

---

# Why It Pays to Be Sceptical About Behaviour Tests



By Felix Jan Nitsch , INSEAD

**Using such techniques to assess individual personality traits such as rationality might overstate their potential.**

Recent years have seen a sharp rise in interest in the field of behavioural economics, which studies how psychological, cultural and social factors can potentially influence people's actions and behaviour. This line of thinking has entered the mainstream through popular books like 2011's bestselling *Thinking Fast and Slow* by Daniel Kahneman, which explains how cognitive systems shape how we think and guide our everyday decision-making.

Academia has also seen a growing body of research, much of it promoting behavioural economics assessment tools as a means of **objectively measuring a range of different traits**. These tools claim to assess everything from levels of risk aversion to trustworthiness, or even the ability to demonstrate sound rational thought – all based on an individual's response to a specific test.

In 2014, researchers conducted a **large-scale study** among the Dutch population, which examined how closely people acted according to the most

objective economic choices. They then related these results to demographic variables (from age to education and wealth) to determine if the choices made differed among different segments of society. The findings revealed that people of lower socio-economic status seemed to perform worse in objective decision making.

Clearly both policymakers and industry have a keen interest in understanding which individuals make rational decisions, be it to design effective policies or improve talent selection processes. But while there has been much attention devoted to the potential of behavioural economic tools, there has been a lot less research into the veracity of the results derived from using such tools.

## **Testing the tests**

My [recent research](#), conducted with colleagues Luca M. Lüpken, Nils Lüscho, and Tobias Kalenscher from the Institute of Experimental Psychology at Heinrich-Heine-University, suggests that such bold claims might actually overstate the case. We ran three experiments to test the ability of individuals to think rationally and in addition, re-analysed data from other studies ranging from risky monetary choices to dietary decisions. Overall, our analysis revealed that the ability of such tests to accurately distinguish between different individuals could only be described as moderate to poor.

The aim was to test their levels of economic rationality, based on the accepted premise that decision makers consistently choose the best option according to their preferences as their budget allows. Indeed, the results showed that most participants (nearly 80%), across all datasets behaved with high consistency.

## **Majority outcomes**

The outcomes were consistent independent of the choice type (social choice, food choice, choice under risk, or choice under ambiguity), the complexity of the choices offered, as well as variables such as study population, sample size, task structure and the duration of the study. The fact that these results also remained consistent regardless of the length of time between measurements was particularly telling.

It wasn't that the results were not accurate, it was simply that they showed that people, with the exception of a few outliers, all acted in very similar

ways. All the experiments really helped to confirm were that most people are pretty rational and make consistently rational decisions. They weren't able to help us decide if one person is smarter or more capable than the other.

This was because the individual differences in rationality were so insignificant, even in large, representative samples, that they couldn't be distinguished from small random variations. Consequently, even with current measurement techniques, it is still difficult to identify individual characteristics linked to rationality. In fact, taking the population average essentially offers a better prediction of individual performance than an individual measurement taken only 30 minutes before.

While our research was specifically focused on economic games that addressed rational thinking, **additional research** analysing other behavioural tests for other traits such as **risk attitudes**, or **loss aversion** appeared to have similar outcomes: the results of the majority of people were close. These studies affirmed that most people behave in very similar ways and scored very similar results in these tests.

Of course, that's not to discount them as having no value at all. Our research was able to show that such economic tests can certainly be used to help identify patterns of behaviour. It's just that our research suggests that claiming that they can help draw conclusions about specific individual's behaviour is overstating things.

### **Consequences of low reliability**

What do our results mean in practice, for example, in selecting candidates for a role? The impact of using low reliability tools can be illustrated by three qualitatively different consequences. First, organisations hire candidates based on less informative data, potentially leading to suboptimal hires. Second, candidate rankings are not reproducible. Third, HR professionals' awareness of their own biases might be wrongfully pacified by the supposed objectivity of the test, which in turn can increase unconscious biases in the selection processes.

The magnitude of these consequences depends on the importance or weighting given to these results in the final selection decision. The awareness of potential shortcomings in any personnel selection tool marks the first step towards their mature use in practice.

Unfortunately, these tests are no silver bullet. Behavioural economic tools, like many other methods out there, can certainly help identify patterns of behaviour. But they should not be used in isolation to make major decisions such as a government deciding on an education policy or to determine which person gets a top job. Instead, they should be used in conjunction with multiple forms of assessment, so that decision makers can compare and contrast the different outcomes and see whether they align. Ultimately, they should be used with caution, with the understanding that it can be tricky to make detailed conclusions about how specific people will behave.

#### **Find article at**

<https://knowledge.insead.edu/career/why-it-pays-be-sceptical-about-behaviour-tests>

---

#### **About the author(s)**

**Felix Jan Nitsch** Felix Jan Nitsch is a Postdoctoral Research Fellow in Marketing at INSEAD, where he studies consumer behaviour under stress. His research focuses on transformative consumer research. Specifically, how stress and mental well-being affect consumption and what interventions and nudges support consumers in their decisions.

---

#### **About the research**

**“On the reliability of individual economic rationality measurements”** is published in the *Proceedings of the National Academy of Sciences* (PNAS).