Estimating Risk Tolerance: The Degree of Accuracy and the Paramorphic Representations of the Estimate

Michael J. Roszkowski[1] **and John Grable**[2]

Using a sample of 386 financial advisors and 458 of their clients, the study sought to determine how effective financial advisors and clients are at estimating risk-tolerance, and to test how well items from a risk tolerance test and demographic information can represent the judgmental process used to formulate these estimates (a "paramorphic representation" of the decision). The client's self rating and the advisor's rating of the client produced a Pearson correlation of .40. Moreover, the advisor's rating correlated at about the same level (r =.41) with the client's score on a test of risk tolerance. The data also showed that when it comes to estimating one's own risk tolerance, clients are better than are advisors at this task. The estimates could be represented paramorphically in terms of a few variables. It was observed that advisors assign too much diagnostic value to certain demographic variables in estimating client risk tolerance.

Keywords: *Risk tolerance, paramorphic representation, financial advisors*

Introduction

Surprisingly little has been written about the effectiveness of financial advisors in arriving at critical judgments germane to their profession. A review of the literature shows only a few studies devoted to examining the accuracy of financial services professionals in decision-making (e.g., Slovic, 1969; Torngren & Montgomery , 2004; Tyszka & Zielonka, 2002; Zielonka, 2002). Rather, attempts to analyze the effectiveness of judgments made by financial advisors have tended to focus primarily on critiques of financial planning and the investment management models used by financial professionals (e.g., Kautt, 2002). Without such information, however, financial advisors have no way of knowing if, as a profession of experts, they are better, worse, or the same as others when forming judgments. As the financial services profession continues to grow it will be increasingly important to determine if the holistic decision-making processes used in the profession actually perform according to expectations.

In particular, very little is known about a specific, but extremely important, judgment all financial advisors need to make in the early stages of their work with clients, namely, estimating a client's level of financial risk tolerance. Accurately assessing a client's level of financial risk tolerance is an important task within the financial planning process because a person's level of risk tolerance impacts on a diverse number of financial decisions, such as portfolio management, type of mortgage, insurance deductibles, emergency fund savings, estate planning, and even divorce mediation (Bottom, Holloway , McClurg , & Miller , 2000; Callan & Johnson, 2002; Cicchetti & Dubin,1994; Dreze, 1981; Finke & Huston, 2003; Hallahan, Faff, & McKenzie, 2004; Hanna & Chen, 1997; Harris, 2004; Moreschi, 2004).

The purpose of this study is multifaceted. The first aim is to determine how effective financial advisors and clients are at estimating risk-tolerance. The second purpose is to see if one can represent the judgmental process through multiple regression models using items from a risk tolerance-test and demographic characteristics (i.e., a " paramorphic representation" of the decision). Specifically, the following research questions were used to guide this study:

a) How well does a financial advisor's estimate of a client's risk tolerance correlate with the client's own estimate of his/her own risk tolerance?
b) How well do financial advisors estimate the risk tolerance of their clients as measured by a valid test?
c) How well do clients and advisors estimate their own risk tolerances as measured by a test?
d) Are advisors any better than clients at estimating their own level of risk tolerance relative to what the risk-tolerance test indicates?
e) Using questions from a risk-tolerance test, to what extent can both an advisor's and a client's judgmental process in estimating risk tolerance be represented paramorphically?

[1] *Michael J. Roszkowski, Ph.D., Peale House, 2nd Floor, Office of Institutional Research, La Salle University, 1900 West Olney Ave., Philadelphia, PA 19141-1199, (215) 951-1428, (215) 951-1783 fax, roszkows@lasalle.edu*

[2] *John Grable, Ph.D., CFP®, 318 Justin Hall, Institute of Personal Financial Planning, College of Human Ecology, Kansas State University, Manhattan, KS 66506, (785) 532-1486, (785) 532-5505 fax, jgrable@ksu.edu*

Review of the Literature

How Should Risk Tolerance be Measured?

Although a number of authors have called for the application of formal procedures and tests to the financial risk tolerance assessment process, there is no consensus on how to best conduct it (Bouchey, 2004; Callan & Johnson, 2003; Grable & Lytton, 1999a, 1999b, 2001, 2003; Hanna & Chen, 1997; Hanna & Gutter, 1998; Hanna, Gutter, & Fan, 2001; Hanna & Lindamood, 2004; Roszkowski, 1992; Roszkowski, Davey, & Grable, 2005; Yook & Evverett, 2003). Techniques for measuring risk tolerance have been devised by economists, psychologists, and decision scientists, but as Grable and Joo (2000) observed, the recommended procedures differ, depending in part on the academic or professional background of the assessor.

The formal assessment of risk tolerance can take many forms. The commonly-used techniques have been classified in a variety of ways (see Callan & Johnson, 2003; Hallahan, Faff, & McKenzie, 2004; Hanna & Chen, 1977; Hanna, Gutter & Fan, 1998; MacCrimmon, Wehrung, & Stanbury, 1986; Roszkowski, 1992). At the broadest grouping, one can differentiate between actual behavior and performance on tests, simulations, and questionnaires of various sorts. At a more detailed level, Hanna et al. (1998, p. 53) note that there are at least four methods: "asking about investment choices, asking a combination of investment measures and subjective questions, assessing actual behavior, and asking hypothetical questions with carefully specified scenarios." Slicing the pie into even thinner slices, Roszkowski (1992) lists the following approaches to gauging financial risk tolerance: proxy measures (such as demographic characteristics, investment objectives, and returns expected from investments), preferences for different investment vehicles, reactions to sample portfolios, life-style characteristics, self-classification; self-ratings of more specific aspects of risk-taking, and probability and payoff preferences.

The type of questions posed in expected utility theory–based questionnaires would be classified by Roszkowski (1992) as the "probability and payoff preferences" approach. Other names used in the literature to identify the probability and payoff approach are "gambles" and "prospects." Besides Utility Theory (von Neumann & Morgenstern, 2005), a number of different theories have been proposed on the basis of trade-offs to explain human behavior under risk, including Subjective Expected Utility Theory (Savage 1954), Rank Dependent Utility Theory (Quiggin, 1982), Cumulative Prospect Theory (Tversky & Kahneman 1992), and Reference-Dependent Subjective Expected Utility Theory (Sugden, 2003).

Each approach has its proponents and detractors. Academics trained in economics generally favor approaches based on expected utility theory and its variants (e.g., Hanna, Gutter & Fan, 1998; Hanna & Lindawood, 2004), whereas psychologists and other professionals with a behavioral science bent are willing to also include attitudinal items in the test, provided that these questions can be shown to be valid. Thus, Callan and Johnson (2003) maintain that a variety of "attitudes," such as spoken and unspoken beliefs, regarding financial risk tolerance need to be considered in the assessment, while Hanna et al. (1998, p. 54) are extremely skeptical about any attitudinal questions that "…are not rigorously linked to the concept of risk tolerance in economic theory."

Roszkowski et al. (2005) are also critical of many of today's risk tolerance questionnaires, but for different reasons than Hanna and his colleagues (1998). They contend that many questionnaires billed as financial risk tolerance tests ask questions that, while relevant for giving sound financial advice, are not really part of the psychological construct of risk tolerance per se (e.g., investment time horizon, financial capacity to absorb a loss, etc.). However, they would accept any question type, even ones not rooted in expected utility theory, as a basis for a sound assessment provided that such questions can stand up to commonly accepted psychometric standards. Also, they believe that questionnaires in use today are generally too short to be valid for assessing individual clients.

There is a growing and persuasive body of evidence to suggest that risk tolerance is more than just cognitive in nature and that feelings need to be considered in understanding people's reactions to risk (Magnan & Hinsz, 2005). Loewenstein, Weber, Hsee, and Welch (2001) review such evidence and propose the "Risk-as-Feelings" theory, summarizing the rationale for their position as follows in the abstract of their article:

> Virtually all current theories of choice under risk or uncertainty are cognitive and consequentialist. They assume that people assess the desirability and likelihood of possible outcomes of choice alternatives and integrate this information through some type of expectation-based calculus to arrive at a decision. The authors propose an alternative theoretical perspective, the risk-as-feelings hypothesis, that highlights the role of affect experienced at the moment of decision making. Drawing on research from clinical, physiological, and other subfields of psychology, they show that emotional reactions to risky situations often diverge from cognitive assessments of those risks. When

> such divergence occurs, emotional reactions often drive behavior. The risk-as-feelings hypothesis is shown to explain a wide range of phenomena that have resisted interpretation in cognitive-consequentialist terms. (p. 267).

In the body of the article (p. 271), they elaborate on this position as follows:

> … people's emotional reactions to risks depend on a variety of factors that influence cognitive evaluations of risk only weakly or not at all. These include the vividness with which consequences can be imagined, personal exposure to or experience with outcomes, and past history of conditioning. Cognitive assessments of risk, on the other hand, tend to depend on more objective features of the risky situation, such as probabilities of outcomes and assessments of outcome severity. Even when feelings about risk are influenced by these objective features, the functional form of such dependence is different. For example, it has been demonstrated that feelings about risk are largely insensitive to changes in probability, whereas cognitive evaluations do take probability into account. As a result, feelings about risk and cognitive risk perceptions often diverge, sometimes strikingly.

Roszkowski (1992), who is of the opinion that no approach works perfectly with each and every client, identifies the advantages and shortcomings of methods currently used by advisors. He concludes that it is perhaps most prudent to "diversify" and use a variety of methods:

> In collecting the information on risk tolerance, you can best understand a client by diversifying the approaches used and comparing the impressions of the client that emerge from one approach with the impressions from another approach. If all indicators point to the same conclusion, the job of assessment is easy. Quite frequently, however, you will obtain discrepant images of the client. Attention should be paid not only to the client's answer on each type of question , but also to the potential reasons why a client may be inconsistent in his her answers from one approach to another. You should discuss with the client why he or she answered a given question a certain way, because the client's stated rationale can provide valuable insights into which type of measurement approach may be the best indicator of the client's level of risk tolerance. Probe and clarify until you are satisfied that you have identified the causes for the discrepancies (p. 10).

Roszkowski (1992) recommends that in the absence of any information regarding which technique is best for a particular client, averaging the answers from different techniques should prove to be the most valid approach because "(s)ome approaches may overestimate the true level of risk tolerance whereas others may underestimate it. By averaging the results, you may be able to cancel out these two errors and arrive at the most accurate impression possible, given the circumstances" (p.10).

Can Personality in General and Risk Tolerance in Particular Be Judged without a Test?

The body of literature devoted to better understanding the determinants of a person's risk tolerance is expansive, but there is very little evidence available to document how well people in general and financial services professionals in particular actually estimate someone else's, or even their own, level of risk tolerance (Hsee & Weber, 1997). When advisors work with clients they need to estimate two aspects of risk tolerance. The one estimate requires advisors to determine how the client perceives himself or herself with respect to propensity for risk. The second and probably more critical appraisal involves classifying the client into a true level of risk tolerance. Since risk tolerance is a personality characteristic, albeit one that may be somewhat elastic (see Grable, Lytton, & O'Neill, 2004; Magnan & Hinsz, 2005; Yao, Hanna, & Lindamood, 2004; Yip, 2000), some guidance can be gleaned from the literature that compares people's estimates of personality characteristics relative to actual scores on standardized tests.

Most of the studies dealing with self-knowledge of one's own personality have been concerned with the operations that people use to understand themselves rather than the accuracy of their self-perceptions (Vogt & Colvin, 2005). The research conducted by Furnham and his colleagues is an exception to this statement. In a provocatively-titled article, "Can people accurately estimate their own personality test scores?" Furnham (1990) suggests that the answer depends on the particular personality characteristic. His results with undergraduate students showed significant positive correlations between the students' estimated and their actual scores on 10 of the 15 personality dimensions he studied. In addition, the undergraduates in Furnham's (1990) research were able to estimate other students' scores on eight of these 15 personality characteristics, but as one might suspect, these approximations were not as accurate as the ones of their own scores on these tests. Chamorro-Premuzic, Furnham, and Moutafi (2004) concluded that certain characteristics are easier to estimate than others. Correlations between one's estimated and one's actual test scores ranged from a low of r =.27 for Agreeableness to a high of r =.58 for Conscientiousness. In a related study, Furnham and Chamorro-Premuzic (2004) determined that people are best at estimating their own degree of depression (r = .58), anxiety (r =.54), hostility (r =.52), assertiveness (r =.51), activity (r =.51), and need for achievement (r =.45). Among the least predictable personality characteristics were impulsivity (r =.06), straight forwardness (r =.12), vulnerability (r =.16), and excitement seeking (r =.26). Although risk tolerance was not one of the characteristics under study, it is noteworthy that constructs related to it (e.g., impulsivity, excitement seeking) were not self-estimated very well. It may well be that risk tolerance is a characteristic that is difficult to gauge, but very few studies have addressed either the lay public's or professionals' ability to estimate risk tolerance in themselves or others.

Estimating Risk Tolerance

Studies by Borkenau and Liebler (1993a; 1993b) indicate that accuracy in judging another person's personality characteristics is a function of variables such as (1) meaning systems shared by observers, (2) the amount of information on the target behavior, and (3) the consistency of the target behavior. Dealing specifically with financial risk tolerance, Hsee and Weber (1997) propose four possible mechanisms that people may use to estimate the risk tolerance of others:

a) *Same as Me*. Assume that others have the same level of risk tolerance as the judge possess.
b) *Risk-as-Value*. Perceive others as less risk tolerant than oneself because risk taking is admired in our society (see Clark, Crockett, & Archer [1971] for a discussion of this concept) and people generally view themselves as more likely than others to have desirable characteristics (see Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg [1995]).
c) *Risk-as-Feelings*. Predict others to have similar levels of risk tolerance to the judge, but more risk neutral than the judge.
d) *Stereotype*. Estimate risk-taking proclivity on the basis of the target's membership in a group with known risk preferences (e.g., the person is a female and females are more risk averse, therefore this individual is risk averse).

To determine which of the four mechanisms is most likely, Hsee and Weber (1997) conducted three related studies in which people had to predict the risk preferences of others under both gain and loss scenarios using a series of questions involving a choice between a fixed risky option and a sure thing option. An example of a gain scenario used by them is the choice between A and B where A = receive $1, 000 for sure, and B = flip a coin; receive $2,000 if heads, or $0 if tails. In the loss questions, the phrasing was identical, but the word " pay" appeared instead of the word "receive." Under both loss and gain scenarios, the individuals participating in this research project had to indicate what (a) choice they would make themselves, and (b) what another (abstract) person would do.

The prospects were presented in a story line, such as the following gain prospect: "Suppose that you bought a lottery ticket a week ago. You are now informed that you have won and have been given two options of how to receive the money." An example of a loss frame is: "Suppose that you violated a traffic rule and hurt somebody a week ago. You are now informed that you will be fined and have been given two options of how to pay the fine." In the other person version of the question, "somebody somewhere in the U.S" appeared in place of "you." in the storylines. On the basis of these questions a risk preference (RP) score was computed that ranged from 1 (selected all sure options) to 8 (chose all risky options), with the scores from 1 to 4 representing risk aversion and scores from 5 to 8 representing risk-seeking. The risk neutrality point was defined as 4.5.

The results of their first study, involving 99 undergraduates, showed that people believed that others would be more risk tolerant than they were themselves. The overestimates of their fellow students' propensity to take risks occurred under both the positive outcomes and negative outcomes. The results were clearly counter to the Risk-as-Value hypothesis. The RP was markedly below 4.5 in the "self" conditions, suggesting that participants were highly risk averse, whereas under the "others" conditions, the RP indices were much closer to 4.5. In other words, the predicted risk preferences of others fell between the participants' own risk preference and risk neutrality, which was consistent with the Risk-as-Feelings hypothesis proposed by Hsee and Weber (1997). These results, however, could not be used to refute the Stereotype explanation for ascribing risk tolerance to others because in the "others" condition the terms "American" and "U.S." were used. According to Hsee and Weber "these terms may have evoked in participants the stereotypical image of Americans--adventurous and courageous--and consequently led them to perceive others (i.e., other Americans) to be more risk seeking than themselves" (pp. 47-48).

In their second study, Hsee and Weber (1997) administered the same risk tolerance measure to another sample of undergraduates from the same college (*n* = 159), but this time they asked them to predict the decisions that would be made by three distinct types of "others": (1) other students in the

United States, (2) other students on campus, and (3) the individual who happened to be sitting next to them at the time of the study (lunch). The researchers were interested in seeing whether the Stereotype hypothesis would be supported, in which case they thought the self-others difference should be greater in the "others-in-U.S." condition than in the "others-on-campus" condition. The results did not conform to the Stereotype hypothesis because the self-other discrepancy was not any different between others in the U.S. and others on campus scenarios. However, one can question the assumption that the one condition should necessarily elicit more stereotyping than the other. After all, most of the other students on campus were probably American, even tough the term itself was not used explicitly in the question.

Hsee and Weber (1997) were also interested in determining whether the Risk-as-Feelings explanation would hold in the second study, which they tested by checking if the self-others difference was larger in the two abstract conditions (i.e., others-in-U.S. and the others-on-campus) relative to the concrete condition (next-person). The results did support the Risk-as-Feeling hypothesis since the participants indicated that the two types of abstract others would be more risk seeking than the study participants were themselves. Under the concrete others (next person) condition, however, the difference between self and others was not significant. As in the first study, the RP index was closer to the risk-neutrality point (4.5) under the two "others" conditions than under the self condition.

Hsee and Weber's (1997) third study involved a sample consisting of 141 students at a different university than the prior two studies. It replicated the previous two designs in most respects, but it differed from them in several respects: (1) an incentive for being correct (cash prize) was provided , (2) the number of questions was increased so that the RP index could range from 1 (most risk averse) to 10 (most risk seeking) with the risk-neutrality point now being 5.5 instead of 4.5, and (3) the person variable was manipulated between subjects rather than within subjects (i.e., participants were randomly assigned to provide a choice under only one of the three conditions: self, other students on campus, or the person sitting nearby). The group making the choice for themselves produced RP scores similar to the group that had to predict what the person sitting nearby would do. Moreover, the scores of both these groups were lower, on average, than the ones elicited from the group that rated the fellow students on campus (in the abstract). In other words, the self-others discrepancy occurred only when participants made predictions for abstract others and disappeared once the other individual was a concrete person that could be seen, but remained a stranger.

Commenting on their findings across the three studies, Hsee and Weber (1997) concluded their results did not support the explanation that people would see others as being the same as themselves on risk tolerance nor did the data conform with the Risk-as-Value explanation since people predicted others to be more risk seeking rather than less risk seeking. The Stereotype explanation was rejected by Hsee and Weber because it failed to account for why people predicted abstract others to be more risk seeking than they were themselves. However, as noted earlier, both targets were American, and the stereotyping may have continued even when the word American did not appear in the phrasing of the risk tolerance measure.

According to their interpretation of the data, Risk-as-Feelings was the explanation most consistently supported by their data. They explain their conclusion in the following terms:

> We have suggested that people's risk preferences depend on their feelings toward risk. When people make a prediction of another person's risk preference, they base their prediction partly on their own feelings and partly on risk neutrality, which reflects lack of particular feelings. How much people base their prediction on their own feelings depends on how vivid the target person is. If the target is vivid, people can empathize with the target, perceive the target to have feelings similar to their own, and consequently predict the target to make the same choices as themselves. If the target is abstract, people are emotionally more distant from the target, would have greater difficulty imagining how the target feels about risk, and, consequently, would resort more to risk neutrality for making the prediction. (pp. 52).

Hsee and Weber (1997) cite the research by Alicke et al. (1995) that also discovered that once a target becomes concrete, the person perceives a smaller difference between the target and herself or himself. However, in contrast to Hsee and Weber, Alicke et al. reported that people generally consider themselves to have more of the valued characteristic. But the characteristics examined by Alicke et al. did not include risk tolerance, so it may be that the direction of the difference varies by personality characteristic, or perhaps that risk-taking is not as valued in our culture as may be assumed.

Although they seem to dismiss the role of stereotyping in the formation of an opinion of a stranger's risk tolerance, in their discussion Hsee and Weber (1997) indirectly acknowledge that it could be a factor. They contend that when forming an impression of a stranger, one draws on two sources of information, which they call personal and distributional, respectively. Personal information is about the particular individual whereas distributional information relates to the class to which that person belongs, including the probability information that a certain characteristic is possessed or not possessed in that class. The final decision is based on both sources, but the relative weight assigned to each source depends on whether the target is abstract or concrete. If it is abstract, then the distributive information is afforded greater weight, whereas with concrete targets, the reverse happens. If one accepts the "kernel of truth" definition of stereotypes (i.e., that stereotypes are based, in part, on true group differences), then stereotypes are essentially overgeneralizations. Locksley, Borgida, Brekke, and Hepburn (1980) reported that when the target is abstract, generalizations (stereotypes) are drawn, but more of the individualized information is used once the target becomes more concrete and familiar. Essentially, this is what Hsee and Weber (1997) are proposing.

Research by Eckel and Grossman (2002) also addressed the public's accuracy in the prediction of the risk tolerance of a generalized, abstract "other" person or a stranger. Their study is of particular interest because, unlike Hsee and Weber (1997), these authors concluded stereotyping does play a significant role in undergraduates' estimates of the risk tolerance of strangers assessed on the basis of mainly visual clues. The students participating in that study had to select to play one of five gambles varying systematically in expected return and variance. In addition, each of the 200 participants was asked to guess which of the five gambles every other participant would pick. (Since they were able to keep the money that was won, the task was quite realistic.) The 200 guesses for each participant were averaged and the mean was correlated to the person's actual gamble. The Pearson correlation between actual choice and average guess was .42. Stereotyping on the basis of sex was shown to be a component in the guess of the other person's preferred gamble. Research by Martin (1987) and Siegrist, Cvetkovich, and Gutscher (2002) also identified a gender-based stereotyping effect in undergraduate's assessments of the risk tolerance of strangers.

The studies reviewed so far dealt with risk tolerance attributions assigned to strangers. A recent study by Bateman and Munro (2005) examined husbands' and wives' ability to predict their partner's selections in lotteries involving binary choices differing in risk. Among the 76 couples, the one partner was able to predict the other partner's choice correctly in 65% of the cases. Bateman and Munro's (p. C185) comment on this finding suggests that is less remarkable than one might first assume: "This is significantly better than 50–50; it is also better than the success rate if they supposed (as a benchmark example) that their partner was a risk neutral income pooler. However, if individuals predict according to how they themselves choose and preferences are not correlated within couples then the predicted success rate is 64.7% - which is not statistically significantly different from the actual value." In other words, the ability to predict risk tolerance may not be very good, even among individuals who live together, and suggests that perhaps people do use themselves as the baseline for the risk tolerance estimate of others.

The systematic underestimation bias of one's risk tolerance relative to abstract others, as identified by Hsee and Weber (1997), was also observed by Hallahan, Faff, and McKenzie (2004), who compared the relationship of self-estimated financial risk tolerance to the risk tolerance assessed by means of a 25-item psychometrically-based test, using a primarily Australian sample consisting of the clients of financial planners (n = 20,415). The self – assessment question was phrased as follows: "*This questionnaire is scored on a scale of 0 to 100. In practice, however, the scores range from around 20 to around 80, with the average being 50. When the scores are graphed they follow the familiar bell-shaped curve of the Normal Distribution (diagram provided). About two-thirds of all scores are within 10 points of the average. What do you think your score will be?*" Their results indicated that generally people tend to underestimate rather than overestimate their own risk tolerance. Only 4% of the sample correctly estimated their own financial risk tolerance, whereas 73% underestimated it and 23% overestimated it. On average, people underestimated their score by about five points. It would have been instructive to also see the correlation between self estimated and the tested risk tolerance, but this statistic was not reported in the article. These results of the study by Hallahan, et al. (2004) are inconsistent with the Risk-as-Value hypothesis, which would suggest that people see themselves as more risk tolerant.

These findings document people's underestimation of their own risk tolerance relative to abstract others, the role of stereotyping in the formulation of the lay person's estimates, couple's ineptness at predicting each other's risk tolerance, and people's inability to even judge their own ability to handle financial risk. However, they are less helpful in understanding whether professional financial advisors are effective at estimating a client's level of risk tolerance. A study by

Snelbecker, Roszkowski, and Cutler (1990) bears some indirect relevance to this issue. Financial planners were presented with various hypothetical client statements thought to convey sentiments about risk tolerance, and these advisors were then asked to assign a numerical value to each statement indicating the degree of risk tolerance or aversion that the statement conveyed to them. While there was considerable inter-judge consistency in the interpretation of some statements, enough variability existed in others to suggest that without a test of risk tolerance, measuring the construct with just advisor opinion posed the danger of applying a rubber yardstick to the process.

The one study directly focusing on the accuracy of financial advisors' estimates of their clients' risk tolerance was conducted in Australia during the development of the FinaMetrica risk profiling system. In this report, organizational psychologists (Elsayed & Martin, 1998) determined that the correlation between risk tolerance as estimated by experienced Australian financial planners and their clients' actual scores on the standardized test of risk tolerance was .38. It may be surprising to some readers that the degree of the accuracy exhibited by the professionals was no greater than the one shown by Eckel and Grossman's (2002) undergraduates, but the research on expert's judgments suggests that perhaps it should not be unexpected.

Experts' Reliance on Holistic Judgments

Yieh and Chen (2003) believe that financial planners often use simple "rules of thumb" to conduct their risk tolerance assessments, relying heavily on demographic factors such as wealth, income, sex, marital status, and age. However, the relationship of even relevant demographic variables with financial risk tolerance is far from perfect as demonstrated by research conducted in the United States (see Chaulk, Johnson, & Bulcroft, 2003; Grable & Lytton, 1998, 1999b; Hanna & Lindamood, 2004; Sung & Hanna, 1996; Xiao, Alhabeeb, Hong, & Haynes, 2001) as well as in other countries, such as Japan (Li, 2005), Taiwan (Yieh & Chen, 2003) and Australia (Hallahan, et al., 2004). Given the small to moderate magnitude of the correlations between financial risk tolerance and certain demographic characteristics, even valid generalizations do not necessarily apply to a specific client. As Li observed: "In some cases an individual fits all of the characteristics for high level of risk tolerance but may absolutely avoid all risks, or on the other hand, other individuals will be in a situation where most people would avoid risk but will engage in risky activities."

It is also unclear how financial advisors combine various pieces of information together to arrive at their judgments of client risk tolerance when they do not use a test, but since the predominant form of decision-making employed by experts is holistic (Ruscio, 2003), it is quite likely that financial advisors rely on this process as well. Ruscio defined holistic judgments as ones formed by evaluating all relevant factors together into a complex whole instead of an independent consideration of factors separately. In effect, a decision maker uses clues from the environment – behavioral, psychosocial, demographic, etc. – as inputs into a subjective process that is based on experience, knowledge, and temperament to arrive at a judgment. Holistic judgments are based on the "notion that everything else influences everything else" (Ruscio, p.2), and that a professional can use his or her knowledge, experience, and mental capacity to integrate sometimes conflicting data into superior decisions. Someone using a holistic approach would argue that he or she is better able to assess a situation than someone who relies on a more formalized decision-making process. Proponents of the holistic approach assume that an expert – based on experience, knowledge, and temperament – is better equipped to "accommodate a wider range of relevant information and integrate it in more sophisticated ways" (Ruscio, p. 2) than someone who relies on a mechanical process to form a judgment.

Over the years, researchers have conducted extensive analyses of how experts – those who have gained their knowledge from a combination of formal education/training and experience – in diverse professions reach holistic decisions and the accuracy of these judgments (see Dawes, Faust, & Meehl, 1989; Grove & Meehl, , 1996). For example, data exists on the effectiveness of physicians (Bornstein & Emler, 2001), medical pathologists (Einhorn, 1974), mental health clinicians (Ruscio, 2003), weather forecasters (Stewart, Roebber, & Bosart, 1997; Tyszka & Zielonka, 2002), mechanics (Fischhoff, Slovic, & Lichtenstein, 1978), venture capitalists (Zacharakis & Meyer, 2000), and auditors (Ettenson, Shanteau, & Krogstad, 1987). In general, the published research suggests that specialists are not especially adept at formulating holistic judgments based on their experience (Camerer & Johnson, 1997). It has been observed that generally experts tend to be neither reliable nor accurate when making holistic judgments (Dawes, 1971; Dawes, Faust, & Meehl; Grove &. Meehl, 1996 ; Ruscio, 2003; Shanteau, 1999), although a review by Shanteau (1995) indicates that some professionals (e.g., auditors) are more accurate than others (e.g., clinical psychologists). Garb (1989) found that while an expert's training positively impacts the accuracy of a diagnosis, experience almost never does.

Paramorphic Representation of Holistic Judgment

While holistic decisions are subjective and non-mechanical, it may be nonetheless possible to model the decision-making process that a decision-maker uses to arrive at the estimate of a client's risk tolerance through the application of stepwise multiple regression, a process that has come to known as "paramorphic representation" (Doherty & Brehmer, 1997; Roszkowski, Spreat, & Isett, 1983). The notion of paramorphic representation was first introduced by Hoffman (1960). Borrowing the term from the study of mineralogy, he argued that a mathematical formula could represent the judgment process in the same way a structured chemical makeup can describe a mineral. Hoffman claimed that a regression formula "helps to account for or explain what is observed concerning certain properties or characteristics of the judge, just as the chemical formula explains many, though not all, properties or characteristics of the substance" (p. 124).

Einhorn, Kleinmuntz, and Kleinmuntz (1979) criticized reliance on statistical models to describe judgments because such models do not allow for an understanding of the actual cognitive process used by experts to reach decisions, but as Doherty and Brehmer (1997) explained, a regression model need not "bear even a semblance of any actual state of affairs" (p. 541); instead, the regression model simply simulates the conclusion. In other words, a paramorphic representation does not necessarily replicate or imitate the actual thought process of someone making a judgment (Camerer & Johnson, 1997). Rather, the paramorphic representation reproduces the end product of an expert's decision-making process. It is useful to know to what extent a paramorphic model can represent advisor and client estimates of financial risk tolerance. A logical choice for potential input variables are risk tolerance test questions and demographic information.

The paramorphic representation technique relies on step-wise regression as a methodology. Readers need to be aware that the use of stepwise procedures has been questioned (see Thompson, 1995). The major concern is that the technique capitalizes on chance relationships in the data, and thus may produce results that are over-fitted and difficult to replicate. For instance, a Monte Carlo simulation by Derksen and Keselman (1992) found that 20% to 74% of the variables entering into stepwise multiple regression were noise. The data driven process inherent in stepwise procedures may not lead to the best set of predictors if the predictors are highly redundant (i.e., correlated). The variable that enters the equation on the first step in stepwise regression is the one in the set of predictors that has the highest simple correlation with the criterion. At each stage after the first one, order of entry is determined by which variable has the highest partial correlation with the dependent variable considering all variables already in the model. Only if a variable increases the *F*-value of the equation by some specified threshold value will it enter the model (called "the *F*-to-enter criterion"). A common misunderstanding is that order of entry shows the importance of the independent variables (Gordon, 2001). Because a number of steps are involved in stepwise regression, experiment-wise (at least one) Type I error rates can be rather high.

Some statisticians would argue that stepwise regressions are therefore never appropriate, but a more moderate position would allow for its use when sifting through large numbers of potential predictors (van Belle, 2002). It has been said that the objectives of the study should determine the method for selecting the predictors. As Armstrong (1971, p. 512) pointed out, "… the exploratory end of the continuum asks for as little input from the researcher as possible and the theory-based end asks for as much as possible." Osborne (2000) commented that "(c)urrent practice clearly favors analyst controlled entry, and discourages entry based on the statistical properties of the variables as it is atheoretical." (pp. 1-2) , but at the same time he acknowledged the value of atheoretical analyses in some circumstances when he wrote: "And while theory is useful for identifying what variables should be in a prediction equation, the variables do not necessarily need to make conceptual sense. If the single greatest predictor of future achievement scores was the number of hamburgers a student eats, it should be in the prediction equation regardless of whether it makes sense…" (p. 1). As noted earlier, for paramorphic representation purposes, it is unimportant whether the variables are causally related or for the model to even be realistic (Doherty & Brehmer, 1997).

Methodology

Sample

The sample consisted of 386 advisors and 458 of their clients, but the number of participants varied by the specific analysis due to missing data. The advisors, who were all graduates of The American College's Master's in Financial Services (MSFS) program, were asked to pick two of their clients and to administer a risk-tolerance questionnaire that was being developed by the college. In addition, they were asked to take the questionnaire themselves.

The clients were primarily males (83.2%), with an average age of 48.44 (SD=11.55) years. Approximately, 83.9% were married, 8.7 % single, 3.1% divorced, and 4.3% widowed. On average, they had 2.82 (SD=1.51) dependents. Their highest level of education was: 1.8% less than high school, 6.5% high school, 16.9% some college, 39.8% bachelors' degree, 19.6% masters' degree, and 15.5% doctorate or law degree. In terms of employment, 44.9% worked in the private sector, 41.5% were self-employed, 3.6% worked for the government, and 9.9% were retired.

The advisors were, on average, 51.67 (SD = 9.14) years old and primarily men (95.6%). With respect to marital status, the distribution was: 90.8% married, 7.3% divorced, 1.1% single, and 0.8% widowed. The mean number of dependents was 3.03 (SD=1.57). The advisors represented a variety of sectors of the financial services industry (see Table 1), although the majority were employed in the life insurance industry as career agents.

Table1
The Sector of the Financial Services Industry Represented by the Advisors

	Frequency	Percent	Valid Percent
Life/Health Insurance	246	63.7	69.5
Financial Planning	66	17.1	18.6
Securities	9	2.3	2.5
Law	4	1.0	1.1
Banking	1	0.3	0.3
Other	28	7.3	7.9
Subtotal	354	91.7	100.0
Missing	32	8.3	
Total	386	100.0	

Measures
The participants completed the developmental form of a risk-tolerance questionnaire called the Survey of Financial Risk Tolerance (SOFRT). This initial form consisted of 66 questions, three of which had sub-parts. The total number of items on the questionnaire was thus 93. On the basis of an item-analysis, only 51 items were retained for scoring. However, in one of the multi-part questions, six additional items were retained to provide a context for the portion of the question that was to be scored. The 51 items were used to calculate the risk-tolerance scores, but 57 items were used in the paramorphic regression analyses reported in this study. The Cronbach's alpha on the developmental sample was .91 (Roszkowski, 1992). Internal consistency estimates on other samples are available, and show alpha values to be between .81 and .86. A 45-day test-retest reliability equaled .81 to .83 (Roszkowski, Delaney, & Cordell, 2004).

The questions on the SOFRT are varied in nature, including: preferences for different investment vehicles, expected returns, reactions to sample portfolios, life style characteristics, probability and payoff preferences, preferences for guaranteed versus probable gambles, minimal required probability of success, and minimal return required to undertake a risky venture. The last question on the SOFRT requires the respondent to classify himself or herself into one of seven financial risk-tolerance categories: Extremely Low Risk Taker; Very Low Risk Taker; Low Risk Taker; Average Risk Taker; High Risk Taker; Very High Risk Taker; and Extremely High Risk Taker.

As part of the validation process, the advisors were requested to provide a rating of the client's risk tolerance using a ten-point scale ranging from 1 (very low risk taker) to 10 (very high risk taker) where only the endpoints were labeled with a verbal anchor. The ratings, available for 290 clients, averaged to 5.41 (SD =1.82). The advisors were instructed to provide a rating only if they felt they were in a position to do so. The ratings were made independent of the scores on the SOFRT.

Procedure
Simple Pearson correlations were used to answer the first three research questions. The remaining two questions were addressed using multiple regression procedures.[a]

Results

How well does a financial advisor's estimate of a client's risk tolerance correlate with the client's self rating of his/her own risk tolerance?
The client's self rating on the seven-point scale was correlated with the advisor's 10-point rating of the client. The two ratings correlated at r =.40 (n = 288, p <.001). Since it has been suggested that people rate others' risk tolerance based on how the see themselves, the advisors ratings of their two clients (on a 10 point-scale) were correlated with the rating that the advisor assigned to himself/herself (on a seven-point scale). Recall that the advisors rated two clients. The Pearson correlation between the advisors' self-ratings and the client's self ratings was .09 (n = 241, p =.176) for the first client and .06 (n = 232, p =.381) for the second client. The correlation between the advisors ratings of his/her two clients was also practically nil (r = -.05, n = 226, p = .445). In other words, no evidence existed for the notion that the advisors were predicting the clients to have similar risk preferences to themselves (i.e., the advisors).

How well do financial advisors estimate the risk tolerance of their clients as measured by a valid test?
The correlation between an advisor's rating and a client's risk tolerance as measured by the SOFRT was .41, based on 288 cases (p < .001). Since the SOFRT includes a self-classification question as the last item, the self-classification question was removed from the SOFRT total score in order to determine what the correlation would be under this circumstance. The correlation between planner rating of the client and the

client's risk-tolerance score without the self-rating also equaled .41 (p < .001). In other words, the advisor's estimate of client risk tolerance correlated with the risk-tolerance score to the same extent irrespective of whether the client's self-rating was retained or removed from the risk-tolerance test score.

This correlation is attenuated to an extent due to unreliability in the criterion (Fan, 2003). Based on the correction for attenuation formula, the correlation was adjusted for unreliability in the risk tolerance measure, using the lowest reported reliability estimate on the SOFRT (.81). Even under this circumstance, the correlation between tested risk tolerance and the advisor's estimate of it went up only slightly, to about .46.

How well do clients and advisors estimate their own risk tolerances as measured by a test, and are advisors any better than clients at estimating their own level of risk tolerance?

As noted earlier, both clients and advisors took the SOFRT. The last item on the SOFRT, the self-rating, was correlated with the composite of the remaining items on the risk-tolerance survey. For the clients (n = 446) the Pearson correlation was .77 (p < .001) whereas for advisors (n = 384), it was .63 (p < .001). The difference between the two correlations is statistically significant (z = -3.99, p < .001). The data suggest that when it comes to estimating one's own risk tolerance, clients are somewhat better than are advisors at this task.

Frankly, these results were puzzling. Additional steps were taken to understand the relationships better. Given the research by Borkenau and Lielber (1993a; 1993b), one possibility is that planners are using different benchmarks than clients. One must therefore consider the standards used to evaluate a given behavior when assessing accuracy. For instance, what one person considers "average" may in fact be "high" to another individual. According to some studies (e.g., Falk & Knell, 2004), people tend to use others who are similar to them as their benchmarks. In that case, advisors and clients with the same exact risk-tolerance levels may view themselves as being different on that characteristic simply because their reference groups differ. For instance, an advisor with a moderate level of risk tolerance may see himself or herself as low risk tolerant because the group being used as the benchmark. i.e., other advisors, will be greater in risk tolerance than the general public. Someone who is low risk tolerant when benchmarked against advisors could be average or even high risk tolerant when the norm group is the general public.

It was, therefore, of interest to determine if advisors and clients use the same standards to ascribe different levels to risk tolerance. After excluding the last item on the SOFRT, the sum score of the remaining items comprising the SOFRT was transformed into a standardized T-score (M = 50, SD = 10) basing the transformations on the combined sample of advisors and clients. Table 2 reports the risk-tolerance scores of the advisors and clients on this system at each level of self-classification. An ANOVA showed that the risk-tolerance score was a function of both role (planner versus client) and self-rating (seven levels). The main effect for role was significant, (F (1,816) = 10.90, p <.001), which indicates that, on average, advisors are more risk tolerant than their clients (53.16 versus 47.29). The main effect for self-rating was likewise significant (F(1,816) = 123.91, p <.001), which means that actual differences in risk tolerance exist by self classification (estimates).

Moreover, the interaction term between role and self-rating was significant (F (1,816) = 3.90, p <.001), which denotes that the extent of the difference in actual risk tolerances between planners and clients depended on which seven self-classifications one considered. An inspection of Table 2 reveals that the size differences are inversely related to level of risk tolerance. At the extremely low risk-tolerance self-classification, there is 13.74 point difference between advisors and clients (in favor of the advisors), whereas at the high risk taker classification the difference is minimal (0.83). In other words, the advisors who classified themselves as low risk takers were underestimating their true level of risk tolerance. Unfortunately, the sample sizes at the extremes were too small to allow for meaningful post-hoc analyses to test the differences for statistical significance.

Table 2
Standardized (*T*) Risk Tolerance Scores as a Function of Self-Rating and Role

Self-Rating	Advisor			Client		
	n	*M*	*SD*	*n*	*M*	*SD*
Extremely Low Risk Taker	3	43.26	17.68	16	29.52	4.95
Very Low Risk Taker	9	45.28	7.87	40	35.13	7.01
Low Risk Taker	55	44.08	6.46	117	41.41	5.76
Average Risk Taker	162	51.59	5.55	159	48.57	7.27
High Risk Taker	141	58.04	6.62	97	57.21	7.09
Very High Risk Taker	12	65.38	6.19	14	63.88	6.43
Extremely High Risk Taker	2	62.07	6.14	3	66.87	7.33
Total	384	53.16	8.24	446	47.29	10.58

In view of the small cell sizes at the extremes, the distribution was collapsed into three levels by combining categories and then re-analyzing the data. The first category consisted of the "extremely low," "very low," and "low" self-classifications. The second category was made up of the people who defined themselves as "average." In the third category were the people who self-estimated their risk tolerance to be either "extremely high," "very high," or "high." The consolidated risk-tolerance groups were labeled, respectively: Low, Medium, and High. The collapsed data, reported in Table 3, resulted in a significant main effect for role (F (2,824) =33.40, p <.001), self-rating (F (2,824) = 340.14, p <.001), and the interaction of the two independent variables (F (2,824) = 7.40, p < .001). A post hoc protected t-test revealed a significant difference between the risk tolerance of advisors and clients at the low (t (238) =5.18, p < 001) and the medium levels (t (295) = 4.18, p < .001) of self-estimated risk tolerance, but not at the high level (t (267) = 0.43, p =.666).

Table 3
Standardized (T) Risk Tolerance Scores as a Function of Collapsed Self-Rating and Role

Self rating Collapsed to Three Levels	Role	*n*	*M*	*SD*
Low	Advisor	67	44.21	7.16
	Client	173	38.86	7.17
Medium	Advisor	162	51.59	5.55
	Client	159	48.57	7.27
High	Advisor	155	58.66	6.85
	Client	114	58.28	7.43
Total	Advisor	384	53.16	8.24
	Client	446	47.29	10.58
		830	50.00	10.00

Using the questions from the risk tolerance test and demographic information, to what extent can an advisor's judgmental process in assigning a risk tolerance rating be represented paramorphically?

First, the 57 items on the client's SOFRT were each correlated with advisor opinion. The size of these correlations ranged from .10 to .39. Most items (67%) correlated with advisor rating on the order of .10 to .19. Approximately 28% of the correlations were in the .20 to .29 range, and only 5% (3 items) fell into the .30 to .39 range. Singly, the three items most predictive of advisor estimate of client risk tolerance were the client's self rating (.39), the client's self reported degree of risk taking in the past (.31), and the client's preferred investment portfolio from the sample ones shown in the question (.31). The 16 items with correlations of .20 and above appear in Table 4.

Complete information on all 57 predictors was available in 233 cases, which served as the sample for the multiple regression. If all 57 items on the SOFRT are entered into a multiple regression equation, then R^2 =.44 (adjusted R^2=.25). The standardized Beta weights showed that the two best items were the clients' self classification and a gamble involving a preference for $30,000 for sure versus a 50%-50% chance to earn $60,000 or nothing. However, based on a stepwise regression approach only 7 SOFRT items were necessary to predict advisor rating of the client's risk tolerance at an R^2 of .31. Because it had the largest correlation with the criterion, the first item to enter the step-wise model was the client's self rating, resulting in R^2 = .18. If the client's self rating was not permitted to be one of the predictors, it would be possible to achieve an R^2 of .28 with just 8 predictors.

It was also of interest to determine how well the advisors' rating of the client's risk tolerance could be represented on the basis of strictly demographic information. Table 5 reports the simple correlations between various demographic characteristics (marital status was dummy coded for each level) and advisors' rating of client risk tolerance. Demographic characteristics significantly related to advisor opinion about the client's risk tolerance were sex (r =.32, males more risk taking), individual income (r =.27), household income (r =.23), ownership of options/futures (r =.22), net wealth (r =.21), widowed status (r = -.18, widowed status associated with lower risk taking), ownership of investment real estate (r =.15), and investing in commodities/futures (r =.22). When the significant demographic variables were

Table 4
SOFRT Items Correlating .20 or Higher with Advisors' Rating of Client Risk Tolerance

Item	*r*
How do you rate your willingness to take investment risks relative to the general population?	0.39
What degree of risk have you assumed on your investments in the past?	0.31
Which of the following investment portfolios do you find most appealing?	0.31
Compared to other people you know, how would you rate your ability to tolerate risk?	0.29
You are faced with a choice between greater job security with a small pay increase, and a high pay rise but less job security. Which would you select?	0.29
Compared to other people you know, how much time do you spend reading about financial and investment matters?	0.29
Compared to other investors, how sophisticated are you about investing money?	0.28
Diversification is typically the soundest investment strategy. However, suppose an eccentric uncle left you an inheritance of $75,000, stipulating in his will that you invest all the money in only one of the following investments> Which one would you select?	0.27
If your friends were interviewed, how would they describe your evaluation of risks?	0.26
Which of the following comes closest to your ideal employment compensation structure involving some mix of salary and commissions?	0.26
Investments can go up and down in value. What is the maximum drop in the value of your total investment portfolio that you could tolerate before feeling uncomfortable?	0.26
What is your general outlook on the eventual outcome of your financial decisions after you make them?	0.26
What percent of your funds are you willing to place in investments that are of above-average risk?	0.24
What type of changes have you made in your investment portfolio in the past?	0.24
Do you consider yourself reflective or impulsive when making investment decisions?	0.23
An investment decision involves the possibility of making an amount of money as well as the possibility of losing all or some portion of the funds invested. Some people focus more on the possibility of making money, whereas others focus more on the possibility of losing money as a result of the decision. When making an important investment decision, what dominates your thinking?	0.22
Have you ever borrowed money in order to make an investment (other than a home-mortgage loan)?	0.21
Suppose you are the beneficiary of a $100,00 life insurance policy from a beloved relative. You are considering various investment possibilities, including some very high risk. How much of this money could you lose forever without feeling that you were betraying your relative's motive for leaving you this sum?	0.21
How do you react to unexpected bad financial news?	0.20

entered into the regression equation in a stepwise manner, only three predictors were needed to account for the advisor ratings. The multiple correlation-squared at each step was: Step 1: .09 with just sex; Step 2: .13 with sex and options/futures ownership (yes/no); and Step 3:.16 with sex, plus options/futures ownership, plus household income.

The next issue was the predictability of advisor's estimate of a client's risk tolerance if it is to be papramorphically represented by the SOFRT items and the significant demographic characteristics. With nine predictors, the model had an R^2 of .39. The predictors were seven SOFRT items and two demographic characteristics. The two demographic characteristics are whether the client does or does not own options/futures and the client's sex.

Using the questions from the risk-tolerance test, to what extent can both the advisor's and the client's judgmental process in estimating their own risk tolerance be represented paramorphically?

The advisor's self-rating of his/her own risk tolerance on the seven-point scale was regressed on his/her answers to the 56 remaining items from the SOFRT. Likewise, the client's self-rating was regressed on the remaining SOFRT items. Using all items, R^2 equaled .73 for clients (adjusted R^2 = .68) and .59 (adjusted R^2= .50) for advisors.

The two highest standardized Betas for advisors' self ratings were for the following questions:

- *What degree of risk have you assumed on your investments in the past? (Answer options: 1=very small, 2=small, 3=medium, 4=large, 5=very large)*
- *Compared to other people you know, how would you rate your ability to tolerate the stress associated with important financial matters? (Answer options: 1=very low, 2=low, 3=average, 4=high, 5=very high)*

Based on the standardized Beta values, the two most predictive items of clients' self ratings were:

- *What degree of risk have you assumed on your investments in the past? (Answer options: 1=very small, 2=small, 3=medium, 4=large, 5=very large)*
- *What percent of your funds are you willing to place in investments that are of above average risk? (Answer options:1=0 %, 2=1%-9%, 3=10%-19%, 4= 20%-29%, 5=30%-39%, 6=40%-49%, 7=50%-59%, 8=60%-69%, 9=70%-79%, 10=80-89%, 11=90%-99%, 12=100%)*

Table 5
Correlation of Advisor Ratings of Client Risk Tolerance with Client Demographic Characteristics

Characteristic	*n*	*r*	*p*
Marital Status			
Single	296	+.02	.784
Married	296	+.05	.379
Divorced	296	+.00	.951
Widowed	296	-.18	.003
Sex (male=0, female=1)	282	-.316	.000
Age	281	-.105	.079
Number of dependents	280	+.08	.213
Highest Education	280	+.04	.523
Years with Same Employer	163	-.06	.417
Individual Income	261	+.27	.000
Household Income	245	+.23	.000
Net Wealth	276	+.21	.000
Employment			
Private Sector	296	-.01	.930
Public Sector	296	-.01	.851
Self-employed	296	+.03	.592
Retired	296	-.07	.231
Investments			
Life Insurance	282	+.10	.091
Savings Account/CD	280	-.09	.141
Money Market Funds	279	-.11	.073
Bonds	276	-.04	.470
Stocks	280	+.05	.417
Real Estate	279	+.15	.015
Options/Futures	270	+.22	.000

It is notable that for both clients and advisors, history serves as a gauge of one's self impression.. Under a step-wise regression fewer than 56 items were necessary. For clients, 12 items were required to reach an R^2 of .70, and for advisors 10 items were needed to obtain an R^2 of .54.

Finally, the limited set of items used in predicting the advisor's rating of his/her own risk tolerance were assessed for how well they could predict the client's estimate of his/her own risk tolerance. Likewise, the items that were found to be predictive of the client's self estimated risk were used to predict the advisor's estimate of his/her risk tolerance. Using the variables predictive of the clients' self estimates, one could predict advisors' self-ratings at an R^2 of .46.

Conversely, using the variables predictive of advisor self-ratings resulted in a R^2 of .62 with clients. In both instances the R^2 was lower when the predictors from the other group were used, but not remarkably so (down from .70 to .62 for clients, and down from .54 to .46. for advisors). These results suggest that variables may be inter-exchangeable in the paramorphic representations to some extent due to their inter-correlations.

Discussion

A number of issues pertaining to the estimation of risk tolerance were addressed in this paper. First, an advisor's estimate of a client's financial risk tolerance and the client's own estimate of his or her financial risk tolerance were compared. It was shown that the client's and advisor's estimates of the client's risk tolerance were only moderately correlated (r =.40). One could excuse these disappointing results by arguing that the criterion against which the advisor's accuracy was assessed, namely the client's opinion, is less than ideal. A more disturbing finding, however, was that the advisors were no more accurate in their estimates relative to the score from a financial risk-tolerance test (r =.41), a reliable and valid standard. In an absolute sense, then, advisors' ability to predict actual risk tolerance is rather faulty, accounting for only about 17% of the variation in the clients' actual risk tolerance. Correcting the criterion measure (SOFRT) for unreliability only raised the correlation slightly (.46). The problem more likely resides with the advisor's rating rather than the criterion.

Notably, the magnitude of the correlation observed here between actual risk tolerance and the judge's estimate of it was quite similar to the one found in the Australian sample of financial planners (r =.38) studied by Elsayed and Martin (1998) and almost of the same magnitude as produced by Eckel & Grossman's (2002) undergraduates (r =.42) who relied on visual cues, such as sex, to form their judgments. Considered together, the findings suggest that financial advisors are not particularly accurate when estimating their client's true level of risk tolerance, despite their training and experience. It would not be prudent to rely solely on a financial advisor's judgment to establish a client's level of risk tolerance. The need for the use of a valid test is indicated by the results of the study.

Furthermore, given the moderate magnitude of the correlation between the client's tested and advisor's estimate of client risk tolerance, it is also quite probable that in estimating risk tolerance advisors are influenced by variables that are either spurious or irrelevant. Experts have a tendency to develop and use heuristic shortcuts to arrive at a judgment. More than likely, the advisors were using "rules of thumb" to form their judgments, but these only work some of the time, if at all. Unless a heuristic rule is based on statistically valid inferences, it is likely that the rule itself will be flawed. One flawed mechanism that advisors thankfully do not seem to be using is the "same as me" attribution discussed by Hsee and Weber (1997). The near zero correlations between the client's self assessed risk tolerance and the advisor's self assessed risk tolerance indicates that advisors were not projecting their own level of risk tolerance unto their clients. However, what other flawed heuristics they may be applying is undetermined. Other research exists to support the contention that experts may assign too much diagnostic value to often meaningless information. For instance, Zielonka (2002) studied the degree of agreement among Polish financial analysts about the impact a particular event is considered to have on the movement of stock prices, and found considerable inter-judge consistency in the assumed importance of various signals, but the agreement was due in large part to the use of heuristics-and-biases so that even useless indicators were viewed to be important indicators.

Even when heuristics are correct, they may not be applied consistently or may be overused. For instance, while sex and wealth are predictive to some extent of risk tolerance because they are correlated with risk tolerance and can thus serve as proxies for risk tolerance, the advisors in this study assigned too much diagnostic value to these variables, as evident from the multiple regressions. Sex and wealth remained predictive of advisor's estimates of risk tolerance even after the variance that these two variables have in common with actual risk tolerance was removed. Stereotyping appears to be a factor in the attribution of a professional advisor makes to a client, contrary to what Hsee and Weber (1997) observed with undergraduates rating strangers, but in line with the findings from Eckel and Grossman (2002), Martin (1987), and Siegrist et al. (2002).

Although some judges may use the correct factors to form their estimate of risk tolerance, they may be unable to do it consistently. Computer programs are therefore able to outperform human judges even when the human judge's decision making process is used to create the program (known as bootstrapping) because factors such as fatigue, headaches, boredom, and work interruptions can distract an expert's ability to arrive at a valid judgment (Dawes, 1971; Grove & Meehl, 1996). If there is no feedback about the accuracy of one's judgments, the process is especially prone to error. It is likely that neither advisors nor clients ever receive feedback about the accuracy of their risk-tolerance judgments. As Dawes et al. (1989, *p*. 1671) noted, "Lacking sufficient or clear information about judgmental accuracy, it is problematic to determine the actual validity, if any, of the variables on which one relies." The practical implication again is the need to use standardized measures, such as risk tolerance questionnaires.

It would not be surprising to find that the advisors in the current study were confident of their ability to accurately peg risk tolerance, despite their questionable performance. We did not address the confidence that the advisors had in their judgments of risk tolerance, but only those advisors who felt they were in a good position to do so gave their opinions regarding their clients' degree of risk tolerance. That is, the advisors provided risk-tolerance estimates on only 63% of the clients. Compared to other occupations, financial services professionals may be overly confident in their abilities, given research comparing Polish financial analysts and weather forecasters (Tyszka & Zielonka, 2002).

The magnitude of the correlations between self-estimated and actual risk tolerance were quite high for both clients and advisors relative to the observed level of accuracy reported in studies examining lay people's ability to estimate all sorts of personality characteristics in themselves (e.g., Chamorro-Premuzic et al.., 2004; Furnham, 1990; Furnham & Chamorro-Premuzic, 2004), where the most predictable characteristics showed correlations in the upper .50 range. Not surprisingly, and in line with Furnham's (1990) findings, the results of the present study show that people are better able to estimate their own level of risk tolerance than to estimate it in others. Thus, the advisors' estimates of client risk tolerance were less accurate than the clients' own estimates. However, a rather puzzling finding was that when it came to estimating one's own level of risk tolerance, clients were better at the task than the advisors. One could achieve a better paramorphic representation of the self estimate for the clients than for the advisors on the basis of the items from the risk tolerance test. From a practical standpoint, taken together, these results suggest that if the choice is between client's and advisor's estimate as the basis for a decision, it might be better to rely on the client's opinion of himself or herself.

The poorer ability of advisors to judge their own risk tolerance may be due to the benchmarks they use. The advisors were generally more risk-taking than their clients. An analysis showed that the greater inaccuracy among advisors occurred primarily because advisors with low to moderate levels of risk tolerance were underestimating their degree of risk tolerance, perhaps because they were using other advisors as their comparison group. The advisors with high levels of risk tolerance were relatively more accurate in their self-judgments, and similar in their degree of accurately to high risk-taking clients' estimates of themselves.

The overriding finding in the various paramorphic representation regressions performed in this study is that relatively few variables were necessary to capture both the advisors' and the client's opinions about their own and others' risk tolerance. Their estimates, of course, were far from perfect. The results on advisors estimating clients could be due to the advisors not having all the necessary information to form a better judgment, but one must wonder whether advisors would use all the information even if presented with the answers to all 56 questions in the risk tolerance survey. Slovic's (1969) seminal study demonstrated that even though stockbrokers had access to a wide variety of specific client data, when making judgments about a client's situation, they relied only six to seven factors, on average, to arrive at a conclusion. It appears that decision makers who use a holistic approach rely only a few cues. Shanteau (1999), who commented on this finding, concluded that "experts make important decisions without adequate attention to all the relevant information" (p. 113).

Conclusion

The results from this study are noteworthy in several respects. First, it was determined that financial advisors are not particularly good judges of their clients' financial risk tolerance. This clearly suggests that financial advisors need to use a valid measure of risk tolerance prior to providing financial advice and guidance. This is particularly true in light of recent Securities and Exchange Commission rule changes which mandate that financial advisory firms use prescribed procedures to assess risk tolerance (McGinnis, 2004). As shown in this study, simply relying on one's own holistic impressions of the client's risk tolerance is not sufficient nor prudent given tightening investment management rules.

The second result should be of particular interest to advisors who feel that their own qualitative judgments are more valid than tests. If not using a valid test of risk tolerance, these advisors would be better served using the client's self-assessment of risk tolerance than their own (i.e., advisor's) estimate of the client's tolerance for risk. This advice is offered because, in general, individuals tend to be better judges of their own risk tolerance than the tolerance of someone else. In fact, financial advisors in this study tended to be relatively worse at assessing even their own risk tolerance compared to their clients.

It appears that advisors rely on heuristic shortcuts that may not account for outside influences when making judgments about their clients. A very striking finding is that a financial advisor's judgment regarding a client's level of risk tolerance can be simulated using a relatively simple regression model – a paramorphic representation- with relatively few input variables.

In summary, its is recommended that future research develop the concept of paramorphic representation as it relates to the way in which financial advisors make all types of judgments. This research should then be followed by attempts to identify how financial advisors actually arrive at judgments when using holistic models of decision making. Determining if judgments based on a combination of experience, knowledge, and temperament are reliable and valid will help fill a wide gap in the existing literature, namely, are financial advisors effective in evaluating the attitudes and preferences of their clients.

Endnotes

[a]In place of stepwise regression, critics recommend hierarchical multiple regression and standard multiple regression (i.e., using all available predictors). Hierarchal regression differs from stepwise regression in that the researcher, rather than the computer, determines the order of entry of the variables on the basis of some theoretical rationale. In this study, the regressions using all possible predictors were run in addition to the stepwise regressions. Hierarchical regressions were not performed because a preset model to test did not exist; the aim was to explore the degree to which the test items would be predictive of the global ratings. The purpose of this research was not so much to identify the best set of predictors as to assess the level of prediction possible using the risk tolerance test items to predict global ratings of risk tolerance. Given the redundancy between items, the models developed to predict the self ratings of both advisors and clients could be represented almost as well using the set of items that the stepwise regression identified as predictive for the other.

Another issue with multiple regression is the minimal sample size necessary. If one has as many predictors as data points, the dependent variable can be predicted perfectly, but the equation will not work very well in another sample. Various rules have been proposed regarding the minimum ratio of study participants to independent variables that one needs to produce meaningful results. However, as Brooks (1998) reported, "Subject-to-predictor conventions have existed for decades with little empirical or mathematical support" (p.3).

A generally cited rule seems to be that in order to avoid "overfitting," the ratio of number of subjects to number of independent variables should be no less than 5:1. In this study, the ratio was slightly over 4:1, falling slightly short of this recommendation.

. The practical implication of using a relatively small number of observations with a relatively large number of predictors is that shrinkage will occur if the equations derived from this study were to be applied to a new sample. In other words, the reported sizes of R^2 are probably inflated. More than likely, the models overestimate the predictability of global ratings from the test items. Stated differently, the items may not predict estimated risk tolerance to the level suggested by our results. A number of formulas have been proposed to estimate the shrinkage, but a very common approach is to calculate an "adjusted R^2" (reported in this study) using the following formula: R^2 adjusted = $1 - (1-R^2)(n-1)/(n-p-1)$, where n = sample size and p = number of predictors. This formula, and others like it, essentially adjust R^2 for the number of predictors that the model uses. As sample size decreases and the number of predictors increases, the penalty becomes larger and larger and the adjusted R^2 is predicted to shrink considerably.

Brooks (1994), however, showed "that researchers cannot ignore effect size in determining sample size in multiple regression analysis any more than they can for any other statistical design " (p.17). The sample size used in this study appears quite adequate to identify significant relationships, given the Hair, Anderson, Tatham, and Black (1998) calculation that with power equaling 80% and alpha set to .05, one can detect a statistically significant R^2 of .23 based on $n = 50$ and a R^2 of .12 with $n = 100$. The same conclusion is reached using Green's (1991) formula that also takes into account effect size (f2). Based on his recommendation, the sample size should be greater than or equal to $(8 / f2) + (p - 1)$, where $f2 = R^2/(1 - R^2)$ and p = number of independent variables.

References

Alicke M. D., Klotz M. L., Breitenbecher D. L., Yurak T. J., & Vredenburg D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology, 68*, 804-825.

Armstrong, J. S. (1971). Exploratory analysis of marketing data: A reply**.** *Journal of Marketing Research, 8,* 511-513.

Bateman, I., & Munro, A. (2005). An experiment on risky choice amongst households. *The Economic Journal, 115,* C176–C189.

Bottom, W. P., Holloway, J., McClurg, S., & Miller, G.J. (2000). Negotiating a coalition: Risk, quota shaving, and learning to bargain. *Journal of Conflict Resolution, 44*, 147-169

Borkenau, P. & Liebler, A. (1993a). Convergence of stranger ratings of personality and intelligence with self-ratings, partner ratings, and measured intelligence. *Journal of Personality and Social Psychology, 65*, 546-553.

Borkenau, P. & Liebler, A. (1993b). Consensus and self-other agreement for trait inferences from minimal information. *Journal of Personality, 61*, 477-496.

Bornstein, B. H., & Emler, A.C. (2001). Rationality in medical decision making: A review of the literature on doctors' decision-making biases. *Journal of Evaluation in Clinical Practice***,** *7,* 97-107.

Bouchey, P. (2004, July 1). Questionnaire quest: New research shows that standard questionnaires designed to reveal investors' risk tolerance levels are often flawed or misleading. *Financial Planning.* Retrieved October 10, 2005, from http://www.financial-planning.com/pubs/fp/20040701024.html

Brooks, G. (1994). A new sample size formula for regression. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA. Retrieved October 10, 2005, from http://oak.cats.ohiou.edu/~brooksg/files/aera94pear.pdf

Brooks, G. P. (1998, October). Precision efficacy analysis for regression. Paper presented at the meeting of the Mid-Western Educational Research Association, Chicago, IL. Retrieved October 10, 2005, from http://oak.cats.ohiou.edu/~brooksg/files/mwera98.pdf

Callan, V. J., & Johnson, M. (2002). Some guidelines for financial planners in measuring and advising clients about their levels of risk tolerance. *Journal of Personal Finance, 1*(1), 31-44.

Camerer, C. F., & Johnson, E. J. (1997). The process-performance paradox in expert judgment: How can experts know so much and predict so badly. In W. M. Goldstein & R. M. Hogarth (Eds.). *Research on judgment and decision making: Currents, connections, and controversies* (pp. 342-364). Cambridge, England: Cambridge University Press.

Chamorro-Premuzic, T., Furnham, A., Moutafi, J. (2004). The relationship between estimated and psychometric personality and intelligence scores**.** *Journal of Research in Personality 38*, 505-513.

Chaulk, B., Johnson, P. J., & Bulcroft, R. (2003). Effects of marriage and children on financial risk tolerance: A synthesis of family development and prospect theory. *Journal of Family and Economic Issues, 24*, 257-279.

Cicchetti, C. J., & Dubin, J. A. (1994). A microeconometric analysis of risk aversion and the decision to self-insure. *Journal of Political Economy, 102*, 169-186.

Clark, R. D., Crockett, W. H., Archer, R. I. (1971). Risk-as-value hypothesis: The relationship between perception of self, others, and the risky shift. *Journal of Personality and Social Psychology*, *2*0, 425-429.

Dawes, R. M. (1971). A case study of graduate admissions: Application of three principles of human decision making. *American Psychologist, 26,* 180-188.

Dawes, R. M., Faust, D. & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, *243*, 1668-1674.

Doherty, M. E., & Brehmer, B. (1997). The paramorphic representation of clinical judgment: A thirty-year retrospective. In W. M. Goldstein & R. M. Hogarth (Eds.). *Research on judgment and decision making: Currents, connections, and controversies* (pp. 537-551). Cambridge, England: Cambridge University Press.

Dreze, J. H. (1981). Inferring risk tolerance from deductibles in insurance contracts. *The Geneva Papers on Risk and Insurance, 20,* 48-52.

Eckel, C. C. & Grossman, P. J. (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution & Human Behavior, 23*, 281-295.

Einhorn, H. (1974). Expert judgment: Some necessary conditions and an example. *Journal of Applied Psychology, 59,* 562-571.

Einhorn, H. J., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Linear regression and process-tracing models of judgment. *Psychological Review, 86,* 465-485.

Elsayed, H. & Martin, J. (1998, February 28). Survey of financial risk tolerance - Australian Technical Brief. Unpublished manuscript, Chandler and Macleod Consultants. Retrieved October 10, 2005, from http://www.myrisktolerance.com/downloads/Survey_RT_Chandler&Macleod.pdf

Ettenson, R., Shanteau, J., & Krogstad, J. (1987). Expert judgment: Is more information better? *Psychological Reports, 60,* 227-238.

Falk, A. & Knell, M. (2004). Choosing the Joneses: Endogenous goals and standards. *Scandinavian Journal of Economics*, *106*, 417-435.

Fan, X. (2003). Two approaches for correcting correlation attenuation caused by measurement error: Implications for research practice. *Educational and Psychological Measurement*, *63,* 915-930*.*

Finke, M. S. & Huston, S. J. (2003). The brighter side of financial risk: Financial risk tolerance and wealth**.** *Journal of Family & Economic Issues, 24,* 233-256.

Fischoff, B., Slovic, E, & Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 330-334.

Furnham, A. (1990). Can people accurately estimate their own personality test scores? *European Journal of Personality, 4,* 319-327.

Furnham, A. & Chamorro-Premuzic, T. (2004). Estimating one's own personality and intelligence scores. *British Journal of Psychology, 95*, 149-160.

Garb, H. N. (1989). Clinical judgment, clinical training, and professional experience. *Psychological Bulletin, 105,* 387-396.

Gordon, H. R. D. (2001). American Vocational Education Research Association members' perceptions of statistical significance tests and other statistical controversies. *Journal of Vocational Education Research, 26*(2). Retrieved October 10, 2005, from http://scholar.lib.vt.edu/ejournals/JVER/v26n2/gordon.html

Grable, J. & Joo, S. (1997). Determinants of risk preference: Implications for family and consumer science professionals. *Financial Economics and Resource Management Biennial*, *2*, 19-24.

Grable, J. & Joo, S. (2000). A cross-disciplinary examination of financial risk tolerance. *Consumer Interests Annual,46.* 151-157.

Grable, J. E., & Lytton, R. H. (1998). Investor risk tolerance: Testing the efficiency of demographics as differentiating and classifying factors. *Financial Counseling and Planning 9*(1)*,* 61-74.

Grable, J. E. & Lytton, R. H. (1999a). Financial risk tolerance revisited: The development of a risk assessment instrument. *Financial Services Review, 8*, 163-181.

Grable, J. E., & Lytton, R. H. (1999b). Assessing financial risk tolerance: Do demographic, socioeconomic and attitudinal factors work? *Family Relations and Human Development/Family Economics and Resource Management Biennial, 3*, 80–88.

Grable, J. E., & Lytton, R. H. (2001). Assessing the concurrent validity of the SCF risk tolerance question. *Financial Counseling and Planning*, *12*(2), 43–52.

Grable, J. E., & Lytton, R. H. (2003). The development of a risk assessment instrument: A follow-up study. *Financial Services Review, 8*, 163-181.

Grable, J., Lytton, R.. & O'Neill, B. (2004). Projection bias and financial risk tolerance. *Journal of Behavioral Finance, 5*, 142-147.

Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research, 26,* 499-510.

Grove, W. M., &. Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law, 2*, 293-323.

Hair, J. F., Anderson R. E., Tatham R .L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Englewood Cliffs, NJ: Prentice Hall.

Hallahan, T. A., Faff, R. W., & McKenzie, M. D. (2004). An empirical investigation of personal financial risk tolerance. *Financial Services Review, 13,* 57-78.

Hanna, S. D., & Chen, P. (1997). Subjective and objective risk tolerance: Implications for optimal portfolios. *Financial Counseling and Planning, 8*(2)*,* 17-26.

Hanna, S., & Gutter, M. (1998). A theory based measure of risk tolerance. *Proceedings of the Academy of Financial Services, 15*, 10-21.

Hanna, S. D., Gutter, M. S., & Fan, J. X. (2001). A measure of risk tolerance based on economic theory. *Financial Counseling and Planning, 12*(2), 54-60.

Hanna, S. D., & Lindamood, S. (2004). An improved measure of risk aversion. *Financial Counseling and Planning, 15*(2)*,* 27-38.

Harris, J. L . (2004). Using value at risk to assess investment risk tolerance. *Journal of Personal Finance*, *3*(3), 113-128.

Hoffman, P. J. (1960). The paramorphic representation of clinical judgment*.* *Psychological Bulletin, 57,* 116-131.

Hsee, C. K., & Weber, E. U. (1997). A fundamental prediction error: Self-others discrepancies in risk preference. *Journal of Experimental Psychology: General, 126*, 45-53.

Kautt, G. (2002). The financial planner as decision maker. *Journal of Financial Planning, 15*(10), 96-106.

Li, H. (2005). Factors that affect the risk tolerance of Japanese consumers. Panel discussion on Cultural Comparisons of Risk Tolerance at The Fifth Biennial Conference of Asian Consumer and Family Economics Association, hosted by College of Social Sciences, National Chengchi University, Taipei 116 Taiwan. Retrieved July 5, 2005, from http://www.socialsciences.nccu.edu.tw/acfea/proceedings/panels/Panel-Cultural%20Comparisons%20of%20Risk%20Tolerance.pdf

Locksley, A., Borgida, E., Brekke, N., & Hepburn, C. (1980). Sex stereotypes and social judgment. *Journal of Personality & Social Psychology, 39*, 821-831.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin, 127*, 267-286.

Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 619-642). London: Oxford University Press.

Martin, C. L. (1987). A ratio measure of sex stereotyping. *Journal of Personality & Social Psychology, 52*, 489-499.

Magnan, R. E. & Hinsz, V. B. (2005). Mood, gender, and situational influences on risk-taking advice for others. *Social Behavior & Personality 33*, 1-10.

MacCrimmon, K. R., Wehrung, D. A., & Stanbury, W. T. (1986). *Taking risks.* New York: The Free Press.

McGinnis, S. K. (2004, August). Cleaning up Dodge. *Financial Planning, 8,* 62-63.

Moreschi, R.W. (2004). Incorporating investor risk tolerance into the financial planning process. *Journal of Personal Finance*, *3*(3), 89-98.

Osborne, J. W. (2000). Prediction in multiple regression. *Practical Assessment, Research & Evaluation, 7*(2). Retrieved July 10, 2005 from http://PAREonline.net/getvn.asp?v=7&n=2

Quiggin, J. (1982).A theory of anticipated utility. *Journal of Economic Behavior and Organization, 3*, 323-43.

Roszkowski, M. J. (1992). How to assess a client's financial risk tolerance: The basics. *Personal financial risk tolerance*. Bryn Mawr, PA: The American College.

Roszkowski, M. J., Delaney, M. M., & Cordell, D. M. (2004). The comparability of husbands and wives on financial risk tolerance. *Journal of Personal Finance, 3*(3), 129-144.

Roszkowski, M. J., Davey, G., & Grable, J.E. (2005). Insights from psychology and psychometrics on measuring risk tolerance. *The Journal of Financial Planning, 18*(4), *66-75.*

Roszkowski, M. J., Spreat, S., & Isett, R. (1983). A paramorphic representation of psychologists' clinical impressions of degree of mental retardation. *Journal of Psychoeducational Assessment*, *1*, 243-251.

Ruscio, J. (2003). Holistic judgment in clinical practice. *The Scientific Review of Mental Health Practice, 2*(1), 38-48. Retrieved October 10, 2005, from http://www.srmhp.org/0201/holistic.html.

Shanteau, J. (1995, June 12-14). Expert judgment and financial decision making. In B. Green (Ed.), *Risk behaviour and risk management*. *Proceedings of the first international Stockholm seminar on risk behaviour and risk management*, pp. 16-32.

Shanteau, J. (1999). Decision making by experts: the GHAHM effect. In J. Shanteau, B. A. Mellers, & D. A. Schum (Eds.). *Decision science and technology: Reflections on the contributions of Ward Edwards* (pp. 105-130). Boston: Kluwer.

Siegrist, M., Cvetkovich, G., & Gutscher, H. (2002). Risk preference predictions and gender stereotypes. *Organizational Behavior & Human Decision Processes, 87,* 91-102.

Slovic, P. (1969). Analyzing the expert judge: A descriptive study of a stockbroker's decision process. *Journal of Applied Psychology, 53,* 255-263.

Snelbecker, G. E., Roszkowski, M. J., & Cutler, N. E. (1990). Investors' risk tolerance and return aspirations, and financial advisors' interpretations: A conceptual model and exploratory data. *Journal of Behavioral Economics, 19*, 377-393.

Stewart, T. R., Roebber, P. J., & Bosart, L. F. (1997). The importance of the task in analyzing expert judgment. *Organizational Behavior & Human Decision Processes, 69,* 205-219.

Sugden, R. (2003). Reference-dependent subjective expected utility. *Journal of Economic Theory, 111,* 172-191.

Sung, J. & Hanna, S. (1996). Factors related to risk tolerance. *Financial Counseling and Planning,* 7, 11-20.

Thompson, B. (1995). Stepwise regression and stepwise discriminant analysis need not apply here: A guidelines editorial. *Educational and Psychological Measurement, 55,* 525–534.

Torngren, G., & Montgomery, H. (2004). Worse than chance? Performance and confidence among professionals and laypeople in the stock market. *Journal of Behavioral Finance, 5,* 148-153.

Tversky, A. & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*, 297-323.

Tyszka, T. & Zielonka, P. (2002). Expert judgments: Financial analysts vs. weather forecasters. *Journal of Psychology and Financial Markets*, *3*(3), 152-160.

Xiao, J. J., Alhabeeb, M. J., Hong, G.-S., & Haynes, G. W. (2001). Attitude toward risk and risk-taking behavior of business-owning families. *Journal of Consumer Affairs*, *35*, 307-325.

van Belle, G. (2002). *Statistical rules of thumb*. New York: Wiley-Interscience.

Vogt, D. S. & Colvin C. R. (2005). Assessment of accurate self-knowledge. *Journal of Personality Assessment, 84*, 239-251.

von Neumann, J., & Morgenstern, O. (2005). *Theory of games and economic behavior* (Commemorative ed.). NJ: Princeton University Press.

Yieh, K. & Chen, L-Y. (2003). The determinants of risk preference-an empirical study in Taiwan. Panel discussion on Cultural Comparisons of Risk Tolerance at The Fifth Biennial Conference of Asian Consumer and Family Economics Association, hosted by College of Social Sciences, National Chengchi University, Taipei 116 Taiwan. Retrieved July 5, 2005 from http://www.socialsciences.nccu.edu.tw/acfea/proceedings/panels/Panel-Cultural%20Comparisons%20of%20Risk%20Tolerance.pdf

Yao, R., Hanna, S. D., & Lindamood, S. (2004). Changes in financial risk tolerance, 1983-2001. *Financial Services Review, 13*(4)*,* 249-266.

Yip, U.Y. (2000). Financial risk tolerance: a state or a trait? A thesis submitted in partial fulfillment of the requirements for the degree of Master of Psychology (Organisational) of The University of New South Wales. November, 2000. Retrieved on May 10, 2005 from http://www.riskprofiling.com/Downloads/AStateOrATrait.pdf

Yook, K.C. & Everett, R. (2003, August). Assessing risk tolerance: Questioning the questionnaire method. *Journal of Financial Planning*, 48-55. Retrieved October 10, 2005, from http://www.fpanet.org/journal/articles/2003_Issues/jfp0803-art7.cfm

Zacharakis, A. & Meyer, D. (2000). The potential of actuarial decision models: Can they improve the venture capital investment decision?" *Journal of Business Venturing*, *15*, 323-46.

Zielonka, P. (2002). How financial analysts perceive macroeconomic, political news and technical analysis signals. *Financial Counseling and Planning, 13*, 87-96.

