10

Taking an Intelligence Test: Does the Context Matter?

Adrian Furnham

Introduction

People talk about intelligence all the time. Listen to the way they describe others. There are many synonyms and slang words for intelligence: ability, acumen, acuteness, agility, aptness, astuteness, braininess, brilliance, canniness, cleverness, comprehension, discernment, foxiness, insightfulness, giftedness, grasp, gumption, perspicacious, perceptive, quick-wittedness, sagacity, smartness, sharpness, talent, thoughtfulness, whiz.

Many of these terms refer to how people deal with the daily problems of life, which of course differ from time-to-time, and place-to-place. The question is whether the appropriate understanding and assessment of intelligence is, and indeed should be, influenced by culture and history.

In *Intelligence in Context: The Cultural and Historical Foundations of Human Intelligence.* Editors: Robert Sternberg and David Preiss. London: Palgrave

A. Furnham (✉)
BI: Norwegian Business School, Oslo, Norway
e-mail: a.furnham@ucl.ac.uk

R. J. Sternberg, D. D. Preiss (eds.), *Intelligence in Context*,
https://doi.org/10.1007/978-3-030-92798-1_10

Is being smart in Wall Street, very different from being smart in Walmart? Is intelligence about adapting to, and thriving in, different environments which have quite different demands and rewards?

Intelligence is a characteristic people quite easily recognize and usually admire. But is the understanding of the concept and the willingness to be tested dependent on time and place? Indeed, is the validity of IQ tests, indeed any form of assessment, completely context- dependent?

This chapter will consider what lay people believe about intelligence and intelligence testing. Do they acknowledge the role of context, and if so, how? How different is the lay view from the standard academic view (if such exists)? And what of alternative voices on this topic, many of which are to be found in this volume?

Some countries and organizational cultures have always favored testing for selection. Most militaries have, and will always use, abilities tests. The same is true of schools and universities. But also, over time, commercial and public organizations have used intelligence and other psychometric tests predominantly in selection. This is how people most frequently have their experiences of tests, though now their availability on the web and usefulness in self-awareness and development exercises mean many more people have experience of tests. However, these popular tests may be very different from the well-known and developed psychometric tests in content, administration, and feedback.

This chapter is divided into three sections. The *first* concerns the difference between one prototypic academic view and the lay view of intelligence. Some academics are strict "universalists" who underplay the role of context in definition and assessment, while others stringently reject this view. It also looks at some recent work on popular beliefs about intelligence and intelligence testing, many of which concur with those of critics of tests.

The *second* concerns issues about the perception and accuracy of intelligence tests, as opposed to other ways of assessment. It demonstrates some of the major problems concerned with using tests in the "real world," and the much about lay beliefs about testing. In short, it shows that may people are skeptical about many famous and current tests and the business of testing.

The *third* section looks at other everyday tests and markers of intelligence and what they mean. If people are skeptical about both concept and measurement of intelligence, are there other ways to measure it? Indeed, do we need to develop different assessments for different skills in different cultures?

A Big Divide?

Are lay people and "psychometrically orthodox experts" on intelligence testing in agreement? What do ordinary people believe about context and how different is this from the classic academic viewpoint?

The strict, orthodox, conservative view about intelligence goes something like this: intelligence can be measured, and intelligence tests measure it well. They are among the most accurate (in technical terms, reliable and valid) of all psychological tests and assessments. While there are different types of tests, they all measure largely the same intelligence, as psychometrically defined. IQ is strongly related, and probably more than any other single measurable human trait, to many important educational, occupational, economic, and social outcomes. A high IQ is an advantage in life because virtually all activities require some reasoning and decision-making. The odds for success in our society greatly favor individuals with higher IQs. That said, there are some confounding factors in measurement, because people with higher IQs often are given more opportunities in life, and hence society gives them more opportunities to achieve.

Some still argue that today intelligence tests are not culturally biased; rather, IQ scores predict about equally regardless of race and social class. Indeed MENSA, the high IQ society, still uses a test developed over 70 years ago to be "culture fair" (Cattell, 1949), although it is not really culture-fair. No test is, because taking a test itself is a cultural act. This test was an attempt to measure cognitive abilities devoid of sociocultural and environmental influences. This, of course, is highly contested, as this volume suggests. Many would argue passionately that social class and culture have a huge impact on educational opportunities and general socialization, which would influence how people approach and succeed in standard, Western-based, timed tests. The skills required to survive and

thrive in one society are clearly very different from one another. In this sense, this orthodox view is misleading.

In other words, this standard view suggests that intelligence is a fixed entity, conceptually represented by *g* and its subfactors, or by IQ, and operationalized by conventional intelligence tests. Further, intelligence is static—fixed across time and culture. Thus, although intelligence tests may need to be modified over time and place, the basic nature of intelligence does not change. The view is not fully adequate, but has generated a great deal of theory and assessment. In the end, no one view, at least currently, is "fully adequate" to explain and measure the clearly complex concept of intelligence.

Work psychologists argue that the single best predictor of success in complex, changing managerial jobs is intelligence. Brighter people learn faster, they have a greater store of knowledge, and they often tend to be intellectually more self-confident. They analyze problems more efficiently. Often, they are more open to new experiences. Moreover, the experts on selection argue that tests are useful, cheap, fast, easy, versatile, scorable, and understandable (Gatewood et al., 2015).

Dilchert (2018) wrote: "*It is the responsibility of IWO* [industrial, work, and organizational] *psychologists to make such evidence available and help organizations make the most responsible decision in a given context. Cognitive ability tests are among the most powerful weapons in the IWO psychology arsenal. The analogy might be crude, but it is apt. We must weigh a variety of factors regarding their deployment: effectiveness, efficiency, and consequences (including applicant reactions and workforce diversity). However, we must also consider the consequences of not deploying a reliable and valid predictor tool at our disposal—including reduced objectivity, lowered productivity, and insufficient societal benefit—especially when resources to be distributed (educational opportunities, jobs) are scarce*" (p268).

However, it is possible to argue that this is a very white, privileged-class view. How much are test scores influenced by growing up in Appalachia or inner-city Newark, NJ; rural Romania versus inner London? The experience of actually trying to assess people in different cultures and different experience is salutary and will be discussed later.

There are, however, a large number of scholars who strongly challenge these assumptions. There are those who write about culture and

intelligence (Berry, 1974; Serpell, 2000) who have data from many very different countries that demonstrate how the concept of intelligence and its appropriate assessment are time- and culture-bound.

Change in Intelligence

The Flynn effect—a worldwide increase of 30 points of IQ during the twentieth century—has caused a great debate as to why intelligence levels have been rising (Flynn, 2010, 2012; Trahan et al., 2014). Many of the arguments are contextual. Furnham (2008) summarized various ideas to account for the Flynn effect.

Education: In most countries, with every generation, people are spending longer at school and with better facilities. Schooling is compulsory and people from all backgrounds are used to learning and being tested. Intelligence is related to learning so as education is better and more widespread, scores get higher. This of course differs from country to country: hence strong context effects. These ideas are reviewed by Baker et al. (2015).

Nutrition: People are now, at least in the developed world, better nourished, particularly in childhood, which reduces the incidence of "backwardness" in the population. There are fewer people who have poor nutrition in youth, so the bottom end of the distribution is removed. This means the IQ scores are linked to the calorie counts. This issue has been discussed by Bratsberg and Rogeberg (2018).

Social trends: In the West, people are all now much more used to timed tests and performing against the clock. People are familiar with tests and testing and so do better overall. The experience of test-taking, however, is not true in many developing countries.

Parental Involvement: The idea is that parents provide richer home environments for their children and express a greater interest in their education than they used to. They have higher expectations and get involved more. The trend, but only in developed countries, is to have smaller families where parents invest more in their children may also be an important factor.

Social Environment: The world is more complex and stimulating. Modernization and new technology mean people have to manipulate abstract concepts more, as well as speedily process and store information which is essentially what traditional intelligence tests often measure. However, this inevitably differs dramatically from country to country (compare China, Chile, Cuba, and Cambodia), and may in part explain national differences when they occur.

"Experts" and the Public

But what do lay people think? And what evidence is there to suppose the experts (at least some of them) are right? There have also been a number of studies on lay theories of intelligence. What has surprised some academic researchers on intelligence is the difference between their assumptions and beliefs and those of "laypeople." There are a number of early studies on this topic (Wellman, 1944) and now an extensive, but scattered, literature on myths and misunderstandings about intelligence (Räty, 2015; Räty et al., 1993).

Sternberg (1985, 1990) proposed that the general population has a different conception, or different implicit theories, of intelligence from most experts. That is, "what psychologists study corresponds to only part of what people mean by intelligence in our [Western] society, which includes a lot more than IQ test measures" (Sternberg et al., 1981, p35). Sternberg (1996), in fact, wrote a paper entitled "Myths, Countermyths and Truths about Intelligence" in response to the reactions to "Bell Curve." He discussed various questions such as, "Can intelligence be taught to any meaningful degree?" and "Do intelligence tests measure pretty much all it takes for success in school and on the job?"

Over the years there has also been a particular interest in cross-cultural studies of lay, or implicit theories of intelligence (Beyaztaş-İlhan & Hymer, 2018; Yamazaki & Kumar, 2013) as well as studies of particular groups, such as gifted children, and of experts (Rindermann et al., 2017). In one cross-cultural study, Swami et al. (2008) asked students from three countries to rate 30 items for agreement about the nature, measurement, between-group differences, and practical importance of intelligence. This

was a 30-item scale derived from a summary of psychological research on intelligence signed by 50 (Western) experts in intelligence and allied fields (reprinted in Gottfredson, 1997). Nearly all the statements were backed by these scientific experts, though a significant number of world-renowned experts disagreed with many of the statements, which they considered misleading. An example of some items were: "IQ is strongly related, probably more so than any other single measurable human trait, to many important educational, occupational, economic and social outcomes," "Intelligence can be measured and intelligence tests measure it well," "While there are different types of intelligence tests, they all measure the same intelligence," "Intelligence tests are among the most accurate (in technical terms, reliable and valid) of all psychological tests and assessments." Again, it should be stressed that this is not a universal view, with a number of contributors to this volume disagreeing vociferously.

An exploratory factor analysis revealed three factors: (1) stability, reliability, and validity of intelligence tests, (2) practical importance of intelligence, and (3) origin and stability of within-group intelligence (i.e., why there are differences in the same group like all men) (Swami et al., 2008). They found their participants agreed strongly with general statements about intelligence being a broad and deep mental capability but the participants appeared to disagree most strongly with items that suggested between-group differences in intelligence and those that suggested that intelligence tests were valid and reliable. In short, they disagreed with many, but by no means all, experts in the field.

Recently, Warne and Burton (2020) devised 85 questions about intelligence and classified the questions into seven groups: (1) existence of intelligence, (2) components of intelligence, (3) biology of intelligence and life outcomes, (4) education and intelligence, (5) interventions to permanently raise IQ, (6) group differences, and (7) plausible causes of group differences. These were supposedly based on the academic literature, though it must be stressed that many are debatable.

They compared American teachers and non-teachers and found participants' responses were generally aligned with research findings regarding the components of intelligence. There was, however, disagreement regarding the broader sense of what intelligence is and what IQ scores represent, yet great confidence in the impact of interventions to raise

IQ. The authors concluded that empirically unsupported beliefs about intelligence were widespread and that people are generally unaware of many of the empirically supported findings from intelligence research. They also noted that one consequence is what the researchers believed to be erroneous beliefs about intelligence, which could result in decreased support for gifted programs, unrealistic expectations for interventions, or incomplete/inaccurate theories of giftedness.

In a recent book, Warne (2020) outlined 35 alleged "myths" about intelligence. Again, many established researchers in the field disagree passionately that they are well expressed or should be called myths. Furnham and Horne (2021) recently tested a number of lay people on Warne's list. They indicated whether they thought each statement was definitely or probably true or false, or whether they did not know. This followed the work of those interested in alleged psychological myths, as documented by Lilienfeld et al. (2010) and tested by Furnham and Hughes (2014).

The results are shown in Table 10.1.

Six statements (1, 11, 12, 23, 34, 35) showed that a majority believe the statement was probably false, which according to the psychometrically orthodox view is correct. These referred to the genetic components of intelligence, as well as to the social consequences of intelligence research, which may please some of those working in the field. In fact, the average total score of "definitely" plus "probably false" was 12.48 (12/35 items), almost exactly a third of the statements.

Many of the items that were thought of as true (namely, supposedly myth-endorsing) concerned IQ testing (2, 7, 8, 9, 22, 27). There also remains the widespread belief that tests are neither reliable nor valid, despite the fact that many psychometricians argue the intelligence tests are amongst the most robust and useful in the whole of psychology (Eysenck, 1998; Furnham, 2021). Study after study show the same thing: people do not trust tests. They are seen as being too narrow; to favor those with education and privilege rather than "actual" intelligence; and to measure something that is not that important.

One statement that attracted a high level of "don't know" responses and a wide spread of reactions was statement 10, which maintained that tests were/are biased against minority groups. There was also evidence that the participants accept the multiple/emotional/practical intelligences

Table 10.1 Frequencies of each answer across Intelligence Myth items. Total *N* = 275

	Definitely False	Probably False	Probably True	Definitely True	Don't Know
1. Intelligence is whatever collection of tasks a psychologist puts on a test	(28%)	**(31%)**	(25%)	(7%)	(8%)
2. Intelligence is too complex to summarize with one number	(1%)	(7%)	(26%)	**(64%)**	(2%)
3. IQ does not relate/ correspond to brain anatomy or functioning	(5%)	(22%)	**(36%)**	(25%)	(13%)
4. Westernized views on intelligences are not relevant in non-Western cultures	**(34%)**	(24%)	(18%)	(7%)	(16%)
5. There are multiple intelligences in the human mind	(1%)	(6%)	(27%)	**(63%)**	(4%)
6. Practical intelligence is a real ability separate from general intelligence	(1%)	(8%)	**(42%)**	(37%)	(12%)
7. Measuring intelligence is difficult	(4%)	(8%)	(25%)	**(60%)**	(4%)
8. Content on intelligence tests is trivial and cannot measure intelligence	(2%)	(25%)	**(40%)**	(22%)	(12%)
9. Intelligence tests are imperfect and cannot be used or trusted	(2%)	(26%)	**(41%)**	(21%)	(9%)
10. Intelligence tests are biased against ethnic minorities/diverse publications	(22%)	(23%)	(19%)	(12%)	**(24%)**
11. IQ only reflects a person's wealth and social status	**(43%)**	(31%)	(14%)	(4%)	(8%)
12. Intelligence's strong genetic links (through heredity) mean that raising IQ is impossible	(24%)	**(36%)**	(18%)	(8%)	(14%)

(*continued*)

Table 10.1 (continued)

	Definitely False	Probably False	Probably True	Definitely True	Don't Know
13. Genes are not important for determining intelligence	(18%)	**(41%)**	(24%)	(8%)	(8%)
14. Environmentally driven changes in IQ mean that intelligence is changeable/ malleable	(1%)	(12%)	**(46%)**	(23%)	(17%)
15. Social interventions can drastically raise IQ	(2%)	(16%)	**(46%)**	(19%)	(17%)
16. Brain training programs can raise IQ	(2%)	(9%)	**(50%)**	(27%)	(11%)
17. Improvability of IQ means intelligence can be equalized	(4%)	(21%)	**(38%)**	(11%)	(27%)
18. Every child is gifted	(13%)	(24%)	**(26%)**	(25%)	(11%)
19. Effective schools can make every child perform well/ proficient academically	(6%)	(20%)	**(47%)**	(25%)	(4%)
20. A pupil's environment and personality have powerful effects on academic achievement	(0%)	(7%)	(32%)	**(56%)**	(5%)
21. Admissions tests are a barrier to college for underrepresented students	(5%)	(16%)	**(44%)**	(21%)	(13%)
22. IQ scores only measure how good someone is at taking intelligence tests	(3%)	(15%)	**(40%)**	(33%)	(9%)
23. Intelligence is not important in the workplace	(30%)	**(41%)**	(17%)	(7%)	(4%)
24. Intelligence tests are designed to create or maintain a current power system	(17%)	(23%)	**(27%)**	(10%)	(24%)

(*continued*)

Table 10.1 (continued)

	Definitely False	Probably False	Probably True	Definitely True	Don't Know
25. Very high intelligence is not more beneficial than moderately high intelligence	7%)	(24%)	**(39%)**	(17%)	(13%)
26. Emotional intelligence is a real ability that is helpful in life	(1%)	(7%)	(26%)	**(59%)**	(7%)
27. IQ scores are distributed evenly between men and women	(7%)	(20%)	**(28%)**	(16%)	(28%)
28. Racial/Ethnic group IQ differences are completely environmental in origin	(8%)	(22%)	**(30%)**	(11%)	(29%)
29. Unique influences operate on one group's intelligence test scores	(1%)	(16%)	**(40%)**	(9%)	(33%)
30. Stereotype threat explains score gaps among demographic groups	(6%)	(17%)	**(35%)**	(12%)	(29%)
31. Controversial or unpopular ideas should be held to a higher standard of evidence	(8%)	(19%)	**(40%)**	(12%)	(21%)
32. Past controversies taint modern research on intelligence	(5%)	(18%)	**(41%)**	(14%)	(22%)
33. Intelligence research leads to negative social policies	(12%)	**(29%)**	(24%)	(11%)	(24%)
34. Intelligence research undermines the fight against inequality	(14%)	**(28%)**	(23%)	(9%)	(26%)
35. Everyone is about as smart as I am	(26%)	(25%)	**(26%)**	(8%)	(15%)

Numbers in bold represent the highest number of responses in that category

model (items 5, 6, 26). Around two thirds rejected the concept of "g" being a parsimonious and accurate summary variable, although they accept the fact that measurement is difficult. Two statements accepted as "probably" or "definitely true" were 5 (90%) and 26 (85%), both of which referred to multiple (emotional) intelligences, which has excited great debate among intelligence researchers for over 20 years.

Most of all, they appear to embrace Dweck's growth model, which suggests you can increase your intelligence by a variety of interventions (items 14 to 19) (Dweck, 2006). This is a very complex concept, namely, whether intelligence changes over time (i.e., through childhood and adulthood) and, if so, what can cause it to increase. It seems that many people want to, and do, believe in the "plastic" rather than in the "plaster" hypothesis about change, namely, that it is possible to actually raise/increase intelligence (as opposed to simply getting higher IQ test scores). It is not certain this refers to fluid as opposed to crystallized intelligence, which is important, as some experts would suggest it is easier to raise the latter as opposed to the former type of intelligence (Furnham, 2021). Also, we know that fluid intelligence changes a lot over age, peaking in the twenties and thirties and showing dramatic decline after the age of 60 (Deary, 2001).

However, as Furnham and Horne (2021) note: "*There remains, however, one very serious issue: namely that the statements are rated as 'false' by Warne as there is no necessary agreement about this, even from experts. It is possible that academics, in some disciplines, actively promote these falsehoods (both in their courses and publications) as if there was incontrovertible evidence to that effect. That is, some myths and misconceptions cannot be an either/or proposition:* i.e.*, some myths are only partially false. As regards the myths in this study it may be that many experts would want to caveat many of them with suggestions as to more specific context in which they apply. Further, it could be that many participants were not familiar with a number of issues yet loath to report 'Don't know'. Similarly, some of the items were also nearly tautological like item 21.*"

The results of this study, indeed, like many others in this area, pose the question as to why the public and the more orthodox experts disagree. There have been over the years many "popular" books written by academic psychologists trying to explain the theories and data on

intelligence, particularly, nature-nurture and group differences (particularly race and sex) (Deary, 2001; Plomin, 2018; Ritchie, 2015). Many of them tend to underplay the role of culture, context, and history, favoring a biological and universalist perspective. Some would argue that on many, though not all of these concepts, lay people are essentially correct, given a wider definition of the concept of everyday intelligence.

Indeed, it is important to state some important caveats about experts. First, experts disagree among themselves. Second, that their classic psychometric view is not necessarily the "correct" view. There are disagreements as to whether there even is a correct view. Third, ideas and opinions about intelligence change, as new data is obtained and processed.

Fairness and Accuracy

How accurate and fair are tests? Are they only accurate in the culture in which they were created? Is the use of them fair? Testing is a very "hot" issue, as demonstrated by the increasing number of court-cases where they are cited (Gatewood et al., 2015).

There is a considerable literature on the perception of test accuracy, which includes IQ tests. Many have been interested primarily in applicants' fairness perceptions of different selection methods (Truxillo et al., 2006), including cognitive ability tests (Chan et al., 1997). Results consistently indicate that applicants tend to favor, and rate as fair, work samples and interviews over paper-and-pencil test methods (Nikolaou & Judge, 2007). Further, cross-cultural replications (Moscoso & Salgado, 2004) demonstrated that applicants universally rate work-sample methods and interviews as the fairest types of selection methods.

Fairness perceptions of selection methods do have an impact on various outcomes, including applicant self-efficacy and self-esteem, job-acceptance intentions, motivation to pursue employment, likelihood of recommending the organization to friends, and test-taking motivation (Sanchez et al., 2000).

In one illustrative study, Furnham and Chamorro-Premuzic (2010) asked students to rate the *Fairness and Accuracy Perceptions* of 17 different selection methods, indicating how well they thought each method assessed

eight different characteristics which universities seek in potential students (bright, conscientious, mature, co-operative, initiative, community service, work experience, and communication skills). For example, participants rated how accurately a face-to-face interview measured brightness, conscientiousness, maturity, etc. A 9-point Likert type scale, ranging from 1 being "extremely accurate" to 9 being "extremely inaccurate," was used. They made two ratings, one for fairness and the other for accuracy.

The results are shown in Tables 10.2 and 10.3. Three things are of interest. The *first* is the close relationship between the two ratings; people certainly believe that what is accurate is fair; in the sense that if tests provide accurate scores of abilities, they *may* be fairly used in assessment. *Second,* the relatively small standard deviation shows considerable agreement between the participants. Third, intelligence tests were rated very low on both criteria. Participants think tests of power are *relatively unfair and inaccurate* while tests of preference are fair. This again illustrates the widespread and long-lasting distrust of the validity of intelligence tests among the general public. This view, of course, is contrary to that of those who create and utilize the tests (Furnham, 2008).

Table 10.2 Descriptive statistics and Cronbach's alphas for all assessment methods—accuracy

Selection method	Cronbach's alphas	*M*	SD
Face-to-face interview	0.79	25.69	8.65
Outdoor leadership exercise	0.76	26.59	9.34
References	0.89	27.43	11.7
Panel interview	0.80	27.64	8.91
Observed group discussion	0.78	28.62	8.56
Oral presentation	0.78	30.78	9.11
Personality test	0.68	32.93	8.21
Telephone interview	0.87	33.77	10.94
Video	0.82	34.54	10.43
Exam condition essay	0.79	35.17	10.39
Situation exam	0.83	35.54	10.78
Assessment center	0.87	36.45	10.97
Unseen course-related exam	0.84	36.45	10.31
Application form	0.86	38.46	12.89
General knowledge test	0.86	42.60	11.81
Intelligence test	0.83	44.08	11.11
Drugs test	0.92	52.17	15.84

Scale: Most accurate 8–Least accurate 72
Note: 10*N* ranges from 185–322

Table 10.3 Descriptive statistics and Cronbach's alphas for all assessment methods—fairness

Selection method	Cronbach's alphas	*M*	SD
Face-to-face interview	0.88	33.28	12.97
Outdoor leadership task	0.76	30.66	10.38
References	0.90	27.49	12.34
Panel interview	0.83	30.18	10.78
Discussion	0.81	30.65	9.84
Oral presentation	0.85	32.26	11.09
Personality test	0.87	35.30	12.03
Telephone interview	0.87	34.93	11.88
Video	0.84	35.37	11.09
Essay	0.85	35.89	11.76
Situation exam	0.88	36.86	12.63
Assessment center	0.89	38.74	12.82
Unseen course-related exam	0.86	42.04	12.53
Application form	0.88	37.42	14.17
General knowledge test	0.90	42.94	13.99
Intelligence test	0.88	43.28	12.97
Drugs test	0.93	54.08	15.89

Scale: Most fair 8–Least fair 72

This area of research on the perceived accuracy of assessment techniques reveals again the gap between many (but by no means all) academic experts and the lay public with regard to IQ tests. The question is the cause. Some academics would say it is the result of ignorance (most people have not seen or understood the data) or that laypeople are being defensive, particularly those who do not score highly. This view, of course, may be seen as condescending and patronizing. That is, the results shown in Table 10.2 and 10.3 may be simply wrong. Lay people may respond by saying that this data is based on their personal experience. This is an important issue that will not go away.

Everyday Tests and Playing Games

But what is, and is not, a marker of intelligence that could be considered an intelligence test? Many people complete intelligence-type tests everyday: crosswords, Sudoku, etc. Many daily newspapers have a page dedicated to quizzes and games, which editors know are popular with readers. Some people literally become addicted to these tests and form clubs and

take part in competitions. Are these really intelligence tests? Is Sudoku simply a test of fluid intelligence and crosswords of crystallized intelligence? Do you get much better with practice? Or is it only the talented who get obsessed? Why are some cultures more "addicted" to some games rather than to others?

Others play games like Bridge, Chess, Scrabble, etc., which have been popular for years and are intellectually demanding. Are they essentially intelligence tests? The question is what is, and what is not an intelligence test in the sense that scores correlate with psychometric tests. Or is that both a deeply conservative and misguided view, if the current tests are themselves inaccurate and invalid?

It has been suggested for many years that "strategic games are an invariant expression of certain universal intellectual traits" (Spitz, 1978). Researchers have suggested that various games and tasks, not formally described as intelligence tests, are actually very good measures of intelligence. Ideally, these measures should be simple, robust, and culturally valid, because they may differ across cultures. The question remains however: does being good at cross-words or Sudoku mean you have "street-smarts" and make wise decisions for yourself and others?

Thus, a test of proofreading has been found to be good substitute for measures for intelligence (Furnham, 2010), whereby participants are tasked to correct errors on a page as quickly and accurately as possible. Indeed, there has been a great deal of interest in the development of new tests to measure both intelligence and personality (Ihsan & Furnham, 2018).

Some researchers have suggested that many computer games that exist, but were designed primarily as entertainment, could serve as an excellent proxy for intelligence tests, as they often measure efficiency of information processing (Foroughi et al., 2016; Gnambs & Appel, 2017). They may offer more accurate results because people respond differently when asked to play a game rather than take a test. The latter usually sounds more complex, serious, and therefore anxiety-provoking, and this provocation of anxiety may have potentially serious consequences; this could induce test anxiety and lower performance (Furnham, 2008). On the other hand, it is also possible that if a task is described as a game, the participant may not fully engage their abilities, and thus not show their

full abilities. The issue is the effect on motivation and performance when a cognitive ability task/test is described as a game or a test or something else.

There has been growing interest in the prospective functions of games aside from entertainment, such as their instructional value (Garris et al., 2002), clinical applications (Griffiths, 1997), and how they can contribute to the understanding of cognitive capacity, plasticity, and other processes (Boot, 2015). The Learning Strategies Programme developed a video game that has been used as a research tool, designed to involve skills such as attention, memory, and multi-tasking (Mané & Donchin, 1989). Sajjadi et al. (2017) have provided evidence for the mapping between dimensions of Gardner's Multiple Intelligences and game mechanics, suggesting that games can be designed using empirical data to suit players with certain abilities and inclinations.

Furthermore, games have the advantage of eliciting greater engagement and intrinsic motivation to perform. Hoffman and Nadelson (2010) found that greater motivational engagement in gaming was partly influenced by more positive responses to failure, particularly in multiple-level games that become increasingly complex and challenging as the player progresses.

A number of recent studies have demonstrated how good computer games are at measuring general intelligence. Quiroga et al. (2015) reported extremely high correlations (r = 0.93) between latent factors from video games and intelligence on a range of different tests. Foroughi et al. (2016) reported correlations of r = 0.65 (N = 35) between a video game and the Ravens Progressive matrices, and r = 0.78 (N = 100) between the video game and latent variable measuring of fluid intelligence. They concluded that it is feasible to create measures of fluid intelligence using their test. Sin and Furnham (2018) got 112 participants to complete a standardized intelligence test along with one spatial and one verbal game to determine the relationship between cognitive ability and game performance. Both games were significant correlates of intelligence, though differences were found in the strength of correlations and contribution of other factors between the two tasks.

The two game tasks chosen were vastly different in terms of the specific skills required for each, yet both tasks correlated with intelligence.

Nevertheless, the results confirm previous findings in that game performance can be a reliable predictor of fluid intelligence.

Of course, there are at least two interpretations of these results. The first is that video games are a good measure of intelligence. The other is that intelligence tests are, as some investigators have suggested (e.g., Sternberg, 1997), measures of a certain kind of game-playing—that the people who are good at them are people who are good at playing games, but maybe not as much else as one might have hoped.

The relationship between game performance and cognitive ability has meaningful implications, both for the understanding of cognitive ability and the value of games. The possibility of using games as a measure goes beyond measuring intelligence and increasing engagement, but also as a dynamic method of assessment that allows greater insight into one's problem-solving strategies, situational behavior, and ability to adapt to novel situations. Granic et al. (2014) argued that video games are becoming increasingly sophisticated and serve functions beyond entertainment. If games can improve existing skills, it is plausible that they could also serve as a measure of such skills and other latent factors. Furthermore, it could be worth making the distinction between describing and guessing ability, as they can be considered two distinguishable tasks.

Would the participants have played the games differently had they known they were being tested? In this study, participants were given no explicit information that the games were a measure of ability, though they were aware that their performance was being measured. However, the adaptation of games in institutional settings may risk violating the voluntary, playful element of games which are central to their engagement, a tension that has also been identified by other researchers of the serious functions of games.

The more important question is not how well the games measure intelligence, but how well they predict outcomes that are relevant to the selection criteria. Future research should also consider the effectiveness of games through its relationship to other more salient outcomes, including occupation-specific performance indicators, using longitudinal designs. Instead of simply examining whether game performance predicts intelligence, we can then examine whether game performance can predict relevant outcomes even better than intelligence can.

Yet, some still argue that this whole enterprise is biased and misleading because all the new tests are validated against current intelligence tests, which themselves are flawed. We need to start again with a much more inclusive definition of intelligence.

Conclusion

Of all the topics in psychology, intelligence is one of the hottest, particularly if group (gender, culture, age) differences are considered as well as the nature-nurture issue. Yet there are a growing band of researchers who take a different view, more aligned with what lay people think.

Few, if anyone, would deny the importance of being intelligent (bright, smart). After all, life is an IQ test, though in the broadest sense. Equally people know from personal experience that it might be necessary, but far from sufficient, to guarantee health, happiness, and success in life. As a consequence, most people take a wider view of intelligence than the sort that most tests measure. Many are skeptical, indeed cynical, about tests because their face validity seems not to match up to their understanding of what intelligence means. But in a rapidly developing world the way in which we store and access knowledge has changed and we need new ways to assess how best to thrive in this world. It will be interesting to see whether new and culturally sensitive intelligence tests will be proven to be both valid and reliable and whether, indeed, people who take them have more faith in their accuracy.

References

Baker, D. P., Eslinger, P. J., Benavides, M., Peters, E., Dieckmann, N. F., & Leon, J. (2015). The cognitive impact of the education revolution: A possible cause of the Flynn effect on population IQ. *Intelligence, 49*, 144–158.

Berry, J. W. (1974). Radical cultural relativism and the concept of intelligence. In J. W. Berry & P. R. Dasen (Eds.), *Culture and cognition: Readings in cross-cultural psychology* (pp. 225–229). Methuen.

Beyaztaş-İlhan, D., & Hymer, B. (2018). An analysis of Turkish students' perceptions of intelligence from primary school to university. *Gifted Education International, 34*(1), 19–35.

Boot, W. R. (2015). Video games as tools to achieve insight into cognitive processes. *Frontiers in Psychology, 6*, 3.

Bratsberg, B., & Rogeberg, O. (2018). Flynn effect and its reversal are both environmentally caused. *Proceedings of the National Academy of Sciences of the United States of America, 115*, 6674–6678.

Cattell, R. B. (1949). *Culture free intelligence test, scale 1, handbook*. Institute of Personality and Ability Testing.

Chan, D., Schmitt, N., DeShon, R. P., Clause, C. S., & Delbridge, K. (1997). Reactions to cognitive ability tests: The relationships between race, test performance, face validity perceptions, and test taking motivation. *Journal of Applied Psychology, 82*, 300–310.

Deary, I. (2001). *Intelligence: A very short introduction*. Oxford University Press.

Dilchert, S. (2018). Cognitive ability. In D. S. Ones, N. Anderson, C. Viswesvaran, & H. K. Sinangil (Eds.), *The SAGE handbook of industrial, work & organizational psychology: Personnel psychology and employee performance* (pp. 248–276). Sage.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.

Eysenck, H. J. (1998). *Intelligence: A new look*. Transaction Publishers.

Flynn, J. R. (2010). Problems with IQ gains: The huge vocabulary gap. *Journal of Psychoeducational Assessment, 28*(5), 412–433.

Flynn, J. R. (2012). *Are we getting smarter? Rising IQ in the twenty-first century*. Cambridge University Press.

Foroughi, C., Serraino, C., Parasuraman, R., & Boehm-Davis, D. (2016). Can we create a measure of fluid intelligence using puzzle creator within portal 2? *Intelligence, 56*, 58–64.

Furnham, A. (2008). *50 ideas you need to know: Psychology*. Quercus.

Furnham, A. (2010). Proofreading as an index of crystallised intelligence. *Educational Psychology, 30*(6), 735–754.

Furnham, A. (2021). *Twenty ways to assess people*. Cambridge University Press.

Furnham, A., & Chamorro-Premuzic, T. (2010). Consensual beliefs about the fairness and accuracy of methods at university. *International Journal of Selection and Assessment, 18*, 417–424.

Furnham, A., & Horne, G. (2021). Myths and misconceptions about intelligence: A study of 35 myths. *Personality and Individual Differences, 181*, 111014.

Furnham, A., & Hughes, D. (2014). Myths and misconceptions in popular psychology: Comparing psychology students and the general public. *Teaching of Psychology, 41*, 256–261.

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning a research and practice model. *Simulation & Gaming, 33*, 441–467.

Gatewood, R., Feild, H. S., & Barrick, M. (2015). *Human resource selection*. Nelson Education.

Gnambs, T., & Appel, M. (2017). Is computer gaming associated with cognitive abilities? *Intelligence, 61*, 19–28.

Gottfredson, L. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history and bibliography. *Intelligence, 24*, 13–23.

Granic, I., Lobel, A., & Engels, R. C. (2014). The benefits of playing video games. *American Psychologist, 69*(1), 66.

Griffiths, M. (1997). Video games and clinical practice: Issues, uses and treatments. *British Journal of Clinical Psychology, 36*(4), 639–642.

Hoffman, B., & Nadelson, L. (2010). Motivational engagement and video gaming: A mixed methods study. *Educational Technology Research and Development, 58*(3), 245–270.

Ihsan, Z., & Furnham, A. (2018). The new Technologies in Personality Assessment: A review. *Consulting Psychology, 70*(2), 147–166.

Lilienfeld, S., Lynn, S., Ruscio, J., & Beyerstein, B. (2010). *50 great: Popular myths of popular psychology*. Wiley-Blackwell.

Mané, A., & Donchin, E. (1989). The space fortress game. *Acta Psychologica, 71*(1–3), 17–22.

Moscoso, S., & Salgado, J. F. (2004). Fairness reactions to personnel selection techniques in Spain and Portugal. *International Journal of Selection and Assessment, 12*, 187–196.

Nikolaou, I., & Judge, T. A. (2007). Fairness reactions to personnel selection and techniques in Greece: The role of core self-evaluations. *International Journal of Selection and Assessment, 15*, 206–219.

Plomin, R. (2018). *Blueprint: How DNA makes us who we are*. Penguin Books Ltd..

Quiroga, M. Á., Escorial, S., Román, F. J., Morillo, D., Jarabo, A., Privado, J., & Colom, R. (2015). Can we reliably measure the general factor of intelligence (g) through commercial video games? Yes, we can! *Intelligence, 53*, 1–7.

Räty, H. (2015). Notion of intelligence and social-educational identity. *Educational Studies, 41*, 272–275.

Räty, H., Snellman, L., & Vornanen, A. (1993). Public views on intelligence: A Finnish study. *Psychological Reports, 72*, 59–65.

Rindermann, H., Becker, D., & Coyle, R. T. (2017). Survey of expert opinion on intelligence: The Flynn effect and future of intelligence. *Personality and Individual Differences, 106*, 242–247.

Ritchie, S. (2015). *Intelligence: All that matters*. McGraw Hill.

Sajjadi, P., Vlieghe, J., & De Troyer, O. (2017). Exploring the relation between the theory of multiple intelligences and games for the purpose of player-centred game design. *Electronic Journal of e-Learning, 15*(4), 320–334.

Sanchez, R. J., Truxillo, D. M., & Bauer, T. N. (2000). Development and examination of an expectancy-based measure of test-taking motivation. *Journal of Applied Psychology, 85*(5), 739–750.

Serpell, R. (2000). Intelligence and culture. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 549–580). Cambridge University Press.

Sin, J., & Furnham, A. (2018). Can commercial games function as intelligence tests? A pilot study. *Computer Games Journal, 7*, 27–37.

Spitz, H. (1978). The universal nature of human intelligence: Evidence from games. *Intelligence, 2*, 371–379.

Sternberg, R. J. (1985). Implicit theories of intelligence, creativity, and wisdom. *Journal of Personality and Social Psychology, 49*, 607–627.

Sternberg, R. J. (1990). *Metaphors of mind: Conceptions of the nature of intelligence*. Cambridge University Press.

Sternberg, R. J. (1996). Myths, countermyths and truths about intelligence. *Educational Researcher, 25*(2), 11–16.

Sternberg, R. J. (1997). *Successful intelligence*. Plume.

Sternberg, R., Conway, B., Ketron, J., & Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology, 41*(1), 37–55.

Swami, V., Furnham, A., Maakip, I., Ahmad, M. S., Naw, N. H. M., Voo, P. S. K., et al. (2008). Beliefs about the meaning and measurement of intelligence: A cross-cultural comparison of American, British and Malaysian undergraduates. *Applied Cognitive Psychology, 22*(2), 235–246.

Trahan, L. H., Stuebing, K. K., Fletcher, J. M., & Hiscock, M. (2014). The Flynn effect: A meta-analysis. *Psychological Bulletin, 140*(5), 1332–1360.

Truxillo, D. M., Bauer, T. N., Campion, M. A., & Paronto, M. E. (2006). A field study of the role of big five personality in applicant perceptions of selection fairness, self and the hiring organisation. *International Journal of Selection and Assessment, 14*, 269–277.

Warne, R. (2020). *In the know: Debunking 35 myths about human intelligence*. Cambridge University Press.

Warne, R. T., & Burton, J. Z. (2020). Beliefs about human intelligence in a sample of teachers and non-teachers. *Journal for the Education of the Gifted, 43*(2), 143–166.

Wellman, E. (1944). Some misconceptions about intelligence. *Childhood Education, 21*(3), 108–112.

Yamazaki, S., & Kumar, V. (2013). Implicit theories of intelligence and creative ability: Relationships with academic risk-taking and academic stress. *International Journal of Creativity and Problem Solving, 23*(2), 25–36.