Aging and Decision-making Competence: An Analysis of Comprehension and Consistency Skills in Older versus Younger Adults Considering Health-plan Options

MELISSA L. FINUCANE,1,* PAUL SLOVIC,2 JUDITH H. HIBBARD,3 ELLEN PETERS,2 C. K. MERTZ2 and DONALD G. MACGREGOR2

1Center for Health Research Hawaii, Honolulu, HI, USA
2Decision Research, Eugene, OR, USA
3University of Oregon, USA

ABSTRACT

Older adults need to maintain strong decision-making capabilities as they age. However, we know little about how age-related physical and psychological changes affect older adults’ judgment and decision processes. This paper reports the results of research comparing older versus younger adults’ performance on evaluation and choice tasks about health-plan options. In particular, comprehension and consistency in judgments (across separate versus joint evaluation contexts) were examined. Results indicated that increasing age was related to greater comprehension errors and inconsistent preferences, even when covariates (education, income, gender, self-perceived skill and health, decision style, and attitude toward delegation) were taken into account. Discussion of the results emphasizes difficulties in interpreting the meaning of age differences in performance on decision tasks and the need for research that ascertains the seriousness of the consequences of age differences in real-life tasks. The implications for providing decision-aiding interventions for older adults are highlighted. Copyright © 2002 John Wiley & Sons, Ltd.

KEY WORDS judgment; decision-making competence; aging; elderly; health; comprehension; consistency

Older adults need to maintain strong decision-making capabilities as they age. People are living longer but are not always able to rely on younger family members for support because of the increasing geographical dispersion of families and the modern emphasis on independence. Fortunately, recent improvements in the methods available for cognitive research offer great promise for learning more about the aging mind and using that knowledge to provide decision support, where necessary, to improve the lives of older people.
(National Research Council, 2000). In this paper we present the results of our first attempts to apply methods from the burgeoning field of judgment and decision research to the study of older adults’ evaluation and choice processes. In particular, we compare younger and older adults’ competence at health decisions. A unique feature of this work is that it brings together the research domains of decision making, aging, and competence.

DECISION-MAKING COMPETENCE: IMPORTANT BUT NEGLECTED

As people get older, their everyday decision-making abilities seem to be under increasing scrutiny (Christensen et al., 1995). Historically, most attention has been focused on declines or losses in abilities with age, but now the positive changes that come with age are also attracting attention (Baltes, 1993). Understanding both the limits and potentials of older adults’ decision abilities (and how they interact) is critical in modeling and assisting successful aging. The development or decline of particular abilities can directly impact an individual’s physical and psychological well-being. Identifying the full range of abilities that contribute to judgments and decisions by older adults is likely to be the most effective way to understand how people facing change can optimize favorable circumstances and avoid self-defeating behaviors.

Despite the importance of examining decision-making competence (DMC) in older people, little research has focused on how this should be done. In fact, a search of the psychological literature shows that researchers have neglected the field of older adults’ decision making. The PsychInfo Database (1887–November 2000) yields only 733 publications when searches on the topics decision-making/judgment (49,475 entries and aging/elderly (35,056 entries) are combined. Of the 733 publications, only 47 (6.4%) are related specifically to competence. Among these, DMC in older adults is typically assessed via interviews; the need for standardized tools is highlighted (e.g. Bean et al., 1994; Searight and Hubbard, 1998). The dearth of research on older adults’ decision processes persists, despite the wealth of knowledge gathered by judgment and decision-making researchers in recent decades that could be used to develop performance-based measures of older adults’ DMC.

EVERYDAY PROBLEM SOLVING AND DECISION MAKING

Unlike the field of decision research, everyday problem solving (EPS) research has given substantial attention to how adults of different ages solve practical problems that arise in everyday living (Berg et al., 1998 Denney, 1989; Willis, 1996; Willis and Schae, 1993). Problem solving involves assessing the present and desired state of affairs and finding ways to move from the former to the latter, whereas decision making involves evaluating the possible solutions and selecting one to use (Reese and Rodeheaver, 1985; Yates 1990). Recent empirical studies of EPS competence have underscored the complex array of characteristic of the individual (e.g. age, education, spatial-verbal skills, experience), task (e.g. information complexity) and context (e.g. cultural values) that may influence how problems are represented and resolved. Complex theoretical frameworks have been called on to organize and explain the role of underlying processes as well as the physical and social contexts in which individuals function. The person–environment fit is considered very important to understanding EPS competence (Schae and Willis, 1999).

Two major theoretical views of aging and EPS dominate the literature. One view takes an ‘expertise development’ approach, arguing that the demands of daily life require individuals to draw on their accumulated knowledge systems that become increasingly selective and domain-specific with age (Baltes and Baltes 1990; Salthouse, 1991) and on components of performance that become increasingly automatized with experience (Denney, 1989). The second theoretical view approaches EPS as a complex, ‘compiled’ form of cognition (Marsiske and Willis, 1995; Park, 1992; Salthouse, 1990; Willis and Schae, 1986) with simila
characteristics to age-related changes found in measures of basic abilities. As yet there is little consistent support for either theoretical perspective from empirical work on older adults’ EPS competence, possibly because patterns of age differences are dependent on an interaction of individual, task, and context factors and on the method of measurement (Willis, 1991). Moreover, while declines in basic abilities like speed of processing are well established, it is unclear whether this necessarily leads to a decline in EPS. Older adults may compensate for loss of speed by using other cognitive strategies, leading to different, but not necessarily worse ways of solving problems. That is, increased ‘wisdom’ may make up for decreased quickness of mind (Baltes, 1993).

One common finding in studies of aging and EPS is that older adults tend to seek less information to solve problems than do younger adults (e.g. Johnson, 1990). Older adults also tend to be more interpretative, focusing more on interpersonal and experiential components of a problem rather than the propositional content (Blanchard-Fields, 1996; Labouvie-Vief, 1982; Sinnott, 1989). One explanation for these differences could be that as people age they depend less on analytic processing and more on heuristic processing when solving problems or making decisions (Park, 1999; Yates and Patalano, 1999) in order to conserve their more limited emotional and physical resources (Berg, Meegan, and Klaczynski, 1999). Analytic processing such as comparing the strengths and weaknesses of alternative solutions places heavy demands on working memory (the capacity for processing, storing, and retrieving information online). Many studies have demonstrated age-related declines in working memory (e.g. Salthouse and Babcock, 1991) and how this may interact with declines in processing speed to reduce performance on complex tasks (e.g. Salthouse, 1992, 1996).

The notion that there are stylistic differences between individuals in the exhaustive versus more heuristic nature of problem solving is consistent with the growing literature on simplification strategies in decision making (Baron, 1994; Finucane et al., 2000; Kahneman, Slovic, and Tversky, 1982; Payne, Bettman, and Johnson, 1993). Heuristic strategies are used (consciously or unconsciously) to simplify information processing in judgments and decisions for various reasons, most typically because an individual has insufficient cognitive capacity or motivation to thoroughly deliberate and analyze all available information in a problem. Heuristics provide shortcuts through complex problems, and in many contexts are adaptive, especially when the increased information processing efficiency comes with minimal cost to accuracy. Individuals may use a variety of different heuristics, potentially leading to different decisions, but not necessarily to worse decisions.

Despite the EPS research findings paralleling some trends in research on human judgment and decision making, the EPS work has not examined the specific abilities individuals rely on to evaluate and choose among options when solving problems. In short, the EPS literature is very informative in terms of the array of factors that play a role in solving everyday problems, but it is not definitive regarding judgment and decision processes.

DEFINING DECISION-MAKING COMPETENCE

There are several views on what DMC means. Most commonly, the term refers to legal or clinical determinations of a person’s ability to make important life decisions (Searight and Hubbard, 1998). This view emphasizes measuring individual differences so as to provide diagnostic criteria for determining where an individual falls on a competence continuum. Another view—and the main (but not sole) focus of this paper—is that competence is a function of aging. This emphasizes measuring group differences and examining whether older people in general are better or worse at decision making than younger people are (e.g. Stanley et al., 1984).

We view DMC as a multidimensional concept. Several abilities are required for good decision making. For the purpose of defining and measuring competence, five commonly cited criteria from research on adolescents and younger adults include the ability to (1) understand and remember relevant information,
(2) structure a decision’s dimensions and alternatives, (3) appreciate the personal significance of the information, (4) temper impulsivity, and (5) rationally integrate the information and reason about it\(^1\) (Appelbaum and Grisso, 1988; Parker and Fischhoff, 1999; Rosenfeld and Turkheimer, 1995; see also Sanfey and Hastie, 2000; and Yates and Patalano, 1999, for alternative classifications).

As a first step toward addressing the need for research on how to reliably and validly measure DMC in the elderly, we focused only on two aspects of DMC: comprehension (the ability to understand information) and consistency (the ability to integrate information in an internally consistent fashion). Ultimately, all aspects of DMC should be researched thoroughly, but focusing first on comprehension and consistency skills is important because of their fundamental roles in sound decision making.

**Comprehension**

Good comprehension in a decision-making context means that an individual understands the available decision options and how each option rates on one or more dimensions (Radvansky, 1999). Comprehension is essential to competent decision making because information must be understood before it can be used sensibly in a judgment or decision.

Comprehension is a complex construct, involving diverse components. Successful comprehension involves many levels of processing and representation. The conceptualization of comprehension in prior research has depended largely on the level of analysis of interest to individual researchers. Some dependent measures of comprehension have focused on individuals’ ability to recall or recognize the surface form, main (abstracted) ideas, or inferences in a text (for reviews see Lidz et al., 1984; Meyer and Rice, 1989; Zwaan and Radvansky, 1998). Other measures have included tests of vocabulary meaning or recognition (e.g. Meyer, Marsiske, and Willis, 1993) and reading time (e.g. Meyer, Talbot, and Florencio, 1999; Stine-Morrow, Loveless, and Soederberg, 1996).

A special case of reading comprehension is document literacy, which involves reading skills necessary to understand and use non-prose printed materials, such as charts, schedules, tables, labels, and forms, as is found in the instructions on medicine bottle labels and other products, the schedules outlined in bus and other timetables, and the contractual obligations of financial documents (Kirsch and Mosenthal, 1990). The readability of everyday documents is affected by text processing variables (see Meyer et al., 1993), as well as the specific abilities involved in inductive and arithmetic reasoning such as problem identification, decomposition, step sequencing, and information manipulation (Allaire and Marsiske, 1999; Tirre and Pena, 1993). Comprehension will suffer as the document’s complexity increases and outstrips the reader’s reasoning skills and, of course, his or her working memory capacity (Kirsch and Mosenthal, 1990). Different interpretations of information may lead to different decisions. Communications will be most effectively and efficiently used as intended if they can be tailored in ways that account for differences across populations in how document information is understood.

The few studies of aging and document comprehension have focused mostly on medical contexts. For instance, Morrell, Park, and Poon (1989) found few comprehension errors by younger and older adults for well-organized prescription labels developed by the experimenters, but a substantial increase in comprehension errors for real-world labels (to 14% for younger adults and 21% for older adults) (see also Morrow and Leerer, 1999). Similarly, Willis and colleagues assessed older adults’ comprehension of prescription labels and medication charts (Park et al., 1994), asking subjects two types of questions. First, literal (factual) questions (e.g. ‘What are the side effects of this medication?’) required respondents to find and state information located in a single place on the label. Second, inferential questions (e.g. ‘For how many days will this supply of medicine last, if taken according to directions?’) required searching for information in several places.

---

\(^1\)The ability to express a choice is also a common definition of decision competence, but is mainly of interest in severely impaired individuals.
(e.g. determining how many pills to take each day and the number of pills in the bottle) and then performing a mathematical operation on the information. Results showed that fewer errors were made by normal older adults (age range not reported) in response to a literal question (3%) compared with an inferential question (16%). For a slightly more complex stimulus (a 4 x 4 medication chart), comprehension errors increased to 13% and 23% for literal and inferential questions, respectively. Thus, increasing the complexity of the stimuli and the question can lead to decreased comprehension among older adults.

Drawing firm conclusions about aging and document comprehension is difficult due to several limitations in the existing research. First, of the few studies that have been done, some reported data for older participants only. Second, the assessment techniques used have been problematic. Many studies have relied on self-report which typically leads to overestimates of competence (Ford et al., 1988); others have relied on recall which does not isolate short-term memory problems from difficulties understanding the meaning of information. Third, the generality of findings from one experimental task (reading medication labels) to functioning in everyday life is arguable and the meaning of age differences is unclear.

Integrating information: internal decision consistency

The second aspect of DMC examined in this paper is information integration. Many professionals (physicians, lawyers, economists, psychologists) have much to say about how information should be integrated because proper integration is a hallmark of rationality from which good decisions are most likely to result. Rational information integration involves several aspects, one of the most important of which is the ability to weigh dimensions in an internally consistent manner. That is, to make good decisions repeatedly, individuals need to hold constant the relative importance of dimensions in like situations or across situations that differ only superficially. An individual who weighs information in a way that results in a good decision will benefit from repeating that weighting. If an individual integrates the same information in different ways, s/he may end up with different (perhaps sub-optimal) decision outcomes, may be unable to reap the same benefits predictably, and/or may be vulnerable to manipulation from unscrupulous parties (see the demonstrations of how people may become ‘money pumps’ through situational manipulations, shown in Lichtenstein and Slovic, 1971).

Internal consistency has been largely ignored in research on aging. One exception is a study by Malloy et al. (1992) on the influence of treatment descriptions on medical decisions. They presented 201 individuals (aged 65–94 years) with descriptions of a life-sustaining intervention described positively (‘device to help you breathe’), negatively (‘machine that controls your breathing’), or exactly as worded in a widely-used advance directive (‘breathing by machine’). Results showed that individuals were less likely to choose the intervention when it was worded negatively compared with when it was worded positively or phrased as in the directive already in use (12%, 30%, 19% choice respectively for negative, positive, and current wordings). Moreover, most subjects appeared inconsistent from one moment to the next: 77% changed their minds at least once when responding to the different descriptions of the three interventions within the context of three clinical scenarios.

Malloy et al. (1992) conclude that doctor/patient consultation is needed when patients execute advance directives, suggesting that the elderly need help with decisions. However, their study does not permit any conclusion to be made about whether older adults’ decisions are any worse than others’ decisions, or whether they would in fact benefit from decision aiding by a doctor. What is indicated, rather, is that preferences may change with a change in information format.

Historically, the effect of manipulating information format has been known as the ‘framing’ phenomenon. Researchers have typically found that when decision options are described positively (e.g. in terms of gains, survival rates), individuals tend to prefer a certain option or ‘sure thing’, and when options are described negatively (e.g. in terms of losses, death rates), individuals tend to prefer a risky option, despite the positive and negative information being identical in expected value (McNeil et al., 1982; Tversky and
Kahneman, 1981). More recently, framing research has been expanded to investigate the effects of ‘context’ where the deep structure of information is held constant, but superficial features of information are manipulated. For instance, Hsee (1995; 1996a,b; 1998) has observed for many different tasks that providing contextual cues (e.g. specifying the possible value range of an attribute) changes the weights given to information in decisions. Similarly, Satterfield, Slovic and Gregory (2000) have shown that presenting information in a narrative format, rather than a dry, statistical format, heightens the salience of qualitative dimensions in a decision problem.

In line with recent trends, the present paper focuses on consistency in evaluations in the face of contextual changes to decision information. From a normative perspective, one could argue that if individuals really knew how to value and combine information, then evaluations should remain constant over contexts. From a prescriptive perspective, however, one could argue that if an individual gains new insight from the context then it makes sense to change his/her evaluation. Examining age differences in sensitivity to contextual information is important because it highlights the extent to which information needs to be standardized to yield consistent judgments and choices.

One reason age differences in decision consistency may occur relates to the use of different decision strategies across the life span. Some research suggests that compared with younger adults, older adults tend to rely more on simple non-compensatory strategies (Johnson, 1990). Furthermore, evidence from Weber et al. (1995) indicates that some preference reversals may occur as the result of simplified encoding of information due to short-term memory limitations. Strong evidence of age-related declines in working memory and the role of memory load in the use of judgmental heuristics also suggests that older adults may be more easily influenced by the decision context.

In sum, considerable research on younger adults has demonstrated inconsistency in decision making as a result of constructive processing that is sensitive to contextual changes (Slovic, 1995). Studies of older adults’ judgments and decisions are less common, despite the potential for older adults to depend heavily on heuristic processing and to display inconsistent decision making. Whether older adults are less internally consistent in their decision making than younger adults has not been established. And whether less consistency translates into different decisions, or into worse decisions, is unclear.

OVERVIEW OF PRESENT RESEARCH

Current US policy aims to make the health-care options for adults 65 years and older similar to those available to employed-age adults. In order to broaden the health-care delivery choices available to older adults increasing amounts of information about health-care coverage, health-plan designs, and the plans’ quality of care is being given to consumers. Policymakers assume that providing more information will help older adults make choices that fit their individual needs and preferences. However, the literature reviewed above has not established whether older adults’ decision processes are similar to those of employed-age adults. Whether older adults are able to process information effectively and to make good decisions in a new and complex domain such as health-care coverage remains unknown. To address this gap in knowledge the present research compared younger and older adults’ abilities to make judgments and decisions when given document-type information (tables and charts) about health-plan options.

Although previous research on comprehension and consistency in older adults provides little guidance about what to expect, studies of age-related changes in basic cognitive abilities provide strong grounds to expect substantial age differences.

First, we tested comprehension abilities using literal and inferential questions without the need to recall information from short-term memory. By leaving all information in front of participants we were able to study their understanding of available information rather than forgetfulness. We tested whether older adult would show more comprehension errors than younger adults, particularly when the number of health-plan
options increased or the task involved identifying and combining more than one piece of information describing a health plan.

Second, we tested internal decision consistency on a task manipulating information context. The task required judgments of health-plan options side-by-side (joint condition) or alone (separate condition). We tested whether older adults would show more inconsistency than younger adults in terms of greater differences in evaluations of options under joint and separate conditions.

Finally, we measured self-perceptions of health, skill in using tables and charts, decision style, and desire for decision delegation in order to explore the relationships between perceived abilities and behaviorally measured judgment and decision-making abilities. Despite the importance of determining the match between perceived and actual capacities for making sound judgments and decisions, research in this domain is scarce, providing little ground for hypotheses and rendering this component of the research exploratory.

METHOD

Participants

Older sample
The ‘older’ participants were 253 people recruited from community centers. Their mean age was 75 years (range 65–94 years); 61% were female. Household income of 85% of the participants was less than $40,000 and 54% had at least some college education. Each participant in this sample was paid $20 for his or her time; an additional $5 per person was donated to the organization from which individuals were recruited.

Younger sample
The ‘younger’ participants were 239 people recruited through flyers mailed internally to the pool of classified (non-faculty) staff at the University of Oregon. Their mean age was 40 years (range 18–64 years); 77% were female. Household income was less than $40,000 for 62% of the participants; 82% had at least some college education. Each younger adult was paid $15 for his/her time.

Overview of materials and procedure
Each participant completed a questionnaire booklet (average completion time was about one hour for older adults and 45 minutes for younger adults). The questionnaire was printed in larger font for the older adult sample to accommodate age-related vision deficits. The questionnaire contained multiple pen-and-paper tasks designed to assess comprehension, consistency, and self-perceptions; sociodemographic information was also collected.

COMPREHENSION: LITERAL AND INFERENTIAL TASKS

Stimuli used by Park et al. (1999) and Parker and Fischhoff (1999) that asked individuals literal and inferential questions about drug labels were adapted to assess comprehension in the present study.

Stimuli and procedure
Individuals were given three problems. Problem C-1 presented individuals with the information about a Health Maintenance Organization (HMO) shown in Exhibit 1. Two multiple-choice questions were asked. The first was a literal question, requiring an individual to identify one piece of information in the table. The
Exhibit 1. Simple one-option information shown in Problem C-1

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1999 monthly premium</td>
<td>$100</td>
</tr>
<tr>
<td>Annual premium increase</td>
<td>2.0%</td>
</tr>
<tr>
<td>Benefits:</td>
<td></td>
</tr>
<tr>
<td>In-hospital services</td>
<td>No co-payment</td>
</tr>
<tr>
<td>Office visits</td>
<td>$10 co-payment</td>
</tr>
<tr>
<td>Treatment quality indicators:</td>
<td></td>
</tr>
<tr>
<td>Members ‘very satisfied’ with physician access</td>
<td>41%</td>
</tr>
<tr>
<td>Members ‘very satisfied’ with availability of preventative care (e.g. immunizations)</td>
<td>39%</td>
</tr>
</tbody>
</table>

literal question asked ‘What percentage of members are ‘very satisfied’ with physician access?’ The answer options were (a) 100%, (b) 2%, (c) 41%, (d) 39%. (Correct answer is 41%.) The second question was an inferential one and required an individual to identify and combine information from two places in the table ‘In the year 2000, the HMO’s monthly premium will be (a) $98, (b) $100, (c) $102, (d) $120’. (Correct answer is $102.)

In Problem C-2, individuals were given information about four HMOs as shown in Exhibit 2. Again individuals were asked two multiple-choice questions. The first, literal question asked ‘Which HMO requires the lowest copayment for a visit with a primary care doctor?’ (Correct answer is HMO B.) The second, inferential question asked ‘Which HMO provides the best treatment quality according to the members’ ratings? The correct answer (HMO D) is determined by identifying which HMO has the highest percentage of members ‘very satisfied’ and lowest percentage of members ‘very dissatisfied’.

Problem C-3 presented information about three HMOs, as displayed in Exhibit 3, with the following instructions: ‘David doesn’t want any HMO that is below average on member satisfaction. He also doesn’t want any HMO below average on access to specialists. Which HMO will David choose?’ (Correct answer: HMO A.)

Exhibit 2. Complex four-option information shown in Problem C-2

<table>
<thead>
<tr>
<th></th>
<th>HMO A</th>
<th>HMO B</th>
<th>HMO C</th>
<th>HMO D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly premium</td>
<td>$50</td>
<td>$75</td>
<td>$48</td>
<td>$63</td>
</tr>
<tr>
<td>Copayment for office visit with primary care doctor</td>
<td>$10</td>
<td>$5</td>
<td>$15</td>
<td>$10</td>
</tr>
<tr>
<td>Percentage of members ‘very satisfied’ with treatment quality</td>
<td>38%</td>
<td>34%</td>
<td>28%</td>
<td>38%</td>
</tr>
<tr>
<td>Percentage of members ‘very dissatisfied’ with treatment quality</td>
<td>12%</td>
<td>14%</td>
<td>10%</td>
<td>10%</td>
</tr>
</tbody>
</table>

Exhibit 3. Instruction-following information shown in Problem C-3

<table>
<thead>
<tr>
<th>Member satisfaction</th>
<th>Preventative care strategies</th>
<th>Access to specialists</th>
<th>Customer service</th>
<th>Premium</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMO A</td>
<td>●</td>
<td>●</td>
<td>○</td>
<td>$60</td>
</tr>
<tr>
<td>HMO B</td>
<td>○</td>
<td>●</td>
<td>○</td>
<td>$60</td>
</tr>
<tr>
<td>HMO C</td>
<td>●</td>
<td>○</td>
<td>●</td>
<td>$60</td>
</tr>
</tbody>
</table>

○ = Below average ● = Average ○ = Above average.
Exhibit 4. Frequencies (and percentages) of older and younger respondents who gave incorrect answers to the five comprehension items

<table>
<thead>
<tr>
<th>Problem C-1: Simple one-option task</th>
<th>Older</th>
<th>Younger</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-1a. Literal question(^a)</td>
<td>45 (18.3%)</td>
<td>15 (6.3%)</td>
</tr>
<tr>
<td>C-1b. Inferential question(^b)</td>
<td>76 (31.8%)</td>
<td>42 (17.8%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problem C-2: Complex four-option task</th>
<th>Older</th>
<th>Younger</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-2a. Literal question(^c)</td>
<td>67 (27.4%)</td>
<td>19 (7.9%)</td>
</tr>
<tr>
<td>C-2b. Inferential question(^d)</td>
<td>137 (56.8%)</td>
<td>66 (27.8%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Problem C-3: Instruction-following task(^e)</th>
<th>Older</th>
<th>Younger</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-3b. Instruction-following task(^c)</td>
<td>39 (15.8%)</td>
<td>19 (7.9%)</td>
</tr>
</tbody>
</table>

\(^a\)\(\chi^2(1, N = 484) = 16.0, p < 0.0001.\)
\(^b\)\(\chi^2(1, N = 475) = 12.5, p < 0.001.\)
\(^c\)\(\chi^2(1, N = 484) = 31.2, p < 0.0001.\)
\(^d\)\(\chi^2(1, N = 478) = 41.1, p < 0.0001.\)
\(^e\)\(\chi^2(1, N = 485) = 7.2, p < 0.01.\)

Results

Exhibit 4 shows the frequencies and percentages of older and younger adults who answered the comprehension questions correctly and incorrectly.\(^2\) The results showed that older adults displayed significantly more error than younger adults on every one of the five comprehension questions (the results of chi-square tests are reported below the table). The biggest percentage difference (29.0%) in older and younger adults’ error rates occurred for the inferential question in Problem C-2; the smallest difference (7.9%) occurred for Problem C-3.

In addition, the results showed that the five comprehension items differed in difficulty for both older and younger adults. For Problems C-1 and C-2, as expected, individuals displayed more error on the inferential than literal questions. The inferential question of Problem C-2 was clearly the most difficult for both older and younger adults. Least difficult for both groups were Problem C-3 and the literal question of Problem C-1. The difference in performances by older and younger individuals was smallest for Problem C-1a and largest for Problem C-2b.\(^3\)

In sum, the above analyses show that both older and younger adults have more difficulty understanding information as the problem format becomes more complex. When only one HMO option is described and the question asked requires only identifying one literal piece of information (i.e. Problem C-1a in Exhibit 4), relatively few comprehension errors are made. In contrast, multiple options and inferential questions (i.e. Problem C-2b in Exhibit 4) are more challenging and lead to more comprehension errors for both older and younger adults. Importantly, however, comprehension is worse for older than younger adults, regardless of the information and question complexity (compare the two columns’ entries for each row of Exhibit 4).

Comprehension error index

In order to obtain a continuous-scale measurement (for later use in correlational analyses), a comprehension error index was created for each individual by summing his or her errors across the five comprehension questions (range 0–5). The mean comprehension error index score was 1.4 (s.d. = 1.2) for older adults.

\(^2\)Participants with missing data (up to \(N = 3\) for younger sample and \(N = 17\) for older sample) were deleted (rather than classified as incorrect) before the data for each problem were analyzed. Participants with missing data had a slightly higher mean error rate than subjects without missing data, particularly in the older adult sample. Thus, the exclusion is likely to make our estimates of comprehension errors conservative and also reduce the differences between younger and older individuals.

\(^3\)Confirming the pattern of results, a 5(item: C-1a/C-1b/C-2a/C-2b/C3) \(\times\) 2(age: younger/older) analysis of variance revealed significant main effects of item (\(F(4, 478) = 48.3, p < 0.0001\)) and age (\(F(1, 481) = 75.2, p < 0.0001\)) as well as a significant item \(\times\) age interaction (\(F(4, 478) = 6.0, p = 0.0001\)).
and 0.7 (s.d. = 0.9) for younger adults. Median scores were 1.0 and 0.0 for older and younger adults, respectively. The frequency (and percentage) of older and younger adults for each comprehension error index score is shown in Exhibit 5. The distributions of scores were significantly positively skewed for both groups. A median test revealed a significant difference between the comprehension error index scores of older versus younger adults, \( \chi^2(1, N = 466) = 63.4, p < 0.0001 \). No significant gender differences were found within the older or younger adult groups.

**CONSISTENCY: THE EVALUABILITY TASK**

Stimuli used by Hsee (1996b) were adapted to assess consistency in the present study. In Hsee’s original task individuals were asked to indicate how much they were willing to pay (WTP) for each of two second-hand music dictionaries: Dictionary A had 10,000 entries and was otherwise as new; Dictionary B had 20,000 entries, but also had a torn cover. When evaluated separately, the mean WTP for A was $24 and for B was $20. When evaluated side-by-side, the mean WTP for A decreased to $19 and for B increased to $27. Shifts in preference seemed to occur because the number-of-entries dimension only became ‘evaluable when comparative information was available. That is, knowing whether 10,000 entries is a good or a bad deal was difficult (and thus the dimension was underweighted) until A and B were shown side-by-side for evaluation. In short, judgments were impacted to the extent that individuals’ evaluations of the options were inconsistent across the different judgment contexts, despite the same information being available about the options being judged.

Hsee’s (1996b) ‘evaluability hypothesis’ suggests that individuals’ judgments will be inconsistent to the extent that the relative importance assigned to particular dimensions changes with variations in the mean ingtness of given information. Internal inconsistency in dimension weighting across judgment contexts means that different preferences or decisions may be displayed from one decision environment to the next. The focus in the present study is not on the ‘irrationality’ of potential inconsistency, but on the extent to which it occurs in older versus younger adults.

In the present study we were interested in how consistent individuals were in their interpretations of the values given for various dimensions (such as ‘percent satisfaction’ or ‘waiting time’) of health-plan options. For instance, consider the two HMOs described on two dimensions in Exhibit 6.

‘Membership satisfaction with treatment quality’ is presumably a relatively more important dimension than ‘doctor promptness’. When an HMO from Problem E-1 is shown alone, however, a value for ‘member satisfaction’ is hard to evaluate (meaning that an individual may not have a precise opinion about how good a given value on the dimension is without being able to compare it with another value). In contrast, a value for ‘doctor promptness’ should be easier to evaluate independently (meaning that the individual knows how
good the value is because people are more familiar with waiting times). In separate evaluation, Hsee (1996b) found that individuals tend to base their evaluation mainly on the easy-to-evaluate dimension because they do not know how to evaluate an option’s value on the less meaningful dimension. Thus, in separate evaluation for Problem E-1, values for doctor promptness should be given more weight causing HMO B to be rated as more attractive. But in joint evaluation, values for member satisfaction become evaluable (meaningful) and therefore more influential in judgment, leading to HMO A appearing more attractive. That is, when we ask our study participants to judge HMO plans A and B separately and also jointly, the effect of evaluation context may shift (or completely reverse) individuals’ preferences for one option over another.

Individuals’ sensitivity to changes in the evaluation context as described above will likely depend on the extent to which they rely on decision strategies that simplify decision tasks by (a) focusing on familiar over unfamiliar information, and (b) focusing on contextual cues to provide meaning to information (rather than retrieving information from memory). Thus, we expected the evaluability effect to be larger for older than younger adults because of older adults’ greater tendency for experiential processing of information and greater limitations in short-term memory capacity.

**Stimuli and procedure**

Study participants were given two problems in which they were asked to evaluate two hypothetical HMOs. The HMO descriptions used in Problem E-1 are shown in Exhibit 6 and those used in Problem E-2 are shown in Exhibit 7.

The task had three within-subject conditions for each problem: separate-evaluation-A, separate-evaluation-B, and joint evaluation. In the separate-evaluation conditions, individuals were given information about one health plan ‘you might join’ and were asked to rate the plan’s attractiveness on a 7-point scale (where 1 = ‘not at all attractive’ and 7 = ‘very attractive’). After several filler tasks, information about the second HMO was presented and its attractiveness rated. After further filler tasks, the two HMOs were presented side-by-side (the joint-evaluation condition) for rating on identical 7-point scales. The orders of presentation of separate-evaluation conditions were counterbalanced, but evaluation context order was not counterbalanced (i.e. separate evaluations always occurred before joint evaluations). The orders of the two problems were also counterbalanced. Note that in each problem Option A is superior to Option B on the first dimension (which is most important but hard to evaluate) and Option B is superior to Option A on the other dimension.

### Exhibit 6. Information shown in Problem E-1

<table>
<thead>
<tr>
<th></th>
<th>HMO A</th>
<th>HMO B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Member satisfaction with treatment quality (percent who are ‘very satisfied’)</td>
<td>75%</td>
<td>60%</td>
</tr>
<tr>
<td>Doctor promptness (average wait time past appointment)</td>
<td>25 minutes</td>
<td>10 minutes</td>
</tr>
</tbody>
</table>

*Hard-to-evaluate dimension when HMOs are shown separately.

### Exhibit 7. Information shown in Problem E-2.

<table>
<thead>
<tr>
<th></th>
<th>HMO A</th>
<th>HMO B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premium rate increases</td>
<td>1% per year</td>
<td>4% per year</td>
</tr>
<tr>
<td>Waiting time for customer service by telephone</td>
<td>5 minutes</td>
<td>30 seconds</td>
</tr>
</tbody>
</table>

*Hard-to-evaluate dimension when HMOs are shown separately.

Exhibit 8. Mean attractiveness (and standard deviations) of HMO A and B when rated separately or side-by-side in Problems E-1 and E-2 by older and younger adults

<table>
<thead>
<tr>
<th></th>
<th>Older</th>
<th>Younger</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Separate</td>
<td>Side-by-side</td>
</tr>
<tr>
<td><strong>Problem E-1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMO A</td>
<td>4.6 (1.7)</td>
<td>5.0 (1.6)</td>
</tr>
<tr>
<td>HMO B</td>
<td>5.2 (1.5)</td>
<td>4.2 (1.5)</td>
</tr>
<tr>
<td><strong>Problem E-2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMO A</td>
<td>4.9 (1.8)</td>
<td>5.5 (1.7)</td>
</tr>
<tr>
<td>HMO B</td>
<td>3.8 (1.9)</td>
<td>3.1 (1.6)</td>
</tr>
</tbody>
</table>

Note: The rating scale ranged from 1 (not at all attractive) to 7 (very attractive).

Results

The results showed that HMO attractiveness differed, as predicted by the evaluability hypothesis, according to the information presentation context. This was true for both older and younger adults, although the pattern of results varied somewhat according to age, problem, and HMO being rated. The mean ratings of each HMO in each judgment context are shown in Exhibit 8 and described in the text below.

Older adults

In Problem E-1 (N = 227), there was a preference reversal between separate and joint evaluations as predicted. In separate evaluation, the mean attractiveness was higher for HMO B (M = 5.2) than for HMO A (M = 4.6), t = 5.0, p < 0.001, but in joint evaluation the mean attractiveness was higher for HMO A (M = 5.0) than for HMO B (M = 4.2), t = 5.5, p < 0.001. The difference in separate and joint evaluations was highly significant both for HMO A (t = 3.6, p < 0.001) and HMO B (t = 8.0, p < 0.001).

In Problem E-2 (N = 227) we again found that HMO attractiveness differed according to the presentation context, as predicted, but it was revealed as a strengthening of preference as opposed to a preference reversal. In separate evaluation, the mean attractiveness was somewhat higher for HMO A (M = 4.9) than for HMO B (M = 3.8), t = 7.31, p < 0.001, but in joint evaluation the mean attractiveness was much higher for HMO A (M = 5.5) than for HMO B (M = 3.1), t = 14.4, p < 0.001. The difference in separate and joint evaluations was highly significant both for HMO A (t = 4.3, p < 0.001) and HMO B (t = 5.3, p < 0.001).

Younger adults

Younger adults showed the same pattern of results as reported above for older adults. In Problem E-1 (N = 234), there was a preference reversal between separate and joint evaluations as predicted. In separate evaluation, the mean attractiveness was higher for HMO B (M = 5.0) than for HMO A (M = 4.3), t = 6.2, p < 0.001, but in joint evaluation the mean attractiveness was higher for HMO A (M = 5.0) than for

---

Footnote: A 2x2 (age: younger/older) x 2 (problem: E-1/E-2) x 2 (HMO: A/B) x 2 (context: separate/side-by-side) analysis of variance revealed that the HMO and problem main effects were significant at p < 0.0001, and that all two-way interactions were significant at p < 0.0001, with the exception of the age x problem interaction which was not significant. The age x HMO x context interaction was also significant, F(1.459) = 9.0, p < 0.01. Follow-up tests on the three-way interaction showed that in the separate evaluation context, the ratings of HMO A and B were similar for both older adults (difference = 0.2, p < 0.05) and younger adults (difference = 0.1, n.s.), but in the side-by-side evaluation context, HMO A was rated as superior to HMO B and the superiority was much greater for ratings by older adults (difference = 1.7, p < 0.0001) than by younger adults (difference = 1.0, p < 0.0001).
HMO B \((M = 4.5, t = 4.2, p < 0.001)\). The difference in separate and joint evaluations was highly significant both for HMO A \((t = 6.7, p < 0.001)\) and HMO B \((t = 5.6, p < 0.001)\).

In Problem E-2 \((N = 234)\) we found again that HMO attractiveness differed according to the presentation context as predicted, but as was found with older adults, younger adults showed a strengthening of preference as opposed to a preference reversal. In separate evaluation, the mean attractiveness was somewhat higher for HMO A \((M = 4.6)\) than for HMO B \((M = 3.7, t = 6.5, p < 0.001)\), but in joint evaluation the mean attractiveness was much higher for HMO A \((M = 5.3)\) than for HMO B \((M = 3.7, t = 10.9, p < 0.001)\). The difference in separate and joint evaluations was highly significant for HMO A \((t = 5.3, p < 0.001)\), but not for HMO B.

**Overall**

Another way of looking at the results is to count the number of respondents who showed the predicted preference reversals or shifts in preference across evaluation contexts. Exhibit 9 shows that for both problems the majority of respondents’ preferences either reversed or shifted in the predicted direction. The pattern of results was similar for older and younger adults, with the exception of a stronger tendency by older adults to display a preference reversal or shift in the predicted direction on Problem E-2 (compare the last two columns on the first and second rows in Exhibit 9).

In sum, the pattern of results suggests that the relative importance of dimensions changed with changes to the evaluation context for both older and younger adults. As predicted according to Hsee’s (1996b) evaluability hypothesis, the effect of evaluation context was so powerful in Problem E-1 that it completely reversed many individuals’ preferences for one option over another. Preference reversals were not as frequent in responses to Problem E-2, possibly because ‘premium rate increases’ (the hard-to-evaluate dimension) were meaningful and relatively more important to individuals regardless of the evaluation context. The significant increase in preference for HMO A when individuals moved from separate to joint evaluation contexts, however, still suggests a change in the evaluability of the premium rate information. Importantly, older and younger adults seemed to be similarly influenced by information presentation context, with the exception of slightly greater inconsistency being displayed by older adults in Problem E-2.

**Exhibit 9. Frequencies (and percentages) of older and younger adults showing each response pattern for the two evaluability problems**

<table>
<thead>
<tr>
<th></th>
<th>Problem E-1(^b)</th>
<th>Problem E-2(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Older</td>
<td>Younger</td>
</tr>
<tr>
<td>Preference reversal in predicted direction</td>
<td>59 (26.0%)</td>
<td>61 (26.1%)</td>
</tr>
<tr>
<td>Preference shift in predicted direction(^a)</td>
<td>96 (42.3%)</td>
<td>94 (40.2%)</td>
</tr>
<tr>
<td>Preference reversal in unpredicted direction</td>
<td>6 (2.6%)</td>
<td>7 (3.0%)</td>
</tr>
<tr>
<td>Preference shift in unpredicted direction(^a)</td>
<td>35 (15.4%)</td>
<td>40 (17.1%)</td>
</tr>
<tr>
<td>No differences between ratings in separate and joint conditions</td>
<td>31 (13.7%)</td>
<td>32 (13.7%)</td>
</tr>
<tr>
<td>Total</td>
<td>227 (100.0%)</td>
<td>234 (100.0%)</td>
</tr>
</tbody>
</table>

\(^a\)A preference shift indicates a change in strength of preference for an option (as opposed to a complete reversal of preference).

\(^b\)Pattern of results for older and younger samples was not significantly different.

\(^c\)Pattern of results for older and younger samples approached significance, \(\chi^2(4, N = 466) = 9.0, p = 0.06\).
Exhibit 10. Frequencies (and percentages) of older and younger respondents for sub-ranges of evaluability index scores

<table>
<thead>
<tr>
<th>Evaluability index scores¹</th>
<th>Older</th>
<th>Younger</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–3</td>
<td>50 (22.0%)</td>
<td>78 (33.3%)</td>
</tr>
<tr>
<td>4–6</td>
<td>76 (33.5%)</td>
<td>99 (42.3%)</td>
</tr>
<tr>
<td>7–9</td>
<td>61 (26.9%)</td>
<td>42 (17.9%)</td>
</tr>
<tr>
<td>10–19</td>
<td>40 (17.6%)</td>
<td>15 (6.4%)</td>
</tr>
</tbody>
</table>

¹χ²(3, N = 461) = 23.9, p < 0.0001.

Evaluability index

An evaluability index was created for each individual by summing the absolute difference between separate and joint evaluations for each option in each problem (range = 0–19). The mean evaluability index score was 6.5 (s.d. = 3.5) for older adults and 4.8 (s.d. = 2.5) for younger adults. Median scores were 6.0 and 4.0, respectively, for older and younger adults. The frequencies (and percentage) of older and younger adults for the evaluability index scores are shown in Exhibit 10. (Twenty-six older adults and 5 younger adults had missing data on one or more of the items and were excluded from these analyses.) The distributions of scores were significantly positively skewed for both groups. A median test revealed a significant difference between the evaluability index scores of older versus younger adults, χ²(1, N = 461) = 27.1, p < 0.0001. No significant gender differences were found within the older or younger adult groups.

SELF-PERCEPTION MEASURES

Following the comprehension and evaluability tasks, individuals responded to questions designed to assess their attitudes toward delegating decisions, personal decision-making style, self-perceived skill at using tables and charts, and self-perceived health.

Attitude toward delegation

Respondents’ attitude toward delegation was assessed on eight items (e.g. ‘When choosing a Medicare health plan I prefer to not have the responsibility for choosing’), each rated on a four-point scale (1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree). An index of attitude toward delegation was calculated by summing across the eight items (range = 0–32). The mean of the delegation index was significantly higher for older adults (M = 18.8, s.d. = 5.6) than for younger adults (M = 16.7, s.d. = 4.0), t(466) = 5.4, p < 0.0001.

Decision-making style

Self-assessments of decision style were elicited using items from the ‘rational-vigilance’ scale of Epstein et al.’s (1996) Rational-Experiential Inventory. The mean scale score was significantly lower for older adults (M = 2.1, s.d. = 0.43) than for younger adults (M = 2.2, s.d. = 0.50), t(466) = 2.6, p < 0.05, indicating that older adults saw themselves as less rational-vigilant in their decision style.

¹Despite the evaluability hypothesis predicting only one direction for the effect, we used absolute differences for the index because we wanted to capture any inconsistency displayed by individuals across evaluation contexts.
Self-rated skill in using tables and charts

Individuals were asked to rate their skill in using tables and charts on a 4-point scale from (1) poor to (4) excellent. The mean skill score was significantly lower for older adults ($M = 2.7, \text{s.d.} = 0.7$) than for younger adults ($M = 3.1, \text{s.d.} = 0.7$), $t(462) = 6.2, p < 0.0001$.

Self-rated health

Individuals were asked to rate their health on a 5-point scale from (1) excellent to (5) poor. The mean health score was significantly higher for older adults ($M = 2.6, \text{s.d.} = 0.9$) than for younger adults ($M = 2.3, \text{s.d.} = 0.9$), $t(478) = 3.9, p < 0.0001$.

CORRELATIONAL RESULTS

In this section the intercorrelations of the behavioral performance indices, demographic variables, and measures of self-perceptions are examined. Exhibit 11 shows that the comprehension error index and evaluability index scores correlate positively ($r = 0.24, p < 0.0001, N = 445$). That is, as individuals made more errors in comprehension, they tended to be more inconsistent in their preferences across information presentation contexts. With the exception of income and self-perceived health, most of the other variables in Exhibit 11 also correlate significantly with the index scores, indicating that comprehension error and evaluability performances may be predictable from several factors. However, there are significant intercorrelations among the demographic and self-perception variables. Given that the contributions of the various predictors are not necessarily independent of each other, we examined partial correlations and conducted regression analyses.

Predicting behavioral performance indices

Covariation between age and other factors means that the differing performances of older and younger adults on the comprehension and consistency tasks may have been due to age per se and/or other variables. Compared with the younger adult sample in the present study, the older adult sample tended to be less well educated, to have a lower income, to want to delegate decisions more, and to perceive themselves to be less analytic in their decision style, to be less skilled in using tables and charts, and to have poorer health. The

Exhibit 11. Intercorrelations among behavioral performance indices, demographic variables, and measures of self-perceptions

<table>
<thead>
<tr>
<th></th>
<th>EL</th>
<th>Age</th>
<th>Educ</th>
<th>Income</th>
<th>Gender*</th>
<th>Deleg</th>
<th>Style</th>
<th>Skill</th>
<th>Health</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehension Error Index (CEI)</td>
<td>0.24</td>
<td>0.35</td>
<td>-0.39</td>
<td>-0.07</td>
<td>-0.14</td>
<td>0.34</td>
<td>-0.27</td>
<td>-0.32</td>
<td>-0.20</td>
</tr>
<tr>
<td>Evaluability Index (EI)</td>
<td>0.27</td>
<td>-0.25</td>
<td>-0.03</td>
<td>-0.04</td>
<td>0.15</td>
<td>-0.11</td>
<td>-0.16</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.21</td>
<td>0.20</td>
<td>-0.13</td>
<td></td>
<td>0.21</td>
<td>-0.16</td>
<td>-0.30</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>Education (Educ)</td>
<td>0.16</td>
<td></td>
<td>0.06</td>
<td>-0.30</td>
<td>0.27</td>
<td>0.39</td>
<td></td>
<td>-0.18</td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>-0.06</td>
<td></td>
<td></td>
<td>-0.16</td>
<td>0.10</td>
<td>0.14</td>
<td></td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td>Gender*</td>
<td></td>
<td></td>
<td></td>
<td>-0.06</td>
<td>0.06</td>
<td>0.03</td>
<td></td>
<td>-0.04</td>
<td></td>
</tr>
<tr>
<td>Attitude toward delegation (Deleg)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.44</td>
<td>-0.40</td>
<td>0.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rational-vigilant decision style (Style)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.43</td>
<td>-0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-rated skill in using tables and charts (Skill)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.29</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*1 = male; 2 = female.

Note: Sample sizes for each correlation ranged between 445 and 487.
Significance levels: for $r \geq 0.10, p < 0.05$; for $r \geq 0.15, p < 0.01$; for $r \geq 0.16, p < 0.001$. 

Exhibit 12. Partial correlations, the standardized regression coefficients ($\beta$), $R$-square, and adjusted $R$-square from a stepwise regression with Comprehension Error Index (CEI) as the dependent variable and education, attitude toward delegation, self-rated skill in using tables and charts, rational-vigilant decision style, self-rated health, and age (forced in last) as independent variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Partial correlation with CEI (DV)</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>$-0.23 \ (p &lt; 0.00001)$</td>
<td>$-0.22 \ (p &lt; 0.0001)$</td>
</tr>
<tr>
<td>Income</td>
<td>$-0.06 \ (\text{n.s.})$</td>
<td>$-0.03 \ (\text{n.s.})$</td>
</tr>
<tr>
<td>Gender (1 = male; 2 = female)</td>
<td>$-0.10 \ (\text{n.s.})$</td>
<td>$-0.09 \ (p &lt; 0.05)$</td>
</tr>
<tr>
<td>Attitude toward delegation</td>
<td>$0.16 \ (p &lt; 0.01)$</td>
<td>$0.15 \ (p = 0.01)$</td>
</tr>
<tr>
<td>Self-rated skill in using tables and charts</td>
<td>$-0.04 \ (\text{n.s.})$</td>
<td>$-0.05 \ (\text{n.s.})$</td>
</tr>
<tr>
<td>Rational-vigilant decision style</td>
<td>$-0.08 \ (\text{n.s.})$</td>
<td>$-0.08 \ (\text{n.s.})$</td>
</tr>
<tr>
<td>Self-rated health</td>
<td>$0.00 \ (\text{n.s.})$</td>
<td>$0.01 \ (\text{n.s.})$</td>
</tr>
<tr>
<td>Age (forced last)</td>
<td>$0.25 \ (p &lt; 0.00001)$</td>
<td>$0.25 \ (p &lt; 0.0001)$</td>
</tr>
</tbody>
</table>

$R^2 = 0.27$; Adjusted $R^2 = 0.26$; $F(8, 407) = 19.2, p < 0.0001$.
For age, Partial $R^2 = 0.05$, $F(8, 407) = 27.3, p < 0.0001$.

Group analyses reported above did not permit examination of performance differences linked to these other factors. Would age effects remain when differences on these other factors were held constant?

In order to better understand what underlies the age differences in performance on the comprehension and consistency tasks used in this study, we conducted two multiple regressions on data from the combined sample. Primarily, we were interested in whether age accounts for variance in performance on the comprehension and consistency tasks beyond that accounted for by other demographic variables and the self-perception measures. All independent variables were entered together, with the exception of age, which was entered last.

Exhibits 12 and 13 display the partial correlations, the standardized regression coefficients ($\beta$), $R^2$, and adjusted $R^2$ after entry of all independent variables. For the comprehension error index, partial correlations with education, attitude toward delegation, and age are significant; for the evaluability index, partial correlations with education and age are significant. Entering age as the last predictor in the regression analyses significantly improves the variance accounted for in the scores of the comprehension error and evaluability indices, beyond that afforded by the other variables. In other words, holding the other factors constant, there is still a reliable effect of age on the indices.

Exhibit 13. Partial correlations, the standardized regression coefficients ($\beta$), $R$-square, and adjusted $R$-square from a stepwise regression with Evaluability Index (EI) as the dependent variable and education, attitude toward delegation, self-rated skill in using tables and charts, rational-vigilant decision style, self-rated health, and age (forced in last) as independent variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Partial correlation with EI (DV)</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>$-0.18 \ (p &lt; 0.00001)$</td>
<td>$-0.19 \ (p &lt; 0.001)$</td>
</tr>
<tr>
<td>Income</td>
<td>$-0.03 \ (\text{n.s.})$</td>
<td>$-0.02 \ (\text{n.s.})$</td>
</tr>
<tr>
<td>Gender (1 = male; 2 = female)</td>
<td>$0.01 \ (\text{n.s.})$</td>
<td>$0.00 \ (\text{n.s.})$</td>
</tr>
<tr>
<td>Attitude toward delegation</td>
<td>$0.05 \ (\text{n.s.})$</td>
<td>$0.05 \ (\text{n.s.})$</td>
</tr>
<tr>
<td>Self-rated skill in using tables and charts</td>
<td>$-0.02 \ (\text{n.s.})$</td>
<td>$0.00 \ (\text{n.s.})$</td>
</tr>
<tr>
<td>Rational-vigilant decision style</td>
<td>$-0.02 \ (\text{n.s.})$</td>
<td>$0.02 \ (\text{n.s.})$</td>
</tr>
<tr>
<td>Self-rated health</td>
<td>$-0.01 \ (\text{n.s.})$</td>
<td>$0.00 \ (\text{n.s.})$</td>
</tr>
<tr>
<td>Age (forced last)</td>
<td>$0.20 \ (p &lt; 0.00001)$</td>
<td>$0.21 \ (p &lt; 0.0001)$</td>
</tr>
</tbody>
</table>

$R^2 = 0.10$; Adjusted $R^2 = 0.09$; $F(8, 406) = 5.9, p < 0.0001$.
For age, Partial $R^2 = 0.04$, $F(8, 406) = 15.9, p < 0.0001$.

DISCUSSION

Summary of results and comparison with previous findings
The present research examined older versus younger adults’ decision-making competence (DMC) through an analysis of comprehension and consistency skills on tasks about health plan options. The findings revealed that compared with younger adults, older adults experienced more comprehension difficulties and slightly more inconsistency across contexts differing in dimension evaluable. Age was found to account for a significant part of the performance variance even when covariates (education, self-rated skill in using tables and charts, self-rated health, desire for delegation of decision making, and decision-making style) were controlled statistically. Although the amount of variance accounted for by age (around 5%) may seem small, it is not necessarily trivial. Partial $R$-square values may underestimate the importance of a relationship—an argument made strongly by Rosenthal (1990), D’Andrade and Dart (1990), Funder and Ozer (1983), and Ozer (1985). Difficulties in interpreting the real-life meaning of age differences are discussed below.

The present demonstration of greater comprehension difficulties in older compared with younger adults is consistent with previous findings in studies on document comprehension (e.g. Park et al., 1994). Unlike previous findings, however, the present research revealed relatively high level of comprehension error. The greater comprehension difficulties displayed in our study may reflect differences in the content of stimuli: our stimuli were about unfamiliar health plan options whereas those used by Park et al. focused on the more familiar task of reading medication labels.

The present finding of inconsistent preferences across evaluation contexts is consistent with Hsee’s (1996a,b) evaluable hypothesis that says preferences will change from the option superior on the less important but easy-to-evaluate dimension in separate evaluation to the option superior on the more important but hard-to-evaluate dimension in joint evaluation. Older adults showed somewhat more inconsistency than younger adults. Importantly, the within-subjects design in our study was a very conservative test of consistency because subjects may have observed the manipulation and deliberately tried to make their responses consistent.

In sum, the present findings suggest that increasing age is related to decreasing comprehension and, to a lesser extent, decreasing consistency in judgment and decision making, at least for the tasks studied here. Finally, it is noteworthy that the age differences reported here were far clearer on the comprehension than the consistency task. Various explanations, discussed below, may account for the age differences we found.

Interpretation of results
One explanation for the age differences may be that they were the result of differences between our older and younger samples on variables other than age. For instance, the samples differed somewhat in education level. Fewer educational opportunities for older adults may have limited their development of the skills necessary for the comprehension and consistency tasks we presented to them. The correlational results are relevant here. While education and age covaried, the partial correlations between education and performance on the comprehension and consistency tasks were significant. That is, educational differences, independent of age, explained some of the discrepancy in performances of younger and older adults.

As one astute reviewer noted, several uncontrolled variables in fact may have been confounded with age because of our use of ‘intact’ groups, rather than a completely randomized design, and such confounds may not adequately be controlled for statistically (Baron and Treiman, 1980; Lord, 1969). However, the finding of an age effect within the older adult group undermines this concern (see Hibbard et al., 2001, for a description of comprehension scores among the older adults by age and education). There is much similarity among the individuals in the older adult group (e.g. all came to the same senior citizen’s center to participate in the same
activities), yet the oldest older adults (those over 80 years) clearly performed worse than other older adults. Furthermore, education did not attenuate the comprehension errors among the oldest old.

Given the above qualifications, we now consider potential explanations for the age differences found in this study. One possible explanation is related to the ‘expertise development’ approach to cognitive aging effects. This position suggests that individuals respond to demands by drawing on knowledge systems that become increasingly selective and domain-specific with age (Baltes and Baltes, 1990; Salthouse, 1991) and on components of performance that become increasingly automatized with experience (Denney, 1989). The sizable age differences on the comprehension task in the present study suggest that judgments and decisions about health plan options may be less compatible with older than younger adults’ knowledge and expertise. However, perhaps a different pattern of results would have been found if the task domain focused on something other than health plans. The demands of daily life offer opportunities for older and younger adults to develop differentially relevant expertise on many evaluation and choice topics, the salience of which depends on one’s age. While schooling may be a relatively salient issue for young parents, inheritance law may be an issue faced more often by grandparents. Information on professional backgrounds and recent and relevant experience on different decision tasks was not collected and thus could not be examined within the correlational analyses here, but some measure of experience or specific expertise would be useful information to gather in future studies.

A second explanation for age differences may be related to a disparate salience of analytic versus experiential information processing modes for younger and older adults. Lifespan theories typically make no predictions about mode salience (e.g. Fredrickson and Carstensen, 1990; although see the two-step control process discussed in Heckhausen and Schulz, 1995), but empirical work suggests age-related declines in deliberative processes and increased salience of associative and automatic processes (Peters et al., 2000). For instance, Adams et al. (1997) showed that younger adults tend to interpret stories analytically, whereas older adults focus less on details and more on gist. Increased dependence on schemas during comprehension (e.g. Hess, 1990) and a wider generation and use of inferences (Hamm and Hasher, 1992) may also reflect reduced emphasis on analytic processing, increased emphasis on experiential processing, or both, with increasing age. A tendency by older adults to move away from analytic processing may make them appear less competent than younger adults particularly on complex problems that require thorough deliberation. Age-related declines in analytic processing would suggest also that older adults, compared with younger adults, face greater difficulties in employing consistent judgment and decision strategies when the evaluability of information varies.

A third explanation for age differences in this study may be that they are a function of age-related changes in basic cognitive abilities (e.g. memory, attentional capacity, and speed of processing). This explanation is similar to the view of everyday problem solving as a complex, ‘compiled’ form of cognition (Marsiske and Willis, 1995; Park, 1992; Salthouse, 1990; Willis and Schaie, 1986). To the extent that good comprehension and consistency rely on basic cognitive abilities, as is claimed by some researchers investigating everyday problem-solving processes, the present results suggest that age-related cognitive declines have a negative impact on individuals’ decision competence. For instance, although we designed stimuli to avoid any demands on short-term memory, the tasks still involved working memory to the extent that information about the question that was asked needed to be stored while the answer information was searched for, identified, manipulated (if necessary), and reported. The clearest understanding of the factors underlying individual differences in comprehension and consistency skills will come from measures designed to isolate the role of specific components.

A fourth explanation for the age differences relates to the motivation of individuals participating in the study. One might argue that compared with younger adults, older adults are not as interested in doing

---

6Note, however, that some measured abilities (i.e. ‘crystallized’ abilities) do not typically decline with age compared with others (i.e. ‘fluid’ abilities).
psychology experiments. To address this concern it is important to examine motivation and to show that judgment and decision performance is unaffected by incentives that have measurable effects. Of note, however, is that anecdotal reports from our two data collectors indicated that older adults seemed no less motivated than younger adults. In general, older adults took the task seriously (mean completion times were approximately 45 minutes for younger adults and 60 minutes for older adults) and were interested, enthusiastic, and engaged in the task. Several older participants commented on the relevance to real-life decisions about health plan choices. In the event that there are motivational differences, however, assessing performance in future studies under realistic conditions would be valuable because the results would be meaningful, despite any differences in motivation.

Given that the stimuli used in the present study are simpler than those used in real-life for communicating information about health plan options, the rate of comprehension errors and inconsistencies found here is worrisome. However, the results of the present study do not address what the age differences mean for the quality and consequences of decision outcomes in real-life. The design of the present study did not permit evaluation of the comprehension and consistency skills of participants relative to an external benchmark. Even though older and younger adults differed in the ways they understood and processed information, we cannot tell whether this would have led to worse decisions, rather than simply to different decisions. Older adults may compensate for potential misunderstanding of information or sensitivity to context by using a variety of strategies, some of which might not be available in a laboratory setting or relevant to the HMO choices we were narrowly focused on. Furthermore, the seriousness of the consequences of the decisions made autonomously, compared to the consequences of decisions made in other ways cannot be determined from the present study. Surrogate decision-making (e.g. by a caregiver) may introduce new sources of variance (e.g. misunderstanding the values of the person in question) that affect decision outcomes.

While we collected information about the extent to which individuals desired to delegate decision making, measures of access to family care giving would be needed to better understand the meaning of responses on our delegation index. A desire for delegation does not necessarily mean that those wanting to subcontract (see Rosen and Olshavsky, 1987) with a caregiver to make decisions need such help or that those who prefer to be independent do not need help. Previous research has shown that actual seeking of assistance is not related to comprehension skills (Hibbard et al., 2001). By sampling from a senior center, we were unable to ascertain the views and skills of those who move voluntarily (or involuntarily) to a retirement center. Furthermore, while comprehension is undeniably a component of decision making, it is not a component unique to decision making. Many aspects of life require the communication and interpretation of information. Thus, performance outcomes in many domains could be affected by differential understandings of information.

In sum, it is important not to interpret group differences in DMC as necessarily reflecting underlying biological capabilities rather than responses to unequal opportunities, experience, motivation, or different life conditions (see Caouce, Coronado and Watson, 1998). The focus of the present research was on demonstrating age differences in comprehension and consistency tasks, rather than examining specific mechanisms underlying the differences. Interindividual variation may be related to neural changes that impact basic cognitive abilities or to psychological, behavioral, and social adaptations designed to mitigate the age-related changes (National Research Council, 2000). At this point, what can be said is that the comprehension and consistency indices in our study relate to other decision behaviors (e.g. as comprehension errors and inconsistency increased, individuals rated themselves as less skilled at using tables and charts and had a higher desire for delegating decisions), suggesting a meaningful pattern that supports the construct validity of our comprehension and consistency measures. In order to isolate the independent contributions of specific factors in age-related changes in judgment and decision skills, measures of diverse variables are needed, including reliable and valid assessments of working memory capacity, speed of processing, affective reactivity, actual and perceived decision support available, wisdom or experience in relevant domains, and various sociodemographic characteristics.
Implications and future directions
At the outset of this paper we emphasized our view of DMC as a multidimensional concept, including the abilities of comprehending, remembering, structuring, and rationally integrating information, appreciating the personal significance of information, and tempering impulsivity. Each of these abilities is expected to tap functionally different areas and our focus only on comprehension and consistency, while an important start, provides only a limited view on how older and younger adults’ DMC may differ.\(^7\) Focusing on one ability over another may result in discrepant conclusions about DMC because a person may be simultaneously capable and incapable with respect to different measures (Silberfeld, Corber, and Checkland, 1995). Moreover, examining a range of abilities that reflect not only declines but also age-related gains such as wisdom (see Baltes, 1993) is critical in modeling successful aging in the domain of judgment and decision making. A more comprehensive understanding of how and why DMC changes, may lead to promising interventions to improve decision support for older adults, where necessary. Future researchers should perhaps focus on evaluating alternative forms of intervention such as simply reducing task complexity, involving relatives in shared decision-making processes, imposing an expert’s choice, or doing nothing.

Furthermore, the importance of examining DMC as an interaction between characteristics of decision makers and decision tasks cannot be overstated. Just as declines in cognitive functioning may lead older adults to strategically avoid a deliberative, analytic information-processing mode, a variety of task-related factors (such as inbuilt complexity or external cues and commands) may lead to a dominance of experiential processing in older adults. For instance, if increasing age leads to a greater reliance on affective heuristic processing (e.g. basing an evaluation on the strength of feeling elicited by a dimension value), but no change in reliance on non-affective heuristics, then we would expect that age differences are more likely when tasks present information that can be assessed affectively. Viewing DMC as an interaction between characteristics of decision makers and decision tasks also implies that an individual may deal better with the demands of complex (compared with simple) tasks when thorough deliberation is the main goal if he or she is naturally more reliant on (or capable of) analytic (versus experiential) processing (see McMackin and Slovic, 2000).

In short, when comparing younger and older adults’ DMC future research needs to focus on both simple and complex decision tasks and on a range of individuals who vary in their information processing styles. Individual differences among people 65 years and older are likely to be striking on several other variables also (including the development of wisdom), suggesting that a systematic examination of a range of decision maker characteristics and their interactions with task characteristics will be essential.

Finally, more consideration of how decisions are made by and for the elderly is needed. Many health decisions are made by older adults in conjunction with industry caregivers or family members (often the oldest or nearest daughter) who contribute key inputs (Baumgarten, Rao, and Ring, 1976; Walker and Pratt, 1991). Age differences in the understanding and processing of health information suggest that caregivers may not necessarily make the same decisions as the individuals for whom they are caring. Future research needs to focus on understanding current practices in joint decision making and how this may facilitate or impede decision quality and satisfaction. Social psychological research has focused on how attitudes and beliefs about what older people should and should not do may determine what in fact they do (National Research Council, 2000), implying that there may be value in directing research towards better understanding how to facilitate age-related adaptation to ongoing environmental demands for competent decision making. Determining the meaning and seriousness of age differences in decision making requires a comparison of outcomes on laboratory tasks done autonomously with performance on criterion measures of real-world decisions made with the help of others.

\(^7\)Future research will benefit from simultaneously examining the possibility that some sort of g factor is involved, or that tasks differ in discriminating power, scaling, or g loading. Should this be the case, the main reason for pursuing multidimensional measures of decision competence is an applied one—the quality and consequences of decision outcomes will vary as a function of how the g factor relates to specific tasks.
Overall, the present research has demonstrated age differences in two skills underlying judgment and decision making—namely comprehension and, to a lesser extent, consistency. Decreasing comprehension and increasing inconsistency when evaluating health-plan options implies that simply providing more information may not help older adults make choices that fit their individual needs and preferences. Future research needs to examine other dimensions of DMC in a variety of familiar and unfamiliar problem domains (e.g. meal preparation, financial planning) and to begin exploring mechanisms that may explain when and why age differences occur. The way information is presented in choice and evaluation tasks will affect who will be able to use it effectively. Identifying the types of tasks most challenging for older adults and how age-related potentials and limits develop will allow judgment and decision aids to be tailored optimally. More empirical research on age differences in judgment and decision making is needed to determine why they occur in general and in decision tasks in particular.

ACKNOWLEDGEMENTS

The authors wish to thank the Robert Wood Johnson Foundation’s Changes in Health Care Financing and Organization initiative, the Public Policy Institute at the American Association of Retired Persons, and the Health Care Finance Administration for the funding that supported this work. We are grateful for help with data collection by Martin Tusler, Elizabeth Scholze Schmidt, and Stephen Johnson and for help with manuscript preparation by Janet Kershner, Leisha Wharfield, and Mona Bronson. We also appreciate the comments and suggestions of three anonymous reviewers.

REFERENCES


Hsee CK. 1998. Less is better: when low-value options are valued more highly than high-value options. *Journal of Behavioral Decision Making* 11: 107–121.


Authors’ biographies:

**Melissa L. Finucane** is a Research Investigator at the Center for Health Research Hawai’i, Honolulu, Hawai’i and a Research Scientist at Decision Research, Eugene, Oregon. Her research interests include judgment, choice, decision-making competence, affect, and risk perception.

**Paul Slovic** is president of Decision Research and a professor of psychology at the University of Oregon. He studies human judgment, decision making, and risk analysis.

**Judith Hibbard** is a professor Health Policy at the University of Oregon and a clinical professor of public health and preventive medicine at the Oregon Health Sciences University. Her research interests focus on consumer use of information in health care choices.

**Ellen Peters** is a Research Scientist at Decision Research in Eugene, Oregon. Her research interests include affect, emotion, deliberation, and age differences in judgments, choices, and risk perceptions.

**C. K. Mertz** is a data analyst at Decision Research, Eugene, Oregon. Her research interests include multivariate statistical methods, risk perception, and affect.

**Donald G. MacGregor** is a Senior Researcher with Decision Research, and a Senior Scientist with Applied Decision Concepts, both of Eugene, OR. His interests include judgment and decision making, risk communication, and decision aiding.