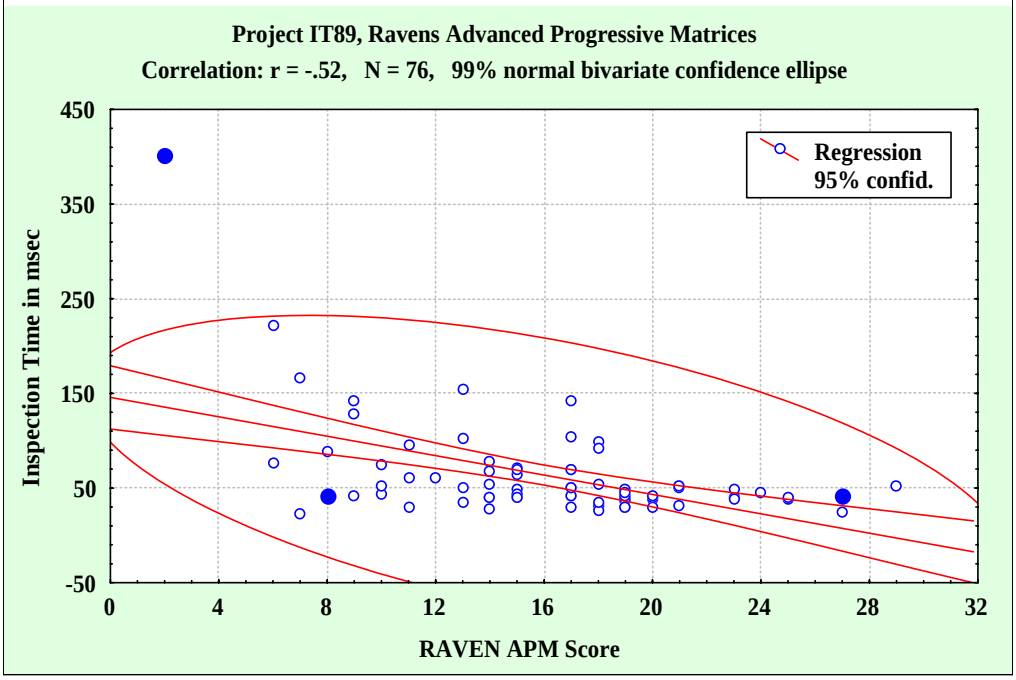
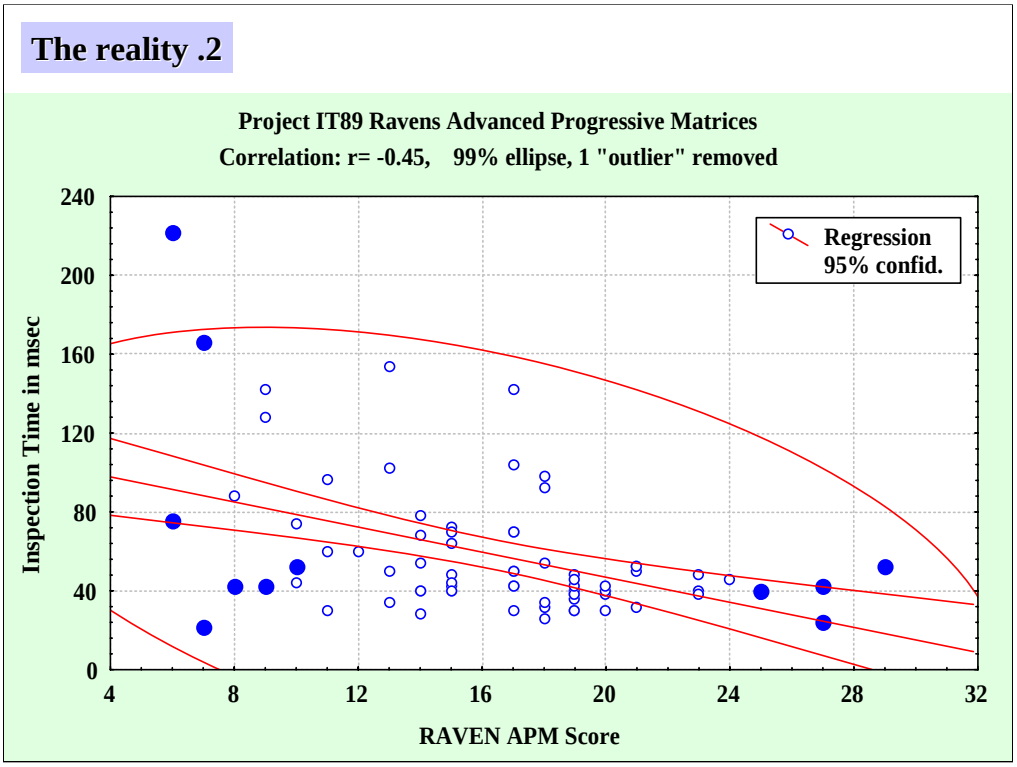
The reality .1



Let us examine what many of the "key" results in individual differences research are really based upon. My own work, with Hans Eysenck at the Biosignal Lab during the 1980s and early 1990s, spanned experimental examination of the **biological correlates of psychometric intelligence** (EEG, AEP, sensory nerve conduction), **chronometrics** (reaction time and inspection time), and **psychometrics** (item analysis, factor analysis, and cross-cultural psychometrics). Our results have been replicated by others, as well as us replicating others' results. However, whilst this effort has served to extend a series of general statements about statistical relationships between psychometric and other kinds of variables, I no longer view this approach as being of any substantive future value. Neither I hope will the audience once they appreciate the reality of these statistical relationships.

The scatterplot above shows the data from an experiment in 1989, our second on the relationship between Inspection Time and IQ scores. I have fitted a 99% bivariate confidence ellipse around the data points - in order to show just how "rare" the single IT of 410msec is. Note the "expected" correlation of -0.52. If I had not plotted the data, I would have been reporting this correlation as given. Note also the two other highlighted points - these pair of respondents score almost at the opposite ends of the intelligence spectrum, yet achieve the same Inspection Time (test-retest reliability for this task is 0.75 (N=72) one-day gap, and 0.83 (N=40) one year gap, the mean absolute discrepancy is 13.3msec).



If I remove the “outlier”, the correlation drops to -0.45. However, I have highlighted some more of the points that seem inexplicable given a model that in some way implicates sensory registration-processing speed as being a substantive causal variable for intelligence. Note also how individuals with very similar and low Ravens scores achieve wildly varying ITs. Do we simply write the “offending” cases off as “measurement error - or should we begin to get very serious about our models, and view these cases as crucial departures from such a model?

This pattern of results replicates across 5 different experiments in which inspection time was available as a variable for such an analysis.

If we talk about inspection time being related to psychometric IQ (where people high in IQ will “tend” to have lower inspection times), we are on safe ground. However, I contend that such ground is about as useful to substantive theory as trying to build the Bide-a-Wee-While Tea Shoppe on a patch of quicksand. No matter that Inspection time has been shown to be related to the “amplitude velocity” of the rising slope of the P200 IT stimulus evoked potential waveform - again, we are in the land of “tendency” and “on average”. I assert that this satisfaction with observing “tendency”phenomena is the hallmark of the current era of individual differences research. It makes for great speculation and wonderful neurocognitive narrative, but is actually bad science. It is one thing to identify a phenomena, quite another to then seek scientific causal explanation of it. Merely continuing with a series of approximating statements and replications hardly seems useful or productive anymore.

The reality .3

Another experiment (Biocog2) … all available data, using the Jackson MAB (group administered WAIS-R equivalent)

Correlations (biocog2.sta)

Continue...

Marked correlations are significant at p < .05000
N=82 (Casewise deletion of missing data)

Variable	FULLIQ	PERFIQ	VERBIQ	IT
FULLIQ	1.00	.91	.93	-.28
PERFIQ	.91	1.00	.70	-.35
VERBIQ	.93	.70	1.00	-.19
IT	-.28	-.35	-.19	1.00

And with "outliers" removed...

Correlations (biocog2.sta)

Continue...

Marked correlations are significant at p < .05000
N=73 (Casewise deletion of missing data)

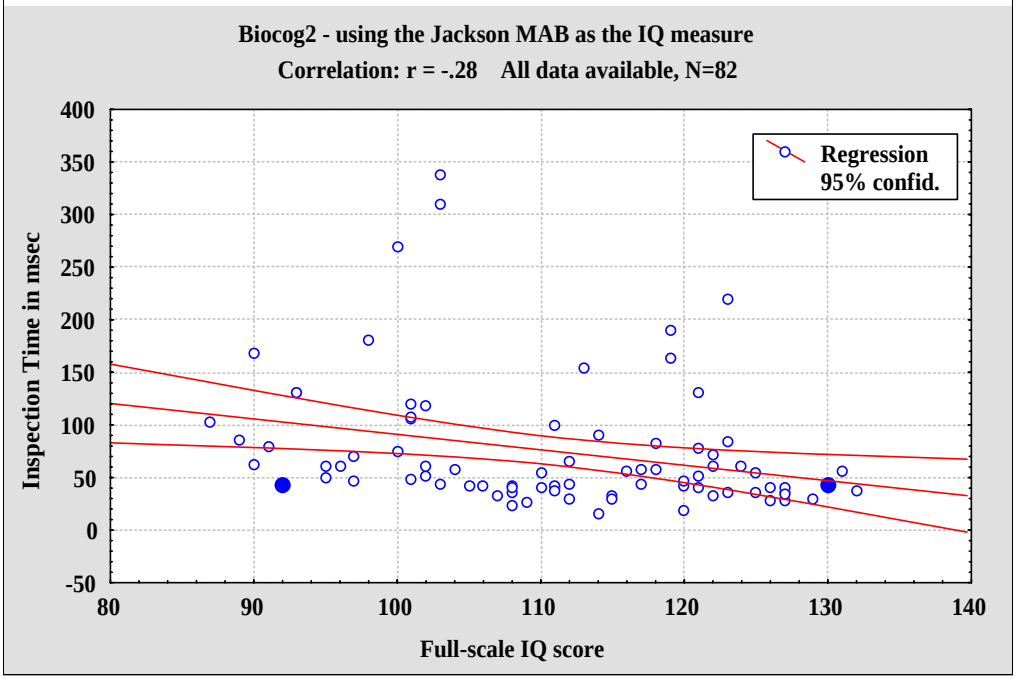
Variable	FULLIQ	PERFIQ	VERBIQ	IT
FULLIQ	1.00	.90	.93	-.41
PERFIQ	.90	1.00	.69	-.42
VERBIQ	.93	.69	1.00	-.35
IT	-.41	-.42	-.35	1.00

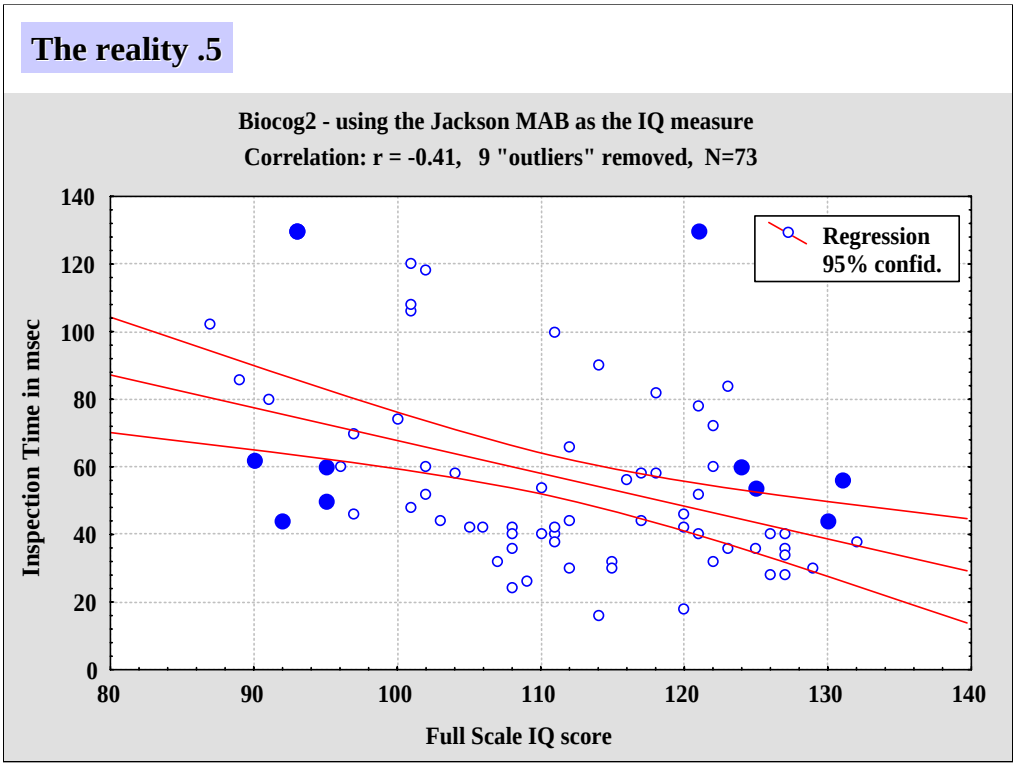
In case some think that this result is a one-off, and perhaps due to the use of the Ravens test instead of a more broad test of ability such as a WAIS-R, I have reported the results from another of our experiments, that from Biocog2. Here you see the correlations between the more familiar Verbal, Performance, and Full-scale IQ and Inspection Times (N=82). I have then removed 9 outliers in order to show how the results change accordingly.

Here, note that instead of decreasing, the correlations between IT and IQ increase with outlier removal. The graphs on the next two slides show the "before" and "after" images, with exactly the same problems.

Again, I reiterate, what kind of "theory" can account for these data. That there is a trend is not in doubt, but this is really part of phenomena identification. A phenomenon (IT correlating with IQ) can be identified and replicated by many independent investigators - but that is really all it is - a phenomenon. We can of course propose that other variable effects are confounding the relationship, hence the "messy" nature of the trend - but now we need to sit down and carefully consider both our initial suppositions about IT and IQ, and just how the other variables might be **deduced** as being responsible for the degredation of a proposed strong IT x IQ relationship. This has the effect of making us think much more carefully about how these variables might achieve degradation of such a proposed relationship, along with the assumption that seems to assert that such a relationship should be strong in the first place. Especially given the **fact** that IQ scores are ordinal, and possess no identifiable unit of measurement.

The reality .4

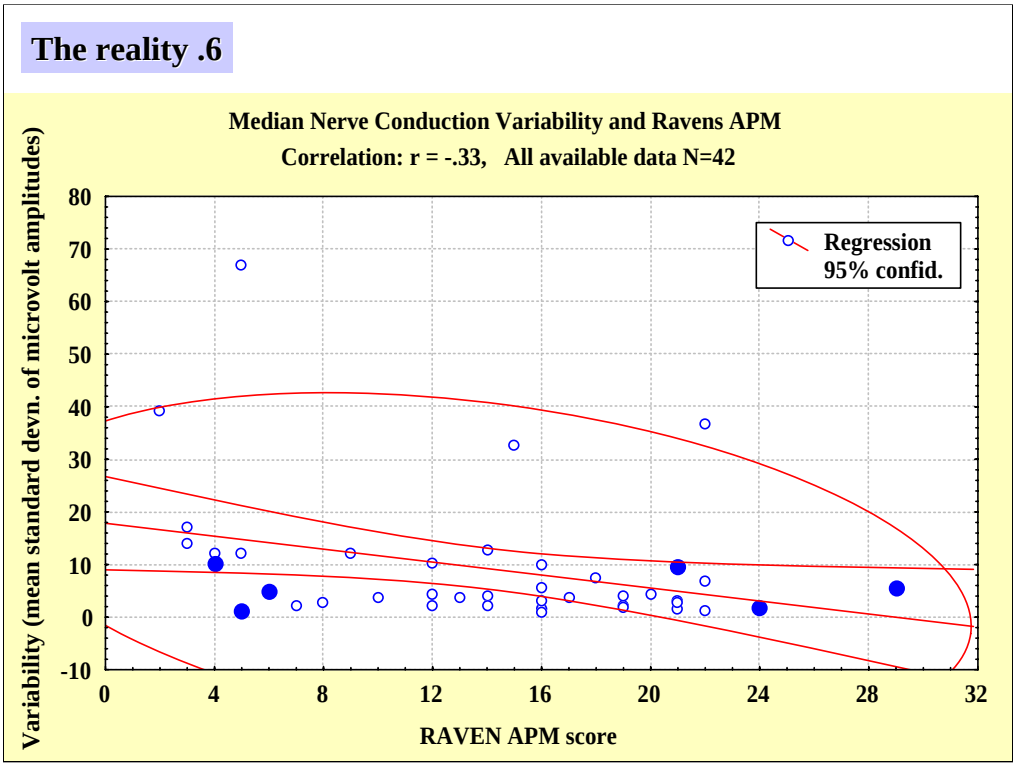




These data exemplify the problem with "cognitive correlates" research. The size of the correlations appear to be substantive, yet, looking at the data, one can see that any kind of coherent causal explanation is not going to be found using such empirical data. You cannot on the one hand insist that "inspection time is a casual variable for IQ", yet look at the blue circles at around IQ 90, and not moderate this statement with "given certain other, as yet unmeasured, variable processes that enable individuals with very low IQ to have Inspection Times (IT) as low as certain individuals with IQ > 125".

Note, I do not deny that a trend might exist – but this is phenomena identification, not science. If I were to propose a causal statement in advance, something like …"IT is causally mediated by brain processes that are also necessary for high IQ", then the above results would contradict such a statement, for we have low IQ subjects attaining low ITs. These subjects, given our theory, do not have sufficient amount or kind of brain processes that permit them to acquire a high IQ score, yet are still capable of performing on the chronometric task as if they had. Something is wrong. Given these results replicate (which they do) on different groups of subjects, then either we have misconstrued the hypothesised causal basis underlying both IT and IQ, or the task is indexing something which is only marginally related (causally) to IQ – in short, it is not a causal variable of IQ, merely one that is related loosely to something which is.

No amount of structural equation modelling or other generalised statistical manipulations can account for these data. The statement "IT is negatively correlated with IQ" is correct, yet actually is almost useless with regard to knowledge acquisition concerning causal explanation of Inspection Time and Intelligence.

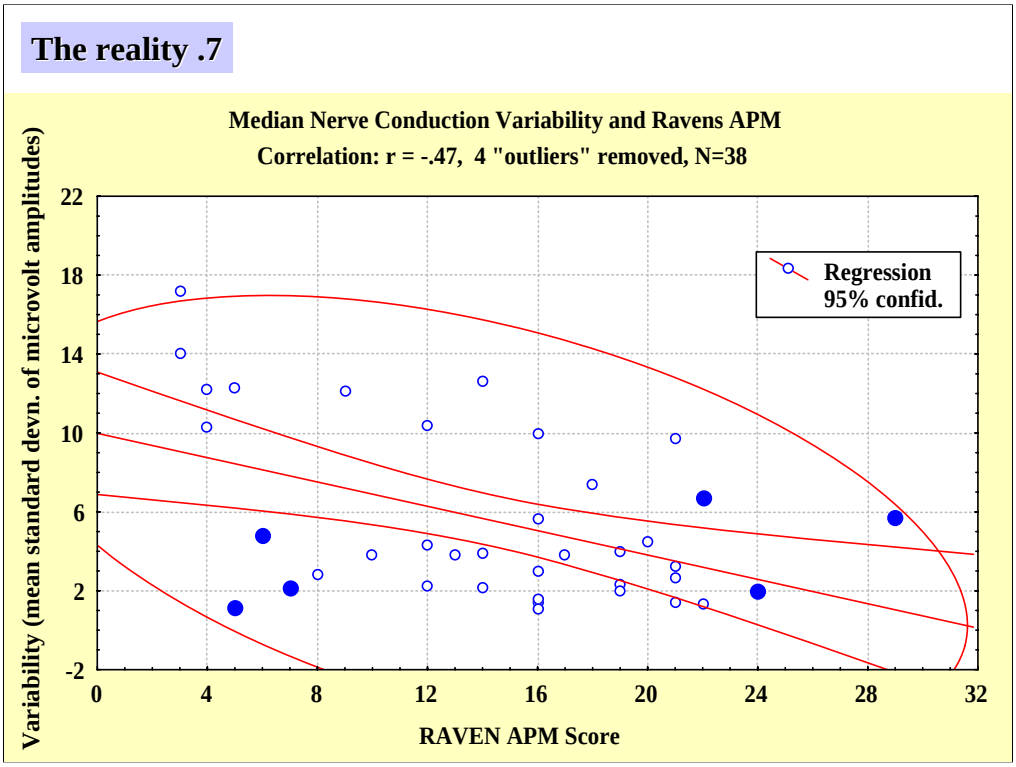


In the next two graphs, I draw upon data from our experiments on peripheral sensory nerve conduction. We implemented two such experiments:

Barrett, P. T., Daum, I., Eysenck, H. J. (1990) Sensory nerve conduction and intelligence: A methodological study. *Journal of Psychophysiology*, 4, , 1-13

Barrett, P.T. and Eysenck, H.J. (1993) Sensory Nerve Conduction and Intelligence: a replication. *Personality and Individual Differences*, 15,3, 249-260

Here we were interested in examining whether sensory nerve action potentials elicited by a constant low current (2ma above threshold) stimulus would provide us with action potential variability data akin to that measured from the scalp EEG using auditory stimulus evoked potentials. The above graph presents data from 42 participants, plotting one of the action potential variability measures against Raven APM scores. The recordings were made 25mm from the wrist, referenced against a neutral electrode at a 90 degree angle to the “active” electrode. The variability measures were computed by calculating the standard deviation of each sample point (sampling at 1.024kHz for 10msec = 1024 sample points, subsequently windowed between ~1.5 and 7.0 msec) across 100 stimulus samples, then taking the mean of these 550 standard deviations. The values of each sample point were expressed in microvolts. The larger SD values represent the contamination of the shock voltage (constant current of just a few milliamps, but voltages of up to 200v in order to achieve this current value - which led to serious potential artefact contamination of some records).



If I remove the 4 “outliers”, I achieve the results above, with a substantive increase in the correlation, but still with “impossible” data under a theory that posits increasing nerve conduction variability causing lower overall IQ.

Maybe I am being unfair in my criticisms here - but what other sensible alternative is there? To posit that conduction variability is but one of several causal variables that account for differences in IQ score is no doubt realistic, but is actually of little help in trying to propose specific theory about the role of nerve conduction variability or these “other causal variables”. In short, these data again highlight the almost “dead-end” nature of such research - we identified a phenomena, replicated it, but are left with description and speculation as the final result. We have no more substantive “causal explanation” of these results than when we first began the experiments. All we have is hazy speculation about a replicable phenomena. As I indicate below, this is all we might ever achieve in individual differences psychology - not because of any inadequacies as investigators, but simply because we are working from an assumption that complex human behaviour is “computable”, or at least can be reduced to certain substantive causal variables that obey a simple, linear and quantifiable form of cause and effect.