



Assessing Risk Tolerance Scientifically



Geoff Davey
Co-founder and Director, FinaMetrica

Today it is generally agreed that planners have a professional, ethical and legal obligation to assess their clients' risk tolerance. However, from that point on there is little agreement about what exactly risk tolerance is and even less on how it should be assessed. Yet these are not two unknowns.

Psychologists have been investigating risk tolerance for more than 50 years. A large body of knowledge based on studies that have been independently refereed and replicated now exists. However, when financial services professionals seek academic/researcher input it is almost invariably from finance and economics. Unfortunately, the silos in academia being as they are, very little of the psychologists' knowledge had made its way into economics and finance.



Risk tolerance is a psychological trait, i.e. a relatively enduring way one individual differs from another.

There are four types of risk tolerance: physical, social, ethical and financial. Individuals behave consistently within type but not across types, i.e. a mountain climber is more likely to be a hang-glider than the man or woman in the street, but may or may not be a financial risk taker.

There are no sub-types of financial risk tolerance, e.g. investment risk tolerance, insurance risk tolerance, borrowing risk tolerance, etc. Factor analysis reveals a single significant factor, which means that there is a single trait that applies across all financial dimensions.

An individual's risk tolerance influences that individual's risk-taking behaviour. An individual is exposed to risk in any situation where there is more than one possible outcome.

The ISO 22222 Personal Financial Planning Standard defines risk tolerance as "the extent to which a consumer is willing to risk experiencing a less favourable financial outcome in the pursuit of a more favourable financial outcome."

Psychometrics

As a personality trait, risk tolerance is not an easy variable to measure. Fortunately there is a scientific discipline, psychometrics, for testing attributes such as risk tolerance. Psychometrics, a blend of psychology and statistics, provides both a discipline for developing valid and reliable tests and standards against which the bona fides of a test can be evaluated.

To meet these standards, a test must go through a rigorous development process. First, a large pool of questions is created and tested on representative samples of the population for which the test is intended, to see if the questions are understandable and answerable by this audience (Usability Trials). Questions that seem straightforward are often revealed to have poor understandability or answerability. Any technical term, even one as simple as "bonds", causes problems. With questions about rates of return the more informed will want to know if this is before or after inflation, and any mention of "after-inflation" in a returns question is too difficult for most.

From the Usability trials questions with apparent promise, based on their understandability and answerability, are tested on further representative samples using statistical criteria (Norming trials). The results are examined to determine if the statistical characteristics of the questions and the scoring algorithm are proper. Upon testing, questions that at first appear insightful are often revealed to have little or no statistical value in differentiating one respondent from another. Typically, question development requires multiple loops through both trial processes.

What Makes a Test Valid?

Broadly defined, a valid test is one that actually measures what it purports to measure. There are various aspects of validity, of which content validity and criterion-related validity are the most common.

If a test has good content validity, the questions it asks are seen to be very relevant by those with expertise in the field.

Criterion-related validity is a measure of behavior related to the construct being tested (the criterion). In the case of risk tolerance, the criterion would be actual behavior reflecting risk-taking propensity, e.g. the proportion of stocks owned within a portfolio.

If the criterion is collected at the same time the test is administered, it is called concurrent validity; if the criterion does not materialize until some later time, it is called predictive validity.

What Makes a Test Reliable?

The score on any test consists of two parts: a true score and an error (that is, test score = true score $\pm$ error of measurement). All tests have some margin of error, so it is a matter of degree. Reliability can be conceptualized as the correlation of the true score to the test score. In other words, reliability tells us what proportion of the test is non-error. If the error component is large, then the test is unreliable and will fail to give consistent results from one testing to the next, even if the client's risk tolerance has not changed. The error generally comes from sources in the test itself (such as ambiguous wording.) Other things being equal, the more questions of the same type one asks, the more reliable an instrument becomes.

For satisfactory reliability, the correlation should be .8 or greater. Once reliability is known we can determine the accuracy of a test using the standard error of measurement which is calculated by the formula $SEM = \sigma \times \sqrt{1-r}$, where σ is the standard deviation and r is the reliability. So for a test with a standard deviation of 10 and a reliability of .9, $SEM = 10 \times \sqrt{1-.9} = 3.1$.

Knowing the SEM means that for a particular confidence level we can calculate the range in which the client's "true" risk tolerance score is located. We can, for example, be 95 percent certain that the client's true risk tolerance lies in a range that is 1.96 times the SEM (because 95 percent of a normal distribution lies within 1.96 standard deviations of the mean.)

Consider a test with mean 50 and standard deviation 10, for which the reliability is .9 and therefore the SEM is 3.1. Suppose a client scores 60 on this test. We can be 95% confident that the client's true score lies within a range of 60 ± 6 (1.96×3.1 rounded.)

Now suppose this test had a reliability of .4 and therefore an SEM of 7.7. Here the 95% confidence range for a test score of 60 would be 60 ± 15 , i.e. a range three standard deviations wide. Such a test would be unsuitable for measuring a client's risk tolerance because it is simply not precise enough.

Psychologists divide behavior into cognitive (intellectual) and affective (emotional) domains. Risk tolerance falls into the affective domain. Years of research have shown that ordinarily it takes more questions to reliably assess affective traits than cognitive ones, typically 20+.

Lest planners be concerned that clients will find a 20+ question psychometric test onerous, it should be remembered that if the questionnaire has been designed appropriately, the understandability and answerability of all questions will have been assured and the process will therefore take less time than

one may think. A 20-question psychometrically designed test should take approximately 15 minutes to complete.

Further, clients need no persuading that it is important for planners to have an accurate understanding of their risk tolerance and the one thing we all want to know more about is ourselves, so the process is an enjoyable one for most clients. Surveys of respondents show that they consider it a worthwhile exercise, which leads to a better understanding of themselves in relation to financial risk (and, in couples, to one another.) In fact, a psychometric risk tolerance test should be a bright spot in the otherwise somewhat burdensome initial fact-find experience.

Finally, because the test can (and should) be taken without input from the adviser the adviser's time is not spent explaining the questions and 'assisting' in the completion of the questionnaire.

A psychometric risk tolerance test will provide an accurate assessment of a client's risk tolerance - with a small known margin of error on a known scale - and a plain language report that will be meaningful to both client and adviser. A psychometric test does not replace discussion with the client but rather provides an objective input to that discussion.

Psychometric testing is complicated. But the complexity resides in the development of the questionnaire, the report and the scoring algorithm, not in its use with clients.

Interestingly, a psychometric test mirrors what many advisers aim to do, either through interviewing or by using non-scientifically derived questionnaires, namely,

- Form a view as to the client's risk tolerance by conducting a question-and-answer discussion about the client's attitudes, values, preferences and experiences in matters involving financial risk.
- Rate the client on some type of low/medium/high scale, usually in comparison with the adviser's other clients.
- Feed back to the client a written summary of this view showing how they are similar and different from others.
- After discussion with the client, amend the summary to reflect the client's feedback (as required), in order to obtain the client's confirmation that the adviser's understanding is accurate and that the adviser may rely on it.

A psychometric test does all the above but with scientific rigour. But does this rigour make a difference? Is it necessary?

Industry-Standard 'Risk' Questionnaires

Our industry is bedevilled by the superficially plausible, of which there is no better example than the industry-standard 'risk' questionnaire.

The problem with industry-standard questionnaires is that they have been arbitrarily constructed without regard to psychometrics. Typically, they contain too many "bad" questions and not enough "good" questions. As a consequence, their results are neither valid nor reliable.

Years ago it was not uncommon to find questions relating to physical risk tolerance in questionnaires designed to

measure financial risk tolerance. Today, the more prevalent problem is that many risk tolerance questionnaires deal with financial matters that are not part of the construct of risk tolerance. This is a legacy from the ubiquitous asset allocation calculators, which were designed to produce an asset allocation (or model portfolio) recommendation based on brief questionnaires about (in addition to risk tolerance) time horizon, withdrawal expectations, investment experience, risk capacity and the like. While time horizon, etc. are relevant to investment advice they are not relevant to risk tolerance.

So the first problem is irrelevant questions. The second is the use of questions that would fail Usability trials. At best, such questions would require explanation by the planner, which will introduce bias and make the result unreliable. At worst, despite any explanation by the planner, the question remains too difficult for the client.

While it is not possible to identify a "good" question without conducting usability and norming trials, questions that are irrelevant or too technical can be identified as "bad" questions by a sight check.

Industry-standard questionnaires have not been subjected to usability or norming trials so it is not possible to say definitively whether any of the questions are "good" questions. However, the number of "bad" questions identified by sight checking is such that even if all the other questions were "good" questions there would not be sufficient to achieve acceptable reliability.

The conceptual flaws in industry-standard questionnaires are evident in practice. Independent studies show advisers using an industry-standard approach make disturbingly inaccurate estimates of their clients' risk tolerance, with estimates correlating to test scores at only ~.4. This level of inaccuracy will give rise to gross errors (2+ standard deviations) in one in six cases. Put another way, advisers would be more accurate if they made no attempt to assess clients' risk tolerance at all and simply assumed everyone was average.

Conclusion

Know-the-client has always been a cornerstone of Financial Planning and knowing the client's risk tolerance is an essential component of that obligation, even more so in a fiduciary environment. A psychometric test ensures that a valid, reliable and accurate assessment is made, allowing the planner to provide a more informed service in the proper discharge of his or her obligations while providing clients with a pleasurable experience in an early demonstration of planning expertise.

The author is a cofounder and director of FinaMetrica. He is the creator of the FinaMetrica Risk Profiling system and manages FinaMetrica's ongoing research activities.

The FinaMetrica system is used by leading advisers in 13 countries in seven languages. Planners can register for a free trial of the system at www.riskprofiling.com.

