RASCH MEASUREMENT

Transactions of the Rasch Measurement SIG
American Educational Research Association

Vol. 27 No. 3 Winter 2013 ISSN 1051-0796

How much item drift is too much?

When equating different forms of an examination across administrations, a common item equating design is often used. It involves anchoring the values of the common items to known difficulty calibrations and then estimating the calibrations of the new items relative to the anchor items. In actual practice, the difficulty of some of the items will change over time, so the previous difficulty calibrations of those anchor items do not always reflect their relative difficulty as found in the current data set that is being equated. This difference between an item's anchored value and the value that would have been estimated had the item been unanchored is called displacement (Linacre 2013).

When equating, it is common practice to unanchor those items that display excessive displacement and instead use the calibration implied by the current dataset. When only a few items must be unanchored and they are symmetrically distributed around zero, there is little cause for alarm; but when many items show an excessive amount of displacement, one becomes concerned that perhaps the particular items they selected to unanchor may have an impact on the equating. To illustrate, if only 2% of the anchor items are unanchored, it seems unlikely to have much impact on the equating, but if 60% of the items are unanchored, the psychometrician may wonder if s/he has unanchored the right items. Although having a substantive theoretical explanation for why an item changed in difficulty is preferred, in practice, working psychometricians are often left with only numerical indices with which to make their decision. In those cases, the question becomes what seems to be a useful threshold for identifying excessive amounts of displacement.

To this issue, Wright and Douglas (1976) found that random displacement of less than .5 logits have little effect on the test instrument. Draba (1977) recommends 0.5 logits based upon the rationale that item difficulties typically range between -2.5 and +2.5 logits, thus a shift of 0.5 logits represents a 10 percent shift within that range. Other studies (Jones & Smith, 2006; Stahl & Muckle, 2007) have found that displacement values symmetrically distributed around zero have very little impact.

At the American Board of Family Medicine (ABFM), we have defined a displacement with an absolute value greater than or equal to 0.6 logits as excessive and found this threshold to be useful. When we implement this criterion we typically find that 10% to 25% of our anchor items are flagged for excessive displacement. At the American Board of Pediatrics (ABP), we also find this displacement threshold useful for identifying items that should be unanchored. When we implement this criterion, we usually find that it flags between 5% and 15% of anchored items on our largest volume examinations (n>500), and between 10% and 30% of anchored items on our subspecialty examinations, which have much lower candidate volume.

Table of Contents

Although the ABFM and the ABP have found the 0.6 logit criterion useful, this does not mean that it will be useful for everyone. Psychometricians considering other examinations may find different thresholds to be useful. There is always a trade-off when unanchoring items, an example being that it suggests there are some differences in the construct across administrations. Of course changes in the construct can happen and should be accommodated.

To illustrate, imagine a question about HIV being given on a test in 1986 and again in 1992. In 1986, the question would be about a rather obscure immunology topic, but by 1992 it would represent a current events topic. Answering the question correctly on those two different occasions would represent two very different levels of immunology knowledge. Clearly, using an item calibration that better reflects the data will improve the data-model fit which is important for interpreting the meaning of a measure. On the other hand, too much "flexibility" in permitting the items to float will cause the substantive understanding of the construct to become fuzzy and perhaps less useful. Finding the correct balance between the stability of the substantive meaning of the construct and the conformity of the respondents to that construct is difficult, and is largely the reason why different thresholds for what is considered excessive displacement exist and are likely to continue. We have found that the 0.6 logit threshold typically restricts the flagging of items to those that might produce a noticeable effect on the test instrument, and usually flags fewer than 15% of the anchored items. This is useful for us. Please tell us what thresholds you use and why you find them useful.

References

Draba, R. (1977). The Identification and Interpretation of Item Bias. Research Memorandum No. 25, Statistical Laboratory, Department of Education, University of Chicago.

Jones, P. & Smith, R. (2006) Item Parameter Drift in Certification Exams and Its Impact on Pass-Fail Decision Making, Paper presented NCME, San Francisco.

Linacre, J. M. (2013). Winsteps® Rasch measurement computer program User's Guide. Beaverton, Oregon: Winsteps.com.

Stahl, J. & Muckle, T. (2007). Investigating Drift Displacement in Rasch Item Calibrations. *Rasch Measurement Transactions*, 2007, 21:3 p. 1126-1127.

Wright, B.D. & Douglas, G. A. (1976). Rasch item analysis by hand. Research Memorandum No. 21, Statistical Laboratory, Department of Education, University of Chicago.

Thomas O'Neill and Michael Peabody, *American Board of Family Medicine*
Rachael Jin Bee Tan and Ying Du, *American Board of Pediatrics*

Rasch Forum Exchange about "Quantitative Attributes"

Andrew Ward:

I've been reading a little of Prof Michell's work and found this comment in one of his papers: "An examination of some relevant textbooks [a list is given, including Bond and Fox, 2001] reveals a consistent pattern: the issue of whether the relevant psychological attribute is quantitative is never raised as a source of model misfit. Other issues, such as the unidimensionality of the underlying attributes, item-discrimination parameters and local independence, are raised, but item response modellers appear never to question that their attributes are quantitative."
(Michell, J., 2004: Item response models, pathological science and the shape of error: Reply to Borsboom and Mellenbergh. *Theory and Psychology*, *14*, 121–129).

One conclusion we might draw is that, while the Rasch model potentially creates measures with useful properties, it may still be useless if an attribute isn't quantitative in the first place.

Rasch Measurement Transactions
www.rasch.org/rmt
Editor: Kenneth Royal
Email Submissions to: Editor \at/ Rasch.org

Mike Linacre:

An interesting, but erroneous, conclusion, Andrew. Rasch needs an attribute that is "ordinal". If we can say that one attribute of an object is "more" than the same attribute of another object in some sense, then that sense defines a latent variable along which Rasch can construct measures. For instance, if an observer says that the Pope is "nearer to God" than Bishop Smith, then we have a latent variable of "nearness to God" along which measures can be constructed. This example comes from http://www.rasch.org/rmt/rmt72d.htm

Rasch fit statistics tell us how well our ordinal observations of attributes of objects conform to the ideal of a unidimensional additive latent variable.

Michell appears to claim that some attributes are inherently quantitative. In all of science, cooking, etc., "quantities" do not exist naturally. They are not inherent. Quantities must always be constructed by the application of some rule. This is made explicit in such rules as "The Treaty of the Metre" - http://en.wikipedia.org/wiki/Metre_Convention. For Rasch, the rule is based on "ordinal comparisons".

On the IMEKO 2013 Joint Symposium in Genoa, Italy

A number of presentations involving Rasch measurement models, methods, and results were made at the joint symposia of the International Measurement Confederation (IMEKO) technical committees on measurement science and metrology education (TC-1 and TC-7) in Annecy (France), London (England), and Jena (Germany) in 2008, 2010, and 2011, respectively (Fisher, 2009, 2010, 2011). A paper based on Mark Wilson's keynote address at the 2011 meeting has recently been published in the IMEKO journal, *Measurement* (Wilson, 2013).

This journal also has a forthcoming celebration of the work of the late Ludwik Finkelstein in press (volume 46, number 8, pp. 2885-2992). Finkelstein made a large number of foundational contributions to educational and conceptual issues in measurement science, and had a special interest in exploring the possibility of a unified science of measurement applicable across the natural and social sciences (see, for instance, Finkelstein, 2003, 2009, 2010).

Wilson's keynote at the Jena IMEKO symposium in 2011 was given at the invitation of Luca Mari, an engineer and philosopher of measurement based at the Universite Cattaneo, in Castellanza, Italy. Wilson reciprocated the invitation by bringing Mari to last summer's International Meeting of the Psychometric Society in Lincoln, Nebraska. Mari gave a well-attended workshop on metrology, and an invited address. He will also be a visiting scholar in the Graduate School of Education at the University of California, Berkeley, in November, 2013. Mari is intensely involved in the ongoing revisions to the International Vocabulary of Metrology (known as the VIM; Joint Committee on Guides in Metrology, 2008), especially as this involves efforts continuing Finkelstein's interest in integrating measurement concepts from all fields into a common frame of reference.

The most recent instance of the IMEKO joint symposium (which now also includes TC-13, the technical committee on measurements in biology and health care) was held in Genoa, Italy, September 4-6, 2013. The papers presented are available in volume 459 of the *Journal of Physics Conference Series* at http://iopscience.iop.org/1742-6596/459/. Mari and Wilson's keynote providing a "gentle introduction to Rasch measurement models for metrologists" will be of special interest. Additional Rasch-oriented presentations were made by Maul, Torres-Irribarra, and Wilson; Camargo and Henson; Bezruczko; Stenner; Massof; Stephanou, Pendrill; and Fisher. RMT readers may also be interested in related work presented by Benoit, Crenna, Rossi, Granovskii, Pavese, Ruhm, Thomas, and others.

An exciting new dialogue between the natural and social sciences is underway. Each has much to learn from the other. Metrology has had little need to attend to the individual-level stochastic processes structuring invariant cognitive and behavioral constructs, and measurement practice in psychology and the social sciences has everything to learn about the value of local traceability to globally uniform units. Everyone interested in contributing to or learning from this dialogue is invited to make their voices heard.

References

Finkelstein, L. (2003). Widely, strongly and weakly defined measurement. *Measurement*, *34*(1), 39-48(10).

Finkelstein, L. (2009). Widely-defined measurement - An analysis of challenges. *Measurement*, *42*(9), 1270-1277.

Finkelstein, L. (2010). Measurement and instrumentation science and technology-the educational challenges. *Journal of Physics: Conference Series*, *238*, doi:10.1088/1742-6596/238/1/012001.

Fisher, W. P., Jr. (2008). Notes on IMEKO symposium. *Rasch Measurement Transactions*, *22*(1), 1147, http://www.rasch.org/rmt/rmt221.pdf.

Fisher, W. P., Jr. (2010). Unifying the language of measurement. *Rasch Measurement Transactions*, *24*(2), 1278-1281 http://www.rasch.org/rmt/rmt242.pdf.

Fisher, W. P., Jr. (2012). 2011 IMEKO conference proceedings available online. *Rasch Measurement Transactions*, *25*(4), 1349 http://www.rasch.org/rmt/rmt254.pdf.

Joint Committee for Guides in Metrology (JCGM/WG 2). (2008). *International vocabulary of metrology: Basic and general concepts and associated terms, 3rd ed*. Sevres, France: International Bureau of Weights and Measures--BIPM. http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2008.pdf. Accessed 17 October 2013.

Wilson, M. (2013). Using the concept of a measurement system to characterize measurement models used in psychometrics. *Measurement, 46,* 3766-3774.

William P. Fisher, Jr., *University of California-Berkeley*

Journal of Applied Measurement
Vol. 14, No. 4, 2013

Application of the Rasch Model to Measuring the Performance of Cognitive Radios, *Edward W. Wolfe, Carl B. Dietrich, and Garrett Vanhoy*

Properties of Rasch Residual Fit Statistics, *Margaret Wu and Raymond J. Adams*

Validating Workplace Performance Assessments in Health Sciences Students: A Case Study from Speech Pathology, *Sue McAllister, Michelle Lincoln, Alison Ferguson, and Lindy McAllister*

Rasch Analysis for the Evaluation of Rank of Student Response Time in Multiple Choice Examinations, *James J. Thompson, Tong Yang, and Sheila W. Chauvin*

Assessing DIF Among Small Samples with Separate Calibration t and Mantel-Haenszel 2 Statistics in the Rasch Model, *Ira Bernstein, Ellery Samuels, Ada Woo, and Sarah L. Hagge*

Application of Latent Variable Model in Rosenberg Self-Esteem Scale, *Shing-On Leung and Hui-Ping Wu*

A Rasch Analysis of the Statistical Anxiety Rating Scale, *Eric D. Teman*

Richard M. Smith, Editor, www.jampress.org

From the Archives: A 1981 Interview with Ben Wright

Interview of Benjamin Drake Wright recorded by David Andrich in Judd Hall Room 438 at the University of Chicago during April, 1981, when David was in Chicago for the first International Objective Measurement Workshop, held in honor of Georg Rasch who left us in August, 1980. (Transcript edited by Ben Wright in 1995).

David: When you first heard Georg talk and were the only one left to listen, it was at an early stage and there were many things that had not been studied or understood. Can you recall what prompted you to think it was worthwhile following his suggestions?

Ben: The stage was set for me because I came into statistics by accident - it happened to be the job that was open. I had never studied statistics as such. The only statistician I had taken courses from was Bill Stephenson of Q-Technique fame. He wasn't a statistician. He taught me how to do factor analysis by hand. He was very clever at it. He was also very independent and very much like Georg. He loved life. Later, when I listened to statistics courses by other social scientists, I was terribly disappointed.

David: What about Feller?

Ben: When I took courses at Cornell from Feller in 1947, I did it for fun. That was probability theory - not statistics. It was a mathematics course. I understood quite a bit of it and enjoyed it. But I didn't think it was ever going to be useful for anything. Later, when I took-up teaching statistics, because it was the job available, I got into trouble. The social science statistics textbooks didn't make sense to me. When I consulted my friend, Jimmy Savage, about the problem, he said: "You're right, these books don't make sense." We got talking about statistics and inference and he taught me his Bayesian approach and likelihood functions and some of Ronald Fisher's work on inference - aspects of Fisher which don't get into text books, especially those in social science statistics. I felt reassured that my common sense and Jimmy, the experts' expert, came to the same conclusion.

While we were talking, I mentioned that I was using factor analysis to analyze some data collected from child care workers at Bettelheim's Orthogenic School. Jimmy was skeptical about the utility of factor analysis. But he was also reasonable and open to discussion. He could see that decomposing matrices into simpler structures, when one did not expect these structures to have any particular inferential status, could be convenient. Jimmy and I and then Dave Wallace and Raj Bahadur, carried on a debate about factor analysis for some time. At the end, we held a series of public meetings, this was in the late fifties, in which Wallace and Bahadur gave their critiques of factor analysis. The ambiguities in factor analysis which Dave and Raj identified were, of course, violently opposed by the cookbook factor analysts left in the Psychology Department (Thurstone had gone to North Carolina). These dogmatic factor analysts, for whom factor analysis was their dignity and identity, were terribly upset by the criticisms that Wallace and Bahadur leveled at their life's work.

I didn't mind the criticisms. Dave and Raj were right. If it were not necessary to think that factor analysis produced some kind of truth, however, but merely a convenient device for summarizing intricate data, then Dave and Raj were perfectly agreeable. That was what I was using factor analysis for. So I came out on their side of that fight.

It was in this context that Georg appeared. Jimmy ran into him at a Biometric Society meeting in Washington (Georg was a founding member). Jimmy had met Georg earlier in

Copenhagen. Georg pressed Jimmy with his need to tell the world of his new discoveries about measurement. Georg was lobbying vigorously on behalf of his new ideas in those days. Five years later when he happened to be flying from Stockholm to Copenhagen next to Ben Bloom, he talked Bloom's ear off, impressing him enough to get him interested in what I was doing with the Rasch model back in Chicago. That's what led Bloom to invite me to give that infamous talk at the 1967 ETS Invitational Conference. I had presented similar, but more detailed material, to the Midwest Psychological Association in 1965 and the Psychometric Society in early 1967. But it was the ETS talk that got the ball rolling.

Georg wanted to tell his news and Jimmy had the money for a visiting professorship. Jimmy said: "Well, Ben, if you tell me to have him come, I'll bring him. I don't see a reason for the Statistics Department to have him. But, if you think the people in Psychology or Education will be interested, then I'll bring him." So I felt obligated to go to Georg's lectures. I might have gone anyway, but it was a personal promise.

So Georg came. His official host was Jimmy. He lived by himself in the Chicago Theological Seminary dormitory behind Robie House. I met him at Jimmy's house for dinner and cocktails. Then I went to the first class, which was heavily attended by most of the Statistics Department and almost all of the statistical people in the Social Science Division.

Georg was bombastic and intolerant, bragging about how smart he was and so was truly obnoxious. People stopped coming. The social scientists couldn't understand the math. The statisticians thought he might be insulting them. Jimmy fell asleep about half way through the first lecture and slept all the way through the second. Then he stopped coming. It was a funny and sad situation. Here was this man who cared so much about what he had to say. And he was driving his audience crazy. He was forcing them to fall asleep or leave in order to defend themselves against what he was doing. Gradually everyone thought up reasons why they needn't come. They "already knew what he was going to say. What he said wasn't really important."

During Georg's opening remarks, I paid more attention to the social situation than his words. It was so dreadful. About all I noticed was that now he seemed critical of factor analysis. This interested me because, as far as I knew then, the only measurement paper he'd published was his 1953 contribution to the Uppsala Symposium on Psychological Factor Analysis, a copy of which he had given me to read.

Many years later, during one of our last conferences, Georg asked me to tell him again exactly what I did in factor analysis. I said, "Not too much since I met you. It's not been satisfying, except in a case where I had the chance to factor the same instrument forty times." This was a well-developed questionnaire for gathering consumer's reactions to product presentations. Was the product appealing? exciting? reassuring? Three semantic differential dimensions were intended. Ralston Purina tested forty different commercials over a four year period. Each time two hundred consumers viewed these presentations and rated them on the instrument. The data had been sent to a firm where I was a consultant. They paid me to factor analyze each of the forty data sets and to report one common set of factor scores for all respondents in all forty studies. When I told that to Georg he became excited. "You must write that up, that's terribly important, getting similar results time after time." So I guess factor analysis was still on his mind.

David: He once told me that he thought it could be formalized to be done objectively.

Ben: His teacher, Nobel economist Frisch, developed a kind of factor analysis called "confluence analysis". Georg learned that from Frisch in the 1930's. That's what Georg wrote about in his 1953 article.

To get back to your first question, it wasn't Georg's great ideas that first caught us. What caught Jimmy's interest was that Georg was a good fellow who had been hospitable to him in Denmark, who was fun to talk with and who was very forward about his need for a forum. I felt obliged to go to Georg's lectures because I promised Jimmy I would. Then, I felt concerned about Georg being deserted by his audience. Although he was a bit obnoxious, I didn't feel he had nothing to say. What he was saying was becoming interesting. Yet everyone was deserting him. Under these circumstances, I couldn't desert him. So I stayed on and listened more closely and began to be interested.

The odd thing is that his ideas didn't come first. First came my promise to Jimmy, second came my distress at a passionate man who had come so far and cared so much but had no audience. What came third was my admiration for Georg's courage. He kept on explaining as though he didn't see there was nobody there. He didn't give-up. He brought in his notebook. He opened it carefully. He gave his lecture, even when there was no one there but me and John Ginther. And after awhile, even John Ginther stopped coming. I certainly couldn't leave him then, when I was his last student.

So we made friends. After his lecture we had lunch. He liked sardines on black bread with lots of pepper and beer. As we chatted over lunch I began to show him some of the semantic differential data that I was using factor analysis on. He took some of it home and spent hours and hours copying it into various formats, making tables and graphs, trying to do something with pencil and paper to analyze it. I couldn't help but become interested in his techniques and, especially, in how hard he was willing to work on a problem. My honest view of those semantic differential data was that they might not be sufficiently compatible with his demanding ideas to support a favorable outcome. I feared he was wasting his time. But I was fascinated that he could work so hard to see whether he could do something with my data, since I myself, at that moment, doubted whether much could be done with them. But he threw himself into it and, as a result of these efforts on his part, I reread his book more carefully and began to appreciate it much more deeply.

After that spring, however, as far as I was concerned, the whole thing disappeared. He went back to Denmark. Later he sent me an autographed copy of his book, long since "borrowed" by someone. I didn't see him again for some time. He called me on the phone once when he was in Washington to ask what I was doing. But I was working on something quite different and had my own methods for dealing with it and didn't think I needed his ideas. It wasn't until I came to the end of my research on teacher development, from which I had lots of semantic differentials to analyze, that I thought about him again. Then, mostly as an excuse to take a trip, I decided to go and see him, bring some of my problems along and see if he had anything to say. I really didn't think he could do anything about those problems. But I did think it would be fun to go to Denmark to visit him. My professional excuse was that I was going to investigate the usefulness of his new methods for these old data that I had.

When I finally got to Denmark in the Spring of 1964, however, I began by being annoyed with him because he was so unhelpful to our coming there. He said he was going to be helpful. But he ended up not finding us any place to live. It was by the skin of our teeth that we finally figured something out. So, I arrived feeling that he was even less interested in my coming there than I was. At that moment I thought the whole visit was going to be a disaster. But the fact is that Georg spent a great deal of time with me. He began by treating me to a splendid, three hour, highly alcoholic Danish Anretning (meaning "everything on the menu") at The Little Prince, his favorite Copenhagen restaurant. Then he had me come out to his rather imposing home in Holte, just outside Copenhagen, three or four mornings a week.

He spent the morning lecturing me on math and statistics - just him and me. He had a blackboard installed in his bedroom for the occasion (that being the only room his wife would let him hang it in). We sat in front of the blackboard installed for this purpose. He lectured and I would took notes. I still have those notes. He told me about his book and what one could read in his book, if one cared to. He also took me with him on his consultations to various Danish institutions.

These consultations took place after an enormous, alcoholic lunch to which Georg always treated me. The consultations were rather amusing. At the military psychology group Georg introduced me to the people there. Major Borking, the "boss", sat at his big desk in command. Georg sat opposite in the "important" chair. Three young fellows sat, one on the couch beside me, and the other two on little chairs along the wall. George asked the young fellows to report on what they had been doing. As soon as they began, Georg went to sleep. They reported to the sleeping Georg for forty-five minutes. Towards the end he awoke, told them what to do next and we left.

On the way out, the Major, a bit distressed at an American visitor seeing the famous Danish mathematical consultant sleep through the consultation, took me aside to explain how miraculous it was that, even though Georg always seemed to be sleeping, he nevertheless heard the reports well enough to know exactly what to advise the group to do next. And, I must say, everyone seemed entirely satisfied with Georg's advice. It was a charming and mysterious experience.

Another place we went was the Danish Pedagogical Institute. That's where I met Georg's son-in-law (Lotte's husband) and author of the Bo Prien's Prove (BPP-N and BPP-F tests) analyzed in Chapters V and VI of Georg's book. Bo showed me the new math test he was working on. At the time it seemed a bit labored to me. Bo was moving at a snail's pace. He had hundreds of multiplication and division items arranged in patterns on sheets of paper to help him think about the order in which they were best done and the effects of digit processions. It was hard for me to see how so much detail could pay off. [Today (1995), however, after Jack Stenner's astonishing success with his Lexile specification equation, I can better appreciate what Bo was trying to do back in 1964.] In spite of my reservations, however, I paid careful attention throughout Bo's lengthy explanations because I wanted to be a good guest and I liked Bo. In addition to Bo's expositions, Gus Leunbach took me through his Rasch model computer programs, every step of the way. And he had lots to show me.

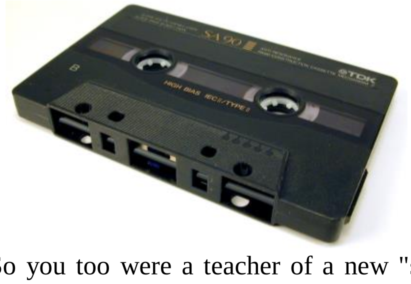
Eric Thompson, the institute director, was the charming leader and diplomat of the group. He was the one who told me how Georg's book ever got written. Georg's weekly impromptu consultations in Erik's office during 1955-1958 were tape recorded by Erik's secretary. The recordings were then transcribed in Danish by Georg's daughter Lotte (Bo's wife), the other mathematician in the family. Finally, Gus Leunbach, "transformed" and later "revised" the Danish manuscript into English (Page xxii of the 1980 edition). Erik also showed us how to smoke a cigar in the shower and we all had fun together chatting, smoking cigars and drinking port. It was an interesting world to be caught up in. The atmosphere was extremely engaging.

David: But there must have been other people in the world to whom you could have gone in the interests of a professional trip. You could have gone to England, New York, San Francisco. Was there an intellectual hang-over, a seed that Georg had planted, that induced you to choose Georg?

Ben: You are right. There was. Georg was unwilling to take traditional clichés for granted. That intrigued me. His impassioned conviction that we are going to think for ourselves, that we are not going to just believe what anybody else says, that we are not going to just do things the way others have done them, but are going to figure things out for ourselves, and only do what makes sense to us, only do what we are able to make sense out of, that really appealed to me. That's the kind of person I am. Georg was a kindred spirit.

At the beginning, however, I didn't really know what he was talking about, or why he was talking about it. I didn't see my own scientific problems in that light. I didn't recognize his ideas as a solution to my problems. I had come to think that what most educational statisticians were doing was screwy. All those correlations computed between all kinds of things, reliability coefficients, validity coefficients, all that gave me a headache. I thought this can never be science. This is a mess. Georg's approach, in contrast, was clean and clear. He didn't get into any of that goofy stuff. He went right to the observation and modeled it. I liked that idea very much. It was clean and clear, fresh and new, sensible and uncluttered.

That was Georg's message. That was the intellectual side of Georg Rasch that appealed to me. And I know that, however I may excuse my going to Denmark, I would not have gone to see just anybody. I was interested in Georg's courage and vision, in his ability to see to the heart of a matter, and in his unwillingness to give up for political reasons. He didn't change his tune when people left his class. Some people might have. He didn't quit and he didn't change. And I didn't think he was wrong. I listened to him and I thought, "This makes sense, in fact, better sense than anything I have heard so far." I had already gotten used to thinking that most of the educational statisticians didn't know what they were talking about. I was completely baffled by what the educational measurement people wrote about. Every time I read one of their books I thought, "This doesn't make sense."



David: So you too were a teacher of a new "statistics", struggling against the establishment. When do you reckon you internalized Georg's principles, so that you would invest so much time in refining and developing them?

Ben: It happened in 1964 and 1965. I was quite tentative when I went to Denmark in 1964. But I was sufficiently intrigued and educated by him in 1964 to go back in 1965, which I arranged to do almost immediately. When I returned in 1965 I took Bruce Choppin with me. Bruce was going home to England for the summer. So I got him to come to Denmark for a few weeks. I was already thinking of being his teacher. Bruce had come to Chicago the year before. He had the same kind of physics background that I had, so we had a lot in common. I asked him to stay with us in Denmark so he could come to some of the lessons with Georg.

Bringing in a student that I valued made a difference. Perhaps I wanted to motivate Bruce to take an interest in Rasch. Perhaps I wanted Bruce's companionship in thinking about Georg's ideas. You know how much better I work when I do it with a student than on my own.

I had tried a few programs on my own between 1964 and 1965. But I did nothing before 1964, except to glance at his book. I certainly did not read it or work on it. I wasn't dealing with dichotomous tests at that time and I wasn't the least bit interested in them. But in taking Bruce to Denmark, now that I think about it, I must have already had something in mind. Bruce and I had lectures from Georg together. When we came back to Chicago we got right to work. We wrote FORTRAN programs for all of Georg's algorithms: LOG (the linear log method he applies in Chapter V), PAIR (the pairwise method he proposes on page 171) and SYMFUN (his fully conditional algorithm based on symmetric functions) and we tested them against simulated data to make sure they worked for us.

We must have worked pretty fast because we organized a MidWest Psychological Association symposium in the fall of 1965 in Chicago. In those days the MidWest Psychological Association had big meetings in Chicago with hundreds of people. Our symposium met in a ballroom and there were at least a hundred people there. We had handouts which may still be in my files. I got Jane Loevinger, a fan of Georg's work, to introduce us and Dave Wallace to discuss the presentations. The papers were Bruce, me, Gary Ramsayer and Richard Brooks, two from Iowa State who had been trying Georg's LOG method in their dissertations because their professor had seen his book. Bruce and I showed that all three algorithms always gave the same answers. That was the debut of Rasch work in this country.

But then Bruce did his thesis on something else. Neither of us can remember how it came about, but he did his thesis on a computer analysis of the psychological structure of children's themes instead of on the Rasch model. Bruce's thesis had nothing to do with Georg's work. If it hadn't been for Nargis Panchapakesan, the third physicist, who was in Chicago with her husband, a nuclear physicist at Argonne National Laboratory, and was looking for something to do, the work might have stopped right there.

Nargis had a physics Ph.D. from Calcutta and came around to MESA because she was interested in education. She began with Ben Bloom. But he didn't know what to do with a physicist so he sent her to me. I talked her into helping me work on the mathematical side of Rasch estimation and to write some better computer programs. She learnt FORTRAN. She was a good mathematician. Finally, I talked her into getting another PhD. She didn't really want to. But she liked the work. That's how the UCON FORTRAN routine (our own unconditional algorithm, still used in nearly all Rasch programs) was born.

It was because I had Bruce and Nargis working together that we got so much done so fast. In those days I spent hours at the computation center. I had an office there. I had free computer time. I was one of the few faculty members who used computers heavily. Most of the other users were physicists, astronomers and meteorologists. This was 1965-67. University computing was just getting started. We had a wonderful IBM 7090. I used to run it myself - the whole thing. It filled an enormous room full of tape drives and memory banks. By today's standards it was crude and primitive. But it had a 32K core! That's half what those little 64K Radio Shacks in the next room have. It took 12 engineers to keep the 7090 going. But we were able to do a lot of things fast that people hadn't done before. In particular we could simulate data with known properties and the use them to test our programs.

There was a lot packed into those few years. I was with Georg in 1964. Then Bruce and I studied with him in 1965. Then we had the Midwest Psychological Association Symposium in 1965. In the spring of 1967 I gave a Rasch paper on conditional estimation at the Psychometric Society in Madison. Then in the fall of 1967 there was that fatal ETS talk in New York. That was meant to be the end of it. That is where I thought I was going to stop. By then I felt I had done everything there was to do.

Sometime in the winter of 1967 we wrote a program for the 2 item parameter Birnbaum model that Darrell became so fond of. We thought it might be worth trying. But after exhaustive investigations, not only by us but also by Bob Ashenhurst, Alex Orden, Hirondo Kuki and anybody else we could get interested, we convinced ourselves once and for all that there was absolutely no chance of Birnbaum's algorithm ever converging on its own. To imitate convergence we had to impose some kind of prior distribution on either the ability parameters or the discrimination parameters.

The trouble, of course, was that while there was no way to determine what the arbitrary constraint should be, the results you got depended upon what constraint you chose. Thus we saw clearly and completely that the Birnbaum model could never qualify as a satisfactory method of inference. There would always be an essential ingredient which would depend entirely on the taste of the analyzer. And how would one know what taste to have? How could people agree on what arbitrary choice they should make?

So we gave Birnbaum up as a bad lot. Then, alas, I had that awful Quad Club lunch with Darrell, Fred Lord and Bill Angoff. I told them what had happened with their favorite Birnbaum model and reminded Fred of his own negative comments in the manuscript which he finally published in Ed.Psych. Measurement in 1968.

That was Lord's first and most famous paper pushing the three parameter model. It contains, at the end, his telling comments about not being able to stop the abilities and discriminations from diverging to infinity, about the need to drop people and items from time to time to prevent divergence, more or less haphazardly, and the need for hundreds of unconverging iterations. These dire comments, which appear at the end of the three parameter paper of 1968, were also in his earlier research memorandum about Birnbaum.

We discussed the problem at great length at that luncheon. He admitted that all of those seemingly intractable problems were real. But he still felt that it must be possible to solve the problem and that it was only a matter of weeks or months at the most before he would be able to prove algebraically that this process must converge. He was absolutely sure that it would be a short step from that proof to actually obtaining convergence in his computer program. He had two top-notch people working for him, a mathematician and a programmer. They would soon solve the problem and Darrell encouraged him. Darrell said he too felt that it would be no trouble to solve this problem.

Later, however, Darrell did admit to me privately that it was "probably impossible" after all to solve that convergence problem, but why not just impose some nice distribution on the ability parameters. "Wouldn't that be all right? What could be the harm? If we can't solve the problem without imposing something arbitrary, since we know discriminations are 'true' and therefore must be estimated, why not just impose the original raw score distribution or some such thing that you actually have, wouldn't that be fair enough?" I objected for the obvious reasons. Raw scores are not linear so their distribution is the wrong shape (which could be fixed by log odds). Far more urgent, the score distribution is sample dependent. But he went ahead, as you know, and continues to program 2 and 3 parameter models and Lord's program still doesn't converge after 14 years (in 1981, 28 years in 1995) of being just around the corner.

To get back to history, Georg was very much with me in the back of my mind through these unpleasant discussions with Lord and Bock. His courage to stick up for what he believed in and to keep talking, even in the face of desertion by his audience, gave me the courage to stick to my guns and to do what I thought was sensible. I don't think I was just being contrary, although I was fighting for my independence and identity as a young scholar. But after this lunch, I certainly did not want to be a follower of Lord or Bock. So Georg gave me a chance to be a different person from the people around me. He also gave me an example of how not to be agreeable to silly ideas but rather to fight against them. But, I guess, I was already inclined that way. I had already been doing that. This event was not the beginning of my career as a trouble-maker. But certainly Georg fit in with that. The interesting thing is that so much was done between 1964 and 1967 with Bruce and Nargis and nobody else - just the three of us.

I was about to let the whole thing drop when Ben Bloom insisted that I give the fatal ETS talk in the Fall of 1967. I felt that the talk would surely finish my Rasch work. I had become obliged to write the Ed.Psych.Meas. paper with Nargis on UCON. It seemed sensible to put into a paper what we had been doing, so we did, forgetting to mention that we were unbiasing the UCON estimates so that they would match the conditional ones. The program unbiases, but the paper didn't mention it, so we had many questions after that. "Why does your program have this (L-1)/L factor when it is not in your paper?" Even so, I felt that article finished my Rasch work.

But then David Farr insisted on my doing that first ever AERA pre-session on the Rasch model (1969). At first, I felt apprehensive about the task and annoyed at being asked. But, again, Ben Bloom said I must do it. It was important to do it, absolutely essential for my career, for the sake of science. Both.

David: It is interesting that in spite his own lack of interest in this kind of thing Bloom was astute enough to see the significance of it.

Ben. If he had not encouraged me, I would not have done it. I wouldn't have given the ETS talk. I wouldn't have done the pre-session.

David: Who was David Farr?

Ben: A nice guy at the Buffalo University of New York who was on the AERA committee for professional training. It was the first year that AERA sponsored pre-sessions. David had heard me talk at the ETS conference and thought it was a great idea. He called me and started persuading. I couldn't resist him. So I did it.

That was the Spring of 1969. The presession was held in Los Angeles for five days. Georg gave the final lectures. He had been working with me in Chicago during 1968-69. I brought him over in the Fall and he and Nille stayed 9 months.

David: This was after 1967, after you thought it was the end, so you must have gotten another impetus.

Ben: I spent August, 1967, at Georg's thatch-roofed cottage on Laeso in order to show him how well our simple, easy-to-use unconditional method (UCON) worked. We had talked about it in 1965. But he insisted then that it would be absolutely wrong. He feared it would not take full advantage of separability and surely that must not be right. When I objected that his log method worked exactly that way, he said, "Yes, but I only used the log method because I didn't have any way to apply the conditional method then.

When I went in 1967, I took a suitcase full of output - very heavy to carry. I went by ship and took my son Chris with me. But Chris didn't go to Laeso. He stayed with friends in Jutland. That was when I showed Georg that the only bias produced by UCON was removed by the factor (L-1)/L. All of my output from simulated and real data of all kinds showed, over and over again, for different length tests, for different patterns of item difficulties, that you got rid of all discernible trouble by shrinking the item difficulties by (L-1)/L. He saw that he couldn't win the argument. But he didn't like it.

That was an important point in our relationship because at that moment he and I separated a little bit. Up until then, as far as he was concerned, I was doing everything exactly the way he told me. I was entirely his creature. And up until then I pretty much let him make all the decisions. I didn't argue with him about anything. But UCON was a new something that we did on our own, not to his liking, which seemed to me plainly convenient, practical and useful. So it was a point in our work where I was becoming myself, in spite of his, indeed, against his wishes.

I also showed him our thorough investigation of the symmetric function program necessary for conditional estimation. But that didn't placate him. He really wanted me to do conditional estimation and no other. He didn't want me to develop UCON further. We continued to be good friends. But from that summer of 1967, there was a bit of a difference between us.

He did, however, come to Chicago for the academic year of 1968-1969. That was how he was able be in Los Angeles for the AERA the Rasch Model presession. The presession was attended by lots of big shots: Bill Angoff, Chester Harris, Henry Kaiser and fifty others. But it didn't make a dent on them.

It did, however, make an impression on Lou Bashaw, Bob Rentz and Charlotte Cox. Rentz and Cox were Bashaw's graduate students. They didn't really understand what was going on. But Bashaw liked it and he pushed them and they finally got a large grant to equate a mass of published test results.

There was a second presession in Minneapolis in 1970, the one George Ingebo sent Fred Forster (of Portland Public School and NWEA fame) to. Once again, after those two pre-sessions, I thought, "That's it." But then you arrived at the end of 1971. I had about signed off on Rasch measurement. I had no more students in the area. I felt that I'd done everything I wanted to do. But then you arrived and so I started again. And then Graham came and then Geoff and then and then and then.

David: Graham's enthusiasm for algebra gave you an opportunity to return to technical questions that you had let slide.

Ben: Yes, Graham and I did Nargis' work all over again making her results even more certain. I also got involved with those English people at NFER about then. That was Brian Start's fault (He also dragged me to Melbourne in 1974). He became head of research at NFER shortly after his participation in the 1969 presession. He asked me to come to NFER to teach Desmond Nuttall and Alan Willmott about the Rasch Model. That must have been 1971. Bruce was working at NFER at that time, but not on the Rasch model, although he had published papers about the PAIR method and item banking.

David: It is interesting to see how many people with a physics background are involved. Georg himself, even though he studied with Fisher and was a mathematician, not a statistician, applied his work to physics problems. Then you, Bruce and Nargis and, even, in a sense, me. I too had to teach statistics without ever having done any.

But why is there so much passion in this topic? Maybe when you're at the edge of an area these things happen. Maybe that's what brings out the personalities of the people involved.

Ben: I feel pretty sure that, had I not gone to Georg's lectures because I promised Jimmy Savage, that Georg would have had no affect at all in this country, even until now. Certainly not in the 60's, because there were so few people who had heard about his book or seen a copy. I ran into some people at the end of the 60's and even more recently who said, "Oh, yes. I saw that book, but I didn't pay any attention to it. Is there really something good in there?" So Georg's book fell on deaf ears. Someone might have picked it up and discovered it, we don't know. But no other American did anything about it except this guy in Iowa with his two students and those guys disappeared.

****

David: Lumsden is an interesting example. He was trying to discover the same things as Georg. It is clear to me that the reason Lumsden has taken so long to warm to Georg is because he was heading for the same thing himself, but nowhere near as efficiently or elegantly as Georg. Lumsden was also dissatisfied with traditional test theory and three parameter models. He was heading in Georg's direction, but on his own. And he really resisted believing that anything he had worked out for himself could possible already be in Georg's book.

Ben: It's his identity, you see. My meeting with him here was terrible, one of the worst experiences I have ever had. I had looked forward to meeting him. I suppose he had looked forward to meeting me. He certainly made a big fuss about it. He brought his girlfriend with him. But the poor woman never got a chance to say a word. He marched in, sat down and bombasted me with pronouncements. He started writing all over my blackboard. I had some messages and equations up there. He wrote right over them. I thought, "Who is this clod! Bursting in. Stamping all over me. What's the matter with him!" But I guess he felt defensive and compensated by being aggressive. Perhaps he was afraid that I wouldn't acknowledge what he had to say, wouldn't give him credit for being the "true" inventor of ideas which both he and I knew were already well developed by Georg.

It was difficult to talk with him. He wouldn't listen to a word I said. As he stamped over to the board, he knocked over the pictures on my desk. I felt like calling the police. I thought, "What have I got here. This guy's wrecking my office and he just came in the door." His poor girlfriend was shrinking into her chair with embarrassment. It was an awful occasion.

Then he insisted I have dinner with them. I had already planned to have dinner with Ross Lambert, so we all went together. When Lumsden found out that Ross was an ophthalmologist. He started lecturing Ross on how to do eye surgery. Ross said to me, "I don't know how much of this guy I can take. He doesn't have any idea what he is talking about."

Then Lumsden got into an argument with the waiter and told him that he didn't like America and didn't like the food. The waiter said, "Right now you’re eating in America, so you better make up your mind what you want." I am hoping to have a nice dinner and Lumsden's picking a fight with the waiter. I thought, "I don't believe this!"

****

David: But despite all these critical incidents, you have invested a lot of time developing and propagating Georg's ideas. You still seem prepared to go on pursuing them. Why?

Ben: I really don't know. There's a great reluctance in me. Part of me has to drag another part of me through it. Part of me doesn't want to do it at all. I always felt that, if I was going to do this right, I should brush up my mathematics. I have always felt delinquent in the mathematics part because I knew that if I would take the trouble to learn a little more mathematics then the whole thing would be much easier to do and I would do it better. But I am lazy. I just don't get around to spending a year or two, or even an hour or two a day to sharpen my mathematical tools. Since I viewed myself as lazy about the mathematics, I always felt guilty, as though I were a bit of an imposter, and I really shouldn't be doing it.

Most of the things I have done about Rasch, I was forced into by circumstance. Each time I came to a place where something had to be done and nobody else was going to do it so I had to do it myself whether I had the mathematics ready or not. So I just plunged in and did the best I could with what often seemed to me the primitive tools I had in hand. And I always felt guilty that I hadn't been responsible enough to go to the store and get good tools.

"Why didn't I?" I don't know whether I wasn't interested or whether I feared I would turn out to be incompetent with them. The conflict inside of me has always been, I don't know enough mathematics to do this, so why am I doing it?

Most of my efforts were motivated by desperation. I try to get some help from my colleague D.B. only to find out that for all his pretensions he doesn't know what he is doing. Then I talk to the "famous" Fred Lord. I was advised that Fred Lord was a genius. He knew everything. So I talk to him and discover that he doesn't know the first things. He talks all kinds of gibberish. He drops all kinds of big mathematical words. But when we get down to the facts of life he turns out to be completely incompetent. That made me angry and, for me, it made it impossible not to act. I had to do something about this shocking mess, ill-equipped as I might be. So I took a chance. I said, "OK. This can't be right. These guys don't know what they're doing. They're lying as well as pretending and I've got to do something about it. So I plunge into the problem with my crude hammer and my dull saw and chop away and it seems to work. But I'm really crude, dreadfully crude, in my mathematics.

David: But there may be an advantage in that. As a result of not being distracted by fancy mathematics and sophisticated equations you are forced to be practical. You are forced to go to the practical because you feel that you can't convince yourself or the world on purely mathematical grounds. In the end you are led into and encouraged by the practical side of things.

Ben: Not led into, but that's what I fall back on, common sense. But that's in desperation not by preference. It's a different kind of experience. I never felt adequately educated to do this work and I never really wanted to do it. It was never my chosen work and yet I spent thousands of hours on it and have been responsible for bringing it into lots of people's lives. Some of them do rather well with it, too. For example, I like the way you talk about it. I like the way Geoff talks about it. I like the way Bruce talked the other day. I'm pleased with the consequences. But in a certain fundamental way it was never what I wanted to do. Part of me feels like I was forced into it by circumstances and by my objection to the sophistry and corruption of those of my colleagues who stood for expertise on this topic.

David: It sounds like a delicate balance. It was probably necessary that you didn't know more mathematics or more statistics. Otherwise you might have pursued the topic with a vested interest in showing off your mathematical sophistication and spoiled your simple solutions. You might have seen through Georg's problem too quickly and so not appreciated its real significance.

Ben: I'm sure there is something in what you say. But, if you mean that, had I a better mathematics education, I might have agreed with Lord and Bock, surely not! I might have viewed the whole thing as not worth bothering with. But I certainly wouldn't have agreed with them, because the holes in their position were so obvious even to me.

Of course, then I might never have been discussing the problem with them at all, that's a possibility. I might never have looked to them for an answer. My initial expectation was that they would understand the problem and know what to do about it and that all I had to do was to bring it to their attention and they would go off and do it right. But instead of doing that they pooh-pooh-ed my observations and said, "No. We have this other thing which is the right way to do it." So I said, "Yes. But I know from my own investigations that your other thing doesn't work." And they said, "Well, don't worry about that. We'll fix it up in no time." I said, "No. You're not going to be able to fix it." and they said, "Oh yes. You can bet we're going to fix it." That got me mad and I became personally involved.

****

David: I hate to put words into your mouth. But what about elegance, simplicity, truth?

Ben: I'm suspicious of "elegance". Elegance is so often merely sophistry, affectation, pretension pursued by people who are shallow and empty. Anyone can dress up. Darrell sticks feathers in his cap and wears London suits. But he's empty. He doesn't solve important problems. He doesn't educate students. You like the word "elegant", but the way you use it is peculiar to you. For many people it is a shallow word.

"Simplicity" is what I like. What I mean by "Truth" is "Utility". I have an appetite for simplicity and utility. I love things to be simple. I love them to work. When something works and is simple, it gives me a thrill. The most exciting thing that happened yesterday morning (at IOMW) was the last part of your paper when you brought that bevy of little relationships into one simple formulation. You said, "Look at this. When we formulate things this way, everything becomes simple." Suddenly all the ideas which those various relationships represented became understandable. Not that I didn't understand them before. But my way of understanding them was burdensome, tiring to work with. After you presented your reformulation, I understood the relationships in a much less expensive way.

****

David: For me "elegance" borders on "simplicity". It means "simple but not trivial".

Ben: I see what you mean. For me simple is beautiful because it implies as few things as possible, with as much reach as possible. That's what I like. The reason I have an allergy to "powerful" and "elegant" is because nincompoops like Ron Hambleton throw those words around to pretend they are saying something important. So many of the people who use those words are phonies that every time I hear them red lights go on in my head. I have to understand your way of using them to appreciate what you mean when you use them. If, when you say them, I think of Ron Hambleton, I have already misunderstood you. You are not talking about what he's talking about.

David: A great deal of the use of mathematics in social science is exactly of that affected form. Mathematics can be powerful, elegant and simple. But if you just put it in there to make an impression of sophistication, you actually confuse things.

Ben: It's not science or scholarship. They're just diddling, a kind of cognitive masturbation. For me the antidote is, "Can I explain it to my grandmother?" If I can explain it to my grandmother, then I understand what I am talking about.

That was what appealed to me about Georg. His model was simple. His first lectures were simple. He started with a Poisson model for reading errors. Then he went on to pairwise estimation. Those were the lectures I liked the best. The misreadings study is in Chapter 2 of his book. Pairwise estimation does not come until page 171. But it was one of his favorite lectures. Mine too.

David: The same simplicity that appealed to you may itself be the reason a lot of other people don't like his model. They think it's too simple. It doesn't appear to be sophisticated enough to justify their reputations as experts.

Ben: Exactly. My nasty explanation of this is that they're looking for something with which to mystify their colleagues. Georg's model is so simple that it won't serve that purpose. So they can't be interested in it. They want something mysterious and complicated. They want models with lots of parameters. They want methods that take lots of time on giant computers. They want computer programs that produce voluminous output. The more computer time the better because that's the hallmark of their stature, the currency they prize. Using lots of computer time, having lots of output, having lots of parameters and doing something that nobody can understand is their heart's desire.

Well those are just the miserable, deceitful things that I am allergic to. I don't like any of that stuff. When people do that to me, it makes me angry. I put a great deal of personal value into making things simple, first for myself and then for anyone else who's interested. My ideal is to take a classroom teacher and show her how she can use the Rasch model in the everyday organization and implementation of her teaching so that it makes her professional life easier and better. That may be unreasonable, but that's what I dream about.

We all have ideals, schemes which may be unrealistic, but nevertheless guide us, our guiding stars. My guiding star is that every school teacher shall understand what we call the Rasch model and be able to use it naturally and easily to keep their classroom clean, to keep their measurements of what their students are doing, simple, orderly, relevant and immediately useful and meaningful. That's why KIDMAPs are particularly important.

A KIDMAP is a graphical presentation of a Rasch analysis in which you make a picture of a child's performance, a picture which "shows" (rather than "tells") all of the information. The child's measure is shown as a vertical position on the measurement ruler. It is rendered in a box, the vertical width of which depicts the error in the measure. Among the 4 quadrants implied by the child's measure level, two should be full of items showing, in the lower left, the difficult levels and content of the items the child succeeded on because they were easy for him and, in the upper right, the levels and content of the items the child has failed because he has yet to learn them. The opposing two quadrants, because of their increasing improbability, should be mostly empty. Any items appearing there, however, have great diagnostic potential with respect to understanding the child as a particular individual. The teacher does not have to understand numbers to see what the child has done and what the child is ready to do. Indeed, my conviction is that any result that cannot be expressed that way is probably not going to be worth knowing.

It took me a great deal of effort and ingenuity to get results expressible in that form and I admire that kind of ingenuity. For me the hallmark of a good outcome is that the results are so obvious and simple that the newcomer cannot see what all the fuss was about. When someone says to me, "Well, that's obvious. I'm glad you told me. But actually I knew it already." and "I can't see what you guys have been working so hard on for 10 years, if this is all you came up with." I might wish they would appreciate my labors. But when I hear those kinds of words, I cannot help but feel rewarded. That's the reward I'm looking for, and not, "Boy, is this complicated" or "Wow, have you mystified us!" When people say that, I get nervous. I realize they don't know what I'm talking about, that I've failed them. That's not what I want. I want things to be simple and obvious and I want the hard work that went into them to gradually disappear. I think the progress of civilization is based on making complicated things simple, and not on making simple things complicated.

That is what Georg was doing, that was the rapport between Georg and me from the beginning. I recognized in him an older and more experienced version of something that I longed for in myself, the determination to make things simple. Georg loved to solve problems and so do I. We both love straight lines. The beauty of a straight line is a good example of simplicity. A straight line is a great achievement. A constant is OK, but a little boring. A straight line, however, is a variable under control and that's the best you can ever do.

There was another aspect about Georg, besides simplicity, which appealed to me. Georg relished independence, not only statistical but also personal. One of the great griefs of Georg's life was his father's cruelty to his mother. His mother would beg his father to be interested in her and he would deny her. She bore his children and that was all she was supposed to do. There would be no further sex between them. She knelt before him and wept and his father turned her away.

Every time Georg told me that story, he cried. I think his father's strength and that his father was cruel and needed to be defied, were the foundations of Georg's strength and defiance. He both identified with his father and had a cause. My own father was a bit like that.

The other thing was that Georg loved life, loved to eat, loved to drink, loved to joke. I like that too. I really enjoyed the lunches we had. All through the sixties, whenever we were together, we would go out to lunch, or supper, and eat and drink and drink and eat. He was never stingy or unresponsive, if you wanted to buy him a good meal, he would eat the best meal possible and whenever he took you out, he always bought the best meal at the best restaurants. I enjoyed his generosity and willingness to live. His joy in life nourished my spirit.

I saw that the first time I had lunch with him at his CTS apartment in 1960. He got out his cans of sardines, his brown bread, his pepper and his beer. He opened the sardines, put them on the bread, mashed them a little, poured on some oil and added lots of pepper. He enjoyed it all so much. He even enjoyed opening the can. He was really into it. His pleasure in something as simple as a sardine sandwich was an inspiration to me. I thought, "That's the way life should be. I like this man and the way he does things. I want to be like him."

The people I had known in my early life either tended to be dry, puritanical, austere, reserved and aloof, or they were childish, immature and simple minded. I didn't want to be like any of them. I didn't like the physicists I knew. These physicists were immature and lived uninteresting lives of khaki and plywood as though in an army camp. Their apartments were devoid of art, just somewhere to sleep. There was no culture or anything interesting in their lives. This terrified me. Georg's life was full of things. His house was full of pictures. His mind was full of ideas. His stomach was full of food. He was an example of living which was good for me. His example warmed me, nourished me in a way that I longed for and that I prized.

Georg had all kinds of interests. All kinds of contacts. All kinds of friends. He wasn't much concerned with ceremony. Not that he couldn't be ceremonious, but he did not feel that somebody was necessarily important because they had a title, whereas somebody else who didn't wasn't. He was just as eager to talk to a teenager who seemed interested, as to a university President or the Queen of Denmark. In fact, he was more eager to talk with a teenager who would be interested, than with the Queen who might not. I always liked that about him. The average person is much influenced by the station of the audience. The more titles, the more hats, the more crowns, the more the average person gets excited about, the more important he thinks the conversation. But Georg wasn't like that at all. Georg was not a stuffed shirt.

To return to how I never wanted to do all this, never felt I did it justice, always felt guilty that I wasn't doing it better, spending more time on it, I know now that it may be the best, the most important thing that I will ever have done. I didn't do it as strenuously as other things that I did. I didn't study it as much and I'm often dubious about the value of what I've done. When I meet fools, I fight against them. But, when I'm not struggling with people whose stupidity they have forced on me, when I'm in a quiet situation, then I might say, "I don't know whether this is really valuable or not." I have to think it is when I do it. But, if you ask me, "Is this really the only answer, is this going to last?" I don't know. I do know that I haven't done it as well as it could be done. I feel sad about that. But I'm unlikely to mend it.

****

Ben: David, why are you doing this interview?

David: I got interested in this side of you when I got my own papers rejected. I was stunned at the reactions of the people who were behaving that way. I was far away from the American controversies in 1977 when I turned up at the Conference with a couple of rejected papers, feeling as though I had been punched around the head by someone. Then I found out that in New York, Fred Lord refused to appear in a session, if you were going to speak too. But you came anyway in the audience, choosing to appear on your own. I began realizing that in fact there was something to this.

Ben: When I stood up in '77 in New York to talk against Fred, I wasn't talking to Fred, you know. I was concerned about the people in the audience who had nothing to ally themselves with and I wanted to give them an antidote to his nonsense. I wanted to give them a rallying point, something that they could, if they wished, hang onto which would cure them from any mental diseases that they might catch from this sophistry of nonsense that was going on around them. That was why I stressed the practical. "You can do it yourself. Even if you don't know what these guys are talking about, that's OK, because you can do your own experiments to convince yourself of whether or not these goofy ideas can be useful to you. Be careful how much you believe this man. First, he claims I misquoted him. Then he misquotes my colleague Haberman. He is not even a master of quotation. He misuses a quote in his own defense which, in fact, contradicts his position. And he doesn't know the difference. I say, 'Don't be too smart for your own good, Fred. You quote Haberman to support your position when, in fact he investigated the Rasch model, not the Birnbaum and in so doing shows that the Birnbaum model does NOT have sufficient statistics or asymptotic consistency or, as a result, the possibility of estimation convergence.'"

Rasch SIG Update from Chair

Greetings Rasch SIG colleagues,

I wanted to provide a brief update on SIG activities as we meander through the holiday season. First and foremost are the upcoming elections. By now you are likely well aware of this, as I had sent out appeals to the membership seeking nominees. This set of elections is based on our new by-laws in which we now move to having three elected posts: Chair, Secretary, and Treasurer. (Note that we are still waiting for final confirmation that our by-laws have been formally approved by AERA, but are directed to operate as if they are fully implemented.) We will have final nominations as of December 15th, and please do look out for postings from AERA as far as voting in the election. Additionally we continue to have a need for someone to volunteer to serve as our webmaster. If you are skilled in this area and willing to help out, please do reach out to us. This entails occasional updates to the website and posting of RMT articles. Not an enormous commitment, but one that would be very much appreciated.

I wanted to acknowledge the ongoing efforts of our current SIG officers and appointees. Kelly Bradley and Jessica Cunningham have served as Program Co-Chairs this year and I am looking forward to sharing the results of their efforts as we move closer to the 2014 AERA annual conference. Kirk Becker has consistently availed himself during his tenure as Secretary/Treasurer and his commitment to the SIG has been appreciated. Lastly, I continue to be impressed with RMT and the efforts of Ken Royal in this regard. Thank you all!

Beyond the holidays we are looking forward to Philadelphia for both AERA and the International Objective Measurement Workshop (IOMW). For our SIG business meeting, I am very happy to report that David Andrich has committed to provide this year's talk. Please make plans to attend this meeting.

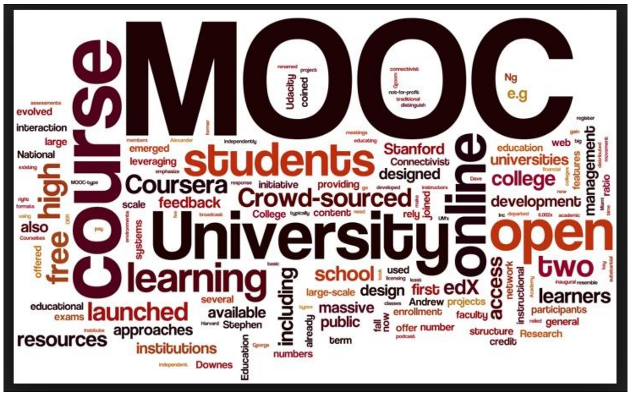
Kind Regards,

Tim O'Neil
Rasch SIG Chair

Is now the time for a Rasch measurement MOOC?

Confusion about Rasch measurement continues to run deep in academia. Similarly, confusion about objective measurement in the social and behavioral sciences is prevalent among non-psychometricians, as the average person assumes latent traits cannot be measured with the same type of rigor as physical measurements. At present, information about Rasch measurement is primarily proliferated in a limited number of graduate programs at a limited number of colleges and universities and by leaders who have conducted online courses and physical workshops throughout the world. The success of these efforts is undeniable as interest in Rasch measurement has never been greater than it is today. However, a relatively new and extremely popular trend in higher education has been the use of MOOCs, which offers an opportunity to expand educational opportunities and reach learners throughout the world in a new way. The question is should Rasch measurement enthusiasts explore the use of MOOCs to expose the fundamentals of objective measurement to an entirely new, and broad, audience of learners? We think so.

So, what is a MOOC? In short, it is a "massive open online course" available free of charge to anyone, anywhere. MOOCs are similar to online university courses in most every aspect, except they currently do not offer academic credit. Numerous top universities and colleges such as Harvard, Stanford, MIT and a host of others are now offering MOOCs to worldwide audiences. To date, more than six million students have already completed courses in this format and numerous, powerful stories about how lives have literally been transformed due to this educational platform have been voiced in student reviews (Fowler, 2013). Many instructors are also very excited about MOOCs as they can now reach an enormous audience of learners. A 2013 Huffington Post article cites a philosophy course co-taught by Duke and University of North Carolina professors that had more than 180,000 students enroll. One of the professors cited 26,000 people were later classified as inactive and 70,000 people never watched a single online video. Despite the attrition, the professor noted he would be lucky to teach 8,000 students over the span of a 40-year career if he averaged 200 students per year. One MOOC has already allowed him the opportunity to teach more than 20 times the number of students he would expect to teach over the course of his entire career.



Rasch measurement history is full of stories about talented people from diverse backgrounds discovering the utility and potential of Rasch models. Many of these individuals experience a dramatic change in perspective and go on to disseminate Rasch measurement theory and applications in their respective "home" disciplines by illustrating new ways in which various types of data can be understood and made meaningful. Imagine the possibilities when a worldwide audience of learners from all walks of life obtains unobstructed access to the fundamentals of objective measurement. Surowiecki's (2004) book on *The Wisdom of Crowds* suggests there could be great promise for a Rasch MOOC as social science research has repeatedly shown that large groups of people are smarter than any handful of elite individuals. Thus, proper exposure to Rasch measurement through a MOOC could be a "game-changer" with regard to how the world comes to understand measurement in the social and behavioral sciences. What's more, it seems a multitude of new ideas and new applications for Rasch measurement could be just around the corner waiting to be explored.

In a quick survey of current MOOC offerings, there were more than 40 courses devoted to statistics and data analysis spanning a variety of disciplines and languages. The world is already tuning in to learn more about data analysis and how to make sense of data. Now would be an optimal time to bring the fundamentals of objective measurement to the forefront and expose thousands of new learners to psychometric science.

Reference

Duke University philosophy professor Walter Sinnott-Armstrong has 180,000 students in MOOC. Available at: http://www.huffingtonpost.com/2013/01/22/duke-university-mooc_n_2525547.html. Retrieved December 2, 2013.

Fowler, G. A. (2013, October 8). An early report card on massive open online courses. *Wall Street Journal*. Retrieved from http://online.wsj.com/news/articles/SB10001424052702303759604579093400834738972.

Surowiecki, J. (2004). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations*. New York: Doubleday.

Kenneth D. Royal, *University of North Carolina at Chapel Hill*
Melanie Lybarger, *Independent Consultant*

Rasch-related Coming Events

Dec. 11-13, 2013, Wed.-Fri. In-person workshop: Introductory Rasch (A. Tennant, RUMM), Leeds, UK,

Jan. 3-31, 2014, Fri.-Fri. Online workshop: Practical Rasch Measurement – Core Topics (E. Smith, Winsteps), www.statistics.com

Jan. 15, 2014, Wed. Submission deadline: IOMW Biennial Meeting, Philadelphia, PA, www.iomw.org

Mar. 12-14, 2014, Wed.-Fri. In-person workshop: Introductory Rasch (A. Tennant, RUMM), Leeds, UK,

Mar. 31-Apr. 1, 2014, Mon.-Wed. IOMW Biennial Meeting. Philadelphia, PA, www.iomw.org

Apr. 3-7, 2014, Thurs.-Mon. AERA Annual Meeting, Philadelphia, PA, www.aera.net

May. 14-16, 2014, Wed.-Fri. In-person workshop: Introductory Rasch (A. Tennant, RUMM), Leeds, UK,

May. 14-16, 2014, Wed.-Fri. In-person workshop: Intermediate Rasch (A. Tennant, RUMM), Leeds, UK,

July 4-Aug. 1, 2014, Sat.-Thurs. Online workshop: Practical Rasch Measurement – Further Topics (E. Smith, Winsteps), www.statistics.com

Call for Submissions

Research notes, news, commentaries, tutorials and other submissions in line with *RMT*'s mission are welcome for publication consideration. All submissions need to be short and concise (approximately 400 words with a table, or 500 words without a table or graphic). The next issue of *RMT* is targeted for March 1, 2014, so please make your submission by Feb. 1, 2014 for full consideration. Please email Editor\at/Rasch.org with your submissions and/or ideas for future content.