

THE 16 PF® NEWSLETTER

Regular attenders at 16PF User Group meetings cannot fail to have noticed over the years that we have a few people who are devoted to using the 4th Edition of 16PF. Their reasons are varied, but several feel that some factors are 'richer' than in the 5th Edition. Such loyalty should not go unrewarded! Accordingly, Rob Bailey, Managing Consultant, Research & Development at OPP addressed us in February. Assistant Editor, Ann Rodrigues reports.

Having had his ears bashed whenever the subject of 16PF 4 v. 16PF 5 (*henceforth referred to as V4 and V5*) has cropped up at a User Group meeting, Rob Bailey decided to take his life in his hands and address the Group on this contentious topic. His talk focussed on the rationale for Version 5 and the main differences between the two versions, both in meaning and psychometric properties; he also looked at the reports available for each.

In a lucid, candid presentation, which even your assistant editor could understand, Rob explained that V5 was developed by a collaboration of Raymond, Karen and Heather Cattell at IPAT, and subsequently adapted for

the UK by ASE in 1994. The rationale was to increase: user friendliness; the questionnaire's ability to relate to test takers in occupational contexts; and precision of measurement i.e. if the meaning of each individual factor is clear and distinct from the other factors, it eases interpretation.

In achieving this however, the questionnaire has been criticised for lacking 'richness' and breadth. [*note from Asst Ed: in a presentation on a previous occasion, the speaker, Pauline Willis, preferred to use V4 as she felt it was better at measuring 'the dark side' of people's personality*]. This is explained by the difference between 'source' and 'surface' traits (the latter being what you see, i.e. behaviours). Cattell opted for breadth and therefore there is more overlapping content in V4. For example, Factor L and Factor E have a high correlation and could be said to measure the same quality twice. V5 eliminated this confusion.

Psychometric properties

The qualities of a good test are appropriate internal consistency, the scales being distinct, reliability (the same results are achieved over time) and validity (it is measuring what it is supposed to measure.)

But it is a juggling act to get the right balance. For example, internal consistency is desirable, but if cohesion of items is very high, the

result would be 'bloated specifics' (*rather as I feel when my husband is trying to explain the finer points of rugby*) with items such as 'I am a lively person', 'my friends would say that I am a lively person', 'other people have described me as lively'. In V4 some of the impulsiveness items in Factor F (liveliness) could also be related to Factor G. This prompted some debate in the Users Group as critics of V5 felt that some of the items were indeed bloated/overly transparent!

Rob then presented us with US data (none available for UK) comparing the internal consistency of Factors A, B, C, E, L, M, N, Q³ and Q¹ - using the 'split half' method rather than test-retest. On the basis that the ideal level of consistency is 0.65 - 0.8, these factors fell below the minimum on V4 and the items were amended in V5 to achieve the desired consistency.

Rob believes that the psychometric properties of the 4th edition do not meet the criteria of a modern psychometric instrument – as for example, those relected in BPS guidelines concerning reliability of items. Hence V4 needs more feedback discussion to get a true interpretation. Therefore requiring a greater degree of skill in the practitioner. It is therefore arguably, better suited to a narrow range of applications e.g. in coaching rather than for selection decisions.

Comparison of meanings

Examining correlation of the factors between the two versions, Factors L, Q1, N and M achieved corrected correlations of 0.3, 0.5, 0.43 and 0.53 respectively, indicating that the two versions are measuring different

things. Rob explained that the core meaning of Factors L is whether we take at face value what others say/do. In V4, L is more concerned with the force of reaction to others (dislike of boss, dislike of people who are conceited), whereas V5 items are phrased more unambiguously around trusting others. The behavioural element of V4 is reduced in V5, and to find the missing themes, you would have to look at Global Anxiety and Global Control.

The core meaning of Factor N is how people use a social mask to hide things from others. V4 attempted to measure judgement, shrewdness, extent of self-disclosure. However, self-reporting of an individual's own judgement and shrewdness tends to be problematic, and therefore V5 looks at the extent of self-disclosure only. Factor M looks at the extent to which people focus on present information or the associations beyond it. In V5 the items refer to specific behaviours such as not liking to talk to conventional people and, in groups, preferring to improve things rather than keep records. V5 focuses more on thoughts and attitudes. Interestingly, in V4 managers score more highly on M than the general population, whereas in V5 it is the opposite. Rob pointed out that in senior managers where the need is to be strategic and see the bigger picture, M+ is more important and this finding has been reflected in other leadership research.

Ways to find the same richness of V4 in V5

Rob suggests looking at clusters of primary factors e.g.

Tension: V4 = Q⁴
V5 = Q4 and C-
Sensitivity: V4 = I
V5 = I, A
Arrogance V4 = L
V5 = H, E, O-

Other updates on V5

- Shorter, less ambiguous questions
- Lower reading age required
- Gender, race, disability bias reduced
- 'Dated' items replaced by more contemporary ones
- All questions now have the same consistent response format, with question mark as the mid response, making the questionnaire easier to complete
- The reasoning questions are grouped together at the end, rather than being spread throughout the questionnaire

Portfolio of reports and services

V4 - a variety of reports, hand scoring, various norms including managerial norms, which should be kept up to date. There is also unique research e.g. by Krug.

V5 - variety of reports, bespoke reporting options, hand scoring, as well as Internet administration, newer norms and new developments from IPAT/OPP

Publications

For those interested in pursuing the topic further, contact OPP for a copy of the white paper Equating and Linking 16PF Editions – advice for practitioners with 4th and 5th Edition

data, by Alan D Mead and Scott Bedwell. Also for a more detailed explanation of V5 factors, “*Essentials Of 16PF Assessment*” by Heather E P Cattell and James Schuerger - available from OPP (and the latter also from Amazon.)

NB A copy of Rob’s presentation can be obtained from Caro Leitzell E-mail admin@leitzell.com

*Ann Rodrigues
Assistant Editor*

DATE FOR YOUR DIARY
Wednesday 25th April
AM – FORENSICS
PM - THE BIG FIVE

CHAIRMAN’S REPORT FEBRUARY 2007

During 2006 we arranged five meetings, but unfortunately the last one had to be cancelled: the ‘social’ format may have run its course and we will revert to a full day’s programme next year.

February - Roy Childs – Team Focus
April – Dr Mike Smith (UMIST)
June – Pauline Willis (BPS)
October – Dr Julian Boon (Leicester University)

Newsletter:

David has continued producing the newsletter this year, and has found sending it by email is now well established and straightforward. It has continued with its very high standard but as always we look to the members for content and interesting stories and articles.

Relationship with OPP

We have continued to build our relationship with OPP and thank them for their support in attending the meetings. I met Penny Moyle in September to have an informal conversation and keep each other up to date. One of the outcomes of that discussion was to invite Rob to talk to us in greater depth about the 16PF 4 and 5 versions.

Administration

During this year Caro Leitzell has continued to manage the administration of the group. On behalf of us all I'd like to thank Caro for all that she has done for us. She continues to be very thorough as well as being friendly, supportive and efficient.

Membership and Communication

The membership has continued to be stable in numbers; we have 83 members. That reflects a number of changes – under the surface – as people have joined and some have left. Where people haven't responded to our requests to be in touch we have now removed them from the members list. We have also asked those we keep in touch with informally as people who occasionally attend but are not members if they want to continue to receive invitations to meetings. That number has now reduced too. So we have a 'clean' database.

Of course we continue to need to attract new members and Xanthy has been doing a great job in promoting the group. We can all do something

by just telling anyone we come across who is a 16PF user about the group.

However our relatively low numbers are still a cause for concern and it has prompted a review of the remit of the group. As a result we sent out a survey in January asking members their views. We have the outcome of the survey and we're going to go into that in detail this afternoon.

The committee

We have continued to benefit from a highly committed group of people on the committee. We have met twice during the year, albeit one time was 'virtual' on a phone conference. These meetings have been very helpful to have a wide ranging discussion about the User Group and how we want to develop and grow it.

As it is the AGM it is time for people to review whether to continue as a committee member.

Of the 7 members of this year's committee, 6 have agreed to be put forward again: David, Hugh, Nicholas, Chris, Ann, and Xanthy. And we have had one nomination, from Michele Williamson.

You'll have gathered the missing name is mine. After 6 years of chairing the committee I believe it is time to step aside and let fresh eyes look at the group and new ideas to come forward.

I have greatly enjoyed my involvement in the committee and would like to thank you for your support and encouragement. I am delighted though to let you know that Ann has agreed to be nominated as chair and I would fully support her nomination and believe she will do a great job.

Looking ahead

As always I would reiterate that the User Group exists only for the benefit of the members – it has no other purpose – and for it to thrive it needs the active interest and involvement of everyone. It would be wonderful if people would let us know if they would like to get more involved in any way – from organising events, to marketing, to getting a website going, to writing articles for the newsletter or indeed editing the newsletter.

We have dates organised for:

April 25 – Lynsey Gozna and
Professor Vince Egan

June 19 – Fiona Patterson and Steve
Woods

October 3 – TBA

November 29 – Karen Blakeley

Belinda Smith
Chairman

§§

NO, IT KARNABINEEZEE

I like reading in bed. Not for long, mind you. Just enough until the eyes close and the book falls to the floor.

Recently, I have been having a clear out. Bibliographically speaking. Among those I have dumped on an unsuspecting Oxfam shop, were three legal books (well, with a reported 3,000 new laws since 1997 alone, my tomes of the 1960s are superfluous), and a number of old books on psychometrics. I have to admit that three of them have been on the bookshelf for a very long time indeed,

with nothing to disturb the even tenor of their days. Certainly not me. They have just sat there, their erudition untapped.

One has the riveting title of Bias in Mental Testing. Another is *The Fourth Book Of Mental Measurements Yearbook* edited by Oscar Buros (1953). The inscription on the flyleaf brought back happy memories. I thought I ought to at least look at it before it went. It had over 1150 pages and weighed about 5 lbs. on the kitchen scales. What sort of friend would wish that on you?

Nevertheless, I decided to skip through it in bed. So I hired a forklift truck and it was deposited at my bedside.

It was of course, impossible to hold whilst lying prostrate, so I sat on the edge of the bed and turned to the index. I'm sure there are people who think that testing has only become a feature of working life since around 1980 – the same people who believe sex was invented around 1963. Not so.

For those not familiar with the Yearbook, it lists all commercially available tests – 793 in this instance - with each being critiqued by up to three academics, (sometimes with what may appear to the layman to be a good dose of envious spite.)

I was surprised at the breadth and variety of the tests in the pre 1950s. Personality and aptitude – of course. Numeracy and verbal ability, certainly. But also etiquette, handwriting, religious education, algebra, physics, music and art. There were 4,417 references to the construction, validity and use of the tests.

Raymond Cattell was well

represented. I noticed his *IPAT Humor Test Of Personality*. Form A included 91 pairs of jokes of which the test taker was asked to choose the funnier. The test reviewer concluded that the test provided a novel basis for interesting research only, at that stage of its development. No doubt that was Cattell's intention. One can hardly imagine it being used as a selection tool. Except perhaps for actuaries, who pass if they score nil? I have it on good authority that the instructions counsel test takers that on no account should they laugh out whilst completing the test, as it might disturb others.

Of course I had to turn to the reports on 16PF. The reviewer noted that the items are of two types, each with three possible answers: the "Yes, No, In Between" with which we are familiar, and a forced choice, for example: "Would you prefer the life of a) an artist, b) a YMCA secretary, c) Uncertain. The reviewer went on to say that at that time, many of the items looked similar to older questionnaires and interest inventories, but the manual did not indicate the method of derivation of items or the method whereby observer-rating factors were translated into questionnaire items. He felt that the items were less subtle and penetrating than the *Guilford-Martin Inventory of Factors*, or the revised *Kuder Preference Record*.

There were no reliabilities reported for the separate forms A and B and the manual suggested that both forms should be given in a single test. "*Hand scoring by transparencies involves repositioning and stapling the sheets of the test booklet. A separate answer sheet can be machine scored with six sets of keys, which fit either Form A or*

B. Positions on the answer sheet do not line up with items on the test booklet."

The reviewer ended by saying that "*in general, the utility of the 16 factors remains to be demonstrated and the neutral observer will probably want better evidence of the purity of the factor scores before he uses the test as an analytical research instrument*". My view is that the "*utility of the 16 factors*" has been demonstrated in abundance over the last 60 years – yes, of course with revisions reflecting societal changes. Cattell's record, and subsequent editions of the 16PF, showed his extraordinary zeal in pursuit of identifying the mainsprings of human behaviour. I for one, have cause to be very grateful for the valuable information provided by 16PF with, of course, all the caveats that must apply to such a statement.

The Yearbook also refers to Cattell's *Culture Free Test* published by the Psychological Corporation in 1944, and the *IPAT Culture Free Intelligence Test* (1950).

A number of other tests were familiar to me. I was encouraged to see that ACER Mechanical Reasoning and the Revised Minnesota Paper Form Board Test were there, both of which I had used extensively over the years, and were regarded as a powerful predictive combination of tests for engineering ability. I can testify to that. ACER Mechanical Comprehension was there, also the MMPI.

The Rorschach test was of course there in the *Character – Projective Section*. Hermann Rorschach devised it in 1924 and the 1953 edition of the Yearbook cited 219 references to this test. And the reviews? I shall

hold that over for a future Newsletter. This book is a treasure chest.

Following the test reviews there is a comprehensive section on books. Again, I turned to see if Cattell featured anywhere. Of course he did! One of his books justified a review of 12 pages. The title rang a bell, *The Description and Measurement of Personality* (1946). I had it on the bookshelf. First edition no less! No Oxfam for you, I said. Turning to the end of the review, I read: *"[this book] is the fruit of immense labour and immense research – the author himself has calculated over 20,000 correlation coefficients and he gratefully acknowledges the labours of numerous assistants. The total effect is rather overwhelming and the bemused reader will sometimes rub his eyes and ask himself what it is really all about. Dr Cattell seems to be carrying the art of taking robots to pieces about as far as it will go."*

Skimming the pages, I saw reviews of three more books by Cattell. I also noted a reference to one book with the ungrammatical title of *This Here IQ*. Not by Cattell I should add.

At this point I must tell you that when I last went to *The Introvert Arms*, where we discuss things like split-half reliability, coefficients, ipsativity, and other things even less interesting than *Gardeners' Question Time*, I saw the wimpish Hubert, and I could not resist telling him about the Yearbook and the Cattell reviews. Mercifully, his companion, the aggressive and overbearing Angus, was not there, and so Hubert's response to my account of Cattell's work was a predictable "Karnabineezee." Every time I marvelled at some piece of Cattell

scholarship, he seemed to say "Karnabineezee".

I see that my edition of the Yearbook is now only available from Buros on microfiche. And it will cost you \$50. So the real thing must be worth a bit. And to think I was going to feed it to Oxfam.

NB The 16th edition of the Yearbook came out in 2005 price \$195

Our long-standing member, Dr Hugh McCredie, has addressed the Group on many occasions, his topics often relating to research he undertook with the managers of his former employer. Here, he deals with executive intelligence.

Several years ago I succeeded in correlating 16PF4 Factor B, Abstract Thinking, and another verbal index of general ability, with a purported measure of managerial intelligence based on the work of the late Elliott Jaques. Aha! I concluded, verbal ability is likely to contribute to managerial effectiveness.

In the intervening years, I carried out much more searching study into the nature of the managerial personality and the predictors of managerial success and I came to question the interpretation of my earlier findings. First of all, both Dave Bartram (1992) and I (2004) discovered that the average IQ of fairly senior managers, as measured by 16PF4, was fairly modest: below 80th percentile. Secondly, I discovered that whilst Factor B correlated with overall managerial performance it did so to a

lesser extent than conventional, longer measures of general ability.

These findings left me puzzled, especially since I had got into the habit of re-testing managerial candidates with low Factor B scores using a non-verbal IQ test: Raven's Standard Progressive Matrices. More often than not, in percentile terms, candidates' Raven scores were much better than those from 16PF4.

A possible answer to the puzzle emerged when I got hold of and read a recent re-issue of a 1949 classic by Hans Eysenck, *Dimensions of Personality*. Eysenck found that extraverts' performance on non-verbal ability tests was better than their performance on verbal tests whilst introverts did best on verbal tests. This finding struck an immediate chord since both Bartram and I had discovered a mean Sten score of nearly eight for our managerial populations.

Could it be, I asked myself, that for managerial work the type of ability measure that best predicted performance was a non-verbal one? I was encouraged to think this way when I read of Raven's non-verbal test that it was measuring the ability to make sense out of confusion which, the manual suggested, was at the heart of managerial work. It would be useful, I thought, to test managers with both verbal and non-verbal tests and then explore which of these correlated strongest with overall managerial performance.

Against this background, I treated myself to a day out at the annual BPS Division of Occupational Psychology Conference in January. (Is there no end to this man's quest for excitement I hear you ask?) In one of the

sessions, Rainer Kurz showed us the results of a battery of three aptitude sub-tests - numerical, verbal and diagrammatic (non-verbal) - correlated with overall managerial performance. Notwithstanding the fact that Rainer's sample was a small one (50), to my delight the clear winner was the non-verbal sub-test.

Intelligence tests were first produced over 100 years ago to predict academic ability and, consequentially, they were verbally saturated. The way in which new tests are first validated, is against the established ones so that the verbal bias gets perpetuated. What I have outlined above are early signs that managerial work might need a different type of intellect where the capacity to sort out confusion, as measured by non-verbal scales, is more important than excellence in verbal ability.

Dr Hugh McCredie is with Coordinates: Management Assessment and Development

§§

Several issues ago, we started to explore management 'fads' - those solutions that were going to revolutionise our working lives, and usually didn't. Inevitably some management thinkers and their ideas have been more durable than others. Here's one.

Most of our readers will have found instances of a lack of motivation in the workplace, and may even have been asked what can be done about it. I

can remember one such example quite clearly. We were having difficulty with a group of workers, who just were not sufficiently well motivated. Taking days off was the most obvious example of this.

The production director called me in. What was I going to do about it?

I recall coming up with a few suggestions. Then he said: "What about the Hawthorne Experiments. Elton Mayo."

I laughed. "Come on Dennis," I said. "For heaven's sake! That was donkey's years ago. The last I heard of all that must have been in the sixties. And Mayo's work was in the 1920s and 30s. We don't want stuff that's now well over 60 years out of date. If we're going to look at this seriously we need to tap into up-to-date stuff."

He seemed somewhat abashed.

"Look," I said, "I'll dig around and see what I can come up with."

In a nutshell, and much to my astonishment, I could find no major study, that threw any more light on motivation, than Mayo's work before the war. Nothing came near it either as a longitudinal study or in quality. Some theories, yes. But no hard evidence of high quality.

Elton Mayo (1880 – 1949) was the Australian born pioneer of industrial sociology and is credited with instilling the "human factor" in management theory.

The legacy of his ideas is evident in many modern management theories that stress communications, corporate culture and leadership.

Although Mayo's most important work took place at Harvard, where he remained until two years before his death, his work on industrial unrest

drew on an eclectic early career - medical training in London and Edinburgh, a partnership in an Adelaide printing company, teaching mental and moral philosophy at Queensland University and pioneering work on the psychoanalytic treatment of shell shock during the First World War.

He transferred some of his ideas relating to nervous problems in industry, arguing that: "So long as commerce specialises in business methods which take no account of human nature and social motives, so long may we expect strikes and sabotage to be the ordinary accompaniment of industry."

His best remembered work arose from the lighting experiments known as the "Hawthorne effect", which took place at the Western Electric company's plant at Hawthorne, Chicago. The researchers set factory lighting at different strengths to discover what would have the greatest effect on productivity. The light was brightened for one group and left unchanged for another. The result was an enigma: both groups increased their production.

Subsequently Mayo found that any change to the workers' conditions led to a rise in output.

He concluded that the employees had gained work satisfaction by the attention lavished on them and by the co-operation between informal groups in the workplace. (*We need to remember that this work was carried out when attitudes to work and to employees were very different from those held to day.*)

Mayo was the first person to reveal the inadequacy of studying the individual worker in isolation and of

only taking into account the purely physical aspects of an industrial environment. He argued that the main objective of management was to encourage co-operation and teamwork among employees.

His early researches however were developed along orthodox “work study” lines.

At the Hawthorne works Mayo set up the Relay Assembly Test Room, a section of six girls who assembled forty intricate parts of a telephone relay system. Production levels were recorded over five years. Most changes to the working conditions resulted in increased output. One of the final changes reverted to the original poor conditions over a period of twelve weeks. Output rose to a record level. It was evident that production had risen because of a change in the girls’ attitude to their situation and to authority.

A further piece of research in the Bank Wiring Room was designed to investigate the social pattern of a group of workers. It concluded that the men regulated their own output. The strength of this group was such that no one attempted to reach the official production targets.

“Worker-management conflict may be due less to ostensible reasons for the dispute such as tea-breaks or insufficient light, than emotional attitudes,” Mayo said. Workers were ruled by the “logic of sentiment”, while managers were moved by the “logic of cost and efficiency”.

Mayo’s stress on communications and good personal relationships appears to be a more humanistic philosophy than for example the “scientific management” principles espoused by Frederick Taylor. But it

can be criticised for its supposition that conflict can be sorted out just by improving communication between management and workers.

Nonetheless, the “Hawthorne effect” remains widely cited by academics.

Management theory and consultancy projects continue to be influenced by the idea that workers want recognition, security and a sense of belonging.

It seems extraordinary now, that at a time of deep depression, the 20s and 30s, a large organisation could commit so wholeheartedly to such a lengthy piece of research.

There had been nothing like it before, and certainly nothing since.

So I went back to Dennis, the production director, and conceded that Hawthorne was still relevant.

He gloated.

Editor

CONSULTING EDITOR

Penny Moyle
OPP Ltd Tel 01865 404500
penny.moyle@opp.eu.com

MEMBERSHIP ADMINISTRATOR

Caro Leitzell Tel 01962 880920
admin@leitzell.clara.co.uk

EDITOR

David Roberts Tel 01509 852870
david@evanroberts.fsbusiness.co.uk

ASSISTANT EDITOR

Ann Rodrigues Tel 07947 159848
ann@acer-hr.co.uk

© 16PF is a registered trademark of the Institute for Personality and Ability Testing, Inc (IPAT).

IPAT is a wholly owned subsidiary of OPP Ltd.