions are summarized in Table 9. Using the comprehension index score as the dependent variable, we found that when age was entered into the regression model by itself, it accounted for 10.6% of the variance. Then, excluding age, significant changes in R^2 were found after entering each of the other groups of variables. Adding age last resulted in a non-significant increment in R^2 of < 0.001 .

Using the dimension-weighting index score as the dependent variable, we found that when age was entered into the regression model by itself, it accounted for 3.0% of the variance. Then, excluding age, significant changes in R^2 were found after entering each of the other groups of variables. Adding age last resulted in a significant increment in R^2 but of only 0.003 (a change of 1.4%).

Using the cognitive-reflection index score as the dependent variable, we found that when age was entered into the regression model by itself, it accounted for 3.3% of the variance. Then, excluding age, significant changes in R^2 were found after entering the social, health, and cognitive skill variables. The attitudinal and decision-style measures did not increase R^2 significantly. Adding age last also resulted in a non-significant increment in R^2 of 0.003 (a change of 0.69%).

We were also interested in examining the extent to which the DMC indices could be explained (mediated) by numeracy above and beyond the influence of basic cognitive abilities (Baron & Kenny, 1986). Many judgment and decision tasks rely heavily on understanding basic mathematical concepts, so we expected a considerable amount of DMC variance to be accounted for uniquely by numeracy. First, we directly regressed the same set of ten cognitive skills used in the Table 9 analyses (see Analysis 2, Step 3 of Table 9) on each of the DMC indices, to confirm that there is an effect that might be mediated. Table 10 (Analysis 1, Step 1) shows the results of this analysis for comprehension, dimension weighting, and cognitive reflection. For comprehension, cognitive skills alone account for 46.7% of the variance. When we regressed the set of cognitive skills on numeracy, we obtained a model R^2 of 0.28, $p < .0001$, which confirms that these are related (Baron and Kenny, 1986). When numeracy is entered in Analysis 1, Step 2, it adds just 3.9% of explained variance in performance on the comprehension index. Thus, after adjusting for cognitive skills, numeracy accounts for a small, but significant amount of comprehension variance and so appears to be a mediator. However, the magnitude of ΔR^2 (see Analysis 2, Step 2, when cognitive skills are added to a model with only numeracy) indicates that numeracy is not a total mediator—after accounting for numeracy, cognitive skills account for an additional 25.3% of variance in comprehension. A comparison of Total R^2 in Step 1 of Analysis 1 and ΔR^2 in Step 2 of Analysis 2 reveals that cognitive skills' explanation of comprehension variance decreases from 46.7% to 25.3% after controlling for numeracy. Thus, numeracy accounts for 45.8% (21.4% out of 46.7%) of the cognitive skills-related variance in the comprehension index.

A similar pattern of results was found when we conducted hierarchical regressions to predict the dimension-weighting and cognitive-reflection indices (Table 10). Cognitive skills alone account for 13.2% (34.8%) of the explained variance in dimension weighting (cognitive reflection). Numeracy alone accounts for 10.4% (28.0%) of the variance. After controlling for cognitive skills, numeracy accounts for an additional 4.5% (7.2%) of the variance. Numeracy accounts for 44.7% (59.8%) of cognitive skills-related variance in the dimension-weighting (cognitive-reflection) indices.

Following the steps outlined by Baron and Kenny (1986), we can conclude that numeracy does not completely mediate the relationship between cognitive skills and DMC indices. Even though cognitive skills are related to both numeracy and the DMC indices and numeracy affects the indices after controlling for cognitive skills, there is still a relationship between cognitive skills and the indices after controlling for numeracy.

Aim 3(c): Structural Equation Modeling—We evaluated the fit between the new data collected in this study and our hypothesized model of DMC, starting with the measurement model (the associations among exogenous variables only). An examination of the initially hypothesized measurement model indicated that risk-benefit correlation was not confirmed as an indicator for decision style and we removed it completely from the dataset, which significantly improved fit. We tested three ways of adding numeracy to the model: as a single-variable factor, as a marker on decision style and as a marker on the DMC factor. All of these approaches resulted in worse fit, and numeracy was omitted from the measurement model. Based on modification indices we evaluated whether combining some of the cognitive skills latent factors might improve fit. The final measurement model was the starting point for the next phase: structural modeling. We removed several paths in response to p -values $> .05$. These included correlation between G_f and G_c , and the path from Decision Style to DMC. The Goodness of Fit Index (GFI) was 0.90 and the root mean square residual (RMR) was 0.08. We were not able to obtain an interpretable solution in the analysis by age group, possibly because the sample size within age groups is too small. The final structural equation model is shown in Figure 5.

Discussion

Our first aim was to assemble performance-based items ranging in difficulty to measure specific DMC skills across the adult life span. The items produced a test yielding an appropriate range of scores and different levels of complexity that are often faced in real life (Tanius et al., 2009). Unlike previous research, we developed multiple new problems for measuring DMC in the domains of health, nutrition, and finance. The present study thus extended the work of Finucane et al. (2005) by developing items with a wider range of difficulty and measuring additional skills (dimension weighting and cognitive reflection) in multiple domains in a standardized manner across the adult lifespan.

Our second aim was to assess the reliability of indices of specific DMC skills. The new index of comprehension showed better internal consistency ($\alpha \geq 0.74$) than was found by Finucane et al (2005), who reported their comprehension error index had coefficient alphas of 0.52 for younger adults (19-50 years) and 0.58 for older adults (65-94 years). Good internal consistency was found for the new dimension-weighting index ($\alpha \geq 0.54$) and cognitive-reflection index ($\alpha \geq 0.77$). Unfortunately, we were unable to improve the internal consistency of the consistency index (≤ 0.44 , similar to the findings of Finucane et al.). Intercorrelations between the items comprising the new indices reflected two related, mutually supportive constructs, long identified as central to decision making (Edwards, 1954; Raiffa, 1968). These results also support Stanovich and West's (2000) suggestion that performance on conventional decision making tasks reflects a positive manifold rather than random performance errors. The indices' reliability suggests potential for developing a normed psychometric DMC test.

Our third aim was to assess the construct validity of the DMC indices. We examined performance on DMC indices across a wide age-span, with special focus on a previously neglected and poorly understood group, namely, individuals over 74 years of age. As predicted, we found that compared with younger adults, young-older and old-older adults showed poorer performance on the DMC indices. We also found, as predicted, better

performance on the DMC indices was related primarily to better cognitive skills. Higher education and income, being male, and being healthier were additional predictors of better DMC. The dominance of cognitive skills as a predictor helps build the case for the utility of these DMC indices, suggesting that they mostly reflect people's ability to understand and use information effectively (as opposed to being primarily a function of broader demographic factors).

The results showed also that numeracy accounted for a significant amount (but not all) of the explained variance in DMC indices, above and beyond the variance accounted for by cognitive skills. The ability to work with numbers and to understand basic probability and mathematical concepts seems to be an applied skill, independent of basic cognitive abilities, that is required for good decision making.

The present results supported the discriminant validity of the DMC indices. Whereas we expect our indices to be positively correlated with related abilities, such as memory or speed of processing, those correlations cannot be too high or they would indicate the tool might be measuring the same construct. Thus, moderate, positive associations with related, but not identical constructs, demonstrated adequate discriminant validity.

The structural equation model portrayed in Figure 5 indicates that across the adult lifespan, individuals with lower crystallized intelligence or poorer performance on other cognitive abilities tend to exhibit poorer performance on our DMC measures. Increasing age also signifies poorer DMC performance. Decision style, however, only signifies change in other key variables. Rather than implying that the relationship of each exogenous factor exerts an independent effect on DMC, the results suggest that we may in fact need fewer independent explanatory mechanisms to account for these effects (c.f. Salthouse & Ferrer-Caja, 2003). While our study demonstrates the feasibility of using structural modeling to examine the existence and relative magnitude of statistically distinct influences on DMC, the generalizability of the results should be tested with data from different samples and different operationalizations of the theoretical constructs.

The findings reported above are restricted to the DMC skills we examined. There may be different relationships between age and other skills (e.g., involving the application of emotional wisdom). For instance, Bruine de Bruin et al. (2007) reported that older adults performed significantly better than younger adults on a task called "recognizing social norms." As suggested elsewhere (Baltes & Baltes, 1990; Denney, 1989; Finucane et al., 2002; Salthouse, 1991), older adults' knowledge and experience may make up for declines in cognitive abilities and allow their problem solving to be more automatized.

In contrast to the A-DMC developed by Bruine de Bruin et al. (2007), the indices in the present study permit examination of DMC at the level of component abilities. Our approach allows a refined analysis of specific skills that may increase or decrease with age and other variables, thereby pinpointing exactly which part of a decision process is going well or poorly. Nonetheless, many real-world decisions require multiple skills to be applied simultaneously and the A-DMC seems to be a good tool for measuring this type of performance. Together, the two tools provide a comprehensive approach. Future research could examine the external validity of both tools.

Limitations

Several limitations of this study should be noted. First, as a rule of thumb, some authors strictly require a reliability of 0.70 or higher before they will use an instrument. This level of reliability was achieved only by the comprehension and cognitive-reflection indices. The rule of thumb should be applied with caution, though, because the appropriate degree of

reliability depends upon the use of the instrument (e.g., we could increase reliability by adding items, but since this tool is designed to be appropriate for older adults, we intentionally kept it short).

A second limitation of this study is that we used convenience samples and (for budget reasons) did not sample 45-64 year olds. Consequently, the results may not generalize to the broader population. A third limitation is that our research did not address activities of daily living besides health, nutrition, and financial decisions. Researchers should assess health-risk behaviors (notably smoking, exercise, and driving) and retirement and death issues (e.g., inheritance tax issues, advance directives), but those were beyond the scope of this study. Fourth, the content of the cognitive-reflection items is somewhat different from that of the comprehension and dimension-weighting items. Future research might benefit from content reflecting more real-world, personally relevant choices for cognitive-reflection items.

Fifth, the cross-sectional rather than longitudinal design means that cohort effects may be interpreted as developmental changes. The better control offered by a longitudinal design would be advantageous, but the cost and logistical complexity of conducting a study that would extend for more than 50 years (if 45 year olds were to be followed until they were 100 years old) is not warranted until a reliable and valid DMC measurement tool is established. The cross-sectional design also poses a challenge for interpreting “age-related variance” because the design confounds individual differences in age and the average between-person differences. The covariance may be at least partly a product of age-related mean trends in the population (Hofer, Berg, & Era, 2003).

Finally, we must underscore that our correlational results do not imply causality. Although speculation about the cognitive processes or neural substrates responsible for variables influencing DMC is tempting, our correlational analyses are best viewed as a classification and not an explanation of these effects. More direct tests of causal relationships include prospective studies examining the effects of decision training on later outcomes (e.g., Downs et al., 2004).

Despite the above limitations, the results indicate progress towards a reliable and valid test of older adults' DMC. This tool is critical for future studies aimed at determining how to facilitate optimal functioning in a rapidly aging society.

Theoretical Implications

In this and previous papers, we conceive of DMC as a multidimensional concept, including the skills of comprehension, consistency, identifying information relevance, and tempering impulsivity. Each of these abilities is expected to tap functionally different areas of decision processes. Although we were unable to develop a reliable index of consistency, reliability of the other indices was adequate and moderate correlations among the indices suggested related yet distinct components of DMC. These results show the importance of distinguishing among decision skills. Other components of DMC need to be investigated to further our understanding of complex constructs like “rational” information integration and how and why DMC changes with age.

A view of DMC as a complex, “compiled” form of cognition (Appelbaum & Grisso, 1988; Park, 1992; Salthouse, 1990; Willis & Schaie, 1986) is supported by our finding that basic cognitive abilities were the strongest predictors of DMC. Age-related changes in DMC are thus likely to have similar characteristics to age-related changes in basic abilities. The structural modeling results suggest that only a small number of explanatory mechanisms may be needed to account for relationships between exogenous factors and DMC performance. Finding that numeracy accounts for additional DMC variance above and

beyond good cognitive abilities suggests that models of sound decision making should include an ability to understand and work with numbers.

Policy Implications—A robust method for measuring individual differences in DMC can enable decision support to be tailored towards those who need it in a timely manner. The present research has advanced a way of measuring exactly when, where, and how older adults need help and whether old-older adults' needs differ from those of young-older adults. Decision support may come in the form of decision agents or optimal information presentation formats and training tools. Longer lifespans and the rapid aging of the world's population (United Nations Population Division, 2002) demand ways of creating decision environments that rely less on qualities typical of youth (e.g., speed) and rely more on qualities typical of age (e.g., crystallized intelligence, wisdom, emotional regulation).

In sum, we have developed and evaluated new, standardized indices of comprehension, dimension weighting, and cognitive reflection, applicable across the adult lifespan. The present results support the overall construct validity of DMC as well as the predictive validity of tasks drawn from behavioral decision research. Our systematically developed battery of items, which were tested on a large, diverse sample that included old-older adults, suggests that the DMC indices have promise for measuring individual differences in decision skills.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This paper was prepared with support from the National Institute on Aging (Grant #R01AG021451-01A2) and the National Science Foundation (Grant #BCS-0525238) awarded to the first author during her tenure as Senior Research Investigator at the Center for Health Research, Hawaii. The authors are indebted to the individuals who agreed to participate in this study. We wish to thank Rebecca Cowan (Project Manager) and the research assistants who helped with data collection. We also thank Lisa Fox, Lin Neighbors, Jennifer Coury, Daphne Plaut, and Catherine Marlow for help with materials and manuscript preparation.

References

- Adams C, Smith MC, Nyquist L, Perlmutter M. Adult age-group differences in recall for the literal and interpretive meanings of narrative text. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 1997;4:187–195.
- Allaire JC, Marsiske M. Everyday cognition: Age and intellectual ability correlates. *Psychology and Aging* 1999;14:627–644. [PubMed: 10632150]
- Appelbaum PS, Grisso T. Assessing patients' capacities to consent to treatment. *New England Journal of Medicine* 1988;319:1635–1638. [PubMed: 3200278]
- Appelbaum PS, Grisso T, Frank E, O'Donnell S, Kupfer D. Competence of depressed patients for consent to research. *American Journal of Psychiatry* 1999;156(9):1380–1384. [PubMed: 10484948]
- Artistico D, Cervone D, Pezzuti L. Perceived self-efficacy and everyday problem solving among young and older adults. *Psychology and Aging* 2003;18:68–79. [PubMed: 12641313]
- Baltes PB. The aging mind: Potential and limits. *The Gerontologist* 1993;33:580–594. [PubMed: 8225002]
- Baltes, PB.; Baltes, MM. Psychological perspectives on successful aging: The model of selective optimization with compensation. In: Baltes, MM.; Baltes, PB., editors. *Successful aging: Perspectives from the behavioral sciences*. Cambridge University Press; Cambridge, England: 1990. p. 1-34.
- Baltes PB, Cornelius SW, Spiro A, Nesselroade JR, Willis SL. Integration versus differentiation of fluid/crystallized intelligence in old age. *Developmental Psychology* 1980;16:625–635.

- Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology* 1986;51:1173–1182. [PubMed: 3806354]
- Betz, NE. Test construction. In: Leong, FT.; Austin, JT., editors. *The psychology research handbook: A guide for graduate students and research assistants*. Sage; Thousand Oaks: 1996. p. 239-250.
- Brown SC, Park DC. Theoretical models of cognitive aging and implications for translational research in medicine. *The Gerontologist* 2003;43(1):57–67. [PubMed: 12637690]
- Bruine De Bruin W, Parker AM, Fischhoff B. Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology* 2007;92(5):938–956. [PubMed: 17484614]
- Cacioppo JT, Petty RE. The need for cognition. *Journal of Personality and Social Psychology* 1982;42:116–131.
- Carroll JB. Psychometric tests as cognitive tasks: A new “structure of intellect.”. *Educational Testing Service* 1974:74–116.
- Carstensen LL. Evidence for a life-span theory of socioemotional selectivity. *Current Directions in Psychological Science* 1995;4:151–156.
- Carstensen, LL. A life-span approach to social motivation. In: Heckhausen, J.; Dweck, CW., editors. *Motivation and self-regulation across life span*. Cambridge University Press; New York: 1998. p. 341-364.
- Cattell, RB. *Abilities: Their structure, growth and action*. Houghton Mifflin; New York: 1971.
- Clark LA, Watson D. Constructing validity: Basic issues in objective scale development. *Psychological Assessment* 1995;7(3):309–319.
- Damasio, AR. *Descartes’ error: Emotion, reason, and the human brain*. Avon; New York: 1994.
- Denes-Raj V, Epstein S. Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology* 1994;66:819–829. [PubMed: 8014830]
- Denney, NW. Everyday problem solving: Methodological issues, research findings, and a model. In: Poon, LW.; Rubin, DC.; Wilson, BA., editors. *Everyday cognition in adulthood and late life*. Cambridge University Press; New York: 1989. p. 330-351.
- Donnelle L, Hoffman-Goetz L, Arocha JF. Assessing health numeracy among community-dwelling older adults. *Journal of Health Communication* 2007;12:651–665. [PubMed: 17934942]
- Downs JS, Murray PJ, Bruine de Bruin W, White JP, Palmgren C, Fischhoff B. An interactive video program to reduce adolescent females’ STD risk: A randomized controlled trial. *Social Science and Medicine* 2004;59:1561–1572. [PubMed: 15279915]
- Edwards W. The theory of decision making. *Psychological Bulletin* 1954;51:380–417. [PubMed: 13177802]
- Ekstrom, RB.; French, JW.; Harman, H.; Derman, D. *Manual for kit of factor-referenced cognitive tests*. Educational Testing Service; Princeton, NJ: 1976.
- Epstein S. Integration of the cognitive and the psychodynamic unconscious. *American Psychologist* 1994;49:709–724. [PubMed: 8092614]
- Epstein, S. Cognitive-experiential self-theory of personality. In: Millon, T.; Lerner, MJ., editors. *Handbook of psychology: Personality and social psychology*. Vol. Vol. 5. Wiley & Sons; Hoboken, NJ: 2003. p. 159-184.
- Epstein S, Pacini R, Denes-Raj V, Heier H. Individual differences in intuitive-experiential and analytical-rational thinking styles. *Journal of Personality and Social Psychology* 1996;71:390–405. [PubMed: 8765488]
- Finucane ML, Alhakami A, Slovic P, Johnson SM. The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making* 2000;13:1–17.
- Finucane ML, Mertz CK, Slovic P, Schmidt ES. Task complexity and older adults’ decision-making competence. *Psychology and Aging* 2005;20(1):71–84. [PubMed: 15769215]
- Finucane ML, Slovic P, Hibbard J, Peters E, Mertz CK, MacGregor DG. Aging and decision-making competence: An analysis of comprehension and consistency skills in older versus younger adults considering health-plan options. *Journal of Behavioral Decision Making* 2002;15:141–164.

- Fitten LJ, Lusky R, Hamann C. Assessing treatment decision-making capacity in elderly nursing home residents. *Journal of the American Geriatrics Society* 1990;38:1097–1104. [PubMed: 2229863]
- Frederick S. Cognitive reflection and decision making. *Journal of Economic Perspectives* 2005;19(4): 25–42.
- Grisso T, Appelbaum P, Hill-Fotouhi C. The MacCAT-T: A clinical tool to assess patients' capacities to make treatment decisions. *Psychiatric Services* 1997;48(11):1415–1419. [PubMed: 9355168]
- Hess, TM. Aging schematic influences on memory. In: Hess, TM., editor. *Aging and cognition: Knowledge organization and utilization: Advances in Psychology*. Vol. Vol. 71. North-Holland; Amsterdam, Netherlands: 1990. p. 93-160.
- Hibbard JH, Peters E, Dixon A, Tusler M. Consumer competencies and the use of comparative quality information. *Medical Care Research and Review* 2007;64(4):379–394. [PubMed: 17684108]
- Hofer SM, Berg S, Era P. Evaluating the interdependence of aging-related changes in visual and auditory acuity, balance, and cognitive functioning. *Psychology and Aging* 2003;18(2):285–305. [PubMed: 12825777]
- Horn, JL.; Hofer, SM. Major abilities and development in the adult period. In: Sternberg, RJ.; Berg, CA., editors. *Intellectual development*. Cambridge University Press; Cambridge, England: 1992. p. 44-49.
- Janofsky J, McCarthy R, Folstein M. The Hopkins Competency Assessment Test: A brief method for evaluating patients' capacity to give informed consent. *Hospital and Community Psychiatry* 1992;43:132–136. [PubMed: 1572608]
- Johnson M. Age differences in decision making: A process methodology for examining strategic information processing. *Journal of Gerontology: Psychological Sciences* 1990;45(2):75–78.
- Kahneman D. A perspective on judgment and choice. *American Psychologist* 2003;58(9):697–720. [PubMed: 14584987]
- Kim S, Karlawish J, Caine E. Current state of research on decision-making competence of cognitively impaired elderly persons. *American Journal of Geriatric Psychiatry* 2002;10(2):151–165. [PubMed: 11925276]
- Kim SYH, Caine ED, Currier GW, Leibovici A, Ryan JM. Assessing the competence of persons with Alzheimer's disease in providing informed consent for participation in research. *American Journal of Psychiatry* 2001;158(5):712–717. [PubMed: 11329391]
- Kitamura F, Tomoda A, Tsukada K, Tanaka M, Kawakami I, Mishima S, et al. Method for assessment of competency to consent in the mentally ill: Rationale, development, and comparison with the medically ill. *International Journal of Law and Psychiatry* 1998;21(3):223–244. [PubMed: 9717087]
- Kitamura T, Kitamura F. Reliability of clinical judgment of patients' competency to give informed consent: A case vignette study. *Psychiatry and Clinical Neurosciences* 2000;54(2):245–247. [PubMed: 10803822]
- Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Medical Care* 2003;41(11):1284–1292. [PubMed: 14583691]
- Kutner, M.; Greenberg, E.; Baer, J. *A First Look at the Literacy of America's Adults in the 21st Century (NCES 2006-470)*. Washington, DC: US Department of Education, National Center for Education Studies (NCES); 2005.
- LeBoeuf TA, Shafir E. Deep thoughts and shallow frames: On the susceptibility to framing effects. *Journal of Behavioral Decision Making* 2003;16:77–92.
- Lipkus IM, Samsa G, Rimer BK. General performance on a numeracy scale among highly educated samples. *Medical Decision Making* 2001;21:37–44. [PubMed: 11206945]
- Löckenhoff CE, Carstensen LL. Aging, Emotion, and Health-Related Decision Strategies: Motivational Manipulations Can Reduce Age Differences. *Psychology and Aging* 2007;22(1): 134–146. [PubMed: 17385990]
- MacPherson SE, Phillips LH, Sala SD. Age, executive function, and social decision making: A dorsolateral prefrontal theory of cognitive aging. *Psychology and Aging* 2002;17(4):598–609. [PubMed: 12507357]
- Marson DC. Determining the competency of Alzheimer patients to treatment and research. *Alzheimer Disease and Associated Disorders* 1994;8:5–18. [PubMed: 7695856]

- Marson DC, McInturff B, Hawkins L, Bartolucci A, Harrell LE. Consistency of physician judgments of capacity to consent in mild Alzheimer's disease. *Journal of the American Geriatric Society* 1997;45:453–457.
- Meyer B, Marsiske M, Willis SL. Text processing variables predict the readability of everyday documents read by older adults. *Reading Research Quarterly* 1993;28:234–249.
- Meyer BJF, Russo C, Talbot AP. Discourse comprehension and problem solving: Decisions about the treatment of breast cancer by women across the life span. *Psychology and Aging* 1995;10:84–103. [PubMed: 7779321]
- Mueller, RO.; Hancock, GR. Best practices in structural equation modeling. In: Osborne, J., editor. *Best practices in quantitative methods*. Sage; Thousand Oaks, CA: 2008. p. 488-508.
- Palmer WB, Nayak GV, Dunn LB, Appelbaum PS, Jeste DV. Treatment-related decision-making capacity in middle-aged and older patients with psychosis: A preliminary study using the MacCAT-T and HCAT. *American Journal of Geriatric Psychiatry* 2002;10(2):207–211. [PubMed: 11925282]
- Park, DC. Applied cognitive aging research. In: Craik, F.; Salthouse, TA., editors. *The handbook of aging and cognition*. Erlbaum; Hillsdale, NJ: 1992. p. 449-493.
- Park DC, Willis SL, Morrow D, Diehl M, Gaines CL. Cognitive function and medication usage in older adults. *The Journal of Applied Gerontology* 1994;13(1):39–57.
- Parker AM, Bruine De Bruin W, Fischhoff B. Maximizers versus satisficers: Decision-making styles, competence, and outcomes. *Judgment and Decision Making* 2007;2(6):342–350.
- Parker AM, Fischhoff B. Decision-making competence: External validation through an individual-differences approach. *Journal of Behavioral Decision Making* 2005;18:1–27.
- Peters E, Dieckmann N, Västfjäll D, Mertz CK, Slovic P. Bringing meaning to numbers: The impact of evaluative categories on decisions. *Journal of Experimental Psychology: Applied* 2009;15(3):213–227. [PubMed: 19751072]
- Peters E, Västfjäll D, Slovic P, Mertz CK, Mazzocco K, Dickert S. Numeracy and decision making. *Psychological Science* 2006;17(5):407–413. [PubMed: 16683928]
- Raiffa, H. *Decision analysis: Introductory lectures on choices under uncertainty*. Addison-Wesley; Reading, MA: 1968.
- Raven, J.; Raven, JC.; Court, JH. *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 4: The Advanced Progressive Matrices*. Harcourt Assessment; San Antonio, TX: 1998.
- Reyna VF, Brainerd CJ. Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences* 1995;7:1–75.
- Reyna, VF.; Lloyd, FJ.; Brainerd, CJ. Memory, development, and rationality: An integrative theory of judgment and decision-making. In: Schneider, S.; Shanteau, J., editors. *Emerging perspectives on decision research*. Cambridge University Press; New York: 2003. p. 201-245.
- Roccaforte WH, Burke WJ, Bayer BL, Wengel SP. Validation of a telephone version of the Mini-Mental State Examination. *Journal of the American Geriatrics Society* 1992;40(7):697–702. [PubMed: 1607586]
- Rönnlund M, Karlsson E, Laggnäs E, Larsson L, Lindström T. Risky decision making across three arenas of choice: Are younger and older adults differently susceptible to framing effects? *The Journal of General Psychology* 2005;132(1):81–92. [PubMed: 15685961]
- Rubin, DB. *Multiple Imputations for Nonresponse in Surveys*. Wiley; New York: 1987.
- Russo JE, Krieser G, Miyashita S. An effective display of unit price information. *Journal of Marketing* 1975;39:11–19.
- Ryan, JJ.; Lopez, SJ.; Sumerall, SW. Understanding test construction. In: Dorfman, W.; Hersen, M., editors. *Understanding psychological assessment*. Kluwer Academic/Plenum; New York: 2001.
- Salthouse, TA. Cognitive competence and expertise in the aging. In: Birren, JE.; Schaie, KW., editors. *Handbook of the psychology of aging*. 3rd ed. Academic Press; New York: 1990. p. 310-319.
- Salthouse TA. Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science* 1991;2:179–183.
- Salthouse TA. Speed and knowledge as determinants of adult age differences in verbal tasks. *Journal of Gerontology: Psychological Sciences* 1993;48:29–36.

- Salthouse TA. The processing-speed theory of adult age differences in cognition. *Psychological Review* 1996;103:403–428. [PubMed: 8759042]
- Salthouse TA, Babcock RL. Decomposing adult age differences in working memory. *Developmental Psychology* 1991;27:763–776.
- Salthouse TA, Ferrer-Caja E. What needs to be explained to account for age-related effects on multiple cognitive variables? *Psychology and Aging* 2003;18(1):91–110. [PubMed: 12641315]
- Schaie, KW.; Willis, SL. Theories of everyday competence and aging. In: Bengtson, VL.; Schaie, KW., editors. *Handbook of theories of aging*. Springer; New York: 1999. p. 174-195.
- Schmand B, Gouwenberg B, Smit JH, Jonker C. Assessment of mental competency in community-dwelling elderly. *Alzheimer Disease and Associated Disorders* 1999;13(2):80–87. [PubMed: 10372950]
- Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine* 1997;127(11):966–972. [PubMed: 9412301]
- Shanteau J. How much information does an expert use? Is it relevant? *Acta Psychologica* 1992;81:75–86.
- Smith, AD. Memory. In: Birren, JE., editor. *The Encyclopedia of Gerontology*. Academic Press; New York: 1996. p. 107-119.
- Smith SM, Levin IP. Need for cognition and choice framing effects. *Journal of Behavioral Decision Making* 1996;9:283–290.
- Stanley B, Guido J, Stanley M, Shortell D. The elderly patient and informed consent: Empirical findings. *Journal of the American Medical Association* 1984;252(10):1302–1306. [PubMed: 6471248]
- Stanovich KE, West RF. Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences* 2000;Vol 23(5):645–726. [PubMed: 11301544]
- Tanius BE, Wood S, Hanoch Y, Rice T. Aging and choice: Applications to Medicare Part D. *Judgment and Decision Making* 2009;4(1):92–101.
- Thurstone LL. Attitudes can be measured. *American Journal of Sociology* 1928;33:529–554.
- United Nations Population Division. World population ageing: 1950-2050. 2002. from <http://www.un.org/esa/population/publications/worldageing19502050/index.htm>
- Vellinga A, Smit J, van Leeuwen E, van Tilburg W, Jonker C. Instruments to assess decision-making capacity: An overview. *International Psychogeriatric Association* 2004;16(4):397–419.
- Weber, EU.; Goldstein, WM.; Barlas, S. And let us not forget memory: The role of memory processes and techniques in the study of judgment and choice. In: Busemeyer, JR.; Hastie, R.; Medin, DL., editors. *Decision making from the perspective of cognitive psychology*. Academic Press; San Diego: 1995. p. 33-81.
- Wechsler, D. Wechsler adult intelligence scale - III. The Psychological Corporation; New York: 1987.
- Willis, SL.; Schaie, KW. Practical intelligence in later adulthood. In: Sternberg, R.; Wagner, R., editors. *Practical intelligence*. Cambridge University Press; New York: 1986. p. 236-270.

COM-3. Consider the information presented below about several brands of yogurt, and answer the following questions. (All the information needed to answer the questions is given in the box.)

| Yogurt Nutrition Facts | | | | | |
|------------------------|---------|---------|---------|---------|---------|
| | Brand A | Brand B | Brand C | Brand D | Brand E |
| Serving Size | 100g | 250g | 200g | 250 g | 200g |
| Calories Per Serving | 220 | 250 | 190 | 200 | 180 |
| Amount Per Serving | | | | | |
| Total Fat | 1.5g | 3.0g | 2.0g | 2.5 g | 3.0g |
| Cholesterol | 20.0mg | 15.0mg | 25.0mg | 10.0 mg | 15.0mg |
| Sugars | 30g | 25g | 20g | 29 g | 30g |
| Protein | 14g | 18g | 14g | 12 g | 13g |
| Net Weight | 500g | 500g | 1000g | 1000 g | 500g |

| | |
|---|--|
| <p>(a) Which brand has the lowest amount of protein per serving? (Mark only one answer.)</p> <p><input type="checkbox"/> Brand A</p> <p><input type="checkbox"/> Brand B</p> <p><input type="checkbox"/> Brand C</p> <p><input type="checkbox"/> Brand D</p> <p><input type="checkbox"/> Brand E</p> | <p>(b) Which brand has the lowest number of servings per container? (Mark only one answer.)</p> <p><input type="checkbox"/> Brand A</p> <p><input type="checkbox"/> Brand B</p> <p><input type="checkbox"/> Brand C</p> <p><input type="checkbox"/> Brand D</p> <p><input type="checkbox"/> Brand E</p> |
|---|--|

Figure 1. Example comprehension problem with (a) literal and (b) inferential questions. The correct answer for (a) is Brand D and for (b) is Brand B.

| | | | |
|---|--|--------------------------------------|-------------------------------------|
| <p>CON-5. Below are facts about some mutual funds. For each fund, the box shows the previous year's gross return, minimum investment required, and years that the fund has been active. Based on these data, which fund would you choose? Note that there is no right answer. Base your answer on your own preference.</p> | | | |
| | <p>Previous Year's Gross Return</p> | <p>Minimum Investment</p> | <p>Years of Activity</p> |
| Fund A | 7.10% | \$1,500 | 15 |
| Fund B | 5.15% | \$1,000 | 8 |
| Fund C | 6.60% | \$1,500 | 7 |
| Fund D | 8.40% | \$2,500 | 10 |
| <p>Which fund would be your first choice? (Mark only one answer.)</p> <p style="padding-left: 40px;"> <input type="checkbox"/> Fund A <input type="checkbox"/> Fund B <input type="checkbox"/> Fund C <input type="checkbox"/> Fund D </p> | | | |

Figure 2. Example consistency problem presenting unordered information (the second presentation shows the same information ordered by gross return).

| <p>DW-2. Imagine you need to choose a health plan based on the member ratings given in the box below. Look carefully at the data and tell us which one you would choose. Note that there is no right answer. Base your answer on your own preference.</p> | | | | | |
|--|--------|--------|--------|--------|--------|
| | Plan A | Plan B | Plan C | Plan D | Plan E |
| Percent of members who believe they are able to get early disease detection | 62% | 74% | 61% | 63% | 60% |
| Percent of members very satisfied overall with plan | 60% | 62% | 74% | 61% | 63% |
| Percent of members who said it was easy to get a referral to a specialist | 61% | 63% | 60% | 62% | 74% |
| Percent of members satisfied with customer service | 63% | 60% | 62% | 74% | 61% |
| Percent of members satisfied with physical therapy facilities | 74% | 61% | 63% | 60% | 62% |
| <p>(a) Which plan would be your first choice? (Mark only one answer.)</p> <p>___ Plan A</p> <p>___ Plan B</p> <p>___ Plan C</p> <p>___ Plan D</p> <p>___ Plan E</p> | | | | | |
| <p>(b) Which item of information was most important to you when you made your choice in Question (a)? (Mark only one answer.)</p> <p>___ Percent of members who believe they are able to get early disease detection</p> <p>___ Percent of members very satisfied overall with plan</p> <p>___ Percent of members who said it was easy to get a referral to a specialist</p> <p>___ Percent of members satisfied with customer service</p> <p>___ Percent of members satisfied with physical therapy facilities</p> | | | | | |

Figure 3.
Example dimension-weighting problem.

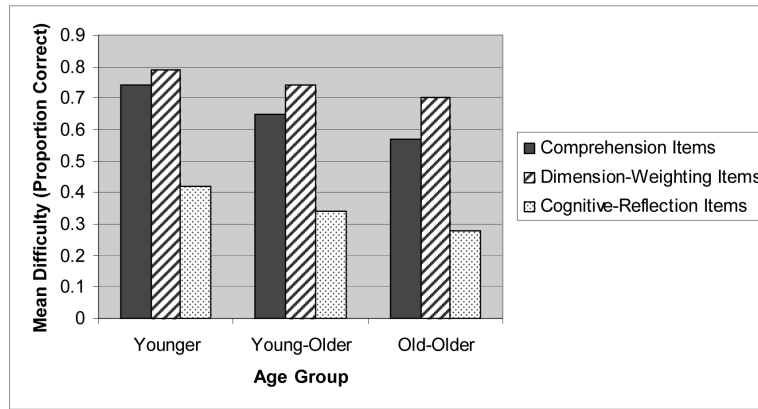


Figure 4. Mean difficulty (proportion correct) for comprehension, dimension-weighting, and cognitive-reflection items for each age group.

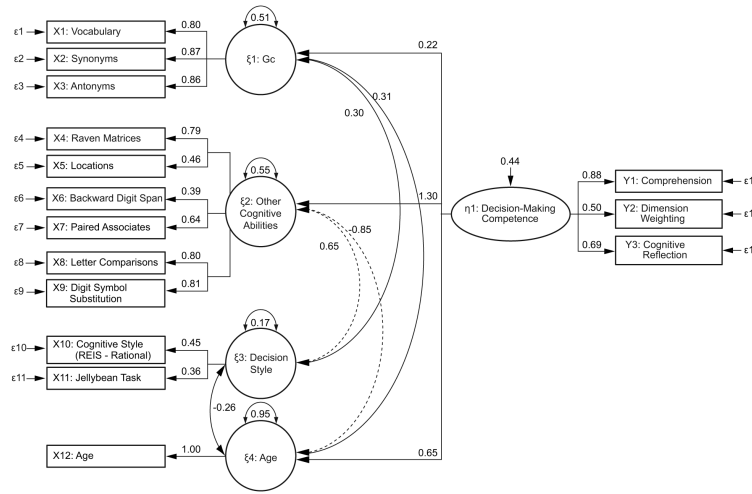


Figure 5. Structural equation model of relationships between basic cognitive abilities, decision style, age, and decision-making competence.

Table 1

Fractional design for administration of cognitive ability marker tests (1=test administered, 0=test missing).

| Group | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Total number of groups in which test administered |
|---|---------------------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|--|
| Measures | Maximum Time (sec) | | | | | | | | | | |
| Crystallized Intelligence | | | | | | | | | | | |
| Synonyms | 150 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 5 |
| Antonyms | 150 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 5 |
| Vocabulary (Part 1, Part 2) | 240 ea | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 5 |
| Fluid Intelligence | | | | | | | | | | | |
| Locations | 360 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 6 |
| Ravens Matrices | 600 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 4 |
| Memory | | | | | | | | | | | |
| Paired Associates (List 1, List 2) | 300 ea | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 5 |
| Backward Digit Span | 600 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 5 |
| Perceptual Speed | | | | | | | | | | | |
| Digit Symbol Substitution | 120 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 6 |
| Letter Comparison (Part 1, Part 2) | 30 ea | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 8 |
| Decision Style | | | | | | | | | | | |
| REI-S Rational, REI-S Experiential | 180 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 5 |
| Risk-Benefit (Part 1, Part 2) | 300 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 7 |
| <i>Total number of tests in each group:</i> | | | | | | | | | | | |
| | 7 | 7 | 5 | 5 | 5 | 5 | 7 | 6 | 7 | 7 | |
| Total Time (mins): | 31 | 27 | 28 | 33 | 34 | 30 | 29 | 34 | 30 | 35 | |

Table 2

Number of participants administered each cognitive ability marker test

| Test | Group | | | | | | | | | | N |
|---------------------------------|-------|----|----|----|----|----|----|----|----|----|-----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | |
| Synonyms | 27 | 32 | 0 | 0 | 0 | 0 | 24 | 0 | 25 | 27 | 135 |
| Antonyms | 27 | 32 | 0 | 0 | 0 | 0 | 24 | 0 | 25 | 27 | 135 |
| Vocabulary – Parts 1 & 2 | 27 | 32 | 29 | 0 | 27 | 0 | 0 | 27 | 0 | 0 | 142 |
| Locations | 0 | 32 | 27 | 24 | 0 | 21 | 24 | 0 | 23 | 0 | 151 |
| Ravens Matrices | 0 | 0 | 28 | 28 | 0 | 0 | 0 | 26 | 0 | 26 | 108 |
| Paired Associates – Lists 1 & 2 | 0 | 0 | 0 | 0 | 27 | 25 | 25 | 27 | 0 | 25 | 129 |
| Backward Digit Span | 27 | 0 | 0 | 28 | 27 | 27 | 0 | 0 | 25 | 0 | 134 |
| Digit Symbol Substitution | 27 | 31 | 0 | 27 | 0 | 0 | 24 | 27 | 0 | 23 | 159 |
| Letter Comparison – Parts 1 & 2 | 27 | 32 | 29 | 0 | 27 | 27 | 25 | 27 | 25 | 0 | 219 |
| REI-S Rational/Experiential | 0 | 0 | 29 | 0 | 0 | 27 | 0 | 27 | 25 | 27 | 135 |
| Risk-Benefit – Parts 1 & 2 | 26 | 32 | 0 | 28 | 27 | 0 | 25 | 0 | 24 | 27 | 189 |

Table 3

Participant characteristics.

| Measure | Overall (N=608) | Younger (N=205) | Young-older (N=208) | Old-older (N=195) | Significance Test |
|--|-----------------|-----------------|---------------------|-------------------|------------------------------|
| Age in years, mean (SD) | 61.4 (19.9) | 34.9 (6.0) | 69.7 (2.9) | 80.3 (4.7) | |
| Age Range (yrs) | 25-97 | 25-45 | 65-74 | 75-97 | |
| Sex (% female) | 62.5 | 66.8 | 58.2 | 62.6 | $\chi^2 = 0.8, p = .3665$ |
| Annual household income (% < \$50,000) | 47.8 | 47.3 | 47.4 | 48.8 | $\chi^2 = 0.11, p = .7532$ |
| Education (% high school or more) | 83.1 | 86.3 | 83.2 | 79.5 | $\chi^2 = 3.3, p = .0681$ |
| Hispanic (% yes) | 5.6 | 13.7 | 1.9 | 1.0 | $\chi^2 = 30.7, p < .0001$ |
| Race (%) | | | | | $\chi^2 = 53.5, p < .0001$ |
| Caucasian | 45.8 | 28.2 | 54.9 | 54.7 | |
| Asian | 31.8 | 32.1 | 29.2 | 34.4 | |
| Native Hawaiian/Pacific Islander | 15.4 | 27.4 | 11.5 | 6.8 | |
| Other | 7.0 | 12.4 | 4.3 | 4.1 | |
| MMSE, mean (SD) | 21.3 (1.0) | 21.6 (0.8) | 21.3 (0.9) | 21.0 (1.1) | $F(2,605) = 16.2, p < .0001$ |

Table 4

Cronbach's coefficient alpha for comprehension, consistency, dimension-weighting, and cognitive-reflection indices for each age group.

| | Index Score Range | Cronbach's α | | | |
|----------------------|-------------------|---------------------|---------|-------------|-----------|
| | | Overall | Younger | Young-Older | Old-Older |
| Comprehension | 0-19 | 0.79 | 0.74 | 0.78 | 0.75 |
| Consistency | 0-9 | 0.41 | 0.41 | 0.32 | 0.44 |
| Dimension Weighting | 0-12 | 0.62 | 0.54 | 0.66 | 0.60 |
| Cognitive Reflection | 0-6 | 0.80 | 0.82 | 0.80 | 0.77 |

Table 5

Principal components solution, with loadings tagged (*) to indicate substantial loading, and portion of variance accounted for in each item (communalities or h^2)

| Item | Factor | | | h^2 |
|---------|--|--|--------------|-------|
| | A (Information Combination/ Transformation) | B (Information Salience/ Relevance) | C (Other) | |
| CR-3 | 0.73 * | -0.01 | 0.33 * | 0.64 |
| CR-6 | 0.72 * | -0.02 | 0.29 | 0.61 |
| COM-12b | 0.69 * | 0.05 | 0.24 | 0.53 |
| COM-3b | 0.65 * | 0.10 | 0.29 | 0.51 |
| CR-5 | 0.63 * | 0.13 | 0.09 | 0.42 |
| COM-13a | 0.54 * | 0.15 | 0.18 | 0.35 |
| CR-1 | 0.52 * | 0.05 | 0.27 | 0.35 |
| COM-13b | 0.51 * | 0.15 | -0.07 | 0.29 |
| COM-7 | 0.47 * | 0.05 | -0.13 | 0.24 |
| COM-14 | 0.44 * | 0.11 | -0.12 | 0.21 |
| COM-5b | 0.42 * | 0.30 | 0.09 | 0.28 |
| COM-10 | 0.41 * | 0.21 | -0.04 | 0.21 |
| COM-4 | 0.41 * | 0.28 | 0.02 | 0.24 |
| COM-2 | 0.28 | 0.23 | 0.01 | 0.13 |
| COM-1 | 0.26 | 0.25 | -0.04 | 0.13 |
| COM-11 | 0.25 | 0.13 | -0.02 | 0.08 |
| COM-8b | 0.24 | 0.23 | 0.06 | 0.11 |
| DW-2 | 0.23 | 0.48 * | -0.02 | 0.28 |
| DW-5 | 0.07 | 0.47 * | 0.14 | 0.25 |
| DW-1 | 0.13 | 0.44 * | -0.02 | 0.21 |
| COM-6 | 0.31 * | 0.41 * | 0.10 | 0.28 |
| DW-9 | 0.05 | 0.40 * | 0.16 | 0.19 |
| DW-7 | 0.13 | 0.40 * | -0.06 | 0.18 |
| COM-9 | 0.34 * | 0.40 * | -0.02 | 0.27 |
| DW-4 | 0.08 | 0.39 * | -0.02 | 0.16 |

| Item | Factor | | | h ² |
|---------|--|--|--------------|----------------|
| | A (Information Combination/ Transformation) | B (Information Salience/ Relevance) | C (Other) | |
| DW-10 | -0.01 | 0.38 * | 0.13 | 0.16 |
| DW-11 | 0.00 | 0.37 * | 0.28 | 0.21 |
| COM-8a | 0.31 * | 0.37 * | -0.17 | 0.26 |
| DW-3 | -0.01 | 0.35 * | 0.12 | 0.14 |
| COM-3a | 0.18 | 0.34 * | 0.05 | 0.15 |
| DW-6 | -0.04 | 0.34 * | 0.29 | 0.20 |
| DW-12 | 0.03 | 0.33 * | 0.11 | 0.12 |
| DW-8 | 0.15 | 0.32 * | -0.07 | 0.13 |
| COM-5a | 0.12 | 0.31 * | -0.05 | 0.11 |
| COM-12a | 0.11 | 0.31 * | -0.15 | 0.13 |
| CR-4 | 0.23 | 0.09 | 0.80 * | 0.69 |
| CR-2 | 0.30 | 0.06 | 0.79 * | 0.71 |

Table 6

Mean score (and standard deviation) for the overall comprehension, dimension-weighting, and cognitive-reflection indices for each age group.

| | Younger | Young-Older | Old-Older |
|---|----------------|--------------------|------------------|
| Comprehension (maximum score = 19) | 14.0 (3.2) | 12.4 (3.8) | 10.9 (3.7) |
| Dimension Weighting (maximum score = 12) | 9.4 (2.0) | 8.9 (2.4) | 8.4 (2.4) |
| Cognitive Reflection (maximum score = 6) | 2.5 (2.1) | 2.1 (2.0) | 1.7 (1.8) |

Table 7

Means (standard deviations) on individual-difference measures and tests of significant differences among age groups.

| | Overall | Younger | Young-Older | Old-Older | Significance Test [†] |
|---------------------------------------|-------------|-------------|-------------|-------------|---|
| Self-reported Health | | | | | |
| Physical Health (% excellent or good) | 77.2 | 67.3 | 80.8 | 83.8 | $\chi^2 = 15.5$, $p < .0001$ |
| Mental health (% excellent or good) | 91.3 | 89.3 | 91.7 | 92.9 | $\chi^2 = 1.7$, $p = .1926$ |
| Number of medications | 1.7 (1.0) | 1.2 (0.6) | 2.0 (1.1) | 2.0 (1.1) | $F(2, 278302) = 50.2$, $p < .0001$ |
| Doctor/nurse visits in previous month | 0.7 (0.9) | 0.5 (0.8) | 0.8 (0.9) | 0.9 (1.0) | $F(2, 138166) = 9.6$, $p < .0001$ |
| Crystallized Intelligence | | | | | |
| Vocabulary (btr [‡]) | 13.5 (4.0) | 11.8 (4.1) | 14.1 (3.8) | 14.7 (3.5) | $F(2, 849.2) = 29.0$, $p < .0001$ |
| Synonyms (btr) | 6.3 (2.9) | 5.2 (2.8) | 6.7 (3.0) | 7.1 (2.7) | $F(2, 360.9) = 21.1$, $p < .0001$ |
| Antonyms (i [*]) | 5.8 (2.9) | 5.0 (2.8) | 6.2 (3.0) | 6.4 (2.7) | $F(2, 1385) = 14.4$, $p < .0001$ |
| Fluid Intelligence | | | | | |
| Ravens Matrices (i) | 6.4 (3.2) | 8.8 (2.7) | 5.9 (2.8) | 4.3 (2.3) | $F(2, 1752.0) = 144.5$, $p < .0001$ |
| Locations (i) | 4.3 (2.7) | 5.6 (2.9) | 3.9 (2.5) | 3.3 (2.0) | $F(2, 210.4) = 38.7$, $p < .0001$ |
| Memory | | | | | |
| Backward Digit Span (btr) | 6.8 (2.6) | 7.6 (3.0) | 6.4 (2.3) | 6.4 (2.4) | $F(2, 490.0) = 12.4$, $p < .0001$ |
| Paired Associates (btr) | 2.1 (1.9) | 3.3 (1.9) | 1.5 (1.5) | 1.4 (1.4) | $F(2, 1213.9) = 84.0$, $p < .0001$ |
| Perceptual Speed | | | | | |
| Letter Comparison (btr) | 9.2 (3.0) | 11.7 (2.9) | 8.3 (2.0) | 7.6 (2.3) | $F(2, 8493.2) = 159.9$, $p < .0001$ |
| Digit Symbol Substitution (btr) | 64.5 (21.3) | 82.4 (20.6) | 60.1 (14.5) | 50.3 (13.9) | $F(2, 168.6) = 153.5$, $p < .0001$ |
| Attitudes and Self-perceptions | | | | | |
| Experience | 2.9 (0.7) | 2.9 (0.7) | 2.9 (0.7) | 2.9 (0.6) | $F(2, 605) = 0.2$, $p = .8071$ |

| | Overall | Younger | Young-Older | Old-Older | Significance Test [†] |
|--------------------------|-----------------|-----------------|-----------------|-----------------|------------------------------------|
| Motivation | 3.4 (0.6) | 3.4 (0.6) | 3.4 (0.5) | 3.4 (0.6) | F(2,605) = 0.1, p = .8945 |
| Skill with tables | 2.9 (0.7) | 3.1 (0.7) | 2.8 (0.7) | 2.7 (0.6) | F(2,145061) = 18.9, p < .0001 |
| Decision support | 1.8 (0.7) | 2.0 (0.7) | 1.7 (0.6) | 1.8 (0.7) | F(2,19E6) = 8.6, p < .0002 |
| Self-reported memory | 3.7 (0.9) | 3.6 (0.9) | 3.7 (0.9) | 3.9 (1.0) | F(2,145024) = 5.3, p < .0049 |
| Driving almost daily (%) | 63.4 | 69.8 | 64.4 | 55.5 | $\chi^2 = 8.7$, p < .0033 |
| Decision self-efficacy | 5.6 (0.9) | 5.6 (0.8) | 5.6 (0.9) | 5.4 (0.9) | F(2,605) = 2.8, p = .0646 |
| Decision Style | | | | | |
| Risk-benefit corr. (btr) | -0.43 (0.37) | -0.47 (0.32) | -0.43 (0.38) | -0.39 (0.40) | F(2, 132.6) = 1.8, p = .1755 |
| REI-S, Rational (i) | 42.5 (7.9) | 43.8 (8.0) | 42.3 (7.6) | 41.5 (8.2) | F(2, 958.2) = 4.0, p = .0189 |
| REI-S, Experiential (i) | 40.3 (8.6) | 41.4 (8.1) | 40.6 (8.2) | 38.8 (9.4) | F(2, 4628.2) = 4.8, p = .0085 |
| Jellybeans (i) | 0.4 (2.1) | 0.7 (2.1) | 0.5 (2.1) | 0.1 (2.1) | F(2, 153976.3) = 3.5, p = .0290 |
| Numeracy | 1.8 (1.0) | 2.0 (0.9) | 1.7 (0.9) | 1.5 (0.9) | F(2, 605) = 16.3, p < .0001 |

[†],"btr" indicates that the variable was transformed to normality for imputation then back-transformed to the original units

[‡],"I" indicates that the variable was not transformed, but does have imputed values

* degrees of freedom are adjusted according to Rubin's rules, as part of combining imputations, the less efficient the imputed values the more the df will be inflated.

Table 8

Correlations of decision-making competence index scores with demographic variables and individual difference measures.

| | Comprehension Index | | Dimension-weighting Index | | Cognitive-reflection Index | |
|-------------------------------------|---------------------|--------|---------------------------|--------|----------------------------|--------|
| | r | p | r | p | r | p |
| Age | -0.33 | <.0001 | -0.17 | <.0001 | -0.18 | <.0001 |
| Gender (female=0, male=1) | 0.08 | 0.0633 | 0.11 | 0.0083 | 0.20 | <.0001 |
| Income | 0.20 | <.0001 | 0.14 | 0.0007 | 0.21 | <.0001 |
| Education | 0.33 | <.0001 | 0.19 | <.0001 | 0.33 | <.0001 |
| Physical health (btr ²) | -0.10 | 0.0156 | -0.09 | 0.0339 | -0.11 | 0.0058 |
| Emotional health (btr) | -0.15 | 0.0002 | -0.10 | 0.0110 | -0.11 | 0.0080 |
| Number of medications (btr) | -0.22 | <.0001 | -0.15 | 0.0003 | -0.12 | 0.0036 |
| Doctor/nurse visits (btr) | -0.09 | 0.0353 | -0.07 | 0.0805 | -0.05 | 0.1885 |
| Vocabulary (btr) | 0.27 | <.0001 | 0.19 | 0.0001 | 0.27 | <.0001 |
| Synonyms (btr) | 0.34 | <.0001 | 0.12 | 0.0030 | 0.35 | <.0001 |
| Antonyms (tr ²) | 0.37 | <.0001 | 0.15 | 0.0013 | 0.35 | <.0001 |
| Ravens Matrices (i) | 0.50 | <.0001 | 0.24 | <.0001 | 0.44 | <.0001 |
| Locations (i) | 0.26 | <.0001 | 0.09 | 0.0370 | 0.24 | <.0001 |
| Backward Digit Span (btr) | 0.30 | <.0001 | 0.22 | <.0001 | 0.26 | <.0001 |
| Paired Associates (btr) | 0.40 | <.0001 | 0.15 | 0.0005 | 0.32 | <.0001 |
| Letter Comparison (btr) | 0.36 | <.0001 | 0.22 | <.0001 | 0.23 | <.0001 |
| Digit Symbol Substitution (btr) | 0.39 | <.0001 | 0.19 | <.0001 | 0.24 | <.0001 |
| Experience (btr) | 0.05 | 0.2004 | 0.10 | 0.0127 | 0.06 | 0.1191 |
| Motivation (btr) | 0.11 | 0.0055 | 0.11 | 0.0076 | 0.08 | 0.0559 |
| Skill with Tables (btr) | 0.33 | <.0001 | 0.22 | <.0001 | 0.29 | <.0001 |
| Decision Support (btr) | -0.01 | 0.7737 | -0.15 | 0.0003 | -0.01 | 0.8976 |
| Self-reported Memory (btr) | -0.05 | 0.2411 | 0.00 | 0.9988 | 0.00 | 0.9450 |
| Driving (btr) | 0.19 | <.0001 | 0.16 | 0.0001 | 0.11 | 0.0054 |
| Decision self-efficacy (btr) | 0.00 | 0.9381 | 0.05 | 0.1899 | -0.01 | 0.8308 |
| Risk-benefit correlation (btr) | -0.06 | 0.2234 | -0.06 | 0.2144 | -0.04 | 0.4324 |
| REI-S, Rational (i) | 0.30 | <.0001 | 0.23 | <.0001 | 0.16 | <.0001 |

| | Comprehension Index | | Dimension-weighting Index | | Cognitive-reflection Index | |
|------------------------|---------------------|--------|---------------------------|--------|----------------------------|--------|
| | r | p | r | p | r | p |
| REI-S Experimental (i) | 0.13 | 0.0041 | 0.11 | 0.0110 | 0.02 | 0.6232 |
| Jellybeans (i) | 0.25 | <.0001 | 0.16 | 0.0001 | 0.21 | <.0001 |
| Numeracy (i) | 0.50 | <.0001 | 0.32 | <.0001 | 0.53 | <.0001 |

[†]“bit” indicates that the variable was transformed to normality for imputation then back-transformed to the original units

[‡]“i” indicates that the variable was not transformed, but does have imputed values

Table 9

Hierarchical regressions with R^2 and increment in R^2 from a series of regression models with the comprehension, dimension-weighting, and cognitive-reflection indices as the dependent variables.

| Independent Variables | Comprehension | | | | Dimension Weighting | | | | Cognitive Reflection | | | |
|---|-----------------|-------------|------------|---------------------------------|---------------------|-------------|------------|---------------------------------|----------------------|-------------|------------|--------------------------------|
| | Change in R^2 | Total R^2 | Adj. R^2 | F test of ΔR^2 | Change in R^2 | Total R^2 | Adj. R^2 | F test of ΔR^2 | Change in R^2 | Total R^2 | Adj. R^2 | F test of ΔR^2 |
| Analysis 1 | | | | | | | | | | | | |
| Step 1: Age by itself | -- | 0.106 | 0.105 | F(1,606) = 71.9, $p < .0001$ | -- | 0.030 | 0.029 | F(1,606) = 18.8, $p < .0001$ | -- | 0.033 | 0.032 | F(1,606) = 20.8, $p < .0001$ |
| Analysis 2 | | | | | | | | | | | | |
| Step 1: Social variables ¹ | -- | 0.154 | 0.146 | F(7,600) = 1142.6, $p < .0001$ | -- | 0.050 | 0.041 | F(7,600) = 1441.9, $p < .0001$ | -- | 0.162 | 0.153 | F(7,600) = 128.0, $p < .0001$ |
| Step 2: Health measures ² | 0.057 | 0.211 | 0.198 | F(4, 596) = 10.8, $p < .0001$ | 0.024 | 0.074 | 0.058 | F(4, 596) = 3.9, $p < .0001$ | 0.016 | 0.178 | 0.164 | F(4, 596) = 2.9, $p = .0214$ |
| Step 3: Basic cognitive skills measures ³ | 0.294 | 0.505 | 0.489 | F(9, 587) = 38.7, $p < .0001$ | 0.083 | 0.157 | 0.130 | F(9, 587) = 6.4, $p < .0001$ | 0.243 | 0.421 | 0.402 | F(9, 587) = 27.4, $p < .0001$ |
| Step 4: Attitudinal and self-perception measures ⁴ | 0.014 | 0.519 | 0.497 | F(7, 580) = 2.4, $p < .0001$ | 0.036 | 0.193 | 0.156 | F(7, 580) = 3.7, $p < .0001$ | 0.005 | 0.426 | 0.400 | F(7, 580) = 0.7, $p = .6536$ |
| Step 5: Decision style measures ⁵ | 0.020 | 0.539 | 0.515 | F(4, 576) = 6.2, $p < .0001$ | 0.027 | 0.220 | 0.179 | F(4, 576) = 5.0, $p < .0001$ | 0.008 | 0.434 | 0.404 | F(4, 576) = 2.0, $p = .0880$ |
| Step 6: Adding age | <0.001 | 0.539 | 0.514 | F(1, 575) = 0, $p = 1.0000$ | 0.003 | 0.223 | 0.181 | F(1, 575) = 2.2, $p = .0002$ | 0.003 | 0.437 | 0.406 | F(1, 575) = 3.1, $p = .0806$ |
| Full model | -- | 0.539 | 0.514 | F(32, 575) = 403.8, $p < .0001$ | -- | 0.223 | 0.181 | F(32, 575) = 338.0, $p < .0001$ | -- | 0.437 | 0.406 | F(32, 575) = 44.1, $p < .0001$ |

¹ Exogenous social variables include gender, income, education, and race

² Health measures include physical health, emotional health, number of medications, and number of doctor/nurse visits

³ Basic cognitive skills measures include Vocabulary, Synonyms, Antonyms, Ravens Matrices, Locations, Backward Digit Span, Paired Associates, Letter Comparison, and Digit Symbol Substitution

⁴ Attitudinal and self-perception measures include experience, motivation, skill with tables, decision support, self-rated memory, driving, and decision self-efficacy

⁵ Decision-style measures include risk-benefit correlation, REI-S Rational, REI-S Experiential, and jellybeans

Table 10

Hierarchical regressions examining contribution of numeracy versus cognitive skills to variance in index scores (participants with imputed data, constant N = 608)

| Independent Variables | Comprehension | | | Dimension Weighting | | | Cognitive Reflection | | | | | |
|---|--------------------------|----------------------|---------------------|--------------------------------|--------------------------|----------------------|----------------------|-------------------------------|--------------------------|----------------------|---------------------|-------------------------------|
| | Change in R ² | Total R ² | Adj. R ² | F test of ΔR ² | Change in R ² | Total R ² | Adj. R ² | F test of ΔR ² | Change in R ² | Total R ² | Adj. R ² | F test of ΔR ² |
| Analysis 1 | | | | | | | | | | | | |
| Step 1: Basic cognitive skills measures ¹ | 0.467 | 0.459 | 0.118 | F(10, 597) = 1083.2, p < .0001 | 0.132 | 0.118 | 0.338 | F(10, 597) = 114.1, p < .0001 | 0.348 | 0.338 | 0.410 | F(10, 597) = 114.1, p < .0001 |
| Step 2: Add Numeracy ² to cognitive skills | 0.039 | 0.506 | 0.497 | F(1, 596) = 47.0, p < .0001 | 0.045 | 0.177 | 0.163 | F(1, 596) = 32.6, p < .0001 | 0.072 | 0.420 | 0.410 | F(1, 596) = 74.0, p < .0001 |
| Analysis 2 | | | | | | | | | | | | |
| Step 1: Numeracy by itself | 0.253 | 0.252 | 0.103 | F(1, 606) = 205.5, p < .0001 | 0.104 | 0.103 | 0.279 | F(1, 606) = 70.6, p < .0001 | 0.280 | 0.279 | 0.410 | F(1, 606) = 235.5, p < .0001 |
| Step 2: Add basic cognitive skills | 0.253 | 0.506 | 0.497 | F(10, 596) = 30.5, p < .0001 | 0.073 | 0.177 | 0.163 | F(10, 596) = 5.29, p < .0001 | 0.140 | 0.420 | 0.410 | F(10, 596) = 14.4, p < .0001 |
| Full model | 0.506 | 0.497 | 0.163 | F(11, 596) = 1082.8, p < .0001 | 0.177 | 0.163 | 0.410 | F(11, 596) = 939.1, p < .0001 | 0.420 | 0.410 | 0.410 | F(11, 596) = 123.6, p < .0001 |

¹Measures of basic cognitive skills include Vocabulary, Synonyms, Antonyms, Raven's Matrices, Locations, Backward Digit Span, Paired Associates, Letter Comparison, and Digit Symbol Substitution.

²Numeracy scores are based on the 3-item measure by Schwartz et al. (1997).