

# ***THE 21<sup>ST</sup> C SCIENTIST- PRACTITIONER: DOING & APPLYING SCIENCE CASE-BY-CASE***

---

Neville M Blampied,  
President, Division 6 IAAP  
Presidential Address,  
IACP, Paris, July, 2014

Supported by the Erskine Fund,  
University of Canterbury  
[Neville.blampied@canterbury.ac.nz](mailto:Neville.blampied@canterbury.ac.nz)

# Can you identify this man?

He has exerted an enduring and profound influence on psychology to the present day

To win the prize text *His name; Your name* to  
Mobile +64276179703



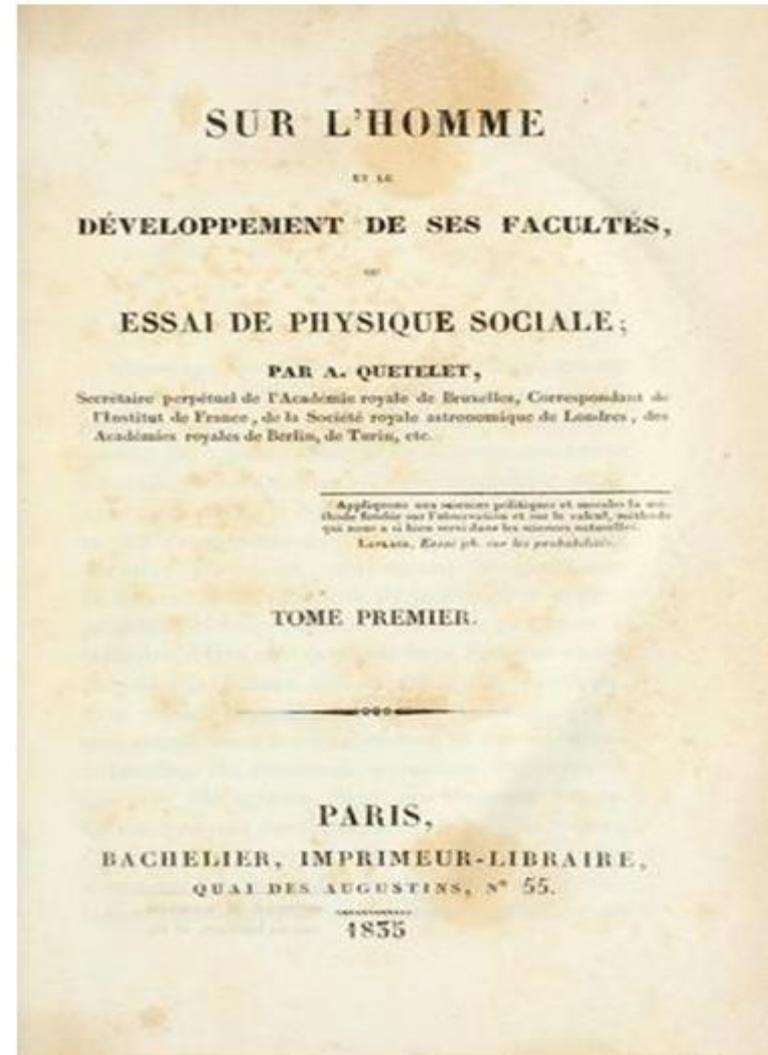
# Adolphe Quetelet – 1796 - 1874



# Quetelet's Famous Book

*A treatise on man and the  
development of his  
faculties*

(English translation, 1842)



# Quetelet's influential ideas (1)

## ***L'homme moyen*** – the average man

*If an individual ... possessed all the qualities of the average man he would represent all that is great, good, or beautiful.*

[Quetelet, 1842, p 100]

# *L'homme moyen* – a cultural meme

*The average man ... still lives in the headlines of our daily papers*

[Steigler, 1986, p 171]

# Men have larger brains than women

BRITAIN

## University study finds significant gender difference in human brains

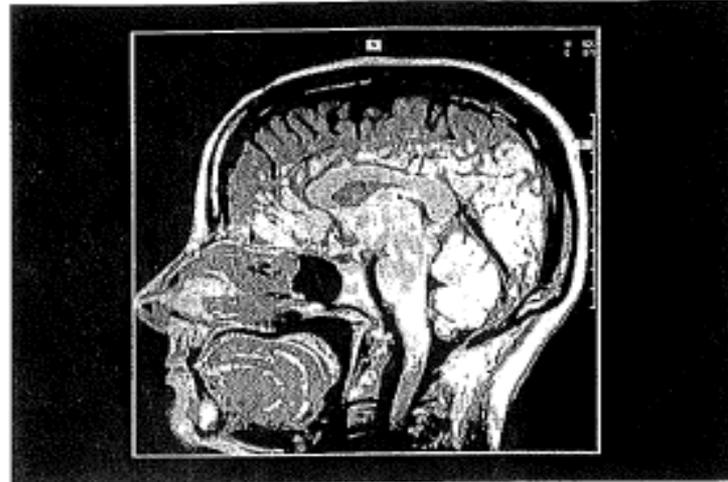
**Billy Kenber**  
in London

Men have larger brains than women and are structurally different in the areas associated with emotional control, short-term memory and risk-taking, according to research.

A study by scientists at the University of Cambridge concluded that the volume of the average male brain was about 8 to 13 per cent larger than a woman's.

It also found significant structural differences in men and women's brains. The study found that in women, the left frontal pole, which is thought to be connected to emotional control, was much denser than in men.

Female brains also had more grey matter in the frontal gyri, which is associated with language, working memory and calculating risks, and the



**Less dense:** The study could shed light on gender-linked psychiatric conditions.  
Photo: GETTY IMAGES

Heschl's gyrus, linked to hearing.

Men's brains showed greater density and area in the cerebellum, linked to co-ordination and balance, and the putamen, associated with

motor skills and movement. The team, which studied more than 126 research articles on brains published over the past 23 years, stressed that they weren't sure how the differences

might affect intelligence, behaviour or mental health.

Amber Ruigrok, who carried out the study for her PhD, said: "For the first time we can confirm that brain size and structure are different in males and females.

"We should no longer ignore sex in neuroscience research, especially when investigating psychiatric conditions that are more prevalent in either males or females," she said.

Professor John Suckling said that differences in the limbic system, the part of the brain that controls emotions, might help to treat psychiatric conditions such as autism, schizophrenia and depression, which affect one gender more than the other.

"This new study may therefore help us understand not just typical sex differences but also sex-linked psychiatric conditions," he said.

The Times

Don 15/02/14

## *Average Male 4,000% Less Effective In Fights Than They Imagine*

WASHINGTON—Contradicting the long-held belief that they would just go off and destroy anyone who tried to mess with them, a Department of Health and Human Services report published Thursday revealed that U.S. males would be on average 4,000 percent less effective in a fight than they imagine. “Despite the typical American male’s conviction that he would viciously beat down anyone who came at him and end the whole thing with one punch, we found that in the event of an actual violent altercation, most adult men would almost certainly injure themselves far worse than any assailant,” read the 80-page report

[*The Onion Report*, Jun 19, 2014]

## The British are slightly happier than the French

Measuring happiness and wellbeing is all the rage among national statisticians, writes Anthony Reuben.

In its [latest release](#) on the subject, the Office for National Statistics compares how happy British people are with how happy people are in the other 27 countries of the European Union. ... In the 2011 European Quality of Life Survey, the UK came 11th out of 28, with 71.8% of over-16s saying they were that satisfied.

That put the UK slightly higher than **the EU average** of 69.3% and a hair's breadth above France in 12th place on 71.6%.

[BBC News, 23 June 2014

# Quetelet's influential ideas (2)

## The normal distribution & human attributes

*L'homme moyen ... was ... a moral ideal. ... "all our qualities, in their greatest deviation from the mean, produce only vices".*

[Porter, 1986, quoting Quetelet, 1853, p 49]

Measures of human physical and 'moral' characteristics are normally distributed

Attributes = ideal (average) +/- error deviation (*vices*)

*It was Quetelet's notion that every individual represented an attempt by nature to achieve perfection, but that, like a person shooting at a target, nature's aim was never exact.*

[Johnson & Pennypacker, 1993]

# Quetelet's legacy lives on in our statistics

## Sir Ronald Aylmer Fisher

- *Statistical Methods for Research Workers* (1925).
- *The Design of Experiments* (1935)
- Devised (or improved on)
  - Control groups
  - Randomization
  - Factorial designs, &
  - Analysis of Variance & Null Hypothesis Tests (NHST)

*Every experiment may be said to exist only in order to give the facts a chance of disproving the null hypothesis. (Fisher, 1935).*



# Fisher's Influence on Psychology

- First use of ANOVA in psychology in 1935
  - 16 more examples by 1940
- Psychology ultimately adopted an “uneasy synthesis” of Fisher's statistics with those of J Neyman and E. Pearson (Gigerenzer, 1987).
- The “inference revolution” from ~ 1950
  - Although their initial understanding of these techniques was often defective, psychologists embraced them with considerable enthusiasm. (Danziger, 1987, p 35).*

# The inference revolution

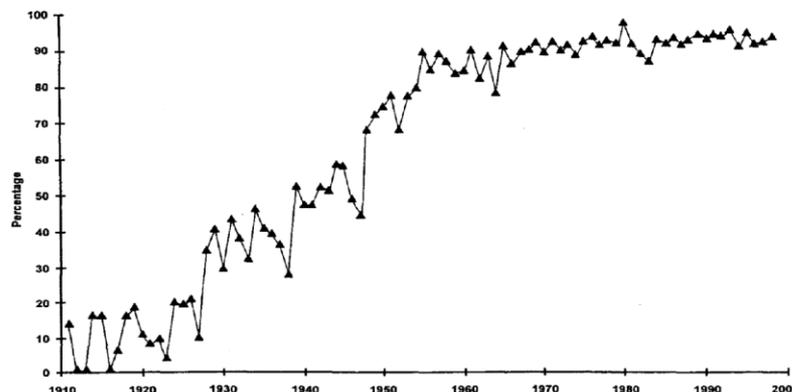


Figure 1.1. Percentage of articles reporting results of statistical tests in 12 journals of the American Psychological Association from 1911 to 1998. From "The Historical Growth of Statistical Significance Testing in Psychology—And Its Future Prospects," by R. Hubbard and P. A. Ryan, 2000, *Educational and Psychological Measurement*, 60, p. 665. Copyright 2001 by Sage Publications. Reprinted with permission.

*By the early 1950's, half of the psychology departments in leading American universities offered courses on Fisherian methods. ... By 1955, more than 80% of experimental articles in leading journals used inferential statistics to justify conclusions ... Editors ... made significance testing a requirement ... and used the level of significance as a yardstick for evaluating quality.*  
(Gigerenzer, 1991, p 255.)

# Major Development # 1

## The *Standard Model* of Research

- Postulate hypotheses about **populations**
  - Recruit as large a sample as possible.
  - Randomly allocate participants to treatment conditions.
  - Aggregate individual data into **group averages &** generate sample statistics.
  - Draw inferences about the population from the sample statistics.
  - Use **Null-hypothesis significance tests** [NHST] to determine the scientific legitimacy of the results (typically, requiring  $p < 0.05$ )
- A scientific fact should be regarded as experimentally established only if a properly designed experiment rarely fails to give this level [0.05] of significance*
- (Fisher, 1926, p 504).

# Major Development #2

## The Scientist-Practitioner ideal

- 1907: *The methods of clinical psychology are ... invoked wherever the status of an individual mind is determined by **observation and experiment*** (Witmer, 1907, p 251).
- 1947: A clinical psychologist has *applied and theoretical knowledge in three major areas: **diagnosis, therapy, and research***. (Committee on Training in Clinical Psychology, 1947, p540)
- 1949: The Boulder Model – The Boulder Conference on Graduate Education in Clinical Psychology (Raimy, 1950): **Training to include training in diagnosis, therapy, research methods, statistics, and a research-based PhD.**

# The two developments converge

#1: Science = Standard Fisherian experimentation & statistical analysis

#2: Therefore, **science** in Scientist-practitioner = Standard Model science

And, therefore, clinical/applied psychology research overwhelmingly done within Standard Model

*Empirically supported psychological treatments ... [are those] that have been exposed to evaluation **using the accepted methods of psychological science.***

[Kendall (1998) p3.]

# Criticisms of the standard model

NHST called

- a *religion* (Cohen, 1990)
- an *addiction* (Schmidt & Hunter, 1997)
- an *enduring tyranny* (Loftus, 1991)
- a *justification for betraying the evidence of the raw data* (Faverau, 1993)
- *is among the most questionable things we do* (Tretweiler & Stricker, 1998)
- *Is the most bone-headedly misguided procedure ever institutionalized in the rote training of science students* (Rozenbaum, 1997)
- *an empty and damaging ceremony* (McCloskey & Ziliack, 2008)
- *a ... kind of sorcery, a form of statistical shamanism* (Lambdin, 2012)

# Paul Meehl on NHST

*... Sir Ronald has befuddled us, mesmerized us, and led us down the primrose path. I believe that the almost universal reliance on merely refuting the null hypothesis as the standard method for corroborating substantive theories [in psychology] is a terrible mistake, is basically unsound, poor scientific strategy, and **one of the worst things that ever happened in the history of psychology.***

Meehl, P.E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, 46, p 817

# Responses to criticism

# Responses to criticism

- There is no comprehensive, definitive refutation of the criticism of NHST!

*Generally speaking, most defenders of NHST believe that it has been misinterpreted and badly used for decades.*

[Balluerka, et al. (2005), p 64]

# Unhappy consequences for applied research

Reviews of applied research have shown major issues

- Numerous violations of test assumptions
- Persistent lack of statistical power
- Misuse and misinterpretation of statistical tests

# Unhappy consequences ...

- Abuse of alpha levels (*marginally significant*, etc)
- Extreme misinterpretation and misuse of  $p$  values
- Belief that small  $p$  values mean that experimental hypotheses are confirmed
- Confusion of statistical with clinical significance
- Failure to consider practical/clinical significance of findings\*

*\*only 40% of articles made any attempt to discuss clinical significance. This is serious. In a major journal dedicated to the research of psychotherapy and other interventions, clinical significance should be relevant to more than 40%... [Fidler, 2005]*

Dar, Serlin, & Omer (1994). Misuse of statistical tests in three decades\* of psychotherapy research. *Journal of Consulting & Clinical Psychology*, 62, p77

- 1967 – 1988

Fidler, Cumming, Thompson et al. (2005). Towards improved statistical reporting in the *Journal of Consulting & Clinical Psychology*. *JCCP*, 73, 136 -143

- 1993 - 2001

# Development #3: Evidence-based practice

*Evidence-based medicine* is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of **individual patients**.

(Sackett, et al., 1996, p 72)

*Evidence-based practice in psychology* is the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences. ... [It] **starts with the patient** and asks what research evidence ... will assist the psychologist in achieving the best outcome ... [and] articulates a decision-making process for integrating multiple streams of research evidence...into the intervention process.

(APA Presidential Taskforce on Evidence-based Practice, 2006, p273)

*Evidence-based practice* is a transdisciplinary, **idiographic** approach

(Spring, 2007, p 611)

# BUT

Standard model research does not yield evidence that fits  
the needs of practice

This is not a recent discovery

*...there is a growing disaffection from traditional  
experimental designs and statistical procedures which are  
held **inappropriate to the subject matter**... [particularly] in  
the area of therapeutic change...*

[Bergin & Strupp, **1970**. *New dimensions in psychotherapy research*].

# Problem is intrinsic to the Standard Model

*Hypotheses are about populations*

Not about

- The sample
- The individual cases in the sample

Yet, Evidence-based Practice

- *starts with the patient*
- *is idiographic*

# The “double standard” in Psychology

*Throughout its history as a science, psychology has been plagued by a double standard in its treatment of the individual subject...*

*In psychological discourse (both scientific and applied) the individual ... is constantly given high relevance. In contrast, **the individual case is usually forgotten** in the practice of psychological research **because it is replaced by samples** of subjects that are assumed to represent some general population (Valsiner, 1986, p1).*

# Getting from the average to the individual

How? Well – we can't, legitimately

So

- We attribute generic, prototypic, ideal status to the population average

*We present our significant result as if it were about all mankind and all time (Gigerenzer, 1987, p 13 – 33).*

- We generalize (inductive generalization) from the average to all members of the putative population

*...characteristics of that abstracted individual may easily become attributed to particular concrete individuals with whom psychologists ... work (Valsiner, 1986, p1-2).*

# Conclusion - for evidence-based practice we need a new way of doing science

*Unanswered questions concerning individual[s]... will continue to puzzle the ... practitioner as he or she works with whomever happens to appear... **This is the one single factor that requires substantial alterations in ... the way we do science.** ... a methodology that highlights the *individual* and, at the same time, *maintains the integrity of an empirical and scientific approach* to the study of behavior. [Barlow, Hayes, & Nelson. (1984). *The scientist-practitioner: Research and accountability in clinical and educational settings*, p 52-53.]*

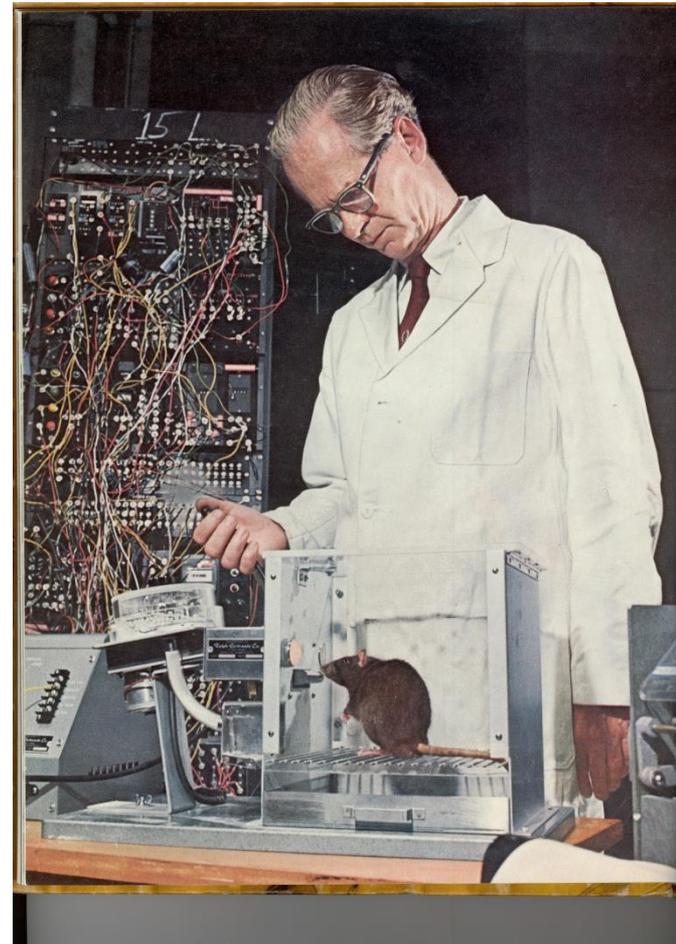
# Single-case research – Origins

*A physician ... is by no means physician to living beings in general, nor even physician to the human race, but rather, physician to a human individual.* [Claude Bernard, 1865/1949, p 92]

# B.F. Skinner – Experimental single-case research

*... the kind of science proposed here belongs on the non-statistical side. ... In placing itself in that position it gains the advantage of a kind of prediction concerning the individual that is necessarily lacking in a statistical science. ... the student of behavior [cannot] predict what a single organism will do if his laws apply only to groups. Individual prediction is of tremendous importance, so long as the individual is to be treated scientifically.*

Skinner (1938). *The behaviour of organisms*, p 443-444



# Single-case research - Application

*The individual is of paramount importance in the clinical science of human behaviour change. ...*

*Recent methodological developments [single-case research] ... provide a base for the establishment of a true science of human behavior change with a focus on the paramount importance of the individual.*

[Hersen & Barlow, 1976]



# Applied Single-case research

What it does not do:

- Rejects sampling theory.
- Does not average over participants.
- Does not use NHST-based inference.

# Single-case research ...

What it does do:

- Maintains a commitment to quantification
- Uses rigorous experimental designs that permit causal inferences
- Uses Visual Analysis as its key data analysis technique
- Uses Replication as its key inference procedure

# The logic of Single-case experimentation

Key tasks of any experimental design

1. **Detect change:** Must permit the detection of any change in the dependent variable coincident with the application of the independent variable.
2. **Permit drawing valid causal inferences.** This involves defending the claim that the application of the independent variable **caused** the change in the dependent variable,

# The logic of Single-case research design

Key tasks of any experimental design

**Detect change:** via **baseline – intervention** phase comparisons

All single-case designs are made up of combinations of baseline – intervention phases

**Permit drawing valid causal inferences:**

via **replication**

Different replication strategies (within and between participants) yield different single-case research designs.

# Example – Withdrawal/Reversal

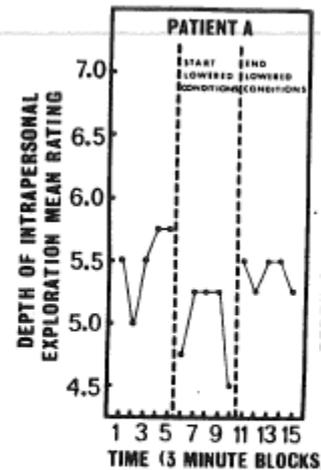


Fig. 5-12. Depth of intrapersonal exploration. (Fig. 4, p. 122, redrawn from: Truax, C.B., and Carkhuff, R.R. Experimental manipulation of therapeutic conditions. *Journal of Consulting Psychology*, 1965, 29, 119-124. Copyright 1965 by the American Psychological Association. Reproduced by permission.)

# Example – Withdrawal/Reversal

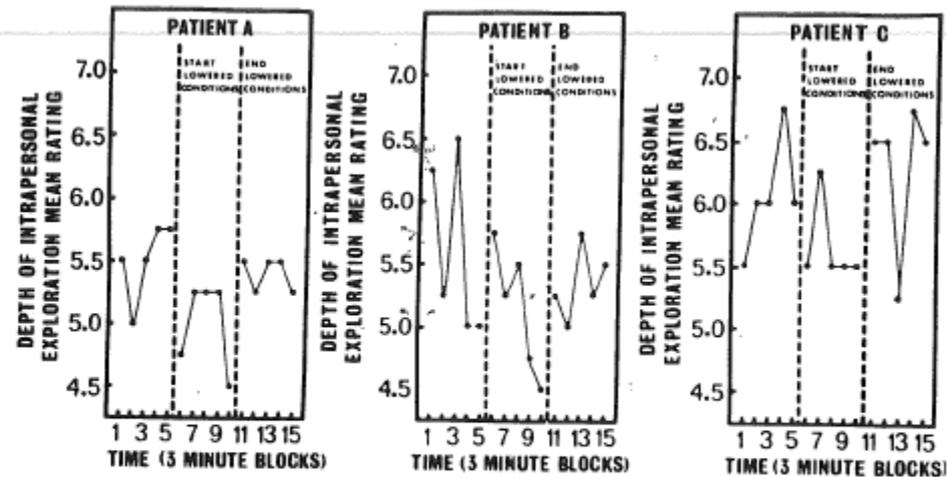


Fig. 5-12. Depth of intrapersonal exploration. (Fig. 4, p. 122, redrawn from: Truax, C.B., and Carkhuff, R.R. Experimental manipulation of therapeutic conditions. *Journal of Consulting Psychology*, 1965, 29, 119-124. Copyright 1965 by the American Psychological Association. Reproduced by permission.)

# Example Multiple- baseline

A multiple-baseline  
across trauma  
scenes

[Saigh, 1989]

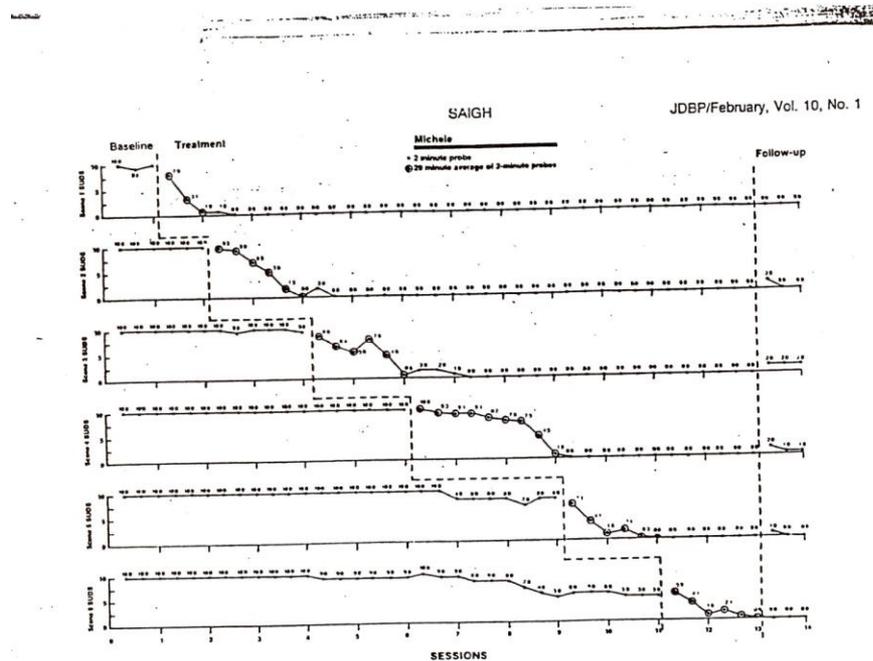


FIGURE 1. Michele's subjective units of disturbance (SUD) levels across scenes.

# Example Multiple- baseline

A multiple-baseline  
across trauma  
scenes, replicated  
across two cases  
[Saigh, 1989]

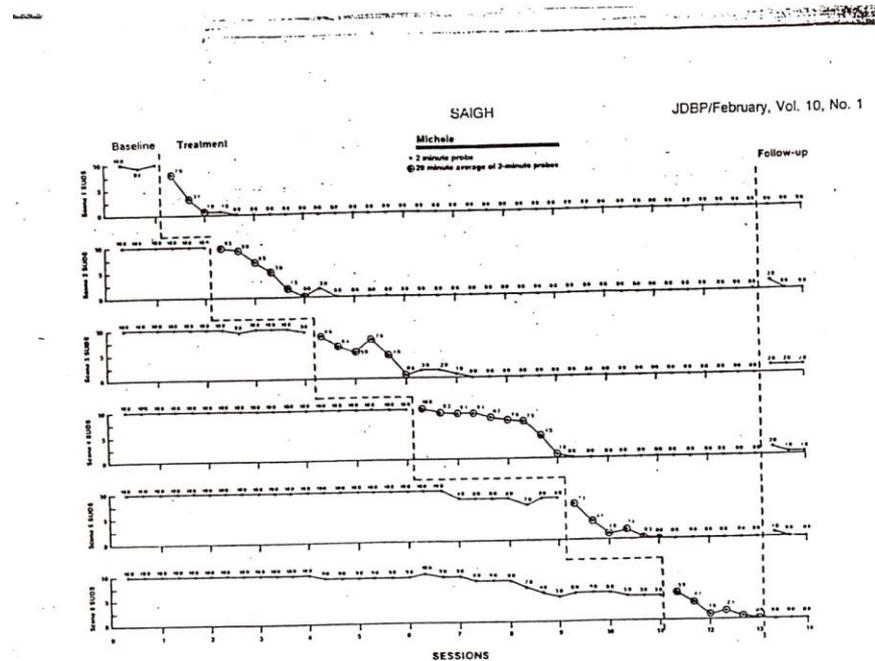


FIGURE 1. Michele's subjective units of disturbance (SUD) levels across scenes.

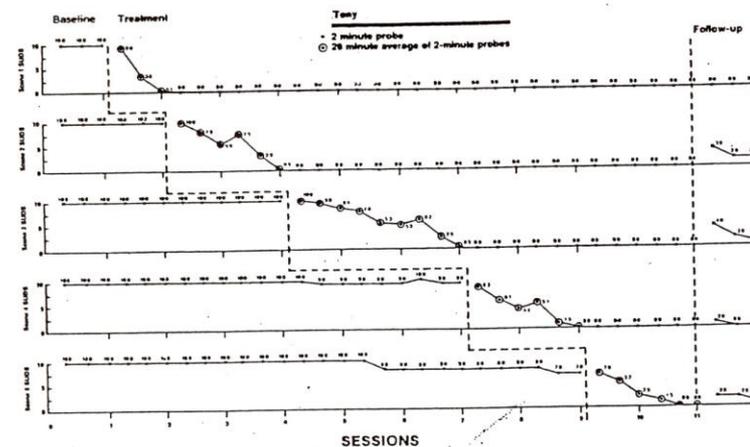
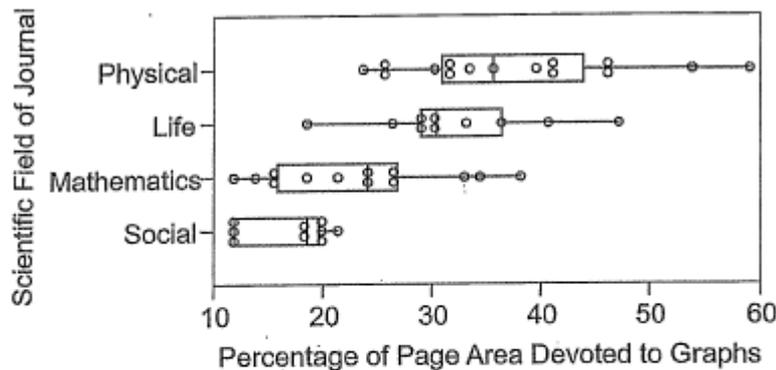


FIGURE 2. Tony's subjective units of disturbance (SUD) levels across scenes.

# The virtues of Visual Analysis

# The virtues of Visual Analysis ...

*you can learn a lot by looking.* Yogi Berra



*Visual analysis ... can simultaneously detect curvilinear trends, repeating patterns or cycles in data, delayed or lagged responses following intervention onset, and within-phase changes in variability. In addition it can detect changes in mean level and trend slope across phases ... **That breath is not equalled by any other analytic technique.***

[Parker & Hagen-Burke, 2007]

FIGURE 4.1. Dot-boxplot of results in Cleveland (1984) on the use of graphics in scientific journals. The horizontal axis represents the percentage of total page area devoted to graphics in 47 articles sampled from the 1980–1981 volumes of selected science journals.

Copper [Editor-in-chief]. *APA Handbook of Research Methods in Psychology: Vol 3* (pp73 – 100). Washington, DC:APA]

# The virtues of Replication

*...replication is at 'the heart of any science'...*

[Schmidt, 2009, p90 -100]

*An ounce of replication is worth a ton of inferential statistics.*

[Steiger, 1990 p 176]

*... we must finally rely, as have the older sciences, on replication.*

[Cohen, 1994, p1002]

# More virtues of Single-case research

- Gives us a science – basic and applied - at the individual level
  - Resolves the problem of drawing inferences from group designs to individuals
  - Avoids the problems of averaging across cases

# More virtues

- Avoids the trap of confusing statistical with clinical significance of outcomes

*...which has been the ruin of empirical research in economics as in medicine, and sociology and psychology.*

McCloskey & Ziliack, 2008, p 53.

# More virtues...

- Permits demonstration of both efficacy and effectiveness of interventions
  - Continuum from development to systematic clinical replication
  - Efficient “mapping” of generality
    - Replicated failure is as informative as replicated success

# More virtues...

- Makes it possible to be a scientist-practitioner/do evidence-based practice in everyday practice

*There are three fundamental questions [about any intervention]: (a) does it work under special, experimental conditions? [Efficacy](b) does it work in practice? [Effectiveness]and (c) **is it working for this patient?** [Progress]The crucial question posed by the practitioner is Question c.*

[Howard, et al., 1996, 1059-64]

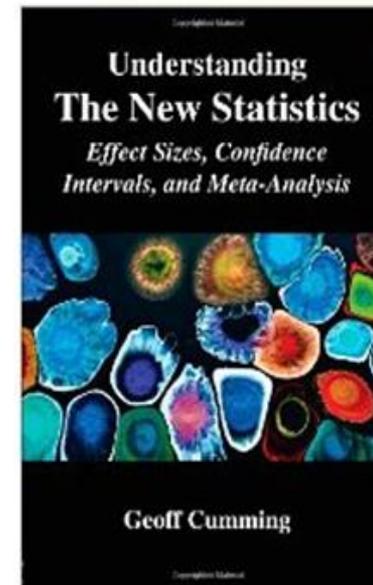
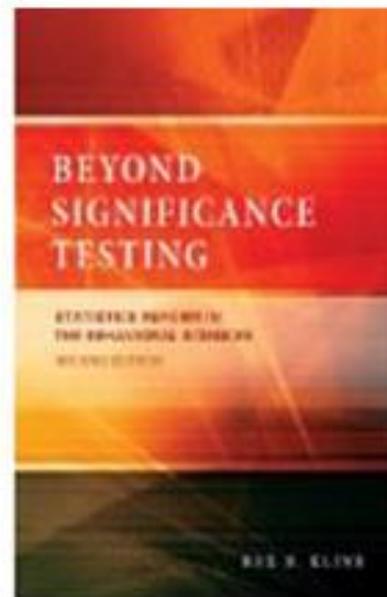
# More virtues...

- Enhances ethics and accountability in applied research
  - avoids exposing large numbers to control or ineffective conditions
  - Enhances accountability – to clients & family, peers, agencies, etc.

# Other virtuous approaches

## The *New Statistics* movement

- Estimation
  - Precision –
    - Confidence intervals
  - Effect sizes
  - Meta-analysis
    - (best evidence synthesis)
  - Replication
- Does not use NHST or  $p < ?$



*friends do not let friends compute  $p$*  [quoted in Klein, 2013).

# Yet more virtues

- Consider measurement error
  - SE - Standard Error of Measurement
- Use a Reliable Change Index
  - $RC = (X_2 - X_1) / S_{DIFF}$

*When RC is greater than 1.96, it is unlikely that the posttest score is not reflecting real change. RC tells us whether change reflects more than the fluctuations of an imprecise measuring instrument.*

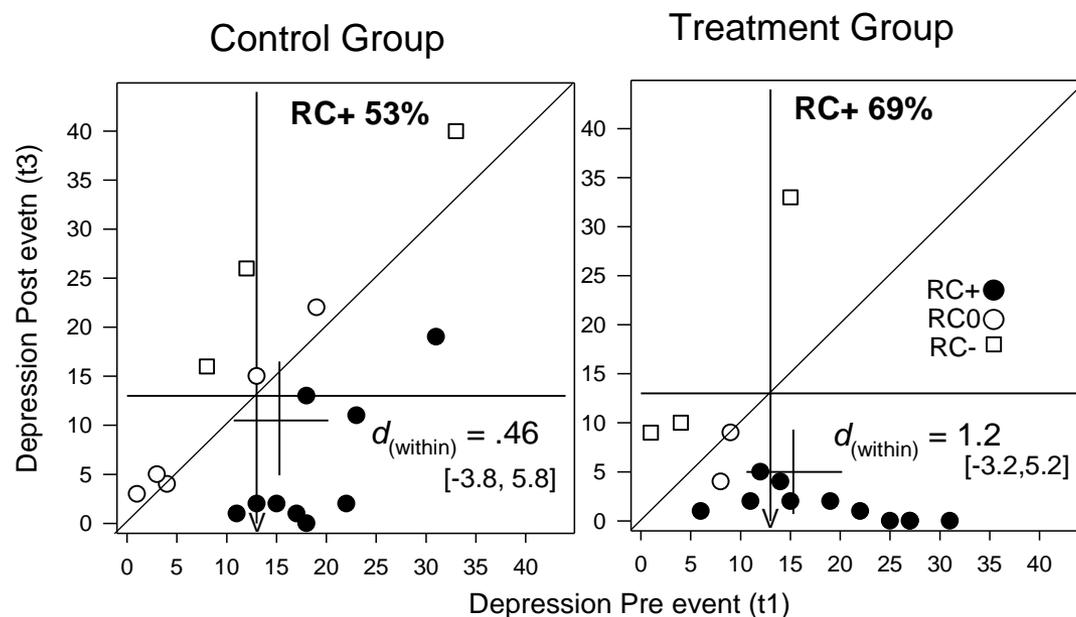
[Jacobson & Truax, 1991]

- Report clinical/practical significance

*Doubling one's thinking is likely to be much more productive than doubling one's sample size*

[McClelland, 2000]

# Example –bringing it together



Modified Brinley plots

[Blampied, 2014, after Rucklidge & Blampied, 2011]

- Visual analysis ✓
- Groups ✓
- Individuals ✓
- Means ✓
  - 95% Confidence intervals ✓
- Reliable Change ✓
- Effect size ✓
  - 95% CI on ES ✓
- % with reliable change ✓
- Clinical significance ✓

# Conclusion

We need Evidence-based **research practice**

AND

Evidence-based **professional practice**

for 21<sup>st</sup> C Applied Psychology

*... in the final analysis, the progress of ... psychology, as of every other science, will be determined by the value ... of its contributions to the advancement of the human race.*

[Witmer, 1907]

# References

For background, see

- Blampied, N.M. (1999) A legacy neglected: Restating the case for single-case research in cognitive-behaviour therapy. *Behaviour Change*, 16, 89 - 104.
- Blampied, N.M (1999). Murray Sidman – An appreciative introduction. *Behaviour Change*, 16, 75 - 78.
- Blampied, N.M. (2000). Single-case research designs - A neglected alternative. *American Psychologist*, 55, 960.
- Blampied, N.M. (2001). The third way - Single-case research, training and practice in clinical psychology. *Australian Psychologist*. 36, 157 – 163.
- Blampied, N.M. (2013). Single-case research and the scientist-practitioner ideal in applied psychology. In G. Madden (Editor-in-chief). *Handbook of Behavior Analysis*, Vol 1. Washington, DC: APA.

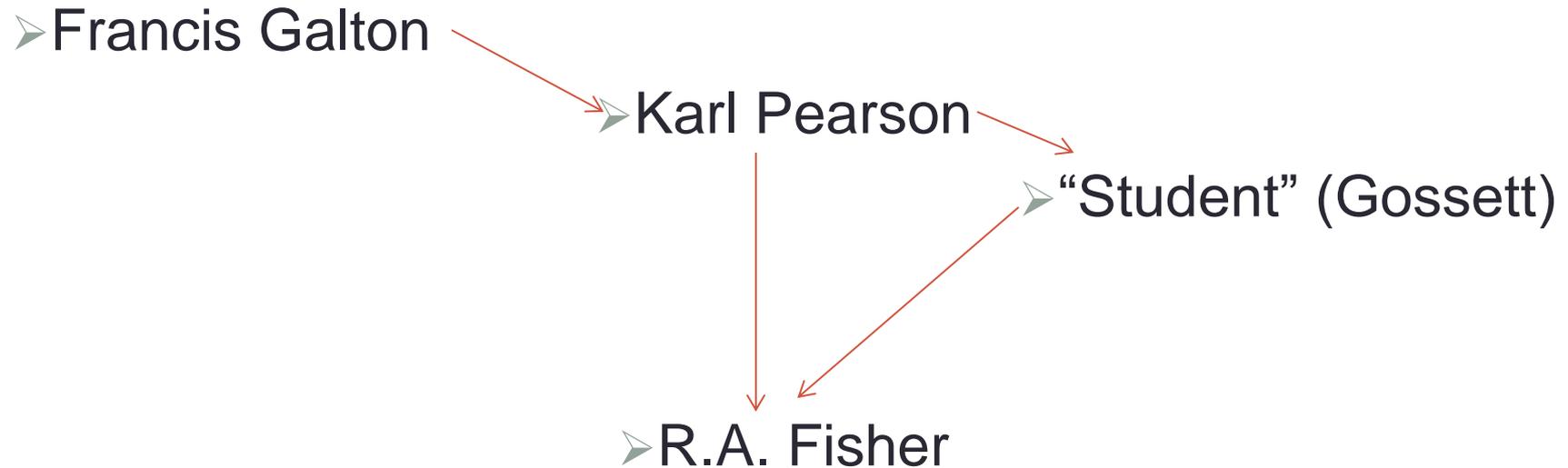
# Abstract

For more than 50 years the concept of the professional psychologist as both a scientist and a practitioner (often called the *Boulder Model*) has been an inspirational ideal that has guided training and informed professional identities across a wide spectrum of applied practice. But, as many have lamented, it has been very difficult to put the ideal into practice. A fundamental reason for this difficulty is the mismatch between the epistemological model of science adopted by psychology – an outdated hypothetical-deductive approach completely dominated by the null-hypothesis statistical testing of between-group mean differences – and the needs of science, and especially, practice. This talk will present the single-case research tradition and its contemporary developments as an alternative approach to both science and practice, one that effectively meets the need for psychological research to be more ideographic, and best-fits the challenges encountered when we apply our science via evidence-based practice to the problems we will encounter in the 21<sup>st</sup> C.

- [Neville.blampied@canterbury.ac.nz](mailto:Neville.blampied@canterbury.ac.nz)

# Quetelet's influence on Psychology

*it is the triumph of scientific men ... to devise tests by which the value of beliefs may be ascertained [Galton, F., 1879]*



The *inference revolution* in psychology

# Challenges for 21<sup>st</sup> C Applied Research

*... in the final analysis, the progress of ... psychology, as of every other science, will be determined by the value and account of its contributions to the advancement of the human race. (Witmer, 1907, p. 4)*

- The paradox of a commitment to evidence-based practice but a refusal to act on criticisms of the standard model of research

# Challenges for 21<sup>st</sup> C Applied Research

*... in the final analysis, the progress of ... psychology, as of every other science, will be determined by the value and account of its contributions to the advancement of the human race. (Witmer, 1907, p. 4)*

- The paradox of a commitment to evidence-based practice but a refusal to act on criticisms of the standard model of research
- The enduring mismatch between the needs of the scientist-practitioner and the standard model of research

# Criticisms of the Standard Model

- Criticism has increased substantially from the 1990's
  - Harlow, L.L., et al. (1997) *What if there were no significance tests?*
  - APA Taskforce on Statistical Inference (1999)
  - Nickerson (2000). Null-hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*, 5, 241-301
  - Ziliak & McCloskey (2008). *The cult of statistical significance.*
- Criticism has focussed on use of NHST as the only decision rule

# But the individual disappears in our research

## *Psychology without individuals*

*The idea of eliminating the individual ... was an innovation [in scientific method] unique to psychology.*

*... we ... investigate 150 freshman from Michigan for 20 minutes of their lifetimes, and think of them as interchangeable physical objects that do not change over time – which allows us to present our significant result as if it were about all mankind and all time.*

(Gigerenzer, 1987. Probabilistic thinking and the fight against subjectivity. In *The probabilistic revolution*. Vol 2. p 13 – 33).

# The problem of averaging

*No one goes to the circus to see the average dog jump through a hoop significantly oftener than untrained dogs. Skinner (1956). American Psychologist, 11, 221-233*

- Averaging across individuals had its origins in a philosophical/religious attempt to remove the effects of original sin from measurement of humans
- It fitted agricultural research fine
- Does it fit psychology?

# No – it just makes psychology a very odd science\*

*Reproducible group data describes some kind of order in the universe, and as such may well form the basis of a science. ... It is a science of averaged behavior of individuals who are **linked together only by the averaging process itself**. Where it fits in the scheme of natural phenomena is a matter for conjecture. My own feeling is that it belongs to the actuarial statistician and not to the investigator of behavioral processes.*

Sidman (1960). *Tactics of Scientific Research*, (p274 – 275)

\* Or at least a very average science!

