Shorter Personality Questionnaires: A User's Guide Part 2

In this two part series, James Bywater and Anna Brown summarise some of the issues involved in determining the correct length of assessment in a personality questionnaire (PQ). In the last edition of *Assessment & Development Matte*rs they discussed the general issues that test designers face, and in this one they cover some more modern solutions to these.

It is aimed at practitioners rather than hard core psychometricians and can not be exhaustive. However wherever possible it attempts to distil out practical messages for the audience.

Introduction

There has always been a compromise between the length of a personality questionnaire and its psychometric properties. There are a number of modern innovations to address this problem but they also contain assumptions and compromises of a different sort.

Innovations in PQ Design

There are a number of innovations in assessment that have become possible with online and computer administration. Old scoring mechanisms designed for manual scoring or basic PC's were necessarily simple and they are only just starting to be replaced by more sophisticated methods. Item Response Theory (IRT), Linear on the Fly Testing (LoFT), Computer Adaptive Testing (CAT) and Dynamic / branching questions are all well-established methods that were not practical or possible to implement until very recently. In general the response from candidates and users to these innovations has been positive. Users like the reduction in completion time, but remain as mystified by how these questionnaires are scored as they ever were.

The first author was doing some feedbacks on the new shorter OPQ32r recently for some students, and asked them how they felt sitting this new, simpler and shorter (25 minutes long) questionnaire. Their feedback was definitely an improvement on the comments we used to get from OPQ32i (more questions, 50 minutes long) but these Gen Y students were still masterfully understated:

"well, I was not completely bored"
"the questionnaire was fair"
"it was good to really think about myself and be forced to choose"

Similarly, despite some benefits in reducing the test length and improving precision by using Computer Adaptive Testing (CAT), the reactions from candidates can be rather negative. It is hard for them to see (not knowing all the technical detail) how it can be fair that every candidate receives a different number of questions: being given fewer questions than your friend might be seen as not being given the same opportunity to express yourself. LoFT techniques can be used to create shorter tests than classical methods that also vary the content between candidates. At the same time they may seem 'fairer' to candidates in that they keep test length and difficulty the same for everyone.

Top Tips:
- You know much more about the questionnaire than your candidates. Just because you know a questionnaire is valid, reliable and fair, does not mean

the candidates think so too.

Under the Bonnet
Although they differ in approach, these new scoring methods are all attempting to do the same thing: they are attempting to get more information from fewer questions. If we accept that this is an emergent trend, then this raises some decisions for the typical HR user of the tools: Which approach works better? How far can it be pushed?

It can be hard for the non-expert user to differentiate between these different methods for shortening questionnaires. Online, automated scoring has the effect of distancing the user from the results. It can feel like moving from a carburettor to a fuel injection system in a car – it pushes the processing and mechanisms so far "under the bonnet" that it can be hard for the lay person to "tinker." Many expert users mourn this loss.

To summarise, the science behind psychometric tests has become highly technical. To make matters worse, no clear guidelines yet exist for tests based on new methodology (such as IRT-scored personality tests) and what requirements they must satisfy to pass certification criteria. Relying on old classical methods for estimating quality of tests created with new methods in many cases simply won't do.

Top Tips:
- Research these new techniques. They are here to stay.
- Review the manuals – ask yourself: "is this claim sufficiently substantiated?". Check references and, if in doubt, request more information.
- Where have the savings been achieved? What is the downside to this approach?
- How will your candidates feel about the tool (e.g. CAT may mean that they sit different length tests from each other)
- Check out the BPS reviews at http://www.psychtesting.org.uk/. What do the experts say?
- Is it a practical HR proposition? – does it possess the languages, IT integration, scalability, legal compliance and IT security that you require? Sit the questionnaire yourself – this can be a helpful way to get to grips with potential issues although it is not a substitute for the relevant technical information.
- Have the scales got good breadth of content or do they suffer from narrow repetition ("bloated specifics")?
- Has the questionnaire been validated across sample/ country/ context?

Finally
Once you have implemented a test or assessment system, go back and review it. It is useful to administer and interpret some assessments yourself, even if you would not normally do so. This is important as it allows you to listen to what modern job applicants say about your tools. There is an assumption that candidates have a greater sense of entitlement. It is true that treating candidates as customers is a very different mindset from typical selection processes (Bywater & Bard, 2009). However, there are arguments to suggest that candidates may get less choosy about the psychometrics and assessments that they are asked to sit due to the global economic downturn and rising unemployment.

We are at an important crossing point of internet technology, demographic change, economic reshuffling, and psychometric revolution, and the only way to keep up with

this is to stay very close to your target talent groups.

The Authors

James Bywater is VP Global Consulting Solutions and Anna Brown is Principal Research Statistician with SHL Group Ltd.

References

Bywater, J. & Bard, E. (2009). Carcinogenic Assessment: Passive aggression in the assessment relationship. British Psychological Society, Assessment and Development Review, 1 (1).