

## Influence of Age and Gender on Estimates of Long-Term Financial Growth Functions\*

Douglas A. Hershey  
George Mason University

### ABSTRACT

A number of studies have shown that when estimating the growth rate of exponentially increasing numerical functions, individuals tend to make linear projections which are substantial underestimates of actual growth rates. The present study was designed to determine whether there are age and gender differences in subjects' ability to accurately estimate exponentially increasing trends. Males and females age 20-79 were asked to estimate the future value of four different savings accounts at six different points in time. Analyses revealed age and gender differences in the shape of subjects' estimated future value functions (i.e., the extent to which they were linear or exponential), and gender differences in the accuracy of those estimates. No age differences in the quality of estimation performance were found. The discussion focuses on how different computational strategies could have led to the observed findings, and how differential levels of basic processing abilities and knowledge of the task may have influenced subjects' performance.

Researchers have shown that people are susceptible to a variety of systematic cognitive biases when making decisions and solving problems (Kahneman, Slovic, & Tversky, 1982). These biases in reasoning lead to suboptimal judgments, choices, and decisions across a wide range of novel and laboratory simulation tasks (Hogarth, 1987; Kaplan & Schwartz, 1977). Although there has been some debate over the real-world significance of these cognitive biases (Beach & Lipshitz, 1993), one bias in particular, the tendency to misperceive the rate of growth of exponentially changing trends, could be expected to have a substantial impact on individuals' personal financial planning decisions.

The fact that individuals systematically misperceive the growth rate of exponentially changing functions has been established in a

number of carefully conducted experimental studies (Dörner, 1983; Shaklee, 1990; Timmers & Wagenaar, 1977; Wagenaar, 1982; Wagenaar & Sagaria, 1975; Wagenaar & Timmers, 1978, 1979). In a typical study of this kind, subjects are asked to make estimates of the rate of growth of some amount of money (say, \$1000), held in a hypothetical savings account earning a specified rate of return (say, 9% interest compounded annually), for a set period of time (say, 30 years). Such an account would actually grow at an exponentially increasing rate over that 30-year period. In all but rare circumstances, however, subjects' perceptions of this type of growth function are linear, and substantial underestimates of the actual financial accumulation.

Savings and investment accounts which accumulate assets on the basis of compound interest

---

\* The author would like to thank Darla Swann and Leslie Burke for their assistance in collecting the data, Jason Allaire for his help in conducting the growth curve analyses, and Ilda T. Hershey for her assistance with the section on estimation strategies. The author is also indebted to Neil Charness, Theodore Gessner, Linda Chrosniak, Alicia H. Farrell, and an anonymous reviewer for insightful comments on an earlier draft of this manuscript. Portions of this paper were presented in November of 1991 at the 32nd annual meeting of the Psychonomic Society in San Francisco. Correspondence may be directed to Douglas A. Hershey, Department of Psychology, George Mason University, Fairfax, VA 22030-4444, USA.

Accepted for publication: July 30, 1995.

payments grow at an exponentially increasing rate over a prolonged period of time. The largest increase in assets occurs over a relatively short period of time, during the later years of the investment term when interest payments skyrocket. Financial planners often refer to this late-term growth as the *miracle of compounding*. Numerous investment vehicles are designed to take advantage of the miracle of compounding, such as savings accounts established for a child's college education, Individual Retirement Accounts (IRAs), 401k plans, and Keogh accounts. Each of these investment vehicles allow individuals to reach their financial goals by setting aside modest amounts of money over a long period of time. Interestingly however, financial institutions fail to emphasize the long-term investment value of these accounts, choosing instead to emphasize annual interest rates, which focus individuals' perceptions on relatively short-term gains (Shaklee, 1990). This leads to a situation where individuals are led to make important long-term investment decisions on the basis of annual interest rates, rather than on the basis of how their investments are likely to grow over a period of decades.

Studies on the misperception of exponential growth (cited above) suggest that we can expect individuals to make suboptimal personal investment decisions based on the tendency to underestimate the long-term growth of savings and investments. Individuals who take a long-term approach to investing but who underestimate long-term yields are prone to making substantial overinvestments, particularly in the early stages of the term when modest investments would suffice. In contrast, the more common scenario involves individuals who wait longer than they should to establish an investment account, and then find themselves "working harder" than they would like to in order to achieve their financial goals. Such individuals have to set aside more money than they can comfortably afford, making relatively large investment contributions over a short period of time. Presumably, most of these individuals would have opted to set aside much smaller amounts of money over the longer term, had they only understood the simple financial principles behind the miracle of compound-

ing. It is interesting to note that a new field of research called *financial gerontology* has emerged within the past few years. It specifically focuses on the way in which personal financial planning behaviors relate to financial independence, perceived well-being, and quality of life issues (Cutler, Gregg, & Lawton, 1992; Gregg, 1990). One of the key underlying premises of this area of research is that a veridical mental model of finance is central to an individual's ability to successfully invest for the long term.

Therefore, from an applied perspective it is important to identify the types of situations in which individuals are likely to make systematic errors in reasoning as a prelude to developing methods for improving decision performance. Many have touted the development of educational and training programs which are aimed at "debiasing" individuals (Evans, 1989; Fischhoff, 1982; Russo & Shoemaker, 1992; Vye, Delclos, Burns, & Bransford, 1988), while others have advocated the use of computer-based decision aides to improve decision making performance (Fox, 1984; Raphael, 1976; Thomas, 1989). These applied training and intervention goals could potentially be of great personal significance given the large number of difficult financial decisions individuals face during the course of their lives, and the late-life implications of having made faulty investment decisions. Unfortunately, a number of studies have demonstrated that individuals are prone to systematic cognitive errors when trying to understand or predict changes in financial trends.

Using a numerical estimation task similar to the financial estimation task described above, Dörner (1983) showed that subjects substantially underestimate the growth rate of exponentially increasing functions. In characterizing his subjects' performance, he wrote, "When individuals *consciously* attempt to evaluate [exponentially increasing] trends, they seem to display a tendency to extrapolate *linearly*; this process, however, doesn't account for the actual exponential developments in economical and ecological systems" (p. 105). Kemp (1984, 1987) also found similar cognitive errors using a conceptually different estimation task. Rather than asking individuals to estimate the future

value of a savings account, he asked subjects to estimate the cost of basic commodities (e.g., a pound of butter, a first class postage stamp) at various points in the past. Subjects in these studies were found to consistently overestimate the past cost of butter and stamps, a mirror image of the types of errors made by subjects in the Dörner (1983) experiment.

In another recent study, Shaklee (1990) reported gender differences on a financial estimation task. She found that males tended to make more accurate estimates of the long-term growth of a savings account than females, although again, the more general finding was that both genders appreciably underestimated the actual growth of the account. Other researchers have sought to determine whether there are gender differences in the ability to project exponentially increasing trends; however, no clear-cut pattern of gender differences has been identified. In two studies no reliable gender differences were found (Kemp, 1984; Wagenaar & Sagaria, 1975), whereas in two other studies, mixed gender effects were found which were dependent upon the nature of the judgment task (Bates & Gabor, 1986; Kemp, 1987). One possible explanation for these equivocal gender-related findings is that these studies weren't originally designed to examine gender effects. That is, gender differences in these experiments were treated as a minor demographic issue, and comparisons between males and females were typically conducted in an exploratory context. Furthermore, these gender-based analyses often failed to account for age, educational level, and task-related experience, three individual difference variables that might also affect estimation performance.

The goal of the present study is to establish whether there are age and gender differences in subjects' ability to accurately estimate the growth of four different savings accounts which earn fixed rates of interest over a period of 30 years. The sampling design and analysis plan employed in the present study will allow for a clear test of gender differences. Furthermore, there are also sound theoretical reasons for anticipating that age differences in performance will be identified.

Current psychometric theories of cognitive functioning suggest that dual intellectual capacities – fluid abilities and crystallized knowledge – jointly contribute to our ability to carry out complex tasks such as estimating numerical trends. Moreover, both types of abilities have been found to display differential growth patterns across the adult life span (Baltes, 1987; Horn, 1968, 1982; Horn & Hoffer, 1992). *Fluid abilities* are the constellation of basic cognitive resources (e.g., memory, attention, reasoning, processing speed) which support high-level information processing. Across a large number of studies, fluid abilities have been found to remain fairly stable throughout young and middle adulthood, and then show marked declines in old age (see Horn, Donaldson, & Engstrom, 1981 for a review). *Crystallized knowledge*, in contrast, is the body of culturally significant knowledge one acquires through life experiences. Crystallized knowledge has been shown to grow steadily throughout young adulthood and middle age, and then plateau in late adulthood, often displaying little or no measurable increases after about 60 years of age. Crystallized knowledge structures are hypothesized to contain both general information about the world, as well as more detailed, domain-specific knowledge related to individuals' particular areas of expertise (e.g., chess, physics, financial planning).

Successful performance on an estimation task such as the one employed in the present study appears to involve a combination of fluid and crystallized abilities. Given the computational nature of the task, both working memory and attention must play a critical role in determining subjects' estimates. Furthermore, both of these fluid abilities have been shown to decline with advancing age (see Craik, Morris, & Gick, 1990; McDowd & Birren, 1990; and Plude & Hoyer, 1985 for reviews). This suggests that younger subjects have a computational advantage when making future value estimates, based on superior levels of fluid processing resources. However, it is also likely that subjects' estimates are strongly influenced by their level of arithmetic abilities, and their prior knowledge of finance and investing, both of which are crystallized abilities. Presumably, the most accurate estimates will be made by individuals

who (a) understand how interest accumulates in a savings account (i.e., exponentially, through the miracle of compounding), (b) have a highly developed repertoire of arithmetic strategies, and (c) are facile in working with numbers. With respect to the latter two points, developmental studies have shown that the arithmetic and numerical abilities of older adults are superior to those of younger individuals (Geary, Frensch, & Wiley, 1993; Geary & Wiley, 1991; Schaie, 1983). Geary et al. (1993) suggest that these age differences stem from the ability of older adults to use a "more mature mix of [computational] problem-solving strategies" (p. 242) when working on complex arithmetic tasks. This being the case, it is conceivable that when making future value estimates, older individuals will have an advantage over younger individuals given a lifetime of investing and financial planning experience, and a strong repertoire of strategies for working with numbers.

The argument presented above suggests that old and young subjects may solve the financial estimation task differently, based on differential levels of fluid abilities and domain-specific crystallized knowledge. Although to date there have been no systematic studies of age differences in the quality of subjects' exponential growth estimates, the theory of fluid and crystallized intelligence (Horn, 1968, 1982; Horn & Hoffer, 1992) provides a strong theoretical framework for hypothesizing that age differences in estimation ability will be identified.

## METHOD

### Participants

Subjects were 96 males and 96 females who ranged in age from 20 to 79 years. Participants were sampled in order to form four age groups: 20-year-olds (20-34 years), 40-year olds (35-49 years), 50-year-olds (50-64 years), and 70-year-olds (65-79 years). Each of the eight age-by-gender subgroups (i.e., 20-year-old males, 20-year-old females, 40-year-old males, etc.) were comprised of 24 participants. The mean age and number of years of formal education for each of the eight subgroups are reported in Table 1. All subjects were residents of the Greater Washington D.C. area and were sampled at various locations throughout the community (e.g., service organization meetings, senior centers, libraries). All subjects willingly participated in the study without remuneration. The rate of attrition was uniformly low for both males and females across the age range; only four subjects chose to withdraw from the experiment once testing had begun.

### Procedure

Subjects estimated the future value of four different hypothetical savings accounts, at six different points in time. Each of the four accounts differed in terms of (a) the amount of money initially placed in the account (\$1,000 or \$2,500), and (b) the annual rate of return associated with the investment (6% APR or 11% APR). Subjects were told that the savings accounts grew solely on the basis of annual compound interest payments made to the account at the end of each calendar year. The task was to estimate the value of the account at five-year intervals over a period of 30 years. The six estimates per account (at 5, 10, 15, 20, 25, and 30 years) in each of the four experimental conditions (\$1,000/6%; \$1,000/11%; \$2,500/6%; and

Table 1. Means and Standard Deviations (in parentheses) for Age and Number of Years of Formal Education as a Function of Age Group and Gender.

	Age		Education	
	Male	Female	Male	Female
20-year-olds	26.9 (4.4)	26.0 (4.6)	15.4 (2.7)	13.7 (1.9)
40-year-olds	41.7 (4.1)	40.7 (4.4)	14.2 (3.2)	14.2 (2.8)
50-year-olds	58.0 (4.3)	57.9 (3.6)	14.2 (2.3)	13.4 (1.7)
70-year-olds	71.2 (4.0)	70.0 (3.9)	14.8 (2.4)	13.1 (2.4)

\$2,500/11%) led to a total of 24 estimates per subject.

Subjects were informed at the outset of the task that they would have an unlimited amount of time to generate the 24 financial estimates, and that they were to make mental estimates of future values, rather than using a pencil and paper to make their calculations.<sup>1</sup> Upon completion of the task, subjects provided demographic information and information regarding their prior experience in making financial investments. All subjects were tested either individually or in small groups of two or three persons.

### Materials

Test booklets included information regarding the characteristics of each particular account (e.g., \$1,000 initial deposit and 6% annual rate of return) at the top of the page, followed by six spaces below labeled in five-year increments. There were four test pages per booklet, one for each experimental condition. The final page of the booklet contained demographic questions and a set of questions which were designed to provide a rough assessment of subjects' prior level of financial/investing experience.<sup>2</sup>

<sup>1</sup> This mental simulation procedure was selected for use in the present study because it is identical to the technique used in several of the studies cited above which served to establish the underestimation bias (e.g., Dörner, 1983; Kemp, 1984, 1987; Shaklee, 1990). Furthermore, it could be argued that many investors (if not a majority) often make critical resource allocation decisions on the basis of "mind's eye simulations." This suggests that in general, mental extrapolation tasks such as the one used in this study possess a reasonable degree of ecological validity. Researchers have yet to determine whether the underestimation bias is also found under conditions where subjects are given the opportunity to make their estimates on the basis of paper-and-pencil computations.

<sup>2</sup> This variable, which was designed to assess individuals' level of financial/investing experience, was based on subjects' self-report of whether or not they had held savings accounts, stocks, bonds, retirement plans, and so forth, at the time of, or prior to testing. It was initially conceived of as a proxy marker of subjects' knowledge of finance and investing. However, as a composite variable it was not found to be correlated with the key measures of decision quality (estimate ratio scores and overall mean error scores). Ostensibly, this was because this variable was not sufficiently fine-tuned to assess individual differences in knowledge. Therefore, it is more reasonable to think of this variable as a marker of one's level of prior financial and investing experience, rather than of one's level of financial knowledge, while at the same time recognizing that these two constructs are not always perfectly correlated with one another.

### Design

A mixed between/within subjects design was employed in which the two between-subject factors were *Gender* (two levels: male, female) and *Age Group* (four levels: 20-year-olds, 40-year-olds, 50-year-olds, and 70-year-olds). A sample size analysis (Cohen, 1988) revealed that a minimum of 44 subjects should be included in each age group, and 64 subjects in each gender group, in order to obtain reasonable power for the primary decision quality analysis (i.e., the MANOVA) reported below.<sup>3</sup> The final sample included 48 subjects in each age group and 96 subjects for each gender group.

The within-subject factors were the *Rate* of return on the account (two levels: 6% and 11%), the initial investment *Amount* (two levels: \$1,000 and \$2,500), and the number of *Years* in the investment period (six levels: 5, 10, 15, 20, 25, and 30 years). Throughout the remainder of the paper these five experimental factors will be referred to as they are italicized above (i.e., *Gender*, *Age Group*, *Rate*, *Amount*, and *Years*). The presentation order of the four different (amount by rate) within-subject conditions was fully counter-balanced in an effort to minimize order effects.

## RESULTS

This section of the paper is divided into three parts. The first part reports the results of a series of regression analyses which were designed to determine whether subjects' growth estimates were linear or exponential. The second set of analyses focuses on the quality of subjects' estimates. The third section contains a description of the various strategies subjects used to generate the future value estimates.

### Shape of the Growth Curve Estimates

Analyses were conducted to determine whether subjects' estimates of financial growth were better fit by exponential or linear functions. As pointed out above, studies have shown that individuals tend to make linearly increasing estimates of accounts that actually grow at an exponentially increasing rate (cf. Dörner, 1983; Shaklee, 1990). In testing whether individuals' estimates of growth curves approximated

<sup>3</sup> These two sample size analyses were based on what Cohen (1988) describes as the "standard conditions," that is, power = .80, alpha = .05 (two-tailed), and  $f = .25$ .

straight-line or exponential functions, two separate regression analyses – one for each type of trend – were computed for each of the four future value functions generated by each subject.<sup>4</sup> The  $R^2$  values associated with these regression analyses were then used as an index of the relative goodness of fit.

The entries in Table 2 indicate the percentage of individuals whose growth estimates were better fit by exponential functions than linear functions. Percentages greater than 50% indicate that a majority of subjects made exponential estimates, whereas values of less than 50% indicate a preponderance of linear estimates. Sign tests were used to establish the statistical significance of the difference in fit for each growth curve condition, based on a comparison of  $R^2$  values derived from the linear and exponential

equations. Asterisks in the table denote conditions where a statistically significant majority of individuals generated exponential estimates. Also reported in the last column of Table 2 is the mean percentage of subjects who were found to have made exponential estimates averaged across all four account types. The same greater than 50% (exponential), less than 50% (linear) relationship described above for the four individual growth curves holds for these mean values. The mean percentages in this last column were also tested against the null hypothesis that  $p = .50$ .

The first line of data in Table 2 reveals that a majority of subjects in the sample made exponentially increasing estimates of financial returns. Fifty-nine percent of subjects made exponential estimates in the \$1,000/6% condition,

Table 2. Percentage of Individuals whose Financial Growth Estimates were Better Fit by Exponential than Linear Functions, Reported as a Function of Age, Gender, and Account Type.

	6% APR		11% APR		Mean for all four accounts
	\$1,000	\$2,500	\$1,000	\$2,500	
All subjects	59*	72**	65**	69**	66**
Males	65**	80**	72**	73**	72**
Females	52	64*	57	64*	59**
20-year-olds	67*	75**	76**	84**	76**
40-year-olds	55	74**	65	67*	66**
50-year-olds	57	71*	67*	69*	66**
70-year-olds	55	66	51	55	57

Note. Asterisks indicate conditions where exponential functions produced significantly better fits than linear functions, based on the results of sign tests using  $R^2$  values as the measure of relative goodness-of-fit. Values in the last column represent the mean percentage of individuals who made exponential estimates averaged across all four account conditions. Each of the percentages contained in the table were tested against the null hypothesis that  $p = .50$ .

\*  $p < .05$ . \*\*  $p < .01$ .

<sup>4</sup> As one reviewer suggested, the strategy of analyzing individuals' growth curves in this fashion provides a more comprehensive view of the data than could be achieved by analyzing average growth curves as represented by various groups (e.g., males, or 20-year-olds). This data analysis strategy resulted in the computation of 1,528 separate regression analyses, half of which involved fitting straight-line functions, and the other half of which were based on exponential functions.

72% made exponential estimates in the \$2,500/6% condition, 65% in the \$1,000/11% condition, and 69% in the \$2,500/11% condition. Moreover, the mean value of 66% for all subjects across all conditions was found to be significantly greater than 50% ( $p < .01$ ). This general finding – that the majority of subjects tended to make exponential estimates of financial returns – is particularly noteworthy in light of the fact that in previous experiments (which

have predominantly relied upon mean estimates to characterize group-based trends), individuals have been shown to make linear projections of financial returns.

The goodness-of-fit data become increasingly interesting as the relative estimation performance of the various subgroups is examined. For example, a clear gender difference was identified in the shape of subjects' growth curve estimates for the two \$1,000 investment conditions. Furthermore, at the level of each of the four accounts, males made significantly more exponential estimates than linear estimates. However, a comparable effect was not found for females. Females were found to have made predominantly exponential estimates only in the two \$2,500 conditions; there was roughly an equal split of females who made linear and exponential estimates in the two \$1,000 conditions. There was also evidence of a gender difference based on the overall mean score values. More males (72%) made exponential estimates than females (59%). A test of independent proportions revealed that this 13% difference was statistically significant at the .01 level.

A pattern of age-related differences is also indicated in Table 2. Among 20-year-olds, a significant majority of subjects made exponential estimates in all four account conditions. In sharp contrast, among 70-year-olds, none of the four accounts reflected estimates which tended toward linear or exponential estimates. A more global assessment of age differences in the shape of the growth curve estimates involved comparing the mean percentage values for each of the four age groups (collapsed across amount and rate) through tests of independent proportions. It was found that significantly more 20-year-old subjects made exponential estimates than 40- or 50-year-olds ( $p < .05$ , both tests). The percentage of subjects who made exponential estimates was lowest among 70-year-olds (57%); however, the mean difference in percentages between this group, and the 40- and 50-year-olds (both 66%) was not found to be significantly different.

Further age group by gender analyses (not shown in the table) were conducted to determine which subgroups contained the largest and

smallest percentage of individuals to have made exponential estimates. Males in the 20- and 50-year-old groups contained the largest percentage of individuals to make exponential estimates (85% and 81%, respectively, both significantly different from  $p = .50$ ), and the 50- and 70-year-old female groups contained the smallest percentage of individuals to make exponential estimates (50% and 48%, respectively, neither significantly different from  $p = .50$ ).

In summary, the above analyses indicate that a majority of individuals made exponential estimates of the financial growth functions in each of the four conditions. More males made exponential estimates than females, and younger subjects made more exponential estimates than older subjects. The group containing the largest percentage of individuals making exponential estimates were 20-year-old males, and the group containing the smallest percentage of individuals making exponential estimates were 70-year-old females. Taken together, these analyses indicate significant influences of age and gender on the growth curve estimates. What these data fail to reveal, however, is how accurate subjects were at making the estimates. In the following section of the paper the quality of subjects' performance is evaluated by comparing their estimated values to the actual financial yields.

#### Quality of the Growth Curve Estimates

In order to establish a common performance metric across the four experimental conditions, subjects' estimates were converted to *estimate ratios* by dividing each of the 24 estimates by the actual values of the four accounts at the six different points in time (Shaklee, 1990). Therefore, estimate ratios of .85 indicate a 15% *underestimate* of the actual account value, ratios of 1.20 indicate 20% *overestimate*, and ratios of 1.0 indicate perfectly accurate judgments. Table 3 contains mean estimate ratios and standard deviations for the eight (age group by gender) subgroups for each of the four (rate by amount) estimation conditions across the 6, five-year periods.

A 2 (Gender) x 4 (Age Group) x 2 (Rate) x 2 (Amount) x 6 (Years) multivariate analysis of variance was computed using the 24 estimate

Table 3. Mean Estimate Ratios (top, row, bold) and Standard Deviations (below) for the Two Between-Subject Factors (Age Group and Gender) and the Three Within-Subject Factors (Percent, Amount, and Years).

6% rate of return												
	\$1000 starting value						\$2500 starting value					
	Years						Years					
	5	10	15	20	25	30	5	10	15	20	25	30
20s females	<b>1.2</b> (.20)	<b>1.1</b> (.49)	<b>1.2</b> (.75)	<b>1.3</b> (1.0)	<b>1.4</b> (1.3)	<b>1.5</b> (2.5)	<b>.97</b> (.18)	<b>.95</b> (.32)	<b>.93</b> (.41)	<b>.90</b> (.46)	<b>.89</b> (.51)	<b>.88</b> (.56)
20s males	<b>.99</b> (.14)	<b>1.1</b> (.23)	<b>.98</b> (.32)	<b>.99</b> (.45)	<b>.96</b> (.51)	<b>.94</b> (.58)	<b>.91</b> (.14)	<b>.89</b> (.18)	<b>.90</b> (.28)	<b>.90</b> (.37)	<b>.90</b> (.44)	<b>.89</b> (.55)
40s females	<b>.99</b> (.16)	<b>1.0</b> (.29)	<b>1.1</b> (.51)	<b>1.3</b> (.74)	<b>1.4</b> (1.0)	<b>1.7</b> (1.5)	<b>.93</b> (.15)	<b>.92</b> (.33)	<b>.96</b> (.52)	<b>1.0</b> (.74)	<b>1.1</b> (1.1)	<b>1.4</b> (1.6)
40s males	<b>.99</b> (.11)	<b>.98</b> (.23)	<b>.97</b> (.33)	<b>.93</b> (.39)	<b>.89</b> (.43)	<b>.84</b> (.47)	<b>.94</b> (.11)	<b>.90</b> (.20)	<b>.87</b> (.27)	<b>.88</b> (.42)	<b>.87</b> (.52)	<b>.86</b> (.62)
50s females	<b>.99</b> (.11)	<b>1.1</b> (.31)	<b>1.1</b> (.45)	<b>1.2</b> (.60)	<b>1.1</b> (.71)	<b>1.1</b> (.89)	<b>.93</b> (.08)	<b>.90</b> (.20)	<b>.89</b> (.31)	<b>.95</b> (.69)	<b>1.0</b> (1.3)	<b>1.0</b> (1.6)
50s males	<b>.94</b> (.11)	<b>.93</b> (.19)	<b>.94</b> (.31)	<b>.95</b> (.40)	<b>.99</b> (.47)	<b>.99</b> (.50)	<b>.91</b> (.10)	<b>.92</b> (.23)	<b>.92</b> (.30)	<b>.91</b> (.41)	<b>.92</b> (.50)	<b>.94</b> (.69)
70s females	<b>1.0</b> (.19)	<b>1.1</b> (.39)	<b>1.4</b> (1.4)	<b>1.5</b> (1.9)	<b>1.5</b> (2.4)	<b>1.5</b> (2.6)	<b>.94</b> (.17)	<b>1.6</b> (3.5)	<b>1.5</b> (3.1)	<b>1.4</b> (2.4)	<b>1.4</b> (2.8)	<b>1.5</b> (3.0)
70s males	<b>1.1</b> (.43)	<b>1.1</b> (.58)	<b>1.2</b> (.82)	<b>1.1</b> (.62)	<b>1.0</b> (.61)	<b>1.0</b> (.60)	<b>.92</b> (.10)	<b>.90</b> (.20)	<b>.92</b> (.28)	<b>.92</b> (.37)	<b>.91</b> (.47)	<b>.91</b> (.56)
<b>TOTAL</b>	<b>1.0</b> (.21)	<b>1.0</b> (.36)	<b>1.1</b> (.70)	<b>1.2</b> (.89)	<b>1.2</b> (1.1)	<b>1.2</b> (1.3)	<b>.93</b> (.13)	<b>1.0</b> (1.3)	<b>.99</b> (1.1)	<b>.98</b> (.98)	<b>1.0</b> (1.2)	<b>1.0</b> (1.4)
11% rate of return												
	\$1000 starting value						\$2500 starting value					
	Years						Years					
	5	10	15	20	25	30	5	10	15	20	25	30
20s females	<b>1.1</b> (1.1)	<b>.94</b> (.74)	<b>.80</b> (.56)	<b>.68</b> (.49)	<b>.57</b> (.45)	<b>.52</b> (.49)	<b>.85</b> (.24)	<b>.73</b> (.37)	<b>.63</b> (.44)	<b>.56</b> (.46)	<b>.47</b> (.42)	<b>.40</b> (.39)
20s males	<b>.94</b> (.21)	<b>.83</b> (.27)	<b>.75</b> (.39)	<b>.69</b> (.48)	<b>.63</b> (.49)	<b>.57</b> (.52)	<b>.87</b> (.19)	<b>.79</b> (.30)	<b>.69</b> (.33)	<b>.63</b> (.39)	<b>.58</b> (.41)	<b>.52</b> (.44)
40s females	<b>.89</b> (.17)	<b>.86</b> (.46)	<b>.82</b> (.46)	<b>.81</b> (.70)	<b>.84</b> (1.2)	<b>1.0</b> (2.3)	<b>.84</b> (.14)	<b>.74</b> (.26)	<b>.64</b> (.29)	<b>.57</b> (.36)	<b>.56</b> (.52)	<b>.62</b> (.95)
40s males	<b>.90</b> (.16)	<b>.78</b> (.23)	<b>.66</b> (.27)	<b>.56</b> (.30)	<b>.48</b> (.31)	<b>.40</b> (.29)	<b>.84</b> (.13)	<b>.69</b> (.18)	<b>.58</b> (.22)	<b>.50</b> (.26)	<b>.42</b> (.26)	<b>.35</b> (.25)
50s females	<b>.87</b> (.15)	<b>.81</b> (.27)	<b>.72</b> (.35)	<b>.62</b> (.35)	<b>.52</b> (.37)	<b>.45</b> (.40)	<b>.86</b> (.12)	<b>.74</b> (.23)	<b>.63</b> (.33)	<b>.53</b> (.37)	<b>.43</b> (.36)	<b>.38</b> (.39)
50s males	<b>.99</b> (.17)	<b>.95</b> (.61)	<b>.84</b> (.57)	<b>.73</b> (.50)	<b>.64</b> (.46)	<b>.56</b> (.47)	<b>.83</b> (.16)	<b>.73</b> (.24)	<b>.63</b> (.27)	<b>.56</b> (.30)	<b>.51</b> (.33)	<b>.47</b> (.37)
70s females	<b>1.0</b> (.55)	<b>1.1</b> (1.3)	<b>1.0</b> (1.4)	<b>.82</b> (1.1)	<b>.62</b> (.85)	<b>.46</b> (.61)	<b>.81</b> (.18)	<b>.76</b> (.40)	<b>.65</b> (.46)	<b>.59</b> (.57)	<b>.53</b> (.75)	<b>.51</b> (.99)
70s males	<b>.91</b> (.17)	<b>.80</b> (.18)	<b>.68</b> (.23)	<b>.59</b> (.23)	<b>.49</b> (.25)	<b>.41</b> (.25)	<b>.86</b> (.23)	<b>.77</b> (.29)	<b>.66</b> (.31)	<b>.56</b> (.33)	<b>.47</b> (.31)	<b>.40</b> (.29)
<b>TOTAL</b>	<b>.95</b> (.49)	<b>.88</b> (.63)	<b>.78</b> (.65)	<b>.67</b> (.60)	<b>.60</b> (.61)	<b>.55</b> (.91)	<b>.85</b> (.18)	<b>.74</b> (.29)	<b>.64</b> (.34)	<b>.56</b> (.38)	<b>.50</b> (.44)	<b>.46</b> (.57)

Note. Mean estimates equal to 1.0 indicate perfect prediction, values of less than 1.0 indicate underestimates, and those greater than 1.0 are overestimates.



ratios as dependent measures (the mean scores associated with this analysis are identical to those that appear in Table 3).<sup>5</sup> The MANOVA revealed significant main effects for Rate,  $F(1, 183) = 85.65, p < .01, MSE = 1.82$ , Amount,  $F(1, 183) = 33.26, p < .01, MSE = .49$ , and Years,  $F(5, 915) = 4.78, p < .01, MSE = .40$ ; three significant two-way interactions: Gender by Rate,  $F(1, 183) = 4.36, p < .05, MSE = 1.82$ , Gender by Years,  $F(5, 915) = 2.86, p < .05, MSE = .40$ , and Rate by Years,  $F(5, 915) = 44.85, p < .01, MSE = .18$ ; and one significant three-way interaction: Gender by Rate by Years,  $F(5, 915) = 3.33, p < .01, MSE = .18$ . Group means and standard errors from this analysis are presented in Figure 1, plotted by percentage rate, over the 6, five-year increments, as a function of gender.<sup>6</sup>

Inasmuch as main effects and two-way interactions cannot be unambiguously interpreted in the company of higher-order interactions, the following discussion will focus solely at the level of the three-way interaction. The significant gender by rate by years interaction can be seen by contrasting the 6% and 11% rate of return conditions (i.e., panels a and b in Fig. 1). In the 6% condition, males were reasonably accurate at judging the value of the account, whereas females tended to overestimate the account value, particularly at the longer terms (e.g., 25 and 30 years). However, in the 11% rate of re-

turn condition the magnitude of the gender difference diminished, and subjects of both sexes were found to substantially underestimate the value of the account. This tendency toward underestimation was greatest at the longer investment periods (e.g., 20, 25, and 30 years), where, on average, subjects' estimates represented only half the amount of the actual value of the account. This interaction clearly shows that subjects underestimate rapidly increasing growth functions – as was the case in the 11% APR condition – and accurately estimate (or overestimate, in the case of females) growth functions which increase at a more moderate pace (i.e., the 6% condition).

The gender-based performance differences identified above should be interpreted in light of the fact that females in the present sample had completed significantly fewer years of education than males ( $p < .05$ ). In order to determine whether performance on the task was associated with educational level, Pearson correlations were computed between subjects' estimate ratios and the number of years of education they had completed. Education was found to be significantly correlated with 15 of the 24 estimate ratio distributions. Each of the significant correlations between education and the estimate ratios were positive, with  $r$  values ranging from .15 to .40 (average  $r = .27$ ). In order to control for influences attributable to prior educational experiences, subjects' education scores were regressed upon each of the estimate ratio distributions, and the resulting 24 sets of residual scores were used as the dependent measures in a five-factor MANOVA which was structurally identical to the MANOVA analysis described above.<sup>7</sup> This education-controlled analysis revealed a main effect of Gender,  $F(1, 183) = 5.37, p < .05, MSE = 7.33$ ; two significant two-way interactions: Gender by Rate,  $F(1, 183) = 4.46, p < .05, MSE = 1.82$ , and Gender by Years,  $F(5, 915) = 5.78, p < .01, MSE = .38$ ; and one

<sup>5</sup> One outlier, a female in the 50-year-old group, generated estimates which were *substantially* larger than those made by any other subject (roughly 14 standard deviations above the mean in each of the four account conditions at the 30-year term). The inclusion of her data in the MANOVA would have contributed an inordinately large amount of error variance to the model. Therefore, the decision was made to exclude this subject from this analysis, and the other decision quality analyses which follow. However, the computational strategy which led this woman to produce such astronomical estimates is considered as a special case in the Strategies for Making the Growth Estimates section.

<sup>6</sup> It is important to note that the higher-order effects related to the age group and amount factors were not found to be statistically significant. On the basis of this finding, and in order to facilitate the interpretation of the data, the group means which appear in Figure 1 have been collapsed over the four age groups and over the two initial investment amounts.

<sup>7</sup> A multivariate analysis of residuals was employed in this situation rather than a multivariate analysis of covariance. The more traditional MANCOVA could not be computed due to limited degrees of freedom once the covariate was added to the equation.

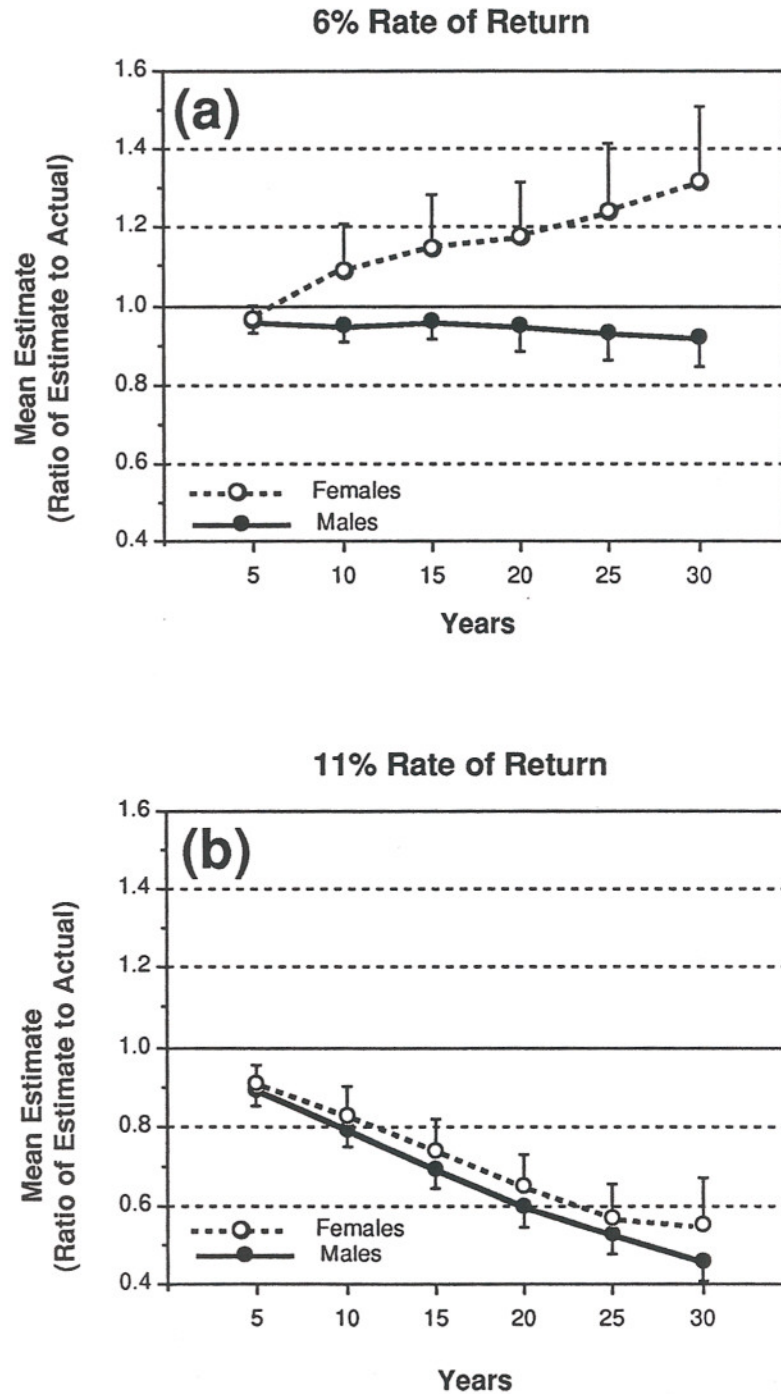


Fig. 1. Panels a and b show mean estimate ratios (and standard error bars), collapsed over the age group and amount factors, and presented as a function of gender across the 6, five-year time intervals. Panel a is based on a 6% rate of return and panel b is based on an 11% annual percentage rate. Mean estimates equal to 1.0 indicate perfect prediction of the growth function, those less than 1.0 indicate underestimates, and those greater than 1.0 are overestimates.

significant three-way interaction: Gender by Rate by Years,  $F(5, 915) = 4.24, p < .01, MSE = .18$ . The gender by rate by years interaction identified in this analysis was the same three-way interaction identified using the unadjusted estimate ratios reported above.

There are two important findings which emerge from the above individual difference analyses. The first is that education is significantly positively correlated with subjects' performance on the task. Subjects with advanced educational backgrounds tend to make *significantly larger* future value estimates than subjects who have completed fewer years of education. The second important finding is that although educational experience significantly correlates with task performance, the influence of education alone does not contribute enough variability to subjects' estimates to modify the basic pattern of age and gender differences established in the MANOVA using the unadjusted estimate ratios.

The above analyses indicate clear gender differences and an absence of age differences in subjects' ability to accurately estimate the long-term growth of four savings accounts. It was also shown that subjects' estimates were systematically related to two of the three within-subject factors: rate and years. However, the above analyses fail to indicate *how accurate* subjects' judgments were in terms of the entire set of estimates.

In order to assess the overall quality of subjects' performance on the task, a single value was derived which represented the cumulative quality of subjects' estimates. This value was arrived at by taking the difference between subjects' 24 estimate ratios and 1.0 (to arrive at an "error-based" difference score for each individual estimate), and then calculating the mean of the difference scores. The resulting overall mean error scores (in percentage points) represents the mean deviation from the set of correct estimates across all 24 trials. Therefore, if a person's error score is .25, then on average, the individual underestimated the value of the account by 25% over the entire set of estimates. The value in a single performance marker of this type is that it accurately represents the overall quality of sub-

jects' estimates through the use of an aggregated error score.

A 4 (Age Group) x 2 (Gender) ANOVA was computed using subjects' mean error scores as the dependent measure. Neither of the two main effects in this analysis were found to be statistically significant: Age Group,  $F(3, 183) < 1, MSE = .33$ ; Gender,  $F(1, 183) = 2.49, ns, MSE = .33$ . Furthermore, the age group by gender interaction also failed to obtain significance,  $F(3, 183) < 1, MSE = .33$ . However, it could be argued that the failure to identify age and gender differences in the quality of subjects' estimates could be attributable to individual differences in the number of years of formal education they had completed (recall that males had reported completing more years of formal education than females).

A Pearson correlation between subjects' educational level and their mean error scores revealed that the two variables were significantly negatively correlated,  $r(189) = -.22, p < .01$ , indicating that higher levels of education were associated with lower error scores. Accordingly, a 4 (Age Group) x 2 (Sex) analysis of covariance was conducted using subjects' educational level as the covariate, and their error score as the dependent measure. As anticipated, the covariate was shown to be significantly related to the dependent measure,  $F(1, 182) = 9.36, p < .01, MSE = .31$ . The more interesting finding, however, was that a main effect of gender emerged,  $F(1, 182) = 5.75, p < .05, MSE = .31$ , once the variance associated with subjects' educational level had been removed from the equation. Again, the main effect of age group and the age group by gender interaction failed to obtain significance:  $F(3, 182) < 1, MSE = .31$ ; and  $F(3, 182) < 1, MSE = .31$ , respectively. An inspection of mean scores revealed that across the four age groups, females' errors ( $M = .20, SE = .078$ ) were, on average, significantly larger than those made by males ( $M = .07, SE = .028$ ). However, the direction of the errors was consistent across both groups; that is, both males and females displayed an overall tendency to underestimate the account values.

The above individual difference analyses revealed two important findings. The first was that

subjects' prior educational level was associated with the overall quality of their estimates, as evidenced by the significant negative correlation between self-reported education and estimation errors. The second important finding was that once the effects of prior educational level had been covaried out of the equation, females were found to produce poorer quality estimates than males.

One final set of analyses was focused on the quality of subjects' estimates. The data were examined to determine which subjects *consistently* made the best estimates. At issue was whether the composition of this "best performers" group was overly represented by individuals of a particular age group or sex. The best performers were defined as those individuals whose estimates were within  $\pm 10\%$  of the actual value for at least half of the 24 estimates.<sup>8</sup> This criteria yielded a group of 16 individuals who comprised roughly the top performing 8% of the sample. The composition of this group revealed a distinctive pattern of age and sex differences. Of the 16 individuals, 13 (81%) were males and only 3 (19%) were females. Furthermore, 7 individuals (44%) were 20-year-olds, four (25%) were 40-year-olds, three (19%) were 50-year-olds, and two (12%) were 70-year-olds. Among the best performers, 6 individuals (38%) were found to be members of the 20-year-old male group.

The curve-fit and estimation quality analyses reported above indicate that some subjects were more accurate than others at making the financial growth estimates. What those analyses failed to reveal, however, were insights regarding the types of strategies subjects used to arrive at their estimates.

<sup>8</sup> This  $\pm 10\%$  criteria was arrived at based on a careful inspection of the raw data. Other more stringent or less stringent criteria could have been adopted to define the set of best performers (i.e.,  $\pm 5\%$  and  $\pm 15\%$  cut offs were considered); however, it was believed that these alternative criteria would have created groups of subjects which were either too small or too large to identify meaningful age or gender-based trends.

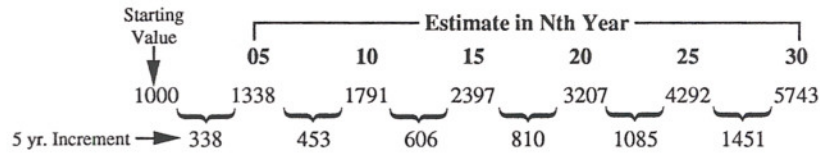
### Strategies for Making the Growth Estimates

In light of the systematic nature of the deviations found in the quality of subjects' estimates, it would have been of value to have collected information about the strategies they used to generate their estimates. Regrettably, reports of subjects' strategies were not systematically collected at the time of testing.<sup>9</sup> However, a number of participants (roughly 30 subjects) described the process they used to calculate their future value estimates during the debriefing. These anecdotal reports, in combination with a search for repeating patterns of values in the raw-score database, allowed for the identification of a variety of computational strategies. Before those strategies are introduced, however, for comparative purposes the correct algorithm for computing compound interest is briefly described.

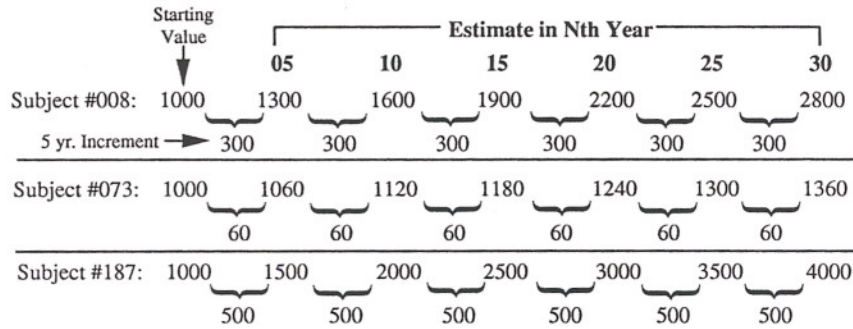
The algorithm used to compute compound interest is not conceptually complex; in practice, however, this type of computation requires an immense amount of cognitive effort. The algorithm specifies that the principle in an account be multiplied by the annual interest rate to determine the first year's interest earnings. Interest earnings are then added to the existing principle and the same process (principle, multiplied by interest rate, and earned interest added to existing principle) is repeated each successive year of the investment term. An example of how a \$1,000 investment accrues compound interest at an annual rate of 6% is shown in the first block of Figure 2. This investment will yield \$60 at the end of the first year. By the fifth year \$338 in interest will have accrued, and the account balance will have grown to \$1,338. By the 30th year the value of the account will have increased nearly fivefold to \$5,743. The miracle of compounding is neatly illustrated in this case by the observation that the amount of the *annual* interest earnings increases by more than 500% over the course of the 30-year term, from \$60 in year 1, to \$325 in year 30.

<sup>9</sup> At the time the study was initially conceptualized, the primary research goal was limited to determining whether there were age and sex differences in the quality of future value estimates.

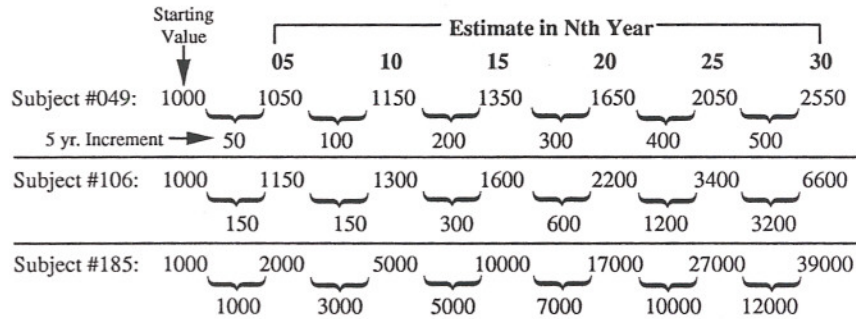
**Values Produced by the Correct Algorithm**



**Add a Constant Amount Strategy**



**Ratchet up the Initial Amount Strategy**



**Multiplicative Strategy**

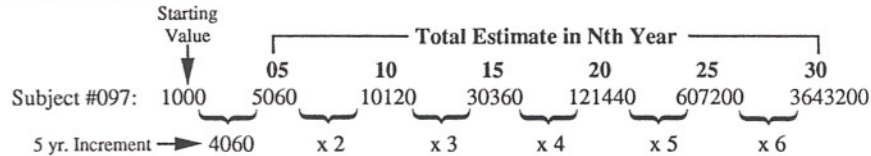


Fig. 2. Actual values for the \$1,000/6% account based on the correct algorithm (block one), and three different strategies subjects used to make the future value estimates: (a) the add a constant amount strategy (block two), (b) the ratchet up the initial amount strategy (block three), and (c) the multiplicative strategy (block four). Each series of estimates shown in blocks 2-4 is based on the \$1,000/6% account condition. The main row of values are subjects' actual estimates, and below those estimates are five-year increment values (i.e., the increase in the total account value over the previous estimate).

One of the more frequently used computational strategies involved a significant deviation from the correct algorithm described above. A number of subjects reported calculating the interest earnings for the first year (say, \$60 in the \$1,000/6% account condition), and then multiplying that value by 5 to derive the fifth year value of the account ( $\$60 \times 5 = \$300$ ). This strategy leads to a total fifth year estimate (principal and interest) of \$1,300. Estimates for each successive five-year term are extrapolated in a similar fashion, based on the amount of the interest earned in the first year of each term. It is easy to see how this *first year extrapolation strategy* could lead to underestimates of the account values, in that this approach fails to account for interest earned on previously earned interest. This strategy results in a \$38 (9%) estimate shortfall in the fifth year, and a \$916 (16%) estimate shortfall in the 30th year. This same strategy produces much larger underestimates in the \$2,500/11% condition, where the 30-year estimate is \$34K dollars – over \$22K less than the \$57K actual account value! An inspection of the data revealed that 27% of subjects used the first year extrapolation strategy, based on the observation that 51 of the 192 individuals tested recorded a fifth year estimate of \$1,300 for the \$1,000/6% account. These same subjects were found to have estimated similarly extrapolated values for the other experimental conditions.

Other more computationally savvy subjects reported using a *modified first year extrapolation strategy*. This involved generating an initial five-year interest amount according to the algorithm for the first year extrapolation strategy, but then “bumping up the total of the interest earnings just a bit” to account for the incremental benefit of earning interest on the previously earned interest. This strategy is *potentially* more accurate than the standard first year extrapolation strategy if the subject’s estimate of the incremental interest earnings is accurate. A low “guess” in this situation would lead to an underestimate of the actual account value, and a high guess would result in an overestimate. Unfortunately, it was difficult to accurately assess how many subjects used the modified first year

extrapolation strategy, because the amount of the interest earnings adjustment (i.e., the additional bump) varied from one subject to the next, and from one five-year estimate to the next for any given subject. However, anecdotal reports suggest that this was one of the more favored strategies; roughly half of the subjects who reported how they generated their estimates claimed to have used this approach.

Further examination of the data revealed two other commonly used strategies, both of which could lead to systematic errors in estimating future values. The *add a constant amount strategy* involves estimating the interest earnings for the initial five-year period (say, \$300, under \$1,000/6% conditions), and then adding that same amount of interest earnings to the account (i.e., \$300) for each additional five-year period. Twenty-one percent of subjects used this strategy. Three different examples of the add a constant amount strategy are shown in block two of Figure 2. Of the 40 subjects who added a constant amount to the \$1,000/6% account each five-year period, the most commonly used constant value was \$300 (see subject #008). This resulted in a \$2,800 30-year estimate, a 51% underestimate of the actual account value. Subject #073 (also shown) estimated \$60 as the constant interest earnings value, which led to a 76% underestimate of the actual 30-year account value. Subject #187 (also shown) estimated a \$500 yield over the first five-year period (a \$162 overestimate), and then continued to increment the account that same amount for each additional five-year period. The 30-year estimate of \$4,000 for this series (a 30% estimate shortfall) illustrates that even in cases where the initial estimate is on the high side, the add a constant amount strategy can lead to a substantial underestimate of the long-term yield.

One other commonly used strategy was identified. The *ratchet up the initial amount strategy* is based on deriving an initial interest amount for the first five-year period, and then adjusting (or ratcheting) that value upward for each additional five-year period. Consider the third block of estimates in Figure 2. Subject #049 added \$50 in interest earnings to the account for the first five-year period, \$100 for the second five years,

then \$200, \$300, \$400, and finally \$500 in interest earnings for the last five-year period. In this case the ratcheting strategy led to a \$3,193 (44%) underestimate of the \$5,743 actual account value. In contrast, the ratcheting strategy employed by subject #106 led to an \$857 (15%) overestimate of the actual account value. Apparently, this subject incorrectly assumed that the amount of interest in the account doubled every five years. He increased the starting value by \$150 in the first five-year period and then by another \$150 in the second five-year period. But then he doubled the interest earnings to \$300 in the third period, doubled them again to \$600 in the fourth period, doubled them a third time to \$1,200 in the fifth period, and then added \$3,200 to the account in the final five-year period. (Perhaps this individual more than doubled the amount for the last five-year period in an effort to take into account the increased earnings associated with late term growth.) Subject #185 also used the ratchet up the initial amount strategy by estimating \$1,000 in interest earnings for the first five-year period, and then incrementing the interest earnings \$2,000 (and in one case \$3,000) for each additional five-year period. In all, just over 6% of subjects were found to have used a version of the ratcheting strategy to generate their estimates.

By far, the most unusual strategy identified was the one used by subject #097. This was the individual who was excluded from the estimation quality analyses on the basis of having made extremely large overestimates of the account values (see footnote 5). Subject #097, whose estimates for the \$1,000/6% account are shown in the fourth block in Figure 2, employed what might best be described as a *multiplicative strategy*. After arriving at an interest earnings estimate of \$4,060 for the first five-year period, the subject multiplied the total value of the account (\$5,060) by 2 to derive the 10-year estimate (\$10,120). She then multiplied that figure by 3 to arrive at the 15-year estimate, 4 to arrive at the 20-year estimate, 5 to arrive at the 25-year estimate, and 6 to arrive at the 30-year estimate. This strategy led to an estimated total yield of approximately \$3.6 million dollars – over 600 times the actual value of the account! This case

illustrates how a flawed mental model could lead to the development of overly optimistic expectations about the potential for earnings. It is important to note, however, that this was the only subject who generated estimates on this order of magnitude.

Despite the fact that a large number of subjects made poor estimates and employed flawed computational strategies, a small subset of individuals, the 16 best performers identified in the previous section of the paper, were identified as having made reasonably good estimates. It is presumed that these individuals could not have consistently generated such high quality estimates had they not been using a close approximation of the correct algorithm.

The data were examined to determine whether subjects' strategy choices were systematically related to age or gender. Of the 51 individuals (27% of the sample) who used the first year extrapolation strategy, roughly one-quarter of the subjects fell into each of the four age groups (22%, 29%, 22%, and 27%, for 20-through 70-year-olds, respectively); and the strategy was equally employed by males (53%) and females (47%). Among the 40 subjects (21% of the sample) who used the add a constant amount strategy, there was no appreciable difference in the percentage of males (45%) and females (55%); however, it appeared that slightly more subjects in the two younger age groups used this method than individuals in the two older groups (25%, 35%, 20%, and 17%, for 20- through 70-year-olds, respectively). Of the 6% of subjects who used the ratchet up the initial amount strategy, the majority (75%) were females, and approximately equal numbers of individuals fell into the three older groups (8%, 33%, 33%, and 25%, for 20- through 70-year-olds, respectively). However, differences in the age and gender patterns associated with the ratcheting strategy need to be interpreted with caution, based on the relatively small number of individuals who employed this approach. Finally, subjects in the best performer group, who were believed to have used either the correct strategy or a close approximation of it, were overrepresented by males and 20-year-olds.

Identification of the various computational

strategies provides some clues as to why the estimation quality data revealed systematic underestimates of account values. In particular, two of the more frequently used strategies were associated with estimate shortfalls: the first year extrapolation strategy and the add a constant amount strategy. Admittedly, the conclusions which can be drawn regarding age and gender differences in strategy usage are limited, due to the fact that a specific strategy could not be identified for each subject across all four growth curve conditions. However, the variety of strategies identified in this preliminary investigation suggest that additional, more controlled studies of strategy selection are warranted.

## DISCUSSION

The overall pattern of results reported above are consistent with a growing body of literature which suggests that people have difficulties in conceptualizing nonlinear growth functions. Under conditions where the exponential change rate was large – in the two 11% APR conditions – subjects were found to reliably underestimate the value of the savings accounts, particularly at the longer investment terms (e.g., 25 and 30 years). Conversely, in cases where the growth rate was more moderately paced – in the two 6% conditions – males were found to accurately gauge the value of the accounts, whereas females overestimated account values. Thus, gender differences, but not age differences, were identified in the quality of subjects' estimates, and the accuracy of those estimates was found to be related to the slope of the exponential function subjects were asked to estimate. Furthermore, both age and gender differences were identified with respect to the shape of the curves subjects generated. Younger subjects were more likely to make exponential estimates of the growth curves than older subjects, and males were more likely to make exponential estimates than females.

### Gender and Estimation Performance

The pattern of gender differences reported above is, for the most part, consistent with the findings

of the Shaklee (1990) study which showed that males are more accurate than females when making long-term financial growth estimates. Shaklee (1990) summed up the general pattern of gender-based findings as follows, "Males [were found to] have a pretty good 'passbook mentality,' with fair accuracy at the lower interest rates which might be linked to a [passbook] savings account. However, their errors at higher interest rates are considerable. Females' judgments show substantial underestimation of investment yield at *all* [italics added] interest levels included in the study" (p. 3). In the present experiment, males were indeed found to have a "pretty good passbook mentality" at the 6% interest rate, and display much larger estimation errors in the 11% condition. Females also made substantial underestimates in the two 11% APR conditions, but unlike females in the Shaklee study, in the present study females overestimated account values in the two 6% conditions.

It is not clear why the performance of males was found to be superior to that of females. One possible reason is that males' mental models (Gentner & Stevens, 1983) of finance and investing are stronger than those of females, based on differential opportunities to control shared family finances. Among married subjects in the sample, males were more than twice as likely as females to report sole control over the family finances, which suggests that they had greater opportunities to learn about investing and financial growth functions. A number of recent studies have shown that one's knowledge of a domain can be a strong determinant of judgment and decision making performance (Hershey, 1990; Hershey, Walsh, Read, & Chulef, 1990; McCloskey, 1983; McCloskey, Caramazza, & Green, 1980). One might expect an experiential benefit in favor of males to be particularly magnified among subjects in the older age groups. This is based on the fact that historically, males were more likely to have been the primary family wage earner, and thus, more likely to have had opportunities to learn about finance and investing.

A second possible reason for the observed gender effects could be that the males and females who participated in this study possessed



different levels of mathematical skills, which presumably would have stemmed from differential gender-related educational experiences (Fennema, 1983; Fennema & Peterson, 1985; Sherman, 1982) or job opportunities. Small but consistent gender differences in mathematical abilities have been well documented in the psychological literature (Halpern, 1992; Maccoby & Jacklin, 1974); however, the reasons for those differences are not well understood, and have been the subject of a good deal of spirited debate (cf. Baker & Jones, 1992; Benbow & Stanley, 1980). Regardless of the reasons for the observed gender differences in mathematical abilities, it is quite likely that one's competence at basic arithmetic plays a role in the ability to make accurate future value projections given the computational demands inherent in the task.

A third possible explanation for the observed gender differences in estimation performance involves positing that the males and females in this study had differential levels of task-specific motivation. A number of studies have shown that men are more motivated than women when solving mathematical problems, and more confident in their overall mathematical abilities in general (see Fennema & Peterson, 1985; and Licht & Dweck, 1983 for a discussion of these points). It has been suggested that both of these effects (i.e., motivation and confidence) stem from a combination of differential gender-based educational experiences and gender-related societal expectations regarding mathematical proficiency. Given the cognitively demanding nature of the estimation task, it is not inconceivable that low levels of motivation could have resulted in computational errors, which in turn could have resulted in overall decrements in performance.

Finally, it is not unreasonable to assume that all three of the gender-specific explanations outlined above – experiential differences, educational and career differences, and motivational differences – could have collectively contributed to the observed gender difference in estimation performance.

#### **Age and Estimation Performance**

A recurring finding across the various decision

quality analyses was that age was not systematically related to the quality of subjects' estimates (one notable exception was that a large majority of the best performers were members of the two younger age groups). The consistent lack of age effects in the quality of subjects' estimates is an intriguing finding in light of the fact that age was found to be systematically related to the *shape* of the growth curves individuals generated. Specifically, only about half (57%) of the subjects in the oldest age group made exponential growth estimates, whereas the remaining 43% of subjects made linear estimates. In contrast, just over three-quarters of subjects in the youngest age group (76%) made exponential estimates, whereas only 24% estimated linear trends. The seemingly contradictory combination of a lack of age effects on the decision quality analyses, and significant age effects regarding the shape of the growth curve analyses, suggests that researchers should examine both types of data when attempting to characterize group-based trends in estimation performance. In the majority of previous estimation studies, investigators have focused solely on the quality of subjects' estimates. In the few studies where the shape of growth curves has been considered, the focus was on how the *average* trajectory of a curve for an entire sample compared with the trajectory of an actual future value function. Unfortunately, the averaging procedures involved in this type of analysis may obscure interesting patterns of subgroup differences, such as the age by gender differences which were identified in this experiment using individual growth curve analyses.

It is unclear why there was a lack of age effects with respect to the *quality* of subjects' estimation performance. One possible explanation for the absence of age-related findings is that age is not systematically related to the way individuals process task information when making future value estimates. This explanation is unlikely, however, based on the assumption that various basic processing abilities (e.g., working memory, attentional resources) underlie the quality of subjects' performance on the estimation task, and the fact that these same basic abilities have been found to exhibit *normative* age-

related declines (Craik & Salthouse, 1992; Salthouse, 1982, 1991).

A more plausible explanation for the lack of age differences in the quality of performance involves positing a *trade-off* between knowledge of the task and basic processing abilities for individuals of different ages. Although the overall quality of the performance for young and old subjects was found to be equivalent, the nature of the cognitive processes which led to those solutions might very well have differed as a function of age. Younger subjects' self-reports indicated that they had significantly less investing experience than their older counterparts ( $p < .05$ ), and based on strong theory, it was assumed that young subjects possessed superior fluid processing resources. Similarly, older subjects, who possessed significantly greater prior financial and investing experience than younger subjects, were in an age range where it was safe to assume that they had experienced a deterioration in fluid abilities (although data were not available to explicitly test this assumption). Therefore, it is conceivable that the young relied more heavily on their fluid processing abilities to generate estimates, whereas the old relied more heavily on their knowledge of the domain. The theoretical notion that task- or domain-specific knowledge can compensate for declines in basic processing abilities has been suggested by a number of cognitive aging psychologists (Baltes & Baltes, 1990; Charness, 1982; Walsh & Hershey, 1993; Welford, 1993). Unfortunately, one limitation of this study is that markers of fluid abilities were not collected, based on a pragmatic decision made during the planning stages of the experiment. To bring nearly 200 subjects into the laboratory to assess levels of fluid abilities would have required a substantial commitment of time and resources, and it was unclear at the outset of the study whether such an effort would have been worthwhile given that there had not yet been an initial study of age differences in estimation performance. Perhaps future studies could be designed to explore the relationship between age, fluid abilities, and estimation performance.

In order to further examine whether the absence of age differences in estimation perfor-

mance stems from a trade-off between knowledge and processing abilities, one might ideally ask subjects to provide a detailed description of their processing strategies at the time of testing. If the trade-off hypothesis is indeed correct, then one would expect to find proportionally more older subjects using accurate estimation strategies than younger individuals. In the present study, no overwhelming age differences were found in the strategies used by members of the different age groups. However, this may have been due to the fact that subjects' strategies were inferred on the basis of a post hoc analysis of the raw data, rather than through retrospective or concurrent verbal reports. It is unknown whether meaningful age differences in strategy use would have been found had detailed self-report data been collected. It also would have been interesting to know whether there were subjects in the sample who were familiar with the correct computational algorithm, but who "satisfied" (Simon, 1955) in generating their estimates, by opting to use a sub-optimal strategy (e.g., the add a constant amount strategy) which was less cognitively demanding than the correct algorithm. Future studies designed to examine subjects' computational strategies could help further our understanding of the underestimation bias by determining how individual differences in domain-specific knowledge mediate performance on this type of task.

Experimental research on cognitive biases appears to have merit beyond the often cited value of identifying situations in which people are prone to making errors in reasoning. Important theoretical advances will be made as we begin to understand why individual difference variables, such as age and gender, are systematically related to complex, real-world cognitive performance. Critical to this line of inquiry will be the ability to identify differences in task-specific information processing strategies, as well as differences in core intellectual abilities, such as fluid intelligence, which have previously been found to be systematically related to age, and in rare circumstances, gender. The theoretical argument advanced in this paper suggests that the interaction between knowledge-based strategies and basic processing abilities will ulti-

mately be shown to account for a substantial portion of the variability found in high-level cognitive performance.

## REFERENCES

- Baker, D. P., & Jones, D. P. (1992). Opportunity and performance: A sociological explanation for gender differences in academic mathematics. In J. Wrigley (Ed.), *Education and gender equity* (pp. 193-203). Washington D.C.: Falmer Press.
- Baltes, P. B. (1987). Theoretical propositions of lifespan developmental psychology: On the dynamics between growth and decline. *Developmental Psychology*, 23, 611-626.
- Baltes, P. B., & Baltes, M. M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In P. B. Baltes & M. M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 1-34). New York: Cambridge University Press.
- Bates, J. M., & Gabor, A. (1986). Price perception in creeping inflation: Report on an inquiry. *Journal of Economic Psychology*, 7, 291-314.
- Beach, L. R., & Lipshitz, R. (1993). Why classical decision theory is an inappropriate standard for evaluating and aiding most human decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsombok (Eds.), *Decision making in action: Models and methods* (pp. 21-35). Norwood, NJ: Ablex.
- Benbow, C. P., & Stanley, J. C. (1980). Sex differences in mathematical ability: Fact or artifact? *Science*, 210, 1262-1264.
- Charness, N. (1982). Problem solving and aging: Evidence from semantically rich domains. *Canadian Journal on Aging*, 1, 21-28.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Erlbaum.
- Craik, F. I. M., Morris, R. G., & Gick, M. L. (1990). Adult age differences in working memory. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short-term memory* (pp. 247-267). New York: Cambridge University Press.
- Craik, F. I. M., & Salthouse, T. A. (1992). *Handbook of aging and cognition*. Hillsdale, NJ: Erlbaum.
- Cutler, N. E., Gregg, D. W., & Lawton, M. P. (1992). *Aging, money, and life satisfaction*. New York: Springer.
- Dörner, D. (1983). Heuristics and cognition in complex systems. In R. Groner, M. Groner, & W. F. Bishof (Eds.), *Methods of heuristics* (pp. 89-107). Hillsdale, NJ: Erlbaum.
- Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences*. Hillsdale, NJ: Erlbaum.
- Fennema, E. (1983). Success in mathematics. In M. Marland (Ed.), *Sex differentiation and schooling* (pp. 163-180). London: Heinemann Educational Books.
- Fennema, E., & Peterson, P. (1985). Autonomous learning behavior: A possible explanation of gender-related differences in mathematics. In L. C. Wilkinson and C. B. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 17-35). New York: Academic Press.
- Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422-444). New York: Cambridge University Press.
- Fox, J. (1984). Formal and knowledge-based methods in decision technology. *Acta Psychologica*, 56, 303-331.
- Geary, D. C., Frensch, P. A., & Wiley, J. G. (1993). Simple and complex mental subtraction: Strategy choice and speed-of-processing differences in younger and older adults. *Psychology and Aging*, 8, 242-256.
- Geary, D. C., & Wiley, J. G. (1991). Cognitive addition: Strategy choice and speed-of-processing differences in young and elderly adults. *Psychology and Aging*, 6, 474-483.
- Gentner, D., & Stevens, A. L. (1983). *Mental models*. Hillsdale, NJ: Erlbaum.
- Gregg, D. W. (1990). Introduction to "financial gerontology" column [Guest editorial]. *Journal of the American Society of CLU & ChFC*, November, XLIV (6).
- Halpern, D. F. (1992). *Sex differences in cognitive abilities* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Hershey, D. A. (1990). The role of knowledge and experience in structuring problem solving performance (Doctoral dissertation, University of Southern California). *Dissertation Abstracts International*, 51, Sec. 7B.
- Hershey, D. A., Walsh, D. A., Read, S. J., & Chulef, A. S. (1990). The effects of expertise on financial problem solving: Evidence for goal directed problem solving scripts. *Organizational Behavior and Human Decision Processes*, 46, 77-101.
- Hogarth, R. M. (1987). *Judgement and choice* (2nd ed.). New York: Wiley and Sons.
- Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, 75, 242-259.
- Horn, J. L. (1982). The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In F. I. M. Craik & S. E. Trehub (Eds.), *Aging and cognitive processes* (pp. 237-278). Boston: Plenum.
- Horn, J. L., Donaldson, G., & Engstrom, R. (1981). Apprehension, memory, and fluid intelligence decline in adulthood. *Research on Aging*, 3, 33-84.
- Horn, J. L. & Hoffer, S. (1992). Major abilities and development in the adult period. In R. J. Sternberg

- & C. A. Berg (Eds.), *Intellectual development* (pp. 44-99). Cambridge, England: Cambridge University Press.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgement under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press.
- Kaplan, M. F., & Schwartz, S. (1977). *Human judgement and decision processes in applied settings*. New York: Academic Press.
- Kemp, S. (1984). Perception of changes in the cost of living. *Journal of Economic Psychology*, 5, 313-323.
- Kemp, S. (1987). Estimation of past prices. *Journal of Economic Psychology*, 8, 181-189.
- Licht, B. G., & Dweck, C. S. (1983). Sex differences in achievement orientations: consequences for academic choices and attainments. In M. Marland (Ed.), *Sex differentiation and schooling* (pp. 72-97). London: Heinemann Educational Books.
- Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.
- McCloskey, M. (1983). Naive theories of motion. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 299-324). Hillsdale, NJ: Erlbaum.
- McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. *Science*, 210, 1139-1141.
- McDowd, J. M., & Birren, J. E. (1990). Aging and attentional processes. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed.; pp. 222-233). New York: Academic Press.
- Plude, D. J., & Hoyer, W. J. (1985). Attention and performance: Identifying and localizing age deficits. In N. Charness (Ed.), *Aging and human performance* (pp. 47-99). New York: Wiley and Sons.
- Raphael, B. (1976). *The thinking computer: Mind inside matter*. San Francisco: Freeman and Co.
- Russo, J. E., & Shoemaker, P. J. H. (1992). Managing overconfidence. *Sloan Management Review, Winter Ed.*, 7-17.
- Salthouse, T. A. (1982). *Adult cognition: An experimental psychology of human aging*. New York: Springer-Verlag.
- Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.
- Schaie, K. W. (1983). The Seattle longitudinal study: A 21-year exploration of psychometric intelligence in adulthood. In K. W. Schaie (Ed.), *Longitudinal studies of adult psychological development* (pp. 64-135). New York: Guilford Press.
- Shaklee, H. (1990, November). *Investment decision making: Short-term interest rates and long-term yield*. Paper presented at the meeting of the Psychonomic Society, New Orleans.
- Sherman, J. A. (1982). Mathematics the critical filter: A look at some residues. *Psychology of Women Quarterly*, 6, 428-444.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69, 99-118.
- Thomas, J. C. (1989). Problem solving by human-machine interaction. In K. J. Gilhooly (Ed.), *Human and machine problem solving* (pp. 317-362). New York: Plenum.
- Timmers, H., & Wagenaar, W. A. (1977). Inverse statistics and misperception of exponential growth. *Perception & Psychophysics*, 21, 558-562.
- Vye, N. J., Delclos, V. R., Burns, M. S., & Bransford, J. D. (1988). Teaching thinking and problem solving: Illustrations and issues. In R. J. Sternberg & E. E. Smith (Eds.), *The psychology of human thought* (pp. 337-365). New York: Cambridge University Press.
- Wagenaar, W. A. (1982). Misperception of exponential growth and the psychological magnitude of numbers. In B. Wegener (Ed.), *Social attitudes and psycho-physical measurement* (pp. 283-301). Hillsdale, NJ: Erlbaum.
- Wagenaar, W. A., & Sagaria, S. D. (1975). Misperception of exponential growth. *Perception & Psychophysics*, 18, 416-422.
- Wagenaar, W. A., & Timmers, H. (1978). Extrapolation of exponential time series is not enhanced by having more data points. *Perception & Psychophysics*, 24, 182-184.
- Wagenaar, W. A., & Timmers, H. (1979). The pond-and-duckweed problem: Three experiments on the misperception of exponential growth. *Acta Psychologica*, 43, 239-251.
- Walsh, D. A., & Hershey, D. A. (1993). Mental models and the maintenance of complex problem-solving skills in old age. In J. Cerella, J. Rybash, W. Hoyer, & M. L. Commons (Eds.), *Adult information processing: Limits on loss* (pp. 553-584). New York: Academic Press.
- Welford, A. T. (1993). The gerontological balance sheet. In J. Cerella, J. Rybash, W. Hoyer, & M. L. Commons (Eds.), *Adult information processing: Limits on loss* (pp. 3-10). New York: Academic Press.