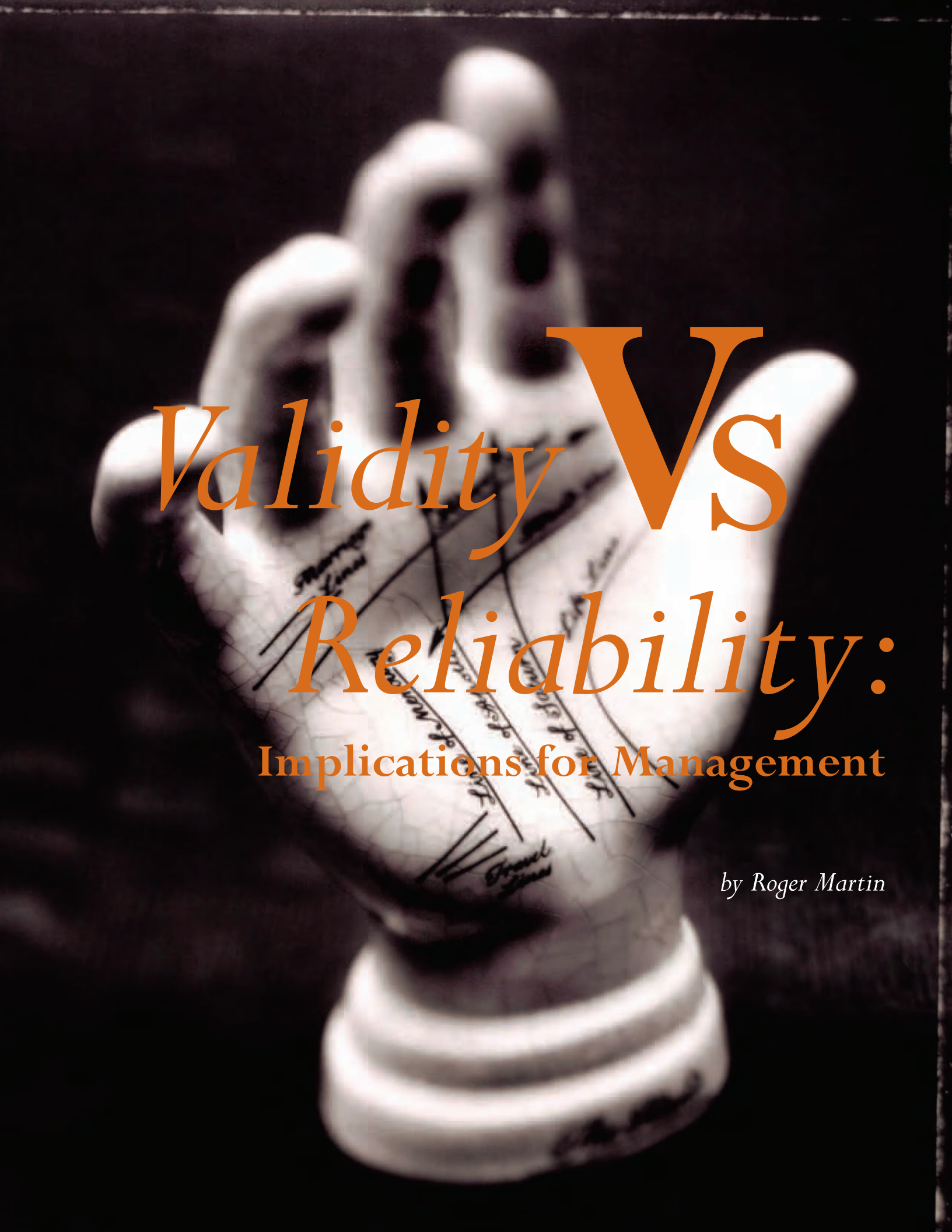

Validity Vs Reliability:

Implications for Management

by Roger Martin

Try as they might, organizations cannot predict the future. Instead of wasting their efforts, they should focus more on employing the vision, judgment and insight of their human capital.

Roger Martin

As the modern economy has evolved, business has increasingly used management science and information technology in an attempt to generate *reliability* – predictable outcomes on a consistent basis.

Examples from recent business history include Enterprise Resource Planning (ERP) systems, which keep track of corporate data in a single database and spit out comprehensive reports on inventory levels, sales by product, etc.; Customer Relationship Management (CRM) systems, which ensure that a company knows exactly who its customers are, what each one is buying, and what more it could sell to them; Six Sigma programs, which knock the waste out of an organization's systems; and Knowledge Management Systems (KMS), which attempt to organize all the knowledge in a corporation. These and other tools are intended to enable the modern corporation to crunch data objectively and make 'scientific' predictions about the future – all part of the quest for reliability.

Sadly, despite these efforts, there is little evidence that the ability of today's organizations to accurately understand the world and predict the future has increased one iota. Massive spending on these systems has not prevented corporations from wandering off the beaten strategic path, or being ambushed by new competitors and changing markets, and I would argue that the reason for this is a natural tension between the pursuits of reliability and validity.

Reliability seeks to produce consistent, predictable outcomes by utilizing a system that is restricted to the use of objective data – for instance, predicting a customer's future purchases by using data collected in their Customer Information File (CIF) in a CRM system. To produce the highest reliability possible, a system must stick to quantitative, objective data and use of the data that does not involve judgment, because blending subjectivity and judgment leads to inconsistency. Considering judgmental factors such as the mood of the customer or their attitude towards new products would be seen as an abomination in a reliable system.

Validity, on the other hand, seeks to produce outcomes that meet the desired objective, even if the system employed can't produce a consistent, predictable outcome. Proponents of ERP, CRM, TQM and KMS would all argue that their systems advance not only the cause of reliability, but of validity. In reality, however, past a certain point, the pursuit of more reliability actually reduces validity, and vice versa.

Pursuit of more reliability entails dropping variables – like customer mood – that can't be objectively measured, and eliminating any judgment from the system that would cause one user of the system to come up with a different answer than another. Pursuit of more validity means adding 'squishy' variables and applying judgment – or 'gut feel'.

Validity and reliability anchor down opposite ends of a spectrum that defines how systems are conceived and solutions are framed. Individuals themselves rarely have a balanced perspective, but rather a pre-disposition toward either validity or reliability. The existence of this spectrum and the pre-disposition of individuals combine to create significant challenges for the management of organizations.

Validity and reliability anchor down opposite ends of a spectrum that defines how systems are conceived and solutions are framed.

Validity vs. Reliability: The Tension

A perfectly reliable system is one that produces an identical output each time if the same inputs are introduced to the system repeatedly. For example, a perfectly reliable blood-testing procedure would produce the same test results each of 100 times if a blood sample were divided into 100 portions and tested successively using the procedure. A perfectly reliable political poll would produce the same result from five different random samples of voters.

A perfectly valid system is one that produces a result that is shown, through the passage of time, to have been correct. A valid blood test, therefore, is one that assesses that the subject has Hepatitis B, and the subject indeed goes on to develop symptoms that confirm the Hepatitis B assessment. A perfectly valid political poll would predict in advance the winner of the election.

Clearly, it would be optimal to have both validity and reliability in any system, and a system that reliably spews out valid answers is the paragon. Unfortunately, in all but the most simple systems and processes, validity and reliability are traded off for one another: to achieve high validity, a system must take into account a high degree of complexity fea-

turing many variables and judgmental measurement of variables; and to achieve high reliability, the number of variables has to be reduced and the measurement of the variables standardized.

The systems for measuring intelligence provide an illustration. The leading intelligence tests – such as **Stanford-Binet** and the **Wechsler Intelligence Scale** – are designed with high reliability in mind. That is to say, if the same person retakes the test multiple times, they should score remarkably consistently each time – an outcome that confirms a highly reliable procedure. However, to achieve the reliability, the tests are restricted to a fairly narrow set of measurements and measurement methods – mainly multiple-choice questions.

Since intelligence is a complicated issue, these 'reliable tests' produce results that are increasingly seen as invalid. Indeed, many universities are dropping them from their application requirements. Harvard College, for one, finds little correlation between aptitude tests and performance in college – or anything else for that matter.

In *Emotional Intelligence*, **Daniel Goleman** argues against the validity of IQ testing and for the validity of a more broad-based measure – emotional intelligence, or 'EQ'. In his view, the much more complicated EQ measure correlates more highly with success in life than IQ. However, the reliability of establishing EQ would be much lower than for IQ, because the EQ testing considers many more variables and requires subjective judgment of those variables.

In all but the most simple systems, reliability drives the exclusion of variables and judgment-free measurement, while validity drives the inclusion of variables and judgmental measurement. The tension is most often resolved in favour of the reliable solution for two reasons: comfort in the idea of 'proof', and aversion to bias.

Reliability is more readily proved than validity because reliability can be demonstrated by looking at the past. If a system has been shown to produce a consistent result over time, it can be judged to be reliable, and the past data can be adduced to confirm the reliability. Validity, on the other hand, can only be demonstrated with certainty by waiting for the future to provide confirmation. A 'track record' of past validity in similar situations can be pointed to (e.g. "My polling firm has called the Presidential winner in the last six elections"), but that still doesn't prove that future validity is ensured.

Reliability drives the exclusion of variables and judgment-free measurement, while validity drives the inclusion of variables and judgmental measurement.

Thus when a system purported to be reliable is pitted against a system purported to be valid, the proponents of reliability can often prove their claim, while the proponents of validity can only assert theirs. Even though the value of the validity claim is inherently higher (i.e. predicting the future) than the value of the reliability claim (i.e. producing a consistent result), the reliability claim often wins because of the greater ability to prove a lower value claim.

The second reason for favouring reliable systems over valid ones is aversion to bias. Valid systems typically require use of subjective or qualitative data in order to improve their predictive power, while reliable systems forego the use of such data in order to maintain reliability. The requirement to use judgment to process the subjective/qualitative data opens up the validity-oriented system to the charge of bias, which can never be fully refuted. In order to eliminate the potential for bias, the system must consider only variables in which judgment-free assessments are possible. This typically undermines the validity of the system, weakening its attractiveness in comparison to the 'reliable' alternative.

An example of the aversion to judgment-based bias is the **Hay System**. The most commonly-used compensation system in corporate America, the Hay System determines compensation levels on the basis of a limited number of quantifiable, verifiable variables, such as number of direct reports, the size of the budget managed, the amount of revenues under authority, and an extensive database of norms.

Remarkably few employees of companies that use it believe that the system produces valid results, i.e. that employees are compensated in proper proportion to the value they create. Despite its lack of popularity, it remains the most prevalent system because it is free of judgmental bias and is deemed reliable: with the same quantifiable inputs, the system will produce an identical compensation output regardless of the time and place. And because the lack of validity is considered a 'lesser sin' than the potential for judgmental bias, the Hay System prevails.

Individual Tendencies

Like systems, individuals also tend to favour either reliability-oriented systems or validity-oriented systems consistently across multiple types of situations. When faced with the trade-off between validity and reliability, some will always sacrifice incremental validity to secure incremental reliability, while others will always sacrifice incremental reliability to secure incremental validity. The latter type of individual is inclined to be bothered by smoothly-running systems that appear unable to consistently produce outputs that are shown to be valid. These validity-oriented individuals will tend to instigate change even if those changes results in the dismantling of a long-standing, even beloved, system. The reliability-oriented individual is likely to see such a move as rash, dangerous and imprudent, and will be inclined to lobby against it.

Validity-oriented individuals exhibit a number of characteristics. They are comfortable processing a wide number and type of variables in the course of making a decision. They are more comfortable with subjective or qualitative data, and in making inferences on the basis of such data. They process data in a more intuitive fashion and are comfortable following a heuristic

process (i.e. a general sense of 'how to proceed' from start to finish, rather than a set algorithm.) They are highly accepting of changes in procedures and processes, and in fact are frequently the instigators of such changes. They tend to be quite comfortable predicting the future based on their processing of data. They collect and process all types of data, not content until they reach a sense of comfort that a valid decision can be made. Thereafter, they are prone to act decisively and confidently.

The validity-oriented individual exhibits a blindness toward the value of reliability and reliable systems. In the quest for validity, he or she is inclined to ignore or forget the need to ensconce the valid solution in a minimally-reliable, replicable system. They are also likely to over-estimate the capability of others to replicate their bias-free assessment of qualitative data and show little concern for the worries of others about bias. These individuals are inclined to be frustrated by the inability of others to see the validity they see, and tend to be impatient with their less validity-oriented colleagues. Not known for building systems and processes, they are inclined to get into trouble by under-investing in control procedures.

Reliability-oriented individuals also exhibit a predictable set of tendencies. They focus on processing a limited number and narrow range of variables in the course of making a decision. They are more comfortable with objective, quantifiable, verifiable data, recognizing the value of such data in avoiding both bias and the ensuing unpleasant arguments to which the appearance of bias can give rise. They process data analytically and prefer to reason in a systematic, rigorous fashion. They work hard to embed their decisions in procedures and processes, which they are loath to change once smoothly running. They tend to insist on waiting until the appropriate process produces its output before making a decision. They build systems and processes with enthusiasm and affection, in an attempt to ensure that their domain is 'in control'.

These individuals often exhibit a blindness toward the shortcomings of their system and will wait overly-long to change a system that produces reliable but increasingly invalid results. That is, in their quest for reliability, they are inclined to ignore the cost of invalidity. They are also likely to exaggerate the tendency of others toward bias and are willing to take extreme measures to guard against bias. They are inclined to be frustrated by those who make what they see as 'snap decisions' based on soft data, who show little respect for smoothly-running systems. Indeed, they are inclined to get into trouble by overprotecting the systems and processes that they create.

Neither extreme validity-orientation nor extreme reliability-orientation is healthy for any organization.

Implications for Management

There are three key implications of the validity versus reliability tension for today's managers.

1. ***A strong management team should have a balance of validity-oriented and reliability-oriented members.***

   Neither extreme validity-orientation nor extreme reliability-orientation is healthy for an organization. A firm needs a balance of both orientations to ensure both that the right things are being pursued and that systems are being built to enable the organization to perpetuate the right decisions. The best way to ensure such a balance is to have a mix of validity and reliability-oriented executives on the senior management team.

   Achieving this balance is not a trivial task, in part because validity-oriented executives and reliability-oriented executives tend to irritate and frustrate one another. They tend to question the motives and even intelligence of one another because they frequently come out on opposite sides of issues, which can result in a leader forming an unbalanced team in his or her own image.

   The first step is for managers to understand both themselves and the natural tendencies of their colleagues in order to take the mystery out of clashes that are in fact rather formulaic and predictable. The second step is to recognize the value of the other orientation in discerning blind spots and taking actions that would seem unnatural given the executive's own orientation. When this level of appreciation is built, a diverse management team can be formed, and it can work

together to produce better results than a highly homogenous team ever could.

2. *The threat of creeping reliability is greater than the threat of excess validity.*

Due to the comfort of 'proof' and the aversion to bias discussed earlier, there is a general tendency in organizations for reliability to outflank validity over time. Reliability can have the creeping effect of establishing systems that err on the side of reliability and become ensconced as the operating systems of the firm. Once in place for long enough, these reliable systems raise the burden of proof for changing them by producing ever more data that demonstrates their reliability. Validity-based arguments can be more easily dismissed in this environment, resulting in the slow-but-steady ossification of the firm's systems.

For this reason, it is incumbent upon top management to act as a consistent, mild advocate for validity. Note that he or she should not be a strong advocate, lest the balance be threatened. Consistent but mild advocacy for validity and against the ossification of reliable systems by a firm's leaders can offset the disequilibrium created by the comfort of proof and the aversion to bias.

3. *Keep in mind that validity and reliability trade off for one another.*

In all but the simplest systems, processes and decisions, additional validity comes only at the price of diminished reliability, and additional reliability at the price of diminished validity. This should be kept in mind as decisions are made and systems built in organizations so that accidental, implicit decisions are not made. It is most important for validity-oriented managers to remind themselves to think about the reliability consequences of added validity, because this is likely to be a natural blind spot for them. Conversely, it is most important for reliability-oriented managers to remind themselves to consider the validity consequences of added reliability, because that is likely to be their natural blind spot.

Exhibit One: Reliability vs. Validity

Reliability	versus	Validity
Enterprise Resource Planning		Robust Strategy
Customer Relationship Management		Customer Intimacy
Six Sigma/ Total Quality Management		Design Excellence
Knowledge Management System		Creativity
Incentive Compensation		Job Meaning
Shareholder Value Maximization		Corporate Social Responsibility
Meeting Analyst Quarterly Targets		Running a Successful Company

Conclusion

Over the past 20 years, corporations have obsessed about reliability so much that it has driven their core procedures away from validity, and there are signs that they should shift the balance more towards validity.

ERP is no substitute for having a robust strategy. CRM is no substitute for what an organization really wants, which is customer intimacy – in fact, it can have the opposite effect, which is customers who think they are buying from Big Brother. Six Sigma/TQM will drive out known waste in the current paradigm, but won't result in business design excellence. KMS will organize all the knowledge in a corporation (sort of), but it won't produce creativity. Having a carefully-crafted incentive compensation system aligned with the corporate goals won't produce meaning in employees' work lives. Pursuing shareholder value appreciation using Economic Value Added (EVA) systems won't result in a corporation that is seen as demonstrating Corporate Social Responsibility. And reliably meeting the analysts' quarterly earnings targets won't result in a successful corporation over time.

For organizations that have been focusing on the left side of **Exhibit One**, above, all of the desired outcomes on the right side of the chart require systems and skills they have not been building and nurturing. Organizations not only have to swing the pendulum toward the more validity-oriented processes on the right side of the chart, they will need to nurture validity-oriented people, which may be even more challenging, because individuals rarely have a balanced perspective, but rather a pre-disposition toward either validity or reliability.

The first step towards achieving balance is for firms to recognize that they have favoured reliability-oriented systems and individuals for too long, unknowingly crowding out subtlety and judgment. To succeed in the next 20 years, they will have to figure out how to become a more welcoming place for validity-oriented systems and for people who are more comfortable with handling 'fuzzy' data and using their judgment.

Harvard Business Review *has named Roger Martin's 'Validity vs. Reliability' concept one of its "Breakthrough Ideas for 2005."*