The predictive stall in selection: Why – and is there anything we can do about it?

Andrew Munro

In the 1980s psychometric testing rediscovered its mojo. The growing evidence base, a combination of greater test usage within the workplace and the emerging wave of meta-analytical studies (Hunter & Hunter, 1984) reassured us that the systematic application of the psychometric method would deliver greater predictive power in selection. The potential gains were significant, delivering sizeable improvements in employee productivity and reductions in recruitment costs.

Thirty years on, however, it is a puzzle that – despite greater sophistication in test design, access to technological innovation, and improvements in statistical methodology – predictive power has not improved. Instead, there is reason to think that in some instances, for example, assessment centres (Thornton, 2009), validity is in decline. So what underpins this predictive stall? This article outlines seven different arguments to point to the possible causes and goes on to propose options to address the stall.

1. We were wrong in the first place
This position suggests that predictive power was in fact never all that impressive, and points to the methodological shortcomings of meta-analytic studies (Shercliffe et al., 2009). This argument also looks to retain the conventional definition of predictive validity (‘a comparison of test data at Time I with an outcome at Time 2’) to highlight the problem of over-generalising from concurrent studies to assume the conclusions will be replicated with applicants in a selection scenario. Concurrent studies may point to genuine predictive validity for some types of test, for example general mental ability, but seem problematic for self-report measures of personality and motivation. The lack of validity evidence from genuinely predictive research designs for these tests continues to represent something of a psychometric embarrassment.

2. Shortcomings in selection practice
This position suggests that the stall is a consequence of a lack of sophistication in the use of psychometric data within selection applications. Psychometric testing has made predictive improvements; unfortunately these potential gains are being washed out through falling standards and poor practitioner practice that are failing to optimise test data in selection decision making.

3. VUCA makes working life less predictable
We shouldn’t expect anything other than a predictive stall in what is fashionably known as Planet VUCA, the world of increasing volatility, uncertainty, complexity and ambiguity. Psychometric testing may have provided predictive power in the relatively stable organisational world of well defined roles and clear performance standards. In today’s fast moving working environments – of shifting structures and changing organisational cultures – it is difficult to know what is being predicted, never mind how best to predict it. As Michael Maubossin (Mauboussin, 2012) has pointed out, ‘complex adaptive systems effectively obscure cause and effect. You can’t make predictions in any but the broadest and vaguest terms.’
4. The savvy candidate throws a spanner in the works
The stall is explained because applicants have shifted from compliant candidates to shrewd game players. This argument suggests that test publishers are being out-maneuvered by a generation of applicants adept in social media to exchange experiences about interview methodology, assessment centre protocols or the goldilocks tactics of personality test completion. It also points to an escalating ‘arms race’ in which any innovation by the test publishers is countered quickly by applicants who find stratagems to mitigate their effect.

5. Flawed assumptions of the right stuff
There is a stall because we assumed that prediction in selection applications was a relatively easy project based on the identification of the ‘right stuff’, a handful of stable attributes transferable across many roles and organisations. And whilst it may be true that a robust measure of general mental ability, conscientiousness, open mindedness and the absence of neuroticism may be a reasonable fall-back position for many roles, it also seems that performance is more contextual than we often assume (Groysberg, 2011). The forecast of performance requires a more sophisticated understanding of the dynamics of work context (e.g. the relative impact of in-depth expertise and knowledge, or access to networks of influence) if we are to improve on current levels of predictive power.

6. The incompetence of competency
Our validation efforts, often frustrated by a lack of access to objective measures with direct organisational impact, have come to rely on competencies – typically a set of 10-15 dimensions of effectiveness – as the criteria of work outcomes. Here the convention has developed that predictive validity is established if reasonable correlations can be demonstrated between test data and competency evaluations. But as the critics have pointed out (Buckingham & Coffman, 2005; Bolden & Gosling, 2006) it is not clear that competency frameworks do in fact predict meaningful long-term organisational outcomes.

7. Reliance on past achievements.
There is a stall because we thought the predictive prize was already won, and little need for the long-term research programmes established in the 1960–1970s (i.e. Bray, 1964). This is the argument that the psychometric industry is resting on the achievements of a relatively small number of old studies. Longitudinal studies are a challenge and if the response is that the research logistics are now too complex we have to revisit the meaning of predictive validity within organisational life in the 21st century. No doubt each of these arguments has more or less merit depending on the specific type of test, its application and selection context. But there is a growing sense of practitioner unease about the widening gap between the popularity of psychometric testing and the evidence base of demonstrated predictive value in selection.

Responding to the predictive stall
Professional humility about the current limitations of psychometric testing would be a good start. This is not to downplay the genuine achievements made and the contribution of well-targeted tests within selection. But it is to note that many of the marketing claims made by the publishers, distributors and psychometric-based consultancies suggest a level of confidence about the selection benefits not supported by a recent evidence base. Secondly, there is a need for better understanding of organisational context and the factors which indicate when tests are more, or less, likely to improve selection. This is not simply the argument for less reliance on the abstractions of validity coefficients from the meta
analyses literature and greater rigour to identify the incremental improvements on selection base rates, within real world applications. It is to propose we need a better understanding of organisational dynamics, and the evaluation of the interplay of strategy, structure and culture (Munro, 2013) that indicate the circumstances when testing can and can’t be expected to have predictive power in selection. This is to think strategically about the nature of work, the organisational climate, and the impact of change and avoid the ‘one size fits all’ application of psychometric testing. Thirdly, greater attention should be given to candidate adaptability and versatility. The focus on identifying the causal attributes of the right predictive stuff largely neglected the impact of the capacity of individuals to change and develop in response to experience. If the predictive enterprise has to accommodate greater levels of future organisational change, we need a shrewder insight into how the dynamics of consequence affect cause, to identify which types of experience, have more or less impact, in shaping the capacity to adapt to new and different challenges. Finally, we should encourage greater use of pilots within selection scenarios. Apart from providing an important check that the predictive claims made from concurrent studies hold up in applicant samples over time, pilots allow the experimentation that generates learning to keep adapting selection systems to optimise the organisational impact.

Summary
Prediction is the goal of the scientific enterprise. From initial ideas to theory testing to explanations of the dynamics of cause and consequence, prediction makes the forecast of future outcomes possible. Improvements in the accuracy of these predictions are a test of the progress a scientific discipline is making. It seems that psychometric testing is stuck and needs to revisit either its predictive claims or the research programmes it is using to build evidence of its predictive power in response to the selection challenges of the 21st century.

References

The author
Andrew Munro MA, C Psychol is a director and founder of AM Azure Consulting.