

Ten things I hated about intelligence research

IT was an honour and a surprise to receive the Society's Book Award for *Looking Down on Human Intelligence: From Psychometrics to the Brain* (OUP, 2000). Especially for a book about reductionist research into the origins of human psychometric intelligence differences, a research monograph larded with multivariate analyses. Why write such a book? Not for money – the key audience is quite small and the price is off-puttingly high. The truth – to have a good moan.

Research on human psychometric intelligence is arguably among the best and most resilient success stories in all of scientific psychology. Part of that story is 'looking down' on the individual differences found in mental test scores: that is, trying to explain them in terms of more tractable, better-understood cognitive and biological constructs. This aspect of research is going well, and has done since the 1970s. There were some edited collections on this area, but no individual had taken stock. Various metaphors are employed in *Looking Down...*: the book is referred to as a spring cleaning of the research area, or as a weeding of a lush but unruly garden. It seemed time to underscore what the area had achieved and to expose and rehabilitate its bad research habits. Some of the following annoyances (there are in total many more than the 10 listed here) are internal complaints about the research area, and some are external, about observers' reactions to the research area. In the rest of this article, whenever chapter numbers are given they refer to the book.

1. What's known, what's not

It is hard to get to first base in discussing cognitive ability differences.

Human intelligence differences can be a controversial area. The heat generated often obscures the century of solid, replicated research. So, before the book discussed the state of research on the information processing origins of human psychometric intelligence differences, chapter 1 outlined what could be agreed on the construct more broadly. Luckily, the American Psychological Association's task force report on human intelligence (Neisser



IAN J. DEARY, winner of the Society's 2002 Book Award, 'looks down' on a controversial area.

et al., 1996) was to hand as a disinterested and expert statement on the issues. I have also tried to describe clearly and even-handedly the knowns and unknowns of human psychometric intelligence differences (Deary, 2001b).

So what am I sure of? Individual differences in human cognitive abilities show a hierarchical structure. A general factor (Spearman's *g*) accounts for about 40–50 per cent of the variance in a battery of varied mental tests administered to

'The heat generated often obscures the century of solid, replicated research'

a large sample of people. Carroll (1993) showed that, ever since Spearman in 1904, this result is regularly replicated – even in data sets like those of Thurstone and Guilford, whose ideas initially appeared to question the existence of a general factor. Further down the hierarchy some variance lies in group factors of ability (like verbal, spatial, memory, speed of processing, and so forth), and some variance lies in even more specific abilities. Psychometric intelligence is thus more than *g*, though *g* is a remarkable and important part of the story (Deary, 2001a).

Psychometric intelligence differences show strong stability across the lifespan: about half the variance is stable from age 11 to age about 80 years (Deary, Whalley *et al.*, 2000, Deary *et al.*, in press). Environment and genes both make substantial contributions to psychometric intelligence differences. The heritability is higher in adulthood than in childhood. The effect of shared environment in adulthood is very low compared with the contribution made by an individual's unique environment.

Psychometric intelligence differences are important predictors of life's outcomes. They have predictive validity for education and workplace success (Deary, 2001; Neisser *et al.*, 1996). Newer research shows that psychometric intelligence differences are predictive of health outcomes, including how long people live. For example, women with one standard deviation lower psychometric intelligence scores at age 11 are only about three quarters as likely to live to 76 as those whose scores are 15 points higher (Whalley & Deary, 2001).

Of course, there is a lot more to life's quality and quantity than psychometric intelligence, and certainly there is more to cognitive differences than can be captured in psychometric intelligence. But it is a principal player among one's mental cast.

2. The missing populations

There are too few representative samples in research on cognitive ability differences. Any cognitive and biological correlates of psychometric intelligence should be estimates of the true effect size in the relevant population. Yet, as good as differential psychologists are at describing the quantitative aspects of their samples, they are often as lazy as most other psychologists when it comes to recruitment. The psychology of the topic being investigated tends to become the psychology of the topic in university undergraduates (often in psychology departments). This recourse to recruiting the most readily available participants does a varying amount of violence to the effect sizes and the generalisability of the findings, depending on the construct being investigated. For psychometric intelligence it is really problematic. Students are abnormal: their mean mental test scores are high and their range is restricted. It is like studying

extraversion differences by recruiting only party animals, or studying neuroticism differences using only Woody Allens.

The upshot is that when quantitative reviews are conducted, the coefficients are underestimates and some upward correction (disattenuation) of the correlations is attempted. For example, this occurred in reviews of the research on the correlations between reaction time and psychometric intelligence (chapter 6). Raw correlations between mean decision time and psychometric intelligence showed a mean of $-.22$, a sharing of about 5 per cent of variance. On the other hand, a study based on a population-representative sample of 55-year-olds found a correlation of $-.49$ between four-choice reaction time and psychometric intelligence, indicating 24 per cent shared variance (Deary, Der *et al.*, 2001). It's a simple message: unless one wants a seriously underestimated effect size, make sure there is appropriate variance in the construct being investigated. Differential and other psychologists have much to learn from epidemiology with regard to their samples' characteristics (Lubinski & Humphreys, 1997).

3. Psychometric tautologies

Researchers can invest too much in the names given to constructs.

In science, near-identical constructs are often given different names and talked about as if they were distinct – this is known as the 'jangle fallacy' (Kelley, 1927). For example, there are innumerable scales purporting to measure differently named character traits that are just 'jangles' for neuroticism. The same happens in intelligence research.

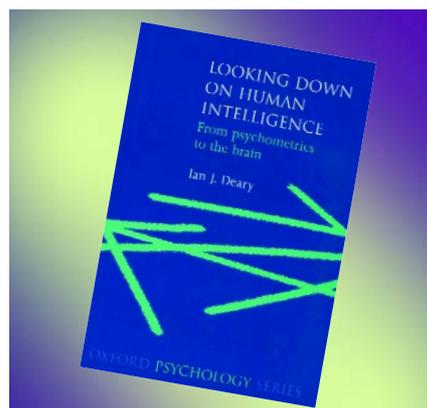
For example, Kyllonen used a model of human information processing to develop a new approach to testing individual differences in mental abilities (chapter 5). The resulting scheme was called the 'Cognitive Abilities Measurement' (CAM) and it assessed processing speed, working memory, declarative knowledge, and procedural knowledge. This sounds refreshingly different from standard psychometric batteries like, say, the ASVAB (Armed Services Vocational Aptitude Battery) which lacked such theoretical foundations. Theoretical foundations notwithstanding, the four domains of the CAM loaded between $.89$ and $.96$ on a general factor, and the CAM's general (g) factor correlated $.994(!)$ with the g factor from the ASVAB.

Moreover, individual differences in 'working memory' from the CAM correlate

almost perfectly with individual differences in the g factor from psychometric test batteries (see also Deary, 2001). This is another jangle, but it is also a superb opportunity for psychometricians interested in g to collaborate with cognitive psychologists studying working memory. They have complementary skills and data on a single construct. Psychologists must invest less in the names they give their constructs and focus more on the nomological network: data speak louder than words.

4. The packaging, not the gift

The research area has been slow to realise that overall reaction times and their variabilities correlate better with psychometric intelligence than do cognitive components extracted from reaction time procedures. Around the early years of the cognitive psychology revolution, some differential psychologists decided they would cherry pick some specific cognitive processes



from cognitive models. Experimental psychologists were isolating the times taken by individual cognitive operations. Largely this was done using the subtractive method: computing the difference between two reaction times, one of which was assumed to contain a specific additional cognitive operation (Posner & Rueda, 2002, provide an up-to-date view of this 'mental chronometry'). Differential psychologists thought that there might be individual differences in the times, and that these might correlate significantly with (i.e. partly explain?) psychometric intelligence differences.

This thinking was applied to, among others, the reaction time procedures developed by Hick, Posner and Sternberg, which could be used to extract cognitive kernels assessing, respectively, rate of gain of information, speed of access to long-term memory stores, and speed of scanning

in short-term memory (chapter 6). In each case the same thing happened. Individual differences in the cognitive kernel (the 'gift') had unimpressive correlations with psychometric intelligence. There were higher and more consistent correlations with the theoretically unexciting overall reaction times (the 'packaging'). Even simple reaction time correlated significantly with psychometric intelligence differences (Deary, Der *et al.*, 2001; Der & Deary, 2003). This warrants more research on why these apparently disparate performances correlate.

5. Performative processing

For many psychologists, a test measures a construct because they say it does.

In *Looking Down...* a principal target for criticism was research that claimed to be investigating psychometric intelligence's foundations in differences in speed of information processing. Operationalising speed of information processing sometimes meant using psychometric tests like the digit-symbol test, experimental psychology procedures like reaction times, psychophysical tasks like inspection time, psychophysiological techniques like event-related potentials, and physiological measures like nerve conduction velocity. For the fair-minded, critical researcher this is a frustrating 'I am Spartacus'-like situation, with a gallimaufry of measures – some substantially correlated, some not – all claiming to get at a construct called 'speed of information processing'.

The complaint here is the damage done to the good, replicable data. There are substantial enough and hugely replicated correlations between reaction times and inspection times and psychometric intelligence (chapters 6 and 7). It devalues these surprising, intriguing findings to lazily state that this is because human intelligence differences are founded in part in 'processing speed' differences. A replicated correlation is a warrant for lots more mechanistic work, not a sojourn in the armchair. If, say, inspection time correlates significantly with psychometric intelligence (it does: Grudnik & Kranzler, 2001), then researchers must find out why. This involves more work on what, in brain-processing terms, inspection time involves (Deary, Simonotto *et al.*, 2001), and investigating the mechanisms of its correlation with general ability and other domains of cognitive function (Burns & Nettelbeck, 2003). Progress is good, but painstaking, and certainly slower than the pseudo-progress made by those whose

correlations are followed by mere verbal incantation of a supposed explanatory mechanism.

6. The truth about brain size

People might guffaw, but brain size really does correlate significantly with psychometric intelligence.

The late S.J. Gould effused a pall of ridicule over attempts to relate brain size to intelligence. But empirical data trump eloquence. Magnetic resonance imaging affords accurate measurement of brain size *in vivo*. A quantitative review showed there is a replicable, modest, significant correlation between brain size and psychometric intelligence (Vernon *et al.*, 2000).

Bigger brains are associated with higher scores on psychometric intelligence tests. Why the correlation occurs is another matter. In *Looking Down...* I argue that modern explanations for the brain size/psychometric intelligence correlations are not much more sophisticated than those that were available in the late Renaissance. Work is progressing. Some data suggest that overall brain size relates to the general factor in mental test batteries (MacLulich *et al.*, 2002). The surprising and interesting correlation between mental test scores and brain size needs investigating further before it can be declared that it has helped us to understand the foundations of psychometric intelligence differences.

Meanwhile attention is gathering around functional brain imaging results whose titles could lead one to suppose that they had revealed neural mechanisms underlying individual differences in *g* (Gray *et al.*, 2003). Any such strong reductionist claims would be premature; they have found yet more interesting correlations between an aspect of the brain's response during mental activity and scores on a mental test. They are some more leads to follow.

7. Ageing: Speed without substrate

Part of cognitive ageing is the slowing of the brain's 'processing speed', but what this

processing speed comprises is not known. Do cognitive domains decline (if they do at all) in parallel with age, or do they age independently, with separate underlying mechanisms? To date, it appears that though there are specific cognitive ageing effects in particular domains, there is a large general component: if one domain

*is declining there is a tendency for others to decline too (Salthouse & Ferrer-Caja, 2003; Wilson *et al.*, 2002). Moreover, much of this general effect on cognitive ageing was accounted for by age-related changes in 'speed of information processing'.*

*It sounds like an impressive set of findings. But the complaint is with what was used to measure it. Sometimes it is psychometric tests like Digit Symbol, and sometimes it is reaction time procedures. This is less impressive. Digit Symbol is a psychometric test, hardly a fundamental measure of 'speed of information processing' (which is just a convenient label). And we do not understand how different brains result in different reaction time means. Though there are regularities in cognitive ageing that are worth investigating, a mechanistic understanding is not yet available. Among other candidates, the decreased connectivity of brain white matter tracts with age should be investigated as a possible candidate. White matter lesions, detectable using structural magnetic resonance imaging, relate to cognitive ageing (Deary *et al.*, 2003). Even more informative perhaps will be new magnetic resonance imaging techniques, such as diffusion tensor imaging, which can provide measures of white matter connectivity and relate to cognition in studies using a small number of participants (Shenkin *et al.*, 2003).*

8. Genes have developed

There is too little appreciation that so-called genetic studies can do far more than provide heritability estimates. Traditional biometric studies – involving

*families, twins and adoptees – provided clear and interesting findings about the environmental and genetic origins of differences in general and specific mental abilities (Plomin *et al.*, 2001). Even when people are aware of these findings, they tend to be less aware that genetic research now does so much more.*

First, traditional biometric methods have been the best sources of evidence for the importance of environmental contributions to psychometric intelligence differences. It tends to be non-shared as opposed to shared environment that is the more important determinant. This renders deficient the application of the term 'behaviour genetics' to the area: it tells only (about) half the story.

Second, genetic covariance is a powerful tool for investigating the correlations between psychometric intelligence and other measures, such as brain size, reaction time, and inspection time (Plomin & Spinath, 2002). For example, there are shared genetic effects on psychometric intelligence and inspection time (Luciano *et al.*, 2001).

Third, molecular genetic studies have the potential to turn heritability estimates into mechanistic understandings (Plomin, 2003). For example, the gene that codes for apolipoprotein E has been informative. One form of the gene is associated with the production of a type of apolipoprotein E that repairs neurons suboptimally after damage. In addition, individual differences in this gene might account for about 1 per cent of the variance in psychometric intelligence in old age (Deary *et al.*, 2002).

9. Glass bead games, cargo cults...

Four behavioural tendencies were identified among some intelligence researchers that might limit progress. 'Glass bead game research' was identified where researchers were led by the aesthetic appeal of a 'big theory' of intelligence, often one that ignores disconfirming data mountains. Guilford's 'structure of intellect' model was one such folly, easily refuted by data. His own data confirmed the three-stratum model of psychometric intelligence rather than illustrating many independent aspects of intelligence (Carroll, 1993). Perhaps Gardner's 'multiple intelligences' is another, though there has been too little operationalisation of the key constructs and data collected on them to tell. Anyone devising a new 'model of intelligence differences' must contend with Carroll's (1993) whole-century summary of data regularities.

KATE GREY

WEBLINKS

Intelligence: A Multidisciplinary Journal:
www.elsevier.com/locate/issn/01602896

International Society for the Study of Individual Differences: www.issid.org

The American Psychological Association's task force on intelligence differences:
www.apa.org/releases/intell.html

'Crying wolf science' is found where an alimentary dislike of a research area is evinced by suggestions that the associations are due to confounding variables. As the adage goes, 'everything can be explained by appealing to unmeasured variables'. Thus, the correlations between, say, psychometric intelligence and inspection time have been 'explained' by appeals to personality, motivation, strategies, concentration, and so forth, none of which have held water, as I discuss in *Looking Down...* (chapter 8). The complaint here is not the search for confounders, it is the tendency only to appeal to confounders.

'Cargo cult science' is the pitiful holding on to a procedure whose theoretical force is spent, yet acting as if it was valid. Researchers still use and write about the Hick reaction time procedure as if its specific ingredient – its slope parameter, which was thought to measure people's 'rate of gain of information' – was useful. It is not. The more boring overall reaction time and its variability are stronger correlates of psychometric intelligence and attention must turn to them (Deary, Der *et al.*, 2001). This can be done without using a Hick procedure.

'Hunter-gatherer science' described the dependency that differential psychologists have upon cognitive scientists, as they await parameters from their cognitive architectures that might correlate with psychometric intelligence. In *Looking Down...* I predicted that the differential psychology of human intelligence differences had been so successful in describing and validating the phenotype that, once the vastly huger community of cognitive scientists developed a taste for individual differences, they would successfully colonise the territory. Results on brain size and functional magnetic resonance imaging are the harbingers of this occupation (Gray *et al.*, 2003).

10. Shot by both sides

To criticise researchers' behaviours within one's research area and also to accuse observers of failing to appreciate the area's achievements attracts dudgeon from both sides.

Looking Down... is a book that can bring people together by providing something to challenge everyone. For the intelligence research diaspora there are chidings about samples and tests, but most of all how it goes about articulating constructs and providing and investigating explanations. For the interested observers there is no surrender on the firm, replicated, yet

provocative findings that put psychometric intelligence high on the list of important human characteristics, and something in which the origins of differences should be investigated.

But the brickbats from both sides are understandable. There is as yet no account of how the brain works. Unlike the kidney, liver and small intestine, the functional units of the brain are not known (Deary, Austin *et al.*, 2000). This is frustrating. It means that differential psychologists have been disappointed by cognitive scientists who did not, after all, produce the cognitive parameters where the key individual differences might lie. One can hardly blame them, then, for using what procedures come to hand and invoking constructs such as 'speed of information

processing', especially when one sees the miscellany of tests and procedures that gather under the umbrellas of 'attention' and, say, 'working memory' elsewhere. And one can hardly blame those with IQ allergies from correctly stating that the correlations are not yet mechanisms. But the surprising and interesting correlations between psychometric intelligence still gather, and investigations on mechanisms continue with impressive force – as evidenced by the references within this piece having appeared largely after the *Looking Down...* was published.

■ Professor Ian Deary is in the Department of Psychology, University of Edinburgh. E-mail: I.Deary@ed.ac.uk.

References

- Burns, N.R. & Nettelbeck, T. (2003). Inspection time in the structure of cognitive abilities. *Intelligence*, 31, 237–255.
- Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. Cambridge: Cambridge University Press.
- Deary, I.J. (2001a). Human intelligence differences: A recent history. *Trends in Cognitive Sciences*, 5, 127–130.
- Deary, I.J. (2001b). *Looking down on human intelligence: From psychometrics to the brain*. Oxford: Oxford University Press.
- Deary, I.J., Austin, E.J. & Caryl, P.G. (2000). Testing versus understanding human intelligence. *Psychology, Public Policy and Law*, 6, 180–190.
- Deary, I.J., Der, G. & Ford, G. (2001). Reaction times and intelligence differences. *Intelligence*, 29, 389–399.
- Deary, I.J., Leaper, S.A., Murray, A.D., Staff, R.T. & Whalley, L.J. (2003). Cerebral white matter abnormalities and lifetime cognitive change. *Psychology and Aging*, 18, 140–148.
- Deary, I.J., Simonotto, E., Marshall, A., Marshall, I., Goddard, N. & Wardlaw, J.M. (2001). The functional anatomy of inspection time: A pilot fMRI study. *Intelligence*, 29, 497–510.
- Deary, I.J., Whalley, L.J., Lemmon, H., Crawford, J.R. & Starr, J.M. (2000). The stability of individual differences in mental ability from childhood to old age. *Intelligence*, 28, 49–55.
- Deary, I.J., Whiteman, M.C., Pattie, A., Starr, J.M., Hayward, C., Wright, A.F. *et al.* (2002). Cognitive change and the APOE e4 allele. *Nature*, 418, 932.
- Deary, I.J., Whiteman, M.C., Starr, J.M., Whalley, L.J. & Fox, H.C. (in press). The impact of childhood intelligence on later life. *Journal of Personality and Social Psychology*.
- Der, G. & Deary, I.J. (2003). IQ, reaction time and the differentiation hypothesis. *Intelligence*, 31, 491–503.
- Gray, J.R., Chabris, C.F. & Braver, T.S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, 6, 316–322.
- Grudnik, J.L. & Kranzler, J.H. (2001). Meta-analysis of the relationship between intelligence and inspection time. *Intelligence*, 29, 523–535.
- Kelley, E.L. (1927). *Interpretation of educational measurements*. Yonkers, NY: World.
- Lubinski, D. & Humphreys, L.G. (1997). Incorporating general intelligence into epidemiology and the social sciences. *Intelligence*, 24, 159–201.
- Luciano, M., Smith, G.A., Wright, M.J., Geffen, G.M., Geffen, L.B. & Martin, N.G. (2001). On the heritability of inspection time and its covariance with IQ: A twin study. *Intelligence*, 29, 443–457.
- MacLullich, A.M.J., Ferguson, K.J., Deary, I.J., Seckl, J.R., Starr, J.M., & Wardlaw, J.M. (2002). Intracranial capacity and brain volumes are associated with cognition in healthy elderly men. *Neurology*, 59, 169–174.
- Neisser, U., Boodoo, G., Bouchard, T.J., Boykin, A.W., Brody, N., Ceci, S.J. *et al.* (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, 77–101.
- Plomin, R. (2003). Genetics, genes, genomics and *g*. *Molecular Psychiatry*, 8, 1–5.
- Plomin, R., DeFries, J.C., McClearn, G.E. & McGuffin, P. (2001). *Behavioral genetics* (4th edn). New York: W.H. Freeman.
- Plomin, R. & Spinath, F.M. (2002). Genetics and general cognitive ability. *Trends in Cognitive Sciences*, 6, 169–176.
- Salthouse, T.A. & Ferrer-Caja, E. (2003). What needs to be explained to account for age-related effects on multiple cognitive variables? *Psychology and Aging*, 18, 91–110.
- Shenkin, S.D., Bastin, M.E., MacGillivray, T.J., Deary, I.J., Starr, J.M. & Wardlaw, J.M. (2003). Childhood and current cognitive function in healthy 80-year-olds: A DT-MRI study. *NeuroReport*, 14, 345–349.
- Vernon, P.A., Wickett, J.C., Bazana, P.G. & Stelmack, R.M. (2000). The neuropsychology and psychophysiology of human intelligence. In R.J. Sternberg (Ed.) *Handbook of intelligence*. New York: Cambridge University Press.
- Whalley, L.J. & Deary, I.J. (2001). Longitudinal cohort study of childhood IQ and survival up to age 76. *British Medical Journal*, 322, 1–5.
- Wilson, R.S., Beckett, L.A., Barnes, L.L., Schneider, J.A., Bach, J., Evans, D.A. *et al.* (2002). Individual differences in rates of change in cognitive abilities of older persons. *Psychology and Aging*, 17, 179–193.