IQ versus RQ
Differentiating Smarts from Decision-Making Skills

May 12, 2015

Authors
Michael J. Mauboussin
michael.mauboussin@credit-suisse.com
Dan Callahan, CFA
daniel.callahan@credit-suisse.com

"Being rational is a moral imperative."
Charlie Munger

- Intelligence quotient (IQ) and rationality quotient (RQ) are distinct. Think of IQ as the horsepower of an engine and RQ as the output.
- We share the results of a classic test of calibration, which is an important facet of rationality. Well calibrated people know what they know and know what they don’t know.
- Consistent with past research, we find that participants overestimate their accuracy as their subjective probability estimates tend to be higher than the actual percent correct.
- Investors and executives can improve their rationality by keeping score, asking about others, using base rates, and updating probabilities.
- A large-scale forecasting project has shown that the best forecasters use inductive and numerical reasoning, have cognitive control and a growth mindset, and are open-minded and effective working as part of a team.
Introduction

Keith Stanovich, a professor of applied psychology at the University of Toronto, distinguishes between intelligence quotient (IQ) and rationality quotient (RQ). Psychologists measure IQ through specific tests, including the Wechsler Adult Intelligence Scale, and it correlates highly with standardized tests such as the SAT.

IQ measures something real, and it is associated with certain outcomes. For example, thirteen-year-old children who scored in the top decile of the top percent (99.9th percentile) on the math section of the SAT were eighteen times more likely to earn a doctorate degree in math or science than children who scored in the bottom decile of the top percent (99.1st percentile).

RQ is the ability to think rationally and, as a consequence, to make good decisions. Whereas we generally think of intelligence and rationality as going together, Stanovich’s work shows that the correlation coefficient between IQ and RQ is relatively low at .20 to .35. IQ tests are not designed to capture the thinking that leads to judicious decisions.

Stanovich laments that almost all societies are focused on intelligence when the costs of irrational behavior are so high. But you can pick out the signatures of rational thinking if you are alert to them. According to Stanovich, they include adaptive behavioral acts, efficient behavioral regulation, sensible goal prioritization, reflectivity, and the proper treatment of evidence.

Your SAT scores shed little light on any of these qualities. So the first lesson in assessing your own decisions or those of others is to consider IQ and RQ separately. Warren Buffett, chairman and chief executive officer (CEO) of Berkshire Hathaway, equates IQ to the horsepower of an engine and RQ to the output. We all know people who are high on IQ but average or low on RQ. Their efficiency is poor. There are others without dazzling IQs but who consistently make sound decisions. They are highly efficient.

Warren Buffett has plenty of horsepower and output. But when asked about his success, Buffett emphasized that it was RQ that made the big difference, not IQ:

How I got here is pretty simple in my case. It’s not IQ, I’m sure you’ll be glad to hear. The big thing is rationality. I always look at IQ and talent as representing the horsepower of the motor, but that the output—the efficiency with which that motor works—depends on rationality. A lot of people start out with 400-horsepower motors but only get a hundred horsepower of output. It’s way better to have a 200-horsepower motor and get it all into output.

Stanovich’s psychological research supports Buffett’s observation. While there is not yet a comprehensive test to measure RQ—Stanovich is working on it—we will look at calibration, one of the important facets of rationality. As part of this research, we measured the calibration of thousands of people. And part of the fun is that you, too, can participate in the exercise and see how you stack up versus others.
Measuring Rationality

Cognitive scientists and philosophers talk about “instrumental” and “epistemic” rationality. Instrumental rationality is behaving in such a way that you get what you want the most, subject to constraints. Expected utility theory, which is based on a series of axioms, provides a normative framework for how to do this. You’ll be instrumentally rational if you follow the axioms. ⁹

Epistemic rationality describes how well a person’s beliefs map onto the world. If you believe in the tooth fairy, for instance, you are showing a lack of epistemic rationality. Here’s a catchier way to remember the two terms: instrumental rationality is “what to do” and epistemic rationality is “what is true.” ¹⁰

We will focus on observations about epistemic rationality. One way to assess this form of rationality is through a test of calibration. Think of a weather forecaster. If it actually rains 70 percent of the time on the days she predicts a 70 percent chance of rain, she is well calibrated. She is poorly calibrated, on the other hand, if it only rains on 30 percent of those days.

Exhibit 1 shows one approach to keeping score. The horizontal axis measures an individual’s subjective forecast (“there’s a 70 percent chance of rain tomorrow”) and the vertical axis captures the actual outcome (“it rained”). You know that someone is well calibrated if their results fall close to the line at a 45-degree angle.

Exhibit 1: Measure of Calibration

Conviction is another important consideration. Conviction measures how well people do when they assign extreme probabilities. ¹¹ Calibration and conviction are related but distinct. For example, it rains about half of the days in an average year in London. So if you wake up every morning, flip a fair coin, and mark your outcome, you will appear well calibrated over one year.

But that doesn’t help you plan picnics. What you want are a series of predictions for sun or rain that correspond with the actual weather that day. Those predictions require higher conviction than what the toss of a coin can offer. Exhibit 2 shows perfect calibration but poor conviction on the left, and perfect calibration and conviction on the right.
Exhibit 2: Calibration and Conviction


Here we present the results of a classic calibration test. Subjects who participated went to the website, http://confidence.success-equation.com, and saw 50 true-false questions. Exhibit 3 is a screenshot of the site.

The subjects then answered either true or false and were asked to register a probability of correctness, from 50 to 100 percent, in increments of 10 percentage points. If you have no idea whether the answer is true or false you should select an answer at random and enter “50%” as your probability of correctness. If you are certain of the answer you provide, you click “100%.”

At the end, the subjects submit their answers and receive their results, which include:

- Mean, or average, confidence and percent correct for all questions
- Mean confidence for correct and incorrect answers
- Number correct and answered for low confidence (50-60 percent), medium confidence (70-80 percent) and high confidence (90-100 percent) submissions
- A calibration graph similar to exhibit 1

We accessed the results of 1,985 participants, all of whom were anonymous.12
Exhibit 3: Calibration Questions and Probability of Correctness


Exhibit 4 shows the results. The pattern is consistent with what researchers have found for decades: Subjective probability estimates are substantially higher, on average, than the actual percent correct. For the whole population, the average subjective probability was 70 percent and the actual percent correct was just under 60 percent.
Note that proper calibration does not require being right all of the time (the dot in the upper right-hand corner indicates that either someone cheated or a deity took the test) but rather being close to the 45 degree line. It’s knowing what you know and knowing what you don’t know.

Exhibit 5 shows the distribution of responses, with the horizontal axis representing the confidence level minus the percent correct and the vertical axis the frequency. A normal distribution, or bell curve, describes the data well, with a mean and standard deviation of about 10 percent.

This distribution allows us to segregate the participants into three groups: those who are overconfident, underconfident, and well calibrated. We define well calibrated as a subjective probability within 1 percentage point of the percent correct.

Based on that criterion, 82.7 percent of the participants were overconfident (subjective confidence exceeded actual percent correct), 11.7 percent were underconfident (subjective confidence less than actual percent correct), and only 5.6 percent were well calibrated.
Amos Tversky, the renowned psychologist, is reported to have said that humans can only distinguish between three levels of probability: “it’s gonna happen,” “it’s not gonna happen,” and “maybe.” Exhibit 6 shows the distribution of subjective probabilities of correctness.

Forty-two percent of the responses were 50 percent, which is the equivalent of saying, “I have no idea what the answer is.” This corresponds to Tversky’s “maybe.” Note that there is likely an additional effect here because 50 percent is the default setting on the site. So if the participant doesn’t change the assigned subjective probability, it registers 50 percent automatically.

The next most popular response, nearly one-quarter of the total, was 100 percent. This is Tversky’s “it’s gonna happen” or “it’s not gonna happen.” So nearly two-thirds of the responses were either “I don’t know” or “I do know” with the balance split between the four choices in between those extremes.
This leads to a logical follow up question: How good were the results for each assigned probability of correctness? Exhibit 7 shows the answer. When the subjects selected 50 percent, their probability of being correct was random. This means they were well calibrated. They didn’t know, knew they didn’t know, and answered as if they didn’t know.

However, as the assigned probability of correctness rose, the subjects became less calibrated. For instance, when the subjects selected 100 percent, they were only correct 77 percent of the time. At 90 percent, they were only correct 65 percent of the time. Overestimation of ability was greatest at the high levels of assigned probability of correctness.¹⁴

Overconfidence can be a problem for a couple of reasons. The first obvious one is if you are highly confident of an outcome and are wrong a relatively high percentage of the time, you will fail to consider alternatives and ultimately make poor decisions.¹⁵ One recent example is Ron Johnson’s tenure as CEO of J.C. Penney.
Johnson, relying on his intuition, quickly repositioned the retailer to poor effect. Following his ouster, the retailer returned to many of its past practices.\textsuperscript{16}

Another problem is that people who think that they know more than they do are less motivated to learn and improve than those who understand their limitations.\textsuperscript{17} Indeed, one study showed that the least capable people have the largest gap between what they think they can do and what they actually achieve.\textsuperscript{18}

**Lessons for Investors and Executives**

The good news in all of this is that we can train ourselves to be more rational. Here are some ideas:

- **Keep score.** To the degree to which you can pose questions that will have a definite answer within a known period of time, you have a basis for keeping score. The classic way to do this is through a Brier score, which we discuss in detail in the appendix. Brier scores were originally developed to help give meteorologists feedback on their predictions for the weather. Through improvements in weather modeling techniques and sharper feedback, meteorologists today are vastly more accurate than they were a generation or two ago.\textsuperscript{19}

- **Ask about others.** Emily Pronin, a professor of psychology at Princeton University, has found that while people recognize biases in the thinking of others, they somehow think that they suffer less from the same biases.\textsuperscript{20} For example, physicians know that gifts from pharmaceutical companies have biasing effects for other doctors but believe they are immune from the effect.

  Here’s a technique to deal with this. Say you are an investor interviewing a company’s management team, and you are considering buying the stock. You should know that the management team will have an optimism bias, and so you have to take what they say with a grain of salt. But their views about other companies are more likely to be accurate. In other words, don’t ask people about themselves, ask them about others.

- **Use base rates.** Perhaps the single most effective de-biasing tool is the use of base rates.\textsuperscript{21} While we all like to think of ourselves as unique, asking what happened when others were in the same situation can be very helpful. Larry Summers, the former president of Harvard University and Secretary of the Treasury of the U.S., had a rule he used with his research assistants. He would ask how long a project would take. And then he would take the assistant’s answer, double the estimate, and move up to the next unit of time.\textsuperscript{22} So “two hours” would be translated as “four days.” Perhaps this was simply a mechanism to avoid disappointment, but the direction of the new estimate was no doubt correct.

- **Update probabilities.** The test we shared with the participants was static. In the real world, probabilities shift all of the time. One of the key challenges in rational thinking is to accurately update probabilities as new information arrives. It turns out that the very best forecasters do this very well and use very granular increments of probability.\textsuperscript{23} Most of us fall into the trap of confirmation bias, preferring to disregard or discount new information than to properly incorporate it into our assessment.
Limitations to This Project

While this project was fun and revealed results that are consistent with past research, we want to be quick to note that this work does not meet the standard of academic research for a few reasons.

To begin, we can define overconfidence in multiple ways. The primary form of overconfidence that this test captures is overestimation—you think you’re right 70 percent of the time but you’re only right 60 percent—but overprecision, the tendency to provide ranges of outcomes that are too narrow, also plays a role. Suffice it to say that overconfidence has no simple definition and hence there’s no uniform way to test it.

How our participants answered also reflected the questions we posed. We attempted to have questions that were varied, but it is likely that an individual may have a different result for a separate set of questions. In other words, an individual’s results might vary from test to test based on the nature of the questions.

Our default percentages were in increments of 10 percentage points, which limited the ability of the participants to provide greater subtlety in their responses. This may have also encouraged the result of the extremes—lots of 50 and 100 percent answers.

Finally, we may have a biased sample of test takers. Most entered the site after having seen it mentioned in social media communication. Participants from four countries constituted more than three-fourths of the sample, including 37 percent from the U.S., 21 percent from Australia, 15 percent from the Netherlands, and 5 percent from the U.K. Nearly all participants took the test on one of four operating systems, including 40 percent on Windows, 27 percent on iOS, 15 percent on Mac OS, and 14 percent on Android.

Characteristics of People with High RQ

As part of a large-scale forecasting project, researchers have identified the characteristics of the very best forecasters. They break down the skills into three variables: dispositional, situational, and behavioral. We believe these are consistent with high RQ. Here are some highlights:

- **Dispositional**
  - Engage in inductive reasoning
  - Exhibit cognitive control
  - Comfortable with numerical reasoning
  - Actively open-minded
  - Have a limited need for closure

- **Situational**
  - Trained in probabilistic reasoning (understand base rates)
  - Effective working as part of a team

- **Behavioral**
  - Growth (versus fixed) mindset
Appendix: Keeping Score with Brier

Psychologists commonly use the Brier score as a method for gauging the accuracy of probabilistic forecasts. Glenn Brier, a meteorologist, developed the score in the 1950s.\(^{28}\) In its simplest form, the Brier score measures the square of the forecast error, or \((\text{forecast} - \text{outcome})^2\). For binary events, the value of the outcome is 1 if the event occurs and 0 if it does not. As in golf, a lower score is better.

You can express a Brier score either on a scale of 0 to 1, or 0 to 2, depending on the calculation. We follow Brier’s original approach and place our results on a scale of 0 to 2. When calculating the Brier score this way, you consider the squared forecast error for both the event and the non-event.

Exhibit 8 shows a meteorologist’s probabilistic forecasts for whether it will rain over the next four days. For example, on Day 2, she forecasts an 80 percent probability that it will rain. Likewise, we can say she forecasts a 20 percent probability that it will not rain. Because it did rain, we place a 1 in the outcome column below “Rain” and a 0 in the “No Rain” column. Her Brier score for that day was 0.08. For multiple forecasts, the overall Brier score is the mean of the scores for each forecast. The meteorologist’s overall Brier score comes to 0.25.

Exhibit 8: Distribution of Subjective Probability of Correctness

<table>
<thead>
<tr>
<th>Day</th>
<th>Rain</th>
<th>Outcome</th>
<th>No Rain</th>
<th>Outcome</th>
<th>Brier Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Forecast</td>
<td>30%</td>
<td>0</td>
<td>70%</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>= (0.3-0)^2 + (0.7-1)^2</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>80%</td>
<td>1</td>
<td>20%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>= (0.8-1)^2 + (0.2-0)^2</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>60%</td>
<td>0</td>
<td>40%</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>= (0.6-0)^2 + (0.4-1)^2</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>100%</td>
<td>1</td>
<td>0%</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>= (1.0-1)^2 + (0.0-0)^2</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Credit Suisse.

The scale from 0 to 2 has a nice feature. Random guesses have a Brier score of exactly 0.50. Exhibit 9 shows the Brier scores for an event that occurs (“Rain”) for subjective probabilities from 0 to 100 percent.
Exhibit 9: Brier Scores of Event That Occurs for Various Subjective Probabilities

Source: Credit Suisse.

Exhibit 10 shows the distribution of the Brier scores for the nearly 2,000 people who took the test. Using this scale, Brier scores below 0.25 are very impressive. But as we can see, the percentage of the population that can operate at that level is small.

Exhibit 10: Brier Scores of 1,985 Participants

Source: Credit Suisse.
Endnotes


6 Stanovich, 15.


12 Thanks to Andrew Mauboussin for building the site and gathering the results.


14 Don A. Moore, Samuel A. Swift, Angela Minster, Barbara Mellers, Lyle Ungar, Philip Tetlock, Heather H.J. Yang, and Elizabeth R. Tenney, “Confidence Calibration in a Multi-Year Geopolitical Forecasting Competition,” Working Paper, April 16, 2015. These researchers found a lower level of overconfidence than in our sample. The subjective probabilities for these participants were 65.4 percent and the outcomes were 63.3 percent. But they did find a larger gap at high levels of confidence.


17 Stanovich, 108.


Important information

This document was produced by and the opinions expressed are those of Credit Suisse as of the date of writing and are subject to change. It has been prepared solely for information purposes and for the use of the recipient. It does not constitute an offer or an invitation by or on behalf of Credit Suisse to any person to buy or sell any security. Nothing in this material constitutes investment, legal, accounting or tax advice, or a representation that any investment or strategy is suitable or appropriate to your individual circumstances, or otherwise constitutes a personal recommendation to you. The price and value of investments mentioned and any income that might accrue may fluctuate and may fall or rise. Any reference to past performance is not a guide to the future.

The information and analysis contained in this publication have been compiled or arrived at from sources believed to be reliable but Credit Suisse does not make any representation as to their accuracy or completeness and does not accept liability for any loss arising from the use hereof. A Credit Suisse Group company may have acted upon the information and analysis contained in this publication before being made available to clients of Credit Suisse. Investments in emerging markets are speculative and considerably more volatile than investments in established markets. Some of the main risks are political risks, economic risks, credit risks, currency risks and market risks. Investments in foreign currencies are subject to exchange rate fluctuations. Before entering into any transaction, you should consider the suitability of the transaction to your particular circumstances and independently review (with your professional advisers as necessary) the specific financial risks as well as legal, regulatory, credit, tax and accounting consequences. This document is issued and distributed in the United States by Credit Suisse Securities (USA) LLC, a U.S. registered broker-dealer; in Canada by Credit Suisse Securities (Canada), Inc.; and in Brazil by Banco de Investimentos Credit Suisse (Brasil) S.A.

This document is distributed in Switzerland by Credit Suisse AG, a Swiss bank. Credit Suisse is authorized and regulated by the Swiss Financial Market Supervisory Authority (FINMA). This document is issued and distributed in Europe (except Switzerland) by Credit Suisse (UK) Limited and Credit Suisse Securities (Europe) Limited, London; Credit Suisse Securities (Europe) Limited, London and Credit Suisse (UK) Limited, authorised by the Prudential Regulation Authority (PRA) and regulated by the Financial Conduct Authority (FCA) and PRA, are associated but independent legal and regulated entities within Credit Suisse. The protections made available by the UK’s Financial Services Authority for private customers do not apply to investments or services provided by a person outside the UK, nor will the Financial Services Compensation Scheme be available if the issuer of the investment fails to meet its obligations. This document is distributed in Guernsey by Credit Suisse (Guernsey) Limited, an independent legal entity registered in Guernsey under 15197, with its registered address at Helvetia Court, Les Echelons, South Esplanade, St Peter Port, Guernsey. Credit Suisse (Guernsey) Limited is wholly owned by Credit Suisse and is regulated by the Guernsey Financial Services Commission. Copies of the latest audited accounts are available on request. This document is distributed in Jersey by Credit Suisse (Guernsey) Limited, Jersey Branch, which is regulated by the Jersey Financial Services Commission. The business address of Credit Suisse (Guernsey) Limited, Jersey Branch, in Jersey is: TradeWind House, 22 Esplanade, St Helier, Jersey JE2 3QA. This document has been issued in Asia-Pacific by whichever of the following is the appropriately authorised entity of the relevant jurisdiction: in Hong Kong by Credit Suisse (Hong Kong) Limited, a corporation licensed with the Hong Kong Securities and Futures Commission or Credit Suisse Hong Kong branch, an Authorized Institution regulated by the Hong Kong Monetary Authority and a Registered Institution regulated by the Securities and Futures Ordinance (Chapter 571 of the Laws of Hong Kong); in Japan by Credit Suisse Securities (Japan) Limited; elsewhere in Asia/Pacific by whichever of the following is the appropriately authorized entity in the relevant jurisdiction: Credit Suisse Equities (Australia) Limited, Credit Suisse Securities (Thailand) Limited, Credit Suisse Securities (Malaysia) Sdn Bhd, Credit Suisse AG, Singapore Branch; and elsewhere in the world by the relevant authorized affiliate of the above.

This document may not be reproduced either in whole, or in part, without the written permission of the authors and CREDIT SUISSE.

© 2015 CREDIT SUISSE GROUP AG and/or its affiliates. All rights reserved