

What's Wrong with Ability Tests

*Dr Tuvia Melamed
ClearWater A&D*

One of the main disappointments in our profession is the current state of commercial Ability tests. After nearly 100 years of first introducing 'scientific' measurements of intellectual capacity, you would have expected something much better than what we currently have.

I do not challenge the available ability tests in terms of their technical capacity – i.e., my criticism is not to do with questions of quality of normative data, nor it is about the validity and reliability of the measures. I take these for granted. What has been frustrating me for the last 25 years of my practice as a business psychologist is the limited output we gain from an ability test.

Most tests requires participants to spend close to an hour responding to demanding test items; yet, the final outcome in most cases does not amount to more than single figure – normally a percentile score comparing the performance of the test taker to that of a norm group (e.g., 'you score above 63% of the norm group). This is a far cry from the rich output and interpretations you get from a personality inventory. No wonder that the British Psychological Society differentiates between Level A certificate of competence in occupational testing (ability tests) and Level B certificate that is dedicated to personality instruments. Relatively to personality instruments – there not much to say about ability tests outcomes.

Do not get me wrong – I am not challenging the value of ability tests, and I am not suggesting that they should be replaced by 'more-value-for-money/effort' personality instruments. What I am opening to discussion is the frustrating state of affairs of limited output we get from ability tests.

The closest attempt to enrich the output of an ability test that I am aware of is the work of Robert Hogan with the HBRI (Hogan Business Reasoning Inventory). The test provides two scores – Strategic Reasoning and Tactical Reasoning. These are combined to create an overall critical reasoning score. Hogan provides an interpretation of each of the two individual scores (e.g., what are the implications of scoring 78% on Strategic Reasoning) as well as a very simple 2x2 classification into a 4-type typology based on the interaction between the two scores. As such, the report is far more informative than the typical practice of providing a single score. Yet, there is so much more that can be obtained from the available data.

The followings are some possible ways to utilise ability tests better:

1. **Wrong versus poor answers:** The raw score on an ability test is made of the number of correct answers. An answer to a question item can be either 'Right' or 'Wrong'. Yet, the multiple choice approach used with the traditional psychometric approach does not differentiate between a wrong answer that is by far the worst option from a wrong answer that is closest to the correct answer. Consider the following simple example. A test item might be: 'Glove' is to 'Hand' as 'Sock' is to ____ (a) Shoe; (b) Arm; (c) Leg; (d) Foot; (e) Cupboard. Although only answer 'd' is correct, some of the answers are closer to it than others. Choosing answer 'e' indicates less ability than choosing answer 'c'. Yet, the dichotomous scoring system of tests will view both answers as equal.
2. **Time to complete test:** Most ability tests are time limited. Two candidates that take the same test and get the same number of correct answers will be considered as equal. Yet, if one took only 20 of the 40 minutes to complete the test; whereas the other took 40 minutes, there is obviously some difference in the ability of these candidates.
3. **Percent of correct answers:** Two candidates that took the same test and got the same number of correct answers will be considered equal. Yet, one may have attempted only 60% of the test items, where the other completed all items. There is a difference in the ability of the two candidates, but this is ignored when interpreting the results.
4. **Level of complexity:** The work of Elliot Jaques suggested 8 levels of complexity of mental processes that are hierarchical, yet conceptually different from one another. Yet, most tests do not differentiate among levels of difficulty or complexity, and cluster all items together for scoring purposes. Thus, the interpretation of someone who scores highly on divergent thinking items and poorly on convergent

thinking items should be different to someone who had a different pattern of scores. Yet, under the single score system; if the total number of correct answers is the same; both candidates will be viewed as equal.

5. **Interaction among sub-scores:** Some tests provide sub-scores based on content (different to complexity and difficulty from the previous point). E.g., we get a breakdown of verbal reasoning, numerical reasoning, and abstract reasoning; or a breakdown of strategic reasoning and tactical reasoning. Most test publishers leave it at that. Hardly any considers the interactions among the sub-scores. This is very different in personality assessment where the ability to interpret the interaction between two factors is considered far more valuable than the interpretation of the simple main effect of each of the two factors.

So, what I am looking for is an ability test that will provide me with richer information than simply telling me how the individual scored in comparison to the norm group.

This frustration, led me to spend the last three years developing an Ability test that utilises these ideas and provide rich and valuable output. I called the final test ***Intellecto ©***. You can find more about it on by clicking [here](#). To see an example of the output from it click [here](#).