

Winter 2015, VOL. 29, NO. 3; ISSN 1051-0796

RMT

RASCH MEASUREMENT TRANSACTIONS

► **Special Issue:**
A Tribute to Benjamin D. Wright

**Transactions of the Rasch Measurement SIG
American Educational Research Association**

In Memoriam: Benjamin Drake Wright

Benjamin Wright passed away on October 25, 2015, aged 89. Ben, whose name needs no introduction among readers of this publication, was one of the most influential psychometricians who ever lived. He was a colleague, mentor and friend to many in the measurement community and will forever remain a significant influence to countless others.

Ben co-founded the Rasch Measurement Special Interest Group (SIG) of the American Educational Research Association (AERA) with Richard Smith in 1988. It was this same collaboration that resulted in the founding of this publication. Twenty-seven years later both the Rasch Measurement SIG and *Rasch Measurement Transactions* (RMT) are still going strong. To honor Ben's memory and his specific contributions to both the SIG and RMT, I have asked a number of Ben's former students and colleagues to share some comments and memories as part of a special issue in his honor. For those that knew Ben personally, I hope this tribute will invoke wonderful memories and feelings of nostalgia. For those that didn't know Ben personally, I hope this tribute will reveal more about Ben's nature and what made him such an incredible talent and human being.

I would also like to reiterate the very eloquent words of Nik Bezruczko who reminds us that while the measurement community is deeply saddened by the loss of Ben, we must remember that we have a responsibility to continue to press forward and advance the ideas and concepts that Ben advocated so passionately. We thank you, Ben, and will continue to demonstrate the enduring significance of your work.

In gratitude,

Kenneth Royal
North Carolina State University



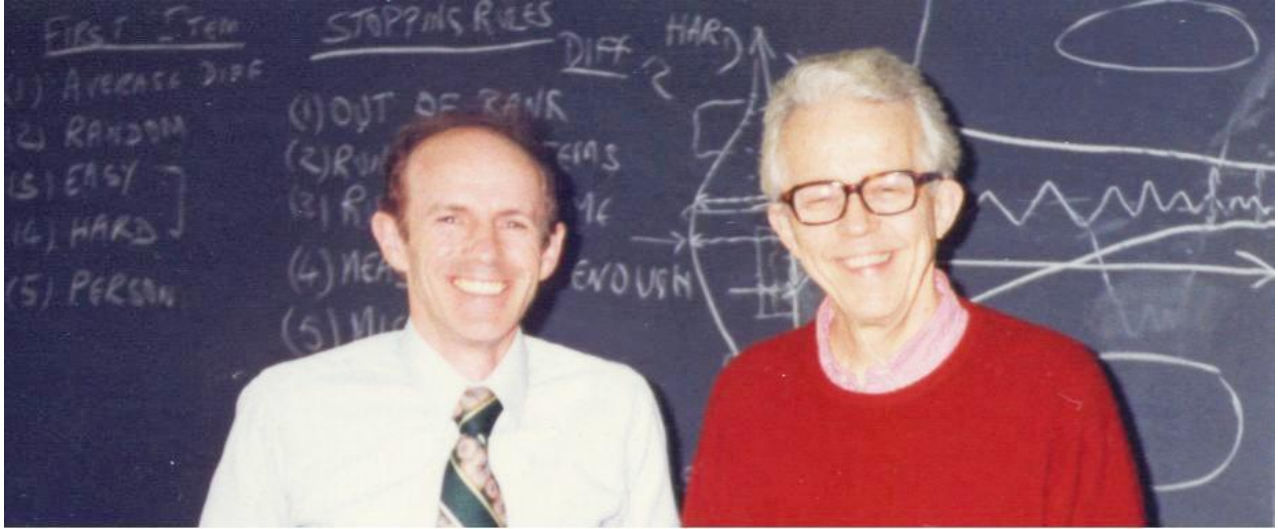
Ben Wright (with photo of Georg Rasch)



Judd Hall at the University of Chicago

Ben Wright in the Psychometrics Classroom

John Michael “Mike” Linacre
Ben’s Assistant from 1986 to 2001



Ben would awake at 4:00 a.m. and start thinking about his special topic for the day’s class. By 8:00 a.m., there would be something fresh and exciting bubbling up in his mind. So on to his bicycle and off to Judd Hall at the University of Chicago. Ben’s classroom was Judd 111 where he had had blackboards installed all the way along one wall in front of the students and on most of another wall to their right.

Ben would begin writing at the extreme left-end of the blackboards. Most of what he wrote was algebra, interspersed with brief wording and occasional diagrams. When he reached the extreme right-end of the front blackboards, he would continue on along the right-hand-side blackboards, sometimes filling all the blackboards with his cogitations.

By now it was 8:20 a.m. and the first students were arriving for the official 8:30 a.m. start time for Ben’s psychometric class. But Ben was so enthused about his insights that he couldn’t wait. He would begin explaining his algebra. The explanations were often somewhat cryptic, because Ben was now so familiar with his material that it had become obvious to him and so must be obvious to us! Students arriving at 8:30 would hear Ben’s last few remarks about his

special topic. Sometimes a late arrival would ask Ben to explain the algebra again. Ben was reluctant to do that, because what was fresh and exciting at 4:00 a.m. was now jaded and routine.

So, onwards to the regular class material. This was often student presentations. Students would bring their educational tests or attitude surveys on paper along with their control and data files on diskette. A laptop computer was connected to a projector and a white screen was lowered over the central part of Ben’s algebra. The remaining visible part of Ben’s algebra would be erased, and the student would outline the test or survey on the blackboard. At least, that was the intention. However, very quickly Ben would take over the student’s explanation, emphasizing the points of particular interest. Meanwhile I would have the student’s data ready to analyze in real time with the analysis projected onto the white screen.

Here is where the fun would start! Almost inevitable, the first attempt at analysis would be a dismal but highly educational failure. There would be something glaringly wrong with the control file for the Rasch analysis or with the data file. If the problem was conceptual, then Ben would lead the discussion. If the problem was technical, then I would. Soon we would have

meaningful output tables and graphs projected onto the screen.

This was where Ben really excelled. He would scrutinize the hierarchy of item difficulties. After some discussion with the student, Ben would have a definition of the latent variable the test or survey was actually measuring. Sometimes this was noticeably different from the claims of the test publisher. Then Ben would investigate the misfitting items and persons. For the students, it was like watching a combination of a detective and a psychoanalyst working together. Why had this seemingly mundane MCQ item provoked some smarter students to respond the way they had? Did they feel that the item was too easy and so had rationalized that one of the superficially wrong answers was really the correct answer? On a survey, were some of the respondents in denial or trying to hide their true attitudes? Ben could discern the mental processes that produced even the foggiest data.

By now, we had overshot the official class time, and students were slipping away to their next classes, but Ben did not want to stop. Finally the instructor of the next class in Judd 111 would grow weary of waiting. At the instructor's command, the next class of students would burst into the room, ejecting Ben and those of us still remaining.

Ben's grading policy for his students was simple. If they made a classroom presentation, followed by a reasonable write-up of their analysis, then they received "A" grades. If they did not, they received no grade at all. Ben would leave those entries on the student grading sheet blank. If such a student later made a decent attempt to meet Ben's criteria, which often happened during the next term, then Ben would contact the University administration and update that blank to an "A".

Ben loved to teach his classes. He would recruit students from all over the University. The more challenging and obscure their projects, the better! He was so proud of the shelves of Ph.D. Dissertations in his office representing over 40 years of students whose lives he had influenced and now continues to influence.

Ben Wright- Teacher, Supporter, and Guide

William J. Boone
Miami University



Anyone who has read Ben's writings or has viewed some "Ben videos" or was lucky enough to meet and study/research with Ben will recall fondly how passionate and expressive he was.

Below are a sample of some of the written comments Ben provided for the class assignments (Ben referred to them as "memos") that I completed while I learned in Ben's E360 & E494 classes from 1989-1991. I know the phrases will be familiar to all who learned from Ben--no matter the teaching venue, be it the classroom of Judd Hall or at Ben's measurement potluck which he graciously hosted at his home each year or through any article of Ben's that one has read! I hope readers will be reminded of Ben the Teacher, Ben the Supporter and Ben the Guide!

A "memo" about an article

Ben: "Yes, his analysis is flawed by his not looking at the data more closely"

A "memo" about validity

Ben: "Clarify your analysis of validity here! What part does fit play for item and person? What part does order play for item and person?"

Four "memo" comments-- supporting and pushing

Ben: "Good. Nice job here. Do more!", "Upward and onward!", "This is progress. Keep going ahead."

A "memo" where I asked "What good did this study do?"

Ben: "None! It made us look foolish, careless and shallow! Why not write to the editor and complain as in this memo."

A "memo" mentioning a by hand analysis

Ben: "Always a good idea"

A "memo" about checking data

Ben: "It is your scientific responsibility to edit the data

A mix of Ben's supportive "memo" comments...

Ben: "How very wonderful! And well deserved!!! Lovely"

On the need to take a stand

Ben: "Take a stand! Be opinionated!"

A "memo" about inventing and Ben's thoughts

Ben: "We invent in order to discover!"

And finally Ben on measurement!!!

Ben: "But when we measure, we must choose an intention".

Warm Memories of Ben

Mark Moulton,
Class of 1996



I remember when Winnie Lopez and I skipped one of Ben's classes to work on a big presentation, something we never did. Ben stormed up to the MESA Lab, shooting lightning bolts: "What's the matter with you?

You really think you have nothing to learn!" And stormed out again. No other professor or teacher, ever, took my presence or absence so personally.

For my dissertation I was interested in modeling highly multidimensional data and worked out this complicated algorithm that looked like the insides of a UFO (and even worked). Ben, never a fan of multidimensional models, nonetheless took it upon himself to have me walk him home every day so that I could explain all the steps, again and again, until he had it firmly in mind. Never in my academic career has someone honored me like that. How often does that happen nowadays?

Ben lavished that care on pretty much all his advisees. He was your doting grandmother. I've met a lot of graduate students who sign up to work with superstars and are sort of ignored. Ben wasn't like that. He didn't give a fig about academic standing or mathematical sophistication -- came to loathe it in fact. He demanded clarity, simplicity, practicality, humanity -- models that acknowledge our ignorance and deal honestly with it.

Ben lost a lot of memory from his stroke, but not his humanity. We just moved into other areas. Our conversations were intimate, searching, spiritual, lustful, lonely, defiant, and joyful. Ben was a man of vivid faults and even more vivid virtues, a great psychometrician and a greater friend.

Top Ten (x2) Remembrances of Ben

Roberta Henderson
Rosalind Franklin University of Medicine & Science



Here are fond remembrances of days in Judd Hall, in no particular order, from one of Ben's graduate students:

1. Yardstick classroom interrogations - no one was safe.
2. The memos – One idea memo & and one critique memo each week.

3. Red ink on assignments – more meant Ben was interested and one might be on to something.
4. Gatherings at Ben's home.
5. Ben's enjoyment of good food.
6. Ben's method of keeping a student on track, "You would like to...., wouldn't you?" – Yes, was the correct answer.
7. Ben's black book, small but mighty – "You are going to want to speak to ... about that."
8. Derive the Rasch Model – "Now".
9. On Factor analysis – "Do not waste your mind. It does not work."
10. The course on 'Ambiguity' & the epiphany of 'The Construction of Measurement'
11. Discussions on what is 'real'.
12. A smile and a bow tie striding rapidly with a purpose.
13. Did you get 'it' to run?
14. Respect for former students and what they achieve.
15. Delight in the return of former students from all corners of the earth.
16. MOMS (Mid-west Objective Measurement Seminars).
17. Anticipation and excitement over a new data set for Rasch Analysis.
18. Impatience (a mild term) with lack of competence.

19. Be concise and perfectly clear.

20. Enduring support through all manner of times.

Working with Ben to Solve Tough Practical Problems

John Stahl
Pearson

I had the good fortune to meet Ben Wright early in my career. I was working for Mary Lunz at the Board of Registry of the American Society of Clinical Pathologists (ASCP).

We had been working with the Rasch model on the other Board of Registry examinations. We explored ways to apply this model to account for examiner differences. The analysis program of that time was BigSteps, but this program could only account for two of the three facets of the examination at one time. We could analyze examiners as one facet but the candidates and the tasks would have to be combined as the second facet. Conversely we could examine the candidates as a separate facet but the examiners and the tasks would be confounded. Lastly we could examine the tasks as the single facet but that left the examiners and the candidates inexplicably combined. We could not put them all together in one analysis as separate elements. The dilemma was presented to Ben Wright, who had a student named J. Michael Linacre. Mike Linacre started working on the project and developed the first version of the multi facet model in 1987, which accounted for all facets of the practical examination at the same time.

Some of the earliest research using the multi-facet model was done with the practical examination described above. This practical examination had facets for candidates, examiners, items, and tasks. We now had a method of analyzing the data, but had to learn how to interpret the results and at the same time contribute to improving the functionality of the multi-facet model. This involved constant interaction with Ben Wright. We did the analysis,

brought it down to the University of Chicago, and Ben would always find another way to look at the data and back we would go to ASCP with more analyses to complete. Through this process we learned how to interpret the data and use it to improve the fairness of the practical examination. (Lunz and Stahl 2006)

I soon learned that this type of interaction and stimulation and innovation and investigation was not uncommon with Ben. I was always amazed at the number of projects in which Ben was simultaneously and enthusiastically involved. Ben delighted in bringing interesting people together and encouraging them to see things differently. He will be sorely missed.

A Humorous Side of Ben

Trevor Bond
James Cook University



I was very lucky that, when I visited Ben after his major health incident, he was quite lucid at times. We reminisced about past meetings and common friends. But, Ben insisted that he had not seen “Bond & Fox”, so, with the aid of his daughter, I went to his bookcase and after a bit of searching, found the signed copy I had sent to him. Ben leafed through some pages stopping occasionally to “um”, “ooh”, or “ah”, as appropriate. Returning to the cover,

Ben said: *You wrote this with Christine Fox?*

I replied: *Yes, from Toledo. Do you remember her?*

Ben said: *Aah, yes. She’s beautiful.*

I then said: *And very smart.*

Then Ben said (smiling): *And very beautiful.*

Remember a Giant in the Measurement Field

Ronald K. Hambleton
University of Massachusetts Amherst



I was very sad to learn the news that Professor Ben Wright had passed away. I did not know Ben nearly as well as many others who will write tributes but I knew of his work in 1967 and I met him for the first time in 1969.

This would make me one of Ben’s longest admirers. Ben had been invited by his colleague, Benjamin Bloom, to give one of the talks at the 1967 ETS Invitational Conference on Testing Problems, an annual meeting of test practitioners, testing specialists, and psychometricians. When I was a young professor at the University of Massachusetts this was a conference I dreamed one day I would be invited to speak at. Presenters were among the top measurement specialists and included Louis Guttman, Fred Lord, and Lee J. Cronbach. In my world at the time, speaking at this conference was a very big event. Ben was invited to speak and of course he spoke about the Rasch model. Not surprising to any of you who knew him, his performance was inspiring, interesting, and emotional, and highly influential. For a long time, his invited paper was one of the most frequently cited papers in the Rasch model literature along with Rasch’s own textbook. My advisor had heard Ben speak and informed me that while he wasn’t convinced of some of Ben’s arguments about model parameter invariance, we were going to need to learn a lot more about this mental model known as the Rasch model because it seemed controversial, interesting, and pioneering, all at the same time. My dissertation was on the topic of the Rasch model and related applications. Indeed, Ben was a pioneer and I think his many talks, workshops, and contributions with his outstanding graduate

students from the University of Chicago inspired many researchers to follow his lead and make their own contributions to the Rasch model literature and apply the model and its extensions to their test development, the study of item bias, and equating. I have often said that Ben, more than anyone else, inspired graduate students and faculty members, and specifically, inspired them to move the model and its applications forward. Our field is all the better for Professor Wright's impact. Today, Professor Wright's contributions can be found around the world and indeed, he was responsible for a paradigm shift in the advancement of measurement. That's something only a very few can claim. Professor Wright was a giant in the measurement field, and his contributions will be long remembered and valued.

Weekly Conversations with Ben

Jack Stenner
MetaMetrics



Throughout the 1990's until his stroke in 2001 Ben Wright and I would talk for one hour every Tuesday morning (the day of the week varied somewhat with his timely demands).

Our affectation was that it did not matter what else was going on (holidays, vacations, conferences, etc.) we found a way to make that call. The format involved one of us presenting for fifteen minutes on a problem, opportunity or upcoming presentation and then a freewheeling discussion would follow. Some topics demanded months' worth of Tuesdays whereas others were over in minutes. In honor of those Tuesdays with Ben, what follows are several brief introductions to topics that are top of mind for me at this moment in time. Each would have formed the kernel of a Tuesday with Ben. You the reader can take the place of Ben and decide whether the topic is worthy of your continued thought and

exploration. I think Ben would have liked that idea.

- The awesome power of the ensemble interpretation
- Causal vs Descriptive Rasch Models
- How to compute fit statistics that are sample independent
- The Fahrenheit method for establishing a unit of measure
- Employing multiple measurement mechanisms to establish the reality (i.e. existence) of an attribute
- Using the trade-off property to test for quantitative status of an attribute

I feel very fortunate to be part of the larger Rasch Community.

Ben and Yardsticks

Trudy Mallinson
George Washington University



A few years ago I came across this old yard stick in a consignment store and it made me smile and think of Ben – we all remember how much Ben loved yardsticks! On the face it, this is a standard ruler with 36 1-inch markings and

1/8-inch markings in between. (Along with the claim that Hochschild, Kohn, & Co. is Baltimore's Best Store!) On the reverse side of the ruler however, it is another story entirely. There are only two markings on this side of the ruler – one indicating a $\frac{1}{4}$ yard, and one indicating $\frac{3}{4}$ yard! Sorry, if you need to measure anything in between, like say a $\frac{1}{2}$ yard, this is not the ruler for you.

This unique, and somewhat puzzling ruler reminded me of Ben's attentiveness to measuring devices and how, above all else, they should be useful. Yes, they should be accurate and consistent but the amount of precision

represented by the device should be practical. You don't need a 36-inch ruler, if the only things you measure are less than 12 inches long. And you don't need a ruler marked off in 288 1/8-inch units, if the only things you need to measure come in lengths of $\frac{1}{4}$ and $\frac{3}{4}$ yards!

Hochschild, Kohn & Co was a department store on the corner of Howard and Lexington in



downtown Baltimore that has since gone out of business and the building burned down. An online-listing of the store directory gives no clues as to what they might have been selling at HK that was needed only in lengths of $\frac{1}{4}$ and $\frac{3}{4}$ yards.

Still, I love this yardstick; it sits on the shelf above my desk, $\frac{1}{4}$ yard and $\frac{3}{4}$ yard markers facing out at me. When I am looking at Rasch output and see a person separation reliability of .86, or a rating scale step with thresholds 1.5 logits apart, or a rating scale with 7 steps but only 4 are actually used by respondents, I am reminded to ask, what am I using this measuring tool for? What kinds of decisions do people need to be able to make with this tool? Our measuring tools should always be useful and practical.



Reflections on Ben

Nikolaus Bezruczko
Linz, Austria

Ben had a deep personal commitment to change educational measurement and bring it more in line with logical scientific thinking. Consequently, when I reflect on Ben over the years, a dominant strand among many was his fiery rejection of arbitrary social research conventions. A well-known example was Ben's resistance early in his career to Department of Education teaching conventions. Ben had migrated to Education in the 1950s from University of Chicago's Department of Physics,

where he was a faculty graduate assistant. Upon accepting duties to teach statistics in Education, he immediately faced rigid methodology conventions concerning analysis of variance (ANOVA), which Ben refused to follow. Based on his mathematical and physics background, the statistics being taught to students were entirely "wrong", and Ben wasted no time changing course presentations and student expectations. Of course, his colleagues in Education were outraged. They wanted Ben fired, and they petitioned the Chairman to do so! Ben had to mobilize substantial support from the top statistician in Department of Statistics, Jimmie Savage, who defended him before the Education chairman. In the end, Ben prevailed though conventions, traditions, and rituals never stopped annoying him. Needless to say, many of Ben's students appreciated his attitude toward mindless conventions and continue to push back against some of those same conventions today.

A direct result of Ben's intolerant attitude toward arbitrary authority and thoughtless habits is much of his life's contributions to science will likely not be appreciated until sometime in the future when others reflect on the seminal events of probabilistic measurement in the mid to late 20th century. Sometime in the future when the social sciences reflect on the methods that have diminished wars, racial prejudice, and economic poverty, as well as improve human development, then the objectivity and mathematical logic that Ben brought to education and social science research may be appreciated more fully. He will then be understood as the beacon showing the way.

Ben forced controversies, and he enjoyed exposing nonsense all around him. It gave him enormous satisfaction to reveal the truth and discredit dishonesty. Nothing gave him more satisfaction than to expose a scientific fraud. I believe Ben was fundamentally successful in pointing the direction to better methods in social science research, and his dedication has "cracked" conventions that become wider with passing years. I would guess Rasch models today are among the widest applied measurement methods in the world, which is certainly a testimony to Ben's determination and persistence

to subvert conventions. Educational measurement throughout the world today is more transparent, and the honesty and fairness of invariant measurement frameworks that Ben advanced now present a higher standard than what was there before. Social science research will never be the same again. Objective science has lost a warrior, while those who worked with him have lost a loyal friend.

A Few Memories of Ben Wright

Thomas O'Neill
American Board of Family Medicine



I first met Ben Wright in person in January of 1995 when I attended his class as one of the requirements of my job at ASCP. I remember that he gave his invariance lecture and discussed the presentation he gave at the 1968 *ETS Invitational Conference on Testing Problems*. I thought that it was really impressive that he found a data set where the invariance property could be demonstrated. He ended the class by saying that it really didn't matter if it worked on *his* data set. It only matters if it works on *ours*. I went back to work and repeated his study on 8 of ASCP's examination data sets, the larger ones. I was amazed that it worked on all of them. It took me a while to realize that results, which seemed miraculous, occurred because it was about structure, not sampling. I was hooked. The Rasch model is how real measures are developed. That was Lesson 1.

Ben was also a foodie. Some combination of doughnuts, bagels, and coffee were always in the back of the classroom. He organized the Chicago Objective Measurement Eating Table (COMET) where people would meet to discuss ideas over a brown bag lunch. Later, this evolved into more formal presentations with everyone going out to a restaurant afterward for dinner. It was usually DAO, a Thai restaurant. It was interesting to talk with people from other fields and get to know

them as people. Ben also hosted the Midwest Objective Measurement Seminar (MOMS) in Judd Hall which was followed by a potluck dinner at Ben's house. This socialization helped us to become a community. Lesson 2: Make learning, exploring, and work fun. Food helps!

Ben was very interested in tackling new measurement problems that were of importance to the real world. For this reason, people from industry were often in attendance. Part of the class format was a stand-up, open-mic psychometrics forum. People would come, present their problem, explain why it was important, and demonstrate what they had done to resolve it. This was limited to a few minutes. Next, Ben and Mike (Linacre) would point out the strengths and weaknesses of the approach, suggest possible next steps, and get the thoughts of other people in the class. This helped everyone in the class see how they would tackle a problem and after a while the students became very good at it too. These suggestions were often executed by the presenter and updates were provided the next week. This helped us to become a network of people with expertise who knew one another. We had people we could discuss our problems with in addition to Ben and Mike. What Ben had done was to provide his analytic approach to many people by making it inexpensive, practical, and entertaining. He then linked us together, socially, just like a distributed computing model. We as a community became intellectually more powerful than we were as individuals. Ben essentially built a supercomputer. That was Lesson 3.

I miss Ben very much and I am thankful that he not only took the time to share his way of thinking with me and so many other students, but he showed us how rewarding and how much fun it could be if there was measurement community to support one another. Ben, thanks for your friendship and the lessons.

Remembering Professor Ben Wright: One Ruler for Everyone, Every Time and Everywhere

George Engelhard
The University of Georgia



Professor Ben Wright had an incredible influence on my ideas about measurement, and the ways that I approach the history and philosophy of measurement. My first encounter with Ben and invariant measurement was during the summer of 1977 when I visited Chicago to discuss my admission in the MESA Program at The University of Chicago. He shared his seminal article (Wright, 1968) with me, and also an article by Choppin (1968). I was hooked...!

Ben was an inspiring teacher, and I would like to share one of his instructional methods that he simply called *Cogitations*. In his seminar on psychometrics at The University of Chicago, Ben required students to keep a weekly reading log:

Part A: 8 weekly memos containing:

- I. An interesting quote on the requirements and/or methods of measurement. These quotes should come about half from Thurstone, Guttman, Loewinger, Torgerson, etc. and about half from articles in current issues of APM, JEM, EPM, and PM.
- II. Your own comments on this quote.

Part B: Essay

- I. A thoughtful essay drawn from your memos (10-15 pages typed, signed and dated).

These weekly memos served as the genesis and inspiration for countless papers on history and philosophy of measurement that I have written over the years comparing the views of key measurement theorists. To my delight, I learned recently that there is even a label for this type of research activity: Complementary Science (Chang, 2004). Here is a small sample of some of the seminal quotes discovered during that long ago seminar that still guide my research and thinking:

The history of science is the history of
measurement
(Cattell, 1893, p. 316)

The scientist is usually looking for invariance
whether he knows it or not
(Stevens, 1951, p. 20)

These weekly memos also led me to recognize the importance of duality as an important concept (Mosier, 1940). Here are some quotes that reflect this duality between persons and items, and its fundamental relationship to invariant measurement.

L.L. Thurstone:

Items: The scale must transcend the group measured. A measuring instrument must not be seriously affected in its measuring function by the object of measurement ... Within the range of objects ... intended, its function must be independent of the object of measurement (Thurstone, 1928, p. 547)

Persons: It should be possible to omit several test questions at different levels of the scale without affecting the individual score ... It should not be required to submit every subject to the whole range of the scale. The starting point and the terminal point ... should not directly affect the individual score (Thurstone, 1926, p. 446)

Georg Rasch:

Items: The comparison between two stimuli should be independent of which particular individuals were instrumental for the comparison; and it should also be independent of

which stimuli within the considered class were or might also have been compared.

Persons: Symmetrically, a comparison between two individuals should be independent of which particular stimuli with the class considered were instrumental for the comparison; and it should also be independent of which other individuals were also compared on the same or on some other occasion (Rasch, 1961, pp. 331-332)

Ben Wright:

Items: First, the calibration of measuring instruments must be independent of those objects that happen to be used for the calibration.

Persons: Second, the measurement of objects must be independent of the instrument that happens to be used for the measuring (Wright, 1968, p. 87)

These quotes and readings lead clearly to requirements for sample-invariant calibration of items and item-invariant measurement of persons with the opportunity to extend these requirements to rater-mediated assessments (Engelhard, 2013). I vividly remember Ben's written comment on my paper on comparing Thorndike, Thurstone and Rasch that was later published (Engelhard, 1984): *Do you want to take this further George? It is a fundamental integration of history and ideas ... (personal communication, Ben Wright, 1980)*

I am still writing memos to Ben ...

References

Cattell, J. M. (1893). Mental measurement. *Philosophical Review*, 2, 316–332.

Chang, H. (2004). *Inventing temperature: Measurement and scientific progress*. Oxford, United Kingdom: Oxford University Press.

Choppin, B. (1968). Item bank using sample-free calibration. *Nature*, 21 (August 24), 870–872.

Engelhard, G. (2013). *Invariant measurement: Using Rasch models in the social, behavioral,*

and health sciences. New York: Routledge.

Engelhard, G. (1984). Thorndike, Thurstone and Rasch: A comparison of their methods of scaling psychological and educational tests. *Applied Psychological Measurement*, 8, 21–38.

Mosier, C. I. (1940). Psychophysics and mental test theory: Fundamental postulates and elementary theorems. *Psychological Review*, 47, 355–366.

Rasch (1960/1980). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research. (Expanded edition,

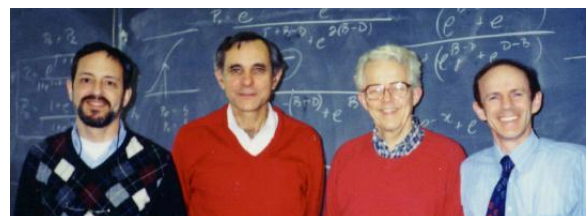
Rasch, G. (1961). On general laws and meaning of measurement in psychology. In J. Neyman (Ed.), *Proceedings of the fourth Berkeley Symposium on mathematical statistics and probability* (pp 321-333). . Berkeley: University of California Press.

Stevens, S. S. (1951). Mathematics, measurement and psychophysics. In S .S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1-49). New York: Wiley.

Thurstone, L. L. (1926). The scoring of individual performance. *Journal of Educational Psychology*, 17, 446–457.

Thurstone, L. L. (1928, July-October). Experimental study of nationality preferences. *Journal of General Psychology*, 405-205.

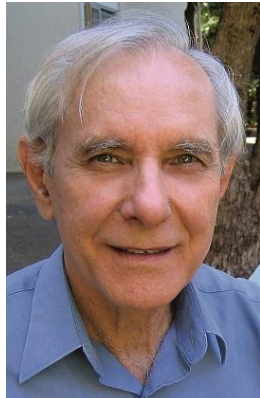
Wright, B.D. (1968). Sample-free test calibration and person measurement. In *Proceedings of the 1967 invitational conference on testing problems* (pp. 85-101). Princeton, NJ: Educational Testing Service.



Stuart Luppescu, David Andrich, Ben Wright, Mike Linacre, ca. 1997

Ben Wright: “Idiosyncrasies of Autobiography and Personality” in Taking up the Rasch Measurement Paradigm

David Andrich
The University of Western Australia



“Individual scientists embrace a new paradigm for all sorts of reasons and usually for several at once. Some of these reasons... lie outside the sphere of science entirely. Others depend upon idiosyncrasies of autobiography and personality.” (Kuhn, 1970, p.152).

This note outlines some of the *idiosyncrasies of autobiography* of Benjamin D. Wright that might help explain why he was the first to fully embrace and advance the work of Georg Rasch (1901 – 1980) on measurement in the social sciences. I have argued elsewhere (Andrich, 2004) that because Rasch’s paradigm for measurement in the social sciences has features which are incompatible with the standard statistical modelling paradigm usually applied to test data, it constitutes a scientific revolution in the sense of Thomas Kuhn (1970).

After he heard Rasch give lectures at The University of Chicago in 1960, Ben Wright of the Department of Education at The University of Chicago, then one of the most well-known Departments of Education in the world, took up the challenge to study and advance the work of Rasch. Up to that point, others had also heard Rasch give lectures, but none of them took it up in the same way that Wright did. They seemed to consider the models that arose from that paradigm as just another class of statistical models.

I suggest that there are two aspects of Wright's intellectual biography that helped make him ready to take up Rasch's work, and in particular

to take it up as a concern with measurement in scientific inquiry as found in the physical sciences, and not primarily as a concern with statistical modelling. First, his early postgraduate studies were in physics, and his first research publication (co-authored) was in physics, making him familiar with the role of measurement in research in the physical sciences. Second, the relevance of which is perhaps less obvious, he had worked for some six years as a child psychoanalyst and thus was familiar with the individual, interrogative methods of psychoanalysis. I elaborate briefly on both of these aspects below.

Significance of physics

The context for the relevance of a background in physics is provided by Kuhn’s analysis of the function of measurement in science:

...large amounts of qualitative work have usually been prerequisite to fruitful quantification in the physical sciences (Kuhn, 1977, p.180).

Only a miniscule fraction of even the best and most creative measurements undertaken by natural scientists are motivated by a desire to discover new laws and to confirm old ones (Kuhn, 1977, p. 187).

What then is the role of measurement in science?

To the extent that measurement and quantitative technique play an especially significant role in scientific discovery, they do so precisely because, by displaying serious anomaly, they tell scientists when and where to look for a new *qualitative* phenomenon. To the nature of that phenomenon, they usually provide no clues (Kuhn, 1977, p.205). (Emphasis added.)

Rasch draws explicit parallels between his models and those of classical physics (Rasch, 1960/1980) and makes explicit the empirical work that was carried out in constructing measurements that appeared in his 1960 book. Rasch would have stressed both aspects in his

lectures in 1960. Being familiar with the role of measurement in research in physics at first hand, Wright understood the integral role of the empirical and qualitative work that was needed to construct instruments to obtain physical measurements. On the other hand, the many other people who would have heard Rasch's lectures, and who were not as advanced in the role of measurement in research in the physics, saw Rasch's work as being primarily about statistical modelling. Wright saw them in the way Rasch intended, operational criteria for the achievement of measurement, whether in the physical or social sciences.

In addition, Rasch had a perspective on the role of modelling data in the design of measuring instruments which was analogous to Kuhn's on the role of theory in constructing measurements. Thus in relation to *anomalies*, (in Kuhn's terms), Rasch argued:

It is tempting, therefore, in the case with deviations of one sort or other to ask *whether it is the model or the test that has gone wrong*. In one sense this of course turns the question upside down, but in another sense the question is meaningful. For one thing, it is not easy to believe that several cases of accordance between model and observations should be isolated occurrences. (Rasch, 1960/1978, p.51). (Emphasis in original.)

In contrast to the above perspective on the relationship between model and data, in the standard statistical paradigm the choice of model is based on whether it accounts for the data, and if it does not then as implied by Rasch above, it is the model that is *wrong*. In Rasch's work, the probabilistic model is a rendition of the requirements of measurement to which data must, not only be valid substantively, but to which it must also conform statistically. In each case of systematic statistical misfit, Rasch (1960) considers further experimentation rather than a modification to his models. From his physics background, Wright would have been comfortable with this position; on the other hand, the typical statistician, or social scientist who had

been taught statistics from the standard statistical paradigm, would not have been.

The relevance of previous training in physics gains added strength when it is realised that Wright's first two students to study Rasch models with him had a background in physics. The first, Nargis Panchapakesan, who completed a PhD with Wright on the Rasch model, had completed a PhD in physics previously, and the second, Bruce Choppin, had completed a master's degree in applied mathematics, a degree that involved a great deal of classical physics. In addition to having a strong background in physics, none of these three people, Wright, Panchapakesan, nor Choppin, had statistics as their main subject area of study

The significance of psychoanalysis

The context of the methods of psychoanalysis is provided by the observation that Freud had a Kuhnian perspective on the methods of research. Freud knew quite as well as we do that nothing *speaks* but men's theories; he opens his paper *Instincts and Their Vicissitudes* (1915) with an almost Kuhnian discussion of theory and observation in scientific method (Malcolm, 1987, p.95):

Even at the stage of description it is not possible to avoid applying certain abstract ideas to the material in hand, ideas derived from somewhere or other but certainly not from the new observations alone. Such ideas - which will later become the basic concepts of the science - are still more indispensable as the material is further worked over. They must at first necessarily possess some degree of indefiniteness; there can be no question of any clear delimitation of their content. So long as they remain in this condition, we come to an understanding about their meaning by making repeated references to the material of observation from which they appear to have been derived, but upon which, in fact, they have been imposed. Thus, strictly speaking, they are in the nature of conventions - although everything depends on their not being arbitrarily chosen but determined by their having significant relations to the empirical

material, relations that we seem to sense before we can clearly recognize and demonstrate.

The early stages of developing an instrument seem to require such a perspective. Then the criterion of a Rasch model for measurement, which is independent of data, is central to breaking the potential circularity.

Wright practised his psychoanalytic work with Bruno Bettelheim, who had studied with Freud and who was one of the most well-known Freudian psychoanalysts of the time. From his experience and study of psychoanalysis, Wright was thus also familiar with the study of individuals, and with qualitative research methods based on individual responses. Rasch stresses in the introduction to his 1960 book the importance of the individual, rather than a population:

In the present work a new approach to test psychology is attempted. Traditionally the properties of a test are defined in terms of variations within some specified population. In the following chapters we are going to deal mainly with three different types of tests and for each of them we shall develop a probabilistic model, in the application of which the role of the population is abolished (Rasch, 1960/1980, p3).

In the late 1960s and early 1970s, Wright taught a class in psychometrics based on Rasch's measurement principles, and a class in psychology based on Freud's psychoanalytic principles. I sat in both classes. Although the contents of psychometrics and psychoanalysis are clearly very different, it seemed to me that where data were connected to theory, Ben Wright used the same principles – taking *anomalies* – *deviations of one sort or another* – (from theory) seriously. In one case where the assessment of an educational attribute subjected to Rasch's multiplicative Poisson model produced an anomaly (a very large statistical residual) its empirical production could be interpreted from a substantive defence mechanism, not just methodological, Freudian perspective.

Conclusion

The suggestions made above cannot possibly explain fully Wright's interest in taking up Rasch's revolutionary measurement principles. His reasons would have been many, including personality factors that were outside *the sphere of science*, one of which was that he did not accept teachings from authority which he did not understand.

Though they often attract only a few scientists to a new theory, it is upon these few that its ultimate triumph may depend. If they had not taken it up for highly individual reasons, the new candidate for paradigm might never have been sufficiently developed to attract the allegiance of the scientific community as a whole (Kuhn, 1970, p.156).

As I have indicated above, I have argued elsewhere that Rasch's (1960/80, 1961) measurement paradigm of can be seen to constitute a scientific revolution from the traditional statistical modeling paradigm for measurement in the social sciences. If this turns out to be a successful revolution, then its *ultimate triumph* will have depended a great deal on the enthusiasm, energy, commitment and teaching of Ben Wright in his taking up the Rasch paradigm.

This note was originally prepared as a talk at a Festschrift in honor of Ben Wright in Chicago, April, 2003.

References

- Andrich, D. (2004) Controversy and the Rasch model: A characteristic of incompatible paradigms? *Medical Care*, 42, 7 – 16. Reprinted in E.V. Smith & R. M Smith, *Introduction to Rasch Measurement: Theory, Models and Application*. JAM Press; Minnesota. Ch. 7 pp 143 – 166.
- Kuhn, T.S. (1970). *The structure of scientific revolutions*. (Second edition enlarged.) Chicago: The University of Chicago Press.

Kuhn, T.S. (1977). *The function of measurement in modern physical science*. Chicago: The University of Chicago Press.

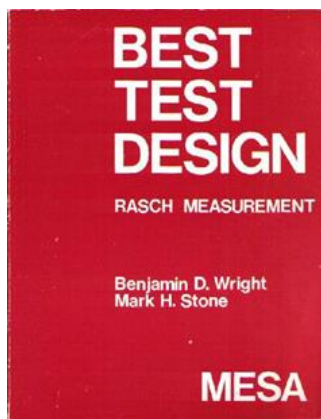
Malcolm, J. (1987, April). Reflections: J'appelle un chat un chat. *The New Yorker*, pp.84-102.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. (Copenhagen, Danish Institute for Educational Research). Expanded edition (1980) with foreword and afterword by B.D. Wright, (1980). Chicago: The University of Chicago Press. Reprinted (1993) Chicago: MESA Press. Available from www.rasch.org/books.htm.

Rasch, G. (1961). On general laws and the meaning of measurement in psychology. In J. Neyman (Ed.). *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*. IV, (pp.321-334). Berkeley CA: University of California Press. Reprinted in Bartholomew D.J. (Ed.) (2006) *Measurement Volume I*, 319 -334. Sage Benchmarks in Social Research Methods, London: Sage

A Humorous Story Behind “Best Test Design”

Mark H. Stone



Through the 1960's, I worked part-time at Social Research Inc. (SRI) in Chicago. This was a research firm founded by U. of C. professors, several of whom were my teachers. Ben came to SRI every Thursday to consult on the large number

of research studies we conducted for advertising and marketing firms nationwide.

Noontimes were spent discussing Rasch. From these noon time discussions came plans to write a book and formalize our activities. Almost every

Sunday I went to the U. of C. campus to meet at his office, or over to Harper Avenue to sit at the kitchen table with Ben.

We decided to have the book printed at the U. of C. and do all the editorial/composition work ourselves. Today one can do these tasks handily thanks to laptops and software, but at that time, compositional typesetting was required, and we needed this service. Our plan by this time was to make the pages 8 ½ by 11 inches to accommodate the tables and figures we thought were required for a step-by-step explanation, and to present the tables and graphs.

I engaged a young woman recently arrived from Germany who was working near SRI. She Americanized her first name as "Sam." For several months, Ben and I worked together each Sunday, and during the week, and I worked with Sam on getting the final pages in correct form. As more than half the chapters took shape, Ben indicated that he appreciated the fine work that Sam was doing and wanted to meet her. Sam was married, working full-time and completing this task for us during her noon hours and after work. She had no time or inclination to visit the campus. Ben finally insisted he must at least talk with her. I gave him her phone number.

The following morning her husband came to my office at SRI. He said that his wife was done with the project and following her conversation with Ben, she had burned the final drafts in their fireplace! Ben did not want to talk about the phone call he had with Sam when I later spoke to him about this incident.

The pages were developed by composing machines and the copy pasted on to make-up boards covered with tracing paper to protect the copy. Fortunately, I kept all the earlier drafts and make-up boards in my basement. I still have them and when I die my children will probably ask, "Why did he save all this junk?"

I was able to "reconstitute" everything lost in the blaze of fire resulting from Sam's apparent rage. I brought a new edition to Ben. Without a word from him about the incident, we continued with the remaining chapters. I never inquired of either

Ben or Sam what happened in that conversation, but anyone who knows Ben knows his (Greek/Socratic) “daemons” and some things are best left alone.

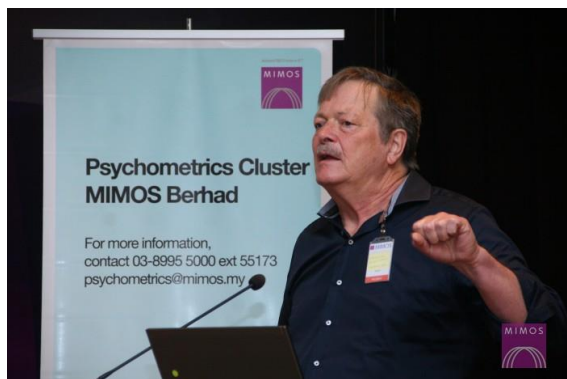
I next engaged Betty Stonecipher to complete the remaining chapters. My wife’s name is Betty and she worked at SRI also. Stonecipher appeared too coincidental to be real, and for a long time Ben wrongly surmised that I was trying to put something over on him. He sometimes asked to meet Mrs. Stonecipher, but I adamantly refused to comply not wishing to have another fire erupt. When *Best Test Design* (BTD) was finished and printed, I invited Mrs. Stonecipher to meet Ben so he could know there actually was such a person, but I am not convinced he did not think I engaged her as a ploy to continue what he thought was a deception.

Out of the fire came BTD when I finally delivered all the plates to the print shop and the book made its appearance. Hardbound copies were printed, but the bulk of BTD have all been paperback. The color of the cover tells the story.

My Thanks to Professor Benjamin Wright: A Personal Perspective

Rense Lange

Integrated Knowledge Systems, Chatham, IL
ISLA – Instituto Politécnico de Gestão e
Tecnologia, Vila Nova de Gaia, Portugal



I first came to know professor Ben Wright during the *Mid-West Objective Measurement Seminars* (MOMS) in the University of Chicago’s Judd Hall at Kimbark Ave., which I started attending in the

mid ‘90s. At that time, I had just become a psychometrician at the *Illinois State Board of Education (ISBE)*, and Wright and Linacre had evaluated the equating of recent *ISBE* tests. To my delight, these seminars proved to be a gathering of enthusiastic students and professionals, all guided by Wright’s insistence that Rasch scaling provides objective measurement in the social sciences. I am very grateful to him for having convinced me of this point of view, and his insights have guided much of my professional life over the last two decades. In the following I take the liberty of summarizing Ben’s influence on my own journey into measurement.

My first impressions of Ben’s approach at *MOMS* were that his enthusiasm brought Rasch models to life – we all know that his ruler was an effective prop - and the presentations there focused on how this approach could be used in practice. This was in sharp contrast to the treatment of the *1-PL* logistic model that I had encountered (much) earlier in an *IRT* oriented class at the University of Groningen, the Netherlands. This course relied heavily on Birnbaum’s chapters in Lord and Novick’s *Statistical Theories of Mental Test Scores*. I remember doing much theorem proving and finding derivatives, but we had no software available to fit any of the models and thus *IRT* seemed impractical to me. As an aside, I note that this was one of the first classes that later Rasch scholar Ivo Molenaar taught at Groningen, and I recall him saying that a simpler model proposed by Georg Rasch actually worked surprisingly well, while being far more tractable. I am sure that Ben would have smiled at this.

It was immediately clear that those Kimbark meetings were different. Here theorems did not need further proving, and Ben together with Mike Linacre had implemented easy to use and reliable software to estimate the model parameters. Ben clearly aimed to make Rasch measurement practically feasible, and being educated as a social psychologist I often heard echoes of Kurt Lewin’s famous dictum that “*there is nothing more practical than a good theory.*” With respect to the Michael Linacre’s *Bigsteps* (now *Winsteps*) and *Facets* this meant that he was interested in enhancing this software. Ben’s concerns

extended to the reporting of results, both to researchers and the public, as he would always insist on inspecting the item and person maps that now carry his name. There was good cause for this as even at *MOMS* some presenters occasionally had reversed their item and / or person dimensions.

Today, in large part to the efforts by Ben Wright, Rasch scaling has become one of the major workhorses in commercial and professional testing. But, at *MOMS* there also was no shortage of research examples that fell outside the standard context of educational assessment. To me the practical importance of objective measurement was evidenced most clearly in his work with Mary Lunz' work on *Computer Adaptive Testing (CAT)* and rater effects, and with Jackson Stenner et al. on the definition and use of *Lexiles*. These applications also drove home the point that the word "scaling" in Rasch scaling probably obscures the most important condition for objective measurement, i.e., the existence of universal item hierarchies, which are the sine qua non of equating. From here it was a short road to appreciating William Fisher's ideas on metrology. I am grateful to have heard Ben discuss some iteration of these ideas at *MOMS*, as this greatly stimulated my own thinking and I'm sure that of others. My own take-away was Rasch scaling could and should be applied where possible, and I have since tried to achieve this.

Future researchers will remain indebted to Ben Wright for showing the way to study a wide variety of topics within the same rigorous framework of objective measurement. This will identify the factors that act universally, thereby allowing the construction of increasingly general theories and explanations. Ben's insistence on the universal rather than the accidental will continue to serve as model for psychometrics and social science research in general.

Forthcoming Book on Ben Wright's Career contributions

In a career spanning more than five decades, Benjamin Drake Wright made foundational contributions to the theory and practice of measurement. His influence extends far beyond education and psychology, where his work in measurement began, into health care and the social sciences at large. Qualitatively meaningful and quantitatively rigorous means of managing constructs measured with ability tests or rating scales are forever indebted in essential ways to Ben's contributions.

Ben's influence on measurement in education and health care continues to resonate around the world. Speaking at the University of Copenhagen in 2010 during the celebration of the 50th anniversary of the publication of Rasch's book, *Probabilistic Models*, Svend Kreiner remarked on the fact that "none of us would be here speaking about the work of an obscure Danish mathematician were it not for Ben Wright." Similarly, at the 2012 Pearson Global Research Conference held in Fremantle, Western Australia, Peter Hill, CEO of ACARA, recalled hearing Ben Wright speak in Australia in the early 1980s on measurement technologies that still have not yet been brought fully into the light of day. Ben addressed not only the technical demands of rigorous theory, models, estimation methods, software, instrument design, and validity assessment, but was also intimately involved in the development of predictive construct theories via his collaborations with Mark Stone on the Knox Cube Test and with Jack Stenner on the Lexile Framework.

On the weekend of April 26 and 27, 2003, a conference in honor of Ben was held at the Rehabilitation Institute of Chicago. A full list of the presenters and titles appeared in *Rasch Measurement Transactions* (Vol. 17, No. 1, pp. 908-909). A selection of papers from that conference focused on some aspect of Ben's personal history, character, and/or accomplishments. William P. Fisher, Jr. and Mark Wilson intend to publish these papers under the title, *Making Psychological and Social*

Measurement More Scientific: Reflections on Benjamin Wright's Career in Physics, Psychology, and Education.

The book will be a collection of memories, histories, and observations from Ben's students and colleagues. The proposed table of contents is shown below. Other papers presented at the conference focused on theoretical advances and applications based in Ben's work were planned for publication in a multi-volume work, but the present book will include only the personal reminiscences. The papers are in hand and will be published with little further editorial delay.

The editors would like to ask the RMT readership to support the publication of the book by helping find a publisher, and by possibly contributing funds. Contact William Fisher at wfisher@berkeley.edu if you have suggestions, ideas, comments, questions, etc.

Proposed Table of Contents

Making Psychological and Social Measurement More Scientific: Reflections on Benjamin Wright's Career in Physics, Psychology, and Education

Ben Wright's Kinesthetic Ventures
Ed Bouchard

Ben Wright: Idiosyncrasies of Autobiography and Personality in Taking up the Rasch Measurement Paradigm
David Andrich

Things I learned from Ben
Mark Wilson

Ben Wright: Provocative, persistent, and passionate
Trevor Bond

Ben Wright: A Multi-Facet Analysis
Mary E. Lunz and John A. Stahl

Reflections on Benjamin D. Wright: Pre and Post-Rasch
Herb Walberg

Reflections: Ben Wright, Best Test Design and Knox's Cube Test
Mark H. Stone

The Influence of Some Family and Friends on Ben Wright
John M. Linacre

Benjamin D. Wright: A Higher Standard
Gregory Ethan Stone

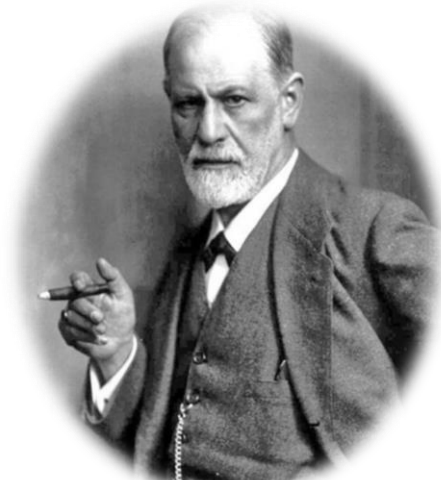
Ben Wright, Rasch Measurement, and Cognitive Psychology
Ryan P. Bowles, Karen M. Schmidt, Tracy L. Kline, Kevin J. Grimm

Provoking Professional Identity Development: The Legacy of Benjamin Drake Wright
William P. Fisher, Jr





*Charles Townes (Nobel Prize winning
physicist for whom Ben worked as an intern)
(1915-2015)*



*Sigmund Freud (influencer of Ben)
(1856-1939)*



*Bruno Bettelheim (Ben's doctoral advisor)
(1903-1990)*



*Georg Rasch (influencer and friend)
(1901-1980)*



*Leonard Jimmie Savage (Ben's close friend
and colleague)
(1917-1971)*



*Benjamin D. Wright
(1926-2015)*