Autumn 2017, VOL. 31, NO. 2; ISSN 1051-0796

RMT

RASCH MEASUREMENT TRANSACTIONS

Special Topic Issue: Teaching Rasch Measurement

Notes:

Transactions of the Rasch Measurement SIG
American Educational Research Association

Teaching Rasch Measurement

Here are the seven pillars of wisdom that are fundamental to my teaching of Rasch measurement:

1. Think like a Physicist, not like a Social Scientist.

Physicists know that they have control of the data. They do not let the data control their thinking. They keep helpful data and reject unhelpful data. It is only when they cannot construct helpful data that they amend their theory. Rasch is a theory of "moreness" and "lessness". This is expressed as measures on an equal-interval straight line, so that everything can be compared with everything else consistently. These measures are the type that physicists, carpenters and cooks rely on. If we cannot obtain helpful data, then we can no longer generalize "moreness" and "lessness" across everything, but are reduced to making local comparisons, perhaps no better than between individual objects.

2. Rasch measures are unidimensional and additive.

Rasch measurement is prescriptive like physical measurement, not descriptive like most Social Science statistical techniques. Descriptive techniques fit statistical models to the data. The best model is the one that fits the data the most efficiently. In contrast, Rasch extracts from inevitably unruly data the consensus dimension, and then estimates additive, interval, linear measures along that one dimension. Aspects of the data that contradict the dimension are reported to the analyst. Often little action is required because, as philosopher-of-science Larry Laudan wrote, "Empirical problems are frequently solved because, for problem solving purposes, we do not require an exact, but only an approximate, resemblance between theoretical results and experimental ones." (Laudan, 1977, p. 224).

3. Rasch measurement isn't complicated.

A common perception: Rasch involves lots of numbers. Numbers are complicated, so Rasch must be complicated. Solution: Start with pictures of variables, not columns of numbers. Rescale logits to avoid negative numbers and decimals. For fit statistics, use mean-squares (size of misfit = amount of disagreement between data and Rasch) rather than t-statistics (probability of fit). Here is a quick summary for mean-squares:

Mean-square Value	**Implication for Measurement**
> 2.0	Distorts or degrades the measurement system. Can be caused by only one or two observations.
1.5 - 2.0	Unproductive for construction of measurement, but not degrading.
0.5 - 1.5	Productive for measurement.
< 0.5	Less productive for measurement, but not degrading. May produce misleadingly high reliability and separation coefficients.

Note: Since mean-squares are constrained to average near 1.0, always investigate high mean-squares before low ones. Remedying high mean-squares often brings low mean-squares into the productive range.

4. The important validities are "Construct Validity" followed by "Predictive Validity". "Reliability" is a distant third.

Construct Validity – before analyzing the data, place the items in difficulty order according to the construct theory that underpinned the writing of the items. After analysis, compare the Rasch difficulty order with the intended order. If they agree, then you have construct validity. If they disagree, either there are items that do not represent the construct theory or the construct theory needs revision. A discovery! Similarly for Predictive Validity. When we look at the distribution of the Rasch measures for the persons, do the higher measures correspond to those we expect to have more of the construct? If we have Construct Validity and Predictive Validity then low person "test" reliability indicates that we need more items, or a person sample with a wider spread on the construct, or both. Rasch also reports item reliability. If this is low, then a bigger sample of persons is needed.

5. Remember your audience.

Rasch is a tool. If reviewers demand a plethora of numbers, then Rasch can produce the numbers to give them. For most audiences, a few pictures (with simple explanations) are enough, and much more memorable! Rasch computations can produce innumerable decimal places, but Rasch estimates, fit statistics and the like are rarely meaningful beyond one decimal place. If the numbers do not produce separate points when plotted, then the numbers are not meaningfully different! Do not shower your audience with "decimal dust".

6. Be ready with your ammunition!

Be ready with answers to questions/criticisms from those with different perspectives on measurement.

For instance, "the Rasch measures only explain 30% of the variance in the data" – This is expected. See "PCA: Variance in Data Explained by Rasch Measures" - www.rasch.org/rmt/rmt221j.htm

"Why didn't you delete more/less persons/items from the data?" – Usually this makes little difference, but see "When to stop removing items and persons in Rasch analysis?" - www.rasch.org/rmt/rmt234g.htm

"Didn't S. S. Stevens write that measurement is merely the assignment of numbers? Why bother with Rasch?" – Of course, Stevens was much more detailed and precise. In Stevens terminology, Rasch produces an interval scale, which can also be a ratio scale if our construct theory enables us to define a substantive zero. See "Stevens Revisited" - www.rasch.org/rmt/rmt111n.htm and "Absolute Zeroes for Reading and Mathematics Abilities" - www.rasch.org/rmt/rmt64a.htm

"Rasch is too complicated/too simple" – the Rasch model has the correct amount of complexity to construct linear measures from ordinal observations – see the seven research notes with titles starting "The Rasch model derived from …." in www.rasch.org/rmt/contents.htm

And so on.

7. Have fun and make discoveries.

Think of a new dataset as an adventure into parts unknown. We expect that most of a Rasch analysis will verge on the routine, but we also expect to gain new insights into the items, the persons, the construct and the world. We are always learning!

John Michael Linacre

Larry Laudan, Progress and its Problems. Berkeley, CA: University of California Press, 1977, p. 224

Journal of Applied Measurement
Vol. 18, No. 2, 2017

Richard Smith, Editor, www.jampress.org

Reflections on Teaching Rasch Measurement Theory

We are excited about the opportunity to respond to Ken Royal's invitation to share our thoughts on teaching Rasch measurement theory (RMT). In the Spring of 2017, we co-taught a course on RMT at The University of Georgia. Based on our teaching philosophy and experiences, we provide our thoughts on the following questions:

- Why do we want to teach RMT?
- Why should graduate students and researchers study RMT?
- What are the key principles that we want students to learn?
- What instructional approach and material do we find useful?

Why do we want to teach RMT?

George: One of the main reasons why I teach RMT is to show students the key role that measurement plays as a critical aspect of methodology in the social, behavioral, health, and economic sciences. I want the students to see the clarity that RMT brings to the apparent complexity of psychometrics. My experience is that RMT provides an approach to measurement that makes sense and that is accessible to emergent scholars. RMT also provides clear instructions for creating latent variables in the substantive theories that guide our research activities. In summary, RMT provides a useful way to introduce students to modern measurement theory as reflected in Rasch's paradigm shift (van der Linden, 2016).

Jue: RMT is a fundamental and modern test theory that includes the key concepts of educational and psychological measurement. My motivation for teaching RMT is to let students get a taste of measurement through the window of RMT. On one hand, Rasch's specific objectivity defines the concept of invariant measurement that is a critical aspect for obtaining the same measures with the use of different instruments. On the other hand, the Wright map provides a yardstick that displays the underlying latent trait on a line. This is a very straightforward and meaningful representation for students. Last but not least, RMT has various applications, such as measuring writing proficiency, musical performance, and student attitudes toward the school environment.

Why should graduate students and researchers study RMT?

George: One of my instructional goals is to motivate students to integrate RMT into their program of research. It is very important to maintain authenticity for students. In essence, I want students to understand how RMT represents a paradigm shift, and to consider its use for solving measurement issues in their own research. Graduate students and researchers should learn RMT because it provides the opportunity to improve measurement in their own field.

Jue: There are three reasons that students should study RMT. First, it reflects a paradigm shift from classical measurement techniques to modern measurement approaches. It provides the opportunity for students to learn about methods developed from the scaling tradition in measurement theory, and to compare these to the methods based on the test-score tradition. Second, students learn a useful methodology and programming skills for solving practical problems, such as test equating/linking, detection of differential item functioning, and score reporting. RMT provides a transparent approach for understanding measurement theory and solving various practical measurement problems.

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

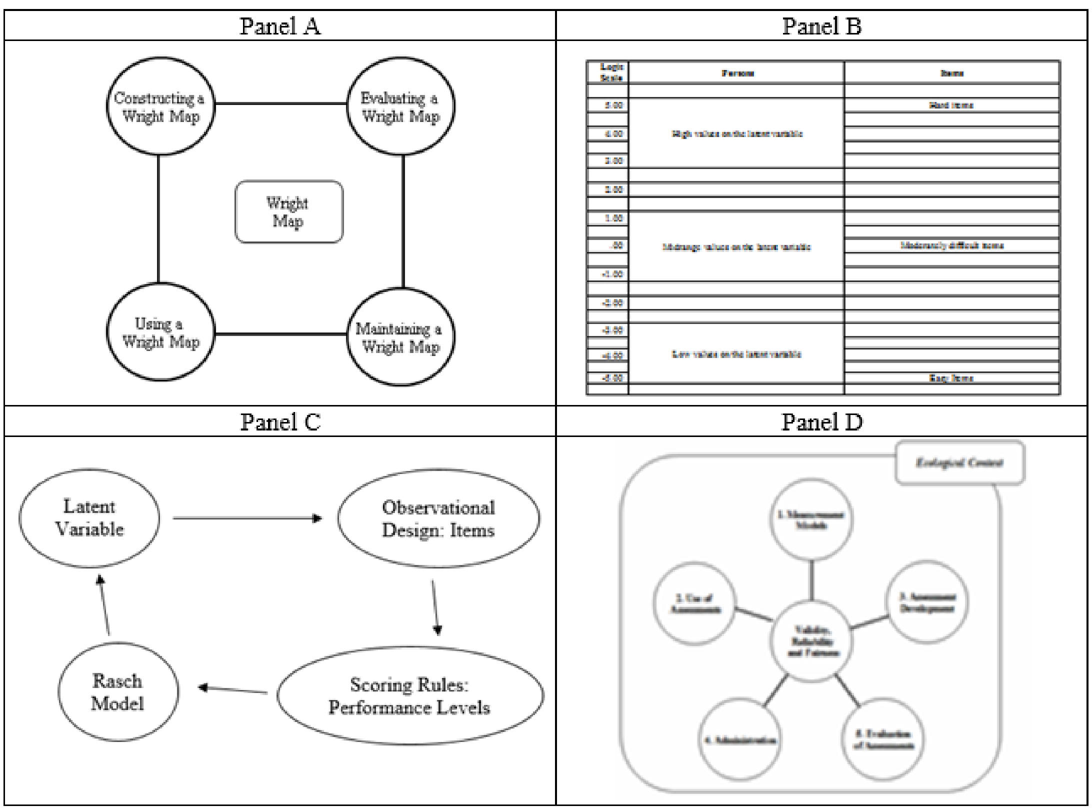
What are the key principles that we want students to learn?

George: The most important principle that I want students to learn is the understanding and interpretation of Wright maps. We have adapted Wilson's constructing measures framework in our classes (Wilson, 2005). The constructing measures framework provides a way of thinking about the plethora of measurement issues that arise in applied measurement area. We also stress the important distinctions between fitting models to data and fitting data to models with an emphasis on the differences between ideal (theory-based) and applied measurement issues.

Jue: Among all features of RMT, I would like students to know the principles of invariant measurement. Engelhard (2013) summarized five requirements for invariant measurement: (a) the measurement of persons is independent of particular items, (b) persons with higher proficiency should always have higher probability of success than those with lower proficiency, (c) the measurement of items is independent of different person groups, (d) a person should always have higher probability of success in answering an easier item than a more difficult one, and (e) the correspondence of item locations and person locations can be mapped along a continuous scale (Wright map). These principles can also be extended to a many-facet Rasch model and rater-mediated assessments.

What instructional approach and material do we find useful?

George: My approach to instruction is project-based with students working on measurement issues that have direct implications for their program of research. This approach provides a level of authenticity that is highly motivating for students. We also use graphical displays to help students visualize the inter-connections between various ideas. Figure 1 lists some of the displays that we have found to be most useful for students. First, Panel A introduces students to the general structure of our course with the Wright Map as the central concept surrounded by a variety of measurement issues related to constructing, evaluation, maintaining and using a Wright Map. Second, Panel B shows a template that we use for



constructing a Wright Map. Next, Panel C is adapted from Wilson (2005), and shows the essential building blocks for constructing measures for a latent trait. Finally, Panel D combines Hattie et al. (1999) with the three foundational areas (reliability, validity and fairness) in the *Test Standards* (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (2014).

Jue: I believe students can learn better when they are more involved in the class. First, we assign key articles for students to read and present to the class. This requires students to think about the content, and to express the major ideas in their own words. Students get very involved in the class activities, and this helps to form a supportive community for collaborative learning. Second, we embrace classroom discussions. In the middle of the semester, we ask students to find a practical problem related to their own research area (e.g., evaluate student performance in a classroom-based oral language test), and to use what they are learning in class to create their own measurement tools. While students are showing their designs, we have great discussions with every student providing input. These instructional approaches can be used to not only monitor student learning progress, but also motivate students to think deeply about their own learning of RMT.

Summary

In summary, our goal is to be invitational, instructional, and supportive for students to learn about RMT. A variety of approaches have been applied in our class to encourage students using RMT as a guidance for solving measurement issues in their program of research. We are currently working on a book that will embody many of our ideas that were briefly mentioned in this note.

George Engelhard, Jr. & Jue Wang
The University of Georgia

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.

Engelhard, G. (2013). *Invariant measurement: Using Rasch models in the social, behavioral, and health sciences*. New York: Routledge.

Hattie, J., Jaeger, R. M., & Bond, L. (1999). Persistent methodological questions in educational testing. In A. Iran-Nejad & P.D. Pearson (Eds.), *Review of Research in Education, 24*, 393–446.Thousand Oaks, CA: Sage.

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

van der Linden, W. J. (Ed.). (2016). *Handbook of item response theory*. CRC Press.

Wilson, M. (2005). Constructing measures: An item response modeling approach (2nd edition). Mahwah, NJ: Erlbaum.

Teaching and Learning Rasch with Google Scholar Alerts

The Rasch classes/workshops that I lead are mostly populated with Rasch learners who are not going to earn a Ph.D. in Psychometrics, rather they wish to apply Rasch to their work and research in fields such as medicine, business, and social science research.

I have found that using Google Scholar Alerts has enabled me to more easily hook students into Rasch and has immensely helped their learning!

Before a class/workshop starts I ask my students to set up an alert using the key words “Rasch” and “Winsteps” (the software I use). This allows students to receive alerts as to Rasch articles, using Winsteps, which have just been published.

Each day/week they will receive an alert with a list of articles!

I find the students are excited to see recent articles which they view as hot off the press. Also by selecting Winsteps in the alert, my students are often able to see article tables and figures which are organized in a similar fashion as the tables and figures from Winsteps which we discuss each day. Another bonus of specifying the specific Rasch program in the Google Scholar Alert is that the statistics and nomenclature will be the same as that which is covered in class!

With each alert---I personally enjoy seeing the range of topics in which Rasch analysis is used. To set up the alert just type in "Google Scholar Alert" in your search engine and you will see how to set up the alert!

William Boone, Miami University

Teaching Rasch through the Development of Conference Proposals and Presentations

As a faculty member, I often work with students who are interested in developing and validating instruments that are designed to measure educational and psychosocial constructs. I also teach a required course in which we focus primarily on the Rasch family of models. During that class, many of my students develop a strong interest in using Rasch in their subsequent research. In order to assist them in gaining practical skills, I have developed a major course project that is designed to facilitate their interests in Rasch and to complement their graduate training. The end product includes a conference proposal, a poster, and a full paper presentation. In essence, the project is a simulation of the proposal and presentation process at a major research conference.

Early in the semester, students begin the project by submitting a short, ungraded mini-proposal that includes a rationale, research question(s), description of the data they intend to use, and preliminary plans for analysis. The data are real and typically come from a project on which they are working with their major professor or from a publicly available data set. Once approved, the students use the proposal as a roadmap to complete the project.

In order to complete the conference proposal portion of the project, the students are instructed to follow the most recent submission guidelines from either the American Educational Research Association or their chosen professional association. In class, we discuss how to adhere to the association's requirements and how to present the proposal in a document that has a relatively short word limit. One of the areas in which we spend a good deal of time is learning to interpret the various statistics and visual outputs from Winsteps, then deciding which findings make for the most compelling and thorough proposal. In my experience, I have found that students tend to take an "everything but the kitchen sink" approach, which gives us the opportunity to discuss the editing and revision process so only the most salient and impactful findings are included.

In preparing the posters and papers, we follow a similar process in which we discuss ways to present findings in a way that is clear, concise, and impactful while providing a full explanation of the project. Students are then expected to produce a poster that is suitable for a research conference, and to write the final paper at a level of scholarship expected of those in the field. In the end, the process is designed so that the student can take their projects from the class and, with minimal editing, present them in a formal scholarly setting.

As the course ends, I offer the students a chance to have their proposal reviewed by some of my friends in the Rasch community. I use a double-blind review process which gives the participating students a different level of insight they would not normally receive. Although this is not a required component of the course, the students who have taken advantage of the blind reviews have found them to be very helpful.

I am happy to report that, as of this writing, every student who has opted to submit their course project as a proposal to a research conference has had the paper or poster accepted for presentation.

R. Trent Haines, Morgan State University

Setting the Tone for Teaching about Measurement

For students developing their skills in quantitative methods the first stop usually is training in statistics. In these courses, students typically are taught a variety of techniques to conceptualize relationships between variables for group-level data (e.g., samples, populations). Measurement, however, largely focuses on relationships within variables and person-level data. So, it is easy to see how such a significant shift in what is emphasized in the inquiry process can result in some initial struggles for students. In order to help students transition from a ‘statistical mindset’, I like to engage students with a topic that likely will be both interesting and relevant to them.

Most courses that I have taught involving Rasch measurement include graduate students in their early 20s to mid-30s. I like to begin by polling students about their musical interests, namely the types of music they like, if they play any musical instruments, and how they would rate their current skill level. Historically, most all of my students enjoy music and most have some experience attempting to play an instrument, usually a guitar or piano.

Given the guitar is the most frequent response I receive, I begin by asking students how many famous musicians they can recall that play lead guitar and sing at the same time? Usually, students can quickly name a few (e.g., Eric Clapton, Brad Paisley, Keith Urban, Jimi Hendrix, etc.), but then abruptly run out of names. I then ask students if anyone has personally tried to sing and play at the same time and how that experience turned out for them. After some initial banter, most agree that they can strum basic chords and sing (although not necessarily well) at the same time. However, virtually everyone agrees that playing complicated scales or sophisticated phrases (‘licks’) in perfect time while also singing in key and hitting each note as intended is exponentially more challenging. This is why so many lead guitar players, many of whom also can sing very well, simply stick to playing guitar only. The ability to both sing and play excellently at the same time is incredibly difficult for most people, including many (if not most) professional musicians.

The musical example involves functions from different parts of the brain and tends to illustrate the challenges associated with doing multiple things well at the same time. More subtly, however, the example shifts the focus away from how well band members might perform with different instruments (group-level, between variables) to how well an individual can perform given his/her personal talents and abilities (person-level, within variable).

A byproduct of the example is that it also illustrates the importance of being able to separate phenomena so one can examine the various elements (e.g., guitar skills, singing ability, etc.) independent of one another. I then reference one of my favorite quotes from Linacre (1996) who stated: “Physical measurement takes great pains to measure one thing at a time. We don't want the patient's temperature reading to be biased by his weight, or height, or blood pressure. It is only when we have clearly isolated one dimension that we can understand the meaning of the measure, and then study how that measure relates to measures on other dimensions.”

I then ask students how we might measure constructs such as musical ability given its multidimensional nature. Although students typically do not have many answers at this point, it usually leads to a fun brainstorming exercise in which we begin to distinguish the requirements for objective measurement. For many, it is the first time they begin to truly think about measurement and how measurement is fundamentally different from statistics.

Kenneth D. Royal, North Carolina State University

References

Linacre, J. M. (1996). The Rasch model cannot be "disproved"! Rasch Measurement Transactions, 10(3), 512-514.

Lessons about Measurement from the Golden Years of Hollywood: "The stuff that dreams are made of"

Engelhard (2013), in his outstanding new book, *Invariant Measurement*, highlights the importance of the work of Louis Guttman in the development of modern measurement theory and approaches. He writes:

"Guttman and Rasch models are based on measurement philosophies that place the models above the vagaries of imperfect data...Underlying the measurement philosophies of both Guttman and Rasch is the key idea that the requirements of our measurement models are falsifiable." (Engelhard, 2013, pages 29 and 30)

"Guttman scaling is important because it lays out in a very obvious way many of the issues and requirements that are necessary for the development of a scale that meets the requirements of invariant measurement. Even though his requirements for a perfect scale are embedded within a deterministic framework, Andrich (1985) has shown how a probabilistic model based on Rasch measurement theory (Rasch 1960/80) can achieve these theory-based requirements. Andrich (1985) has also pointed out the close connections between Guttman's deterministic model and Rasch's probabilistic model." (Engelhard, 2013, p. 40)

Now, before I found the Rasch Fellowship, I was familiar with the work of Louis Guttman only in relation to the use of Guttman Scaling with the development and application of classic functional ability assessments - like Katz's Index of Activities of Daily Living, and Lawton & Brody's Physical Self Maintenance Scale and Instrumental Activities of Daily Living Scale (see Katz, et al., 1963; Lawton & Brody, 1969; McDowell, 2006; LaPlante, 2010; Streiner, Norman & Cairney, 2015). But where had I heard the name of Guttman before? Ah Yes! Gutman, aka "The Fat Man"!

Gutman was the name of one of the bad guys in the movie, The Maltese Falcon. The movie is a classic, black and white film starring Humphrey Bogart, Mary Astor, Sydney Greenstreet and Peter Lorre. This 1941 Warner Bros production, was written for the screen by its first time director, John Huston. Today, about 75 years after its initial release the film holds a legendary place in film history and ranked #31 in the American Film Institute's current list of top 100 "greatest American films of all time" (see http://afi.com/100Years/movies10.aspx).

The Maltese Falcon was based on the Sam Spade detective novel of the same name by Dashiell Hammett. British actor, Sydney Greenstreet played the sinister character, Kasper Gutman with a mixture of menace, relish and desperation. This was Greenstreet's first film role, at the age of 61. He gave a powerful performance as Gutman. Being a film noir detective drama, where you do not know which character, or which story to trust, the script for The Maltese Falcon (see Luhr, 1995) is peppered with many references to probability, chance, certainty, positivity, confidence, counting, answering questions, and passing tests. For example: "I must say you passed the test with flying colours." (p. 90), or "Six, two and, even, they're selling you out, sonny." (p. 85).

The link with the name Gutman started me thinking. If the films and novels of Sherlock Holmes can be used as an introduction to clinical diagnosis, deductive reasoning and hypothesis testing (see Rapezzi, Ferrari & Branzi, 2005; McCrory, 2006; Levine, 2012; Kempster & Lees 2013; Walker, 2013), can the films and novels about Sam Spade be used as an introduction to measurement? Could references from the movie be turned into teachable moments? So, with this in mind, I recently had the pleasure of re-watching The Maltese Falcon and I did indeed find many useful references to measurement concepts and applied psychometrics. Here are some actual quotes from the movie, they are only *slightly* out of context.... (See the movie if you can, and you be the judge.)

QUOTE 1: A statement outlining the relativistic / contextual nature of measurement.

```
BRIGID: ... What difference
   does it make?
CAIRO: It might make a world
   of difference.
```

"Let's imagine you are standing on the top of the mountain and I am standing at the foot of the mountain. Then someone says 'What is the height to the top of your head?' If we both measure our heights from the soles of our feet, then we are measuring 'Relative' to our feet (relative to item difficulty). Our heights will be about the same. If we both measure our heights from sea level, then we are measuring in an 'Absolute' way. This is 'relative to the latent trait'. Our heights will be considerably different." (Linacre, 2012, point 28)

"Physicists are unequivocal as to the quantitative length of an inch, but its substantive implication, its qualitative meaning, depends on the context: one inch added to the height of a mole-hill has different meaning than one inch added to the height of a mountain." Linacre & Wright, 1989, p. 54)

QUOTES 2 & 3: Demonstrate the essential starting point of counting valid observations or events when developing measures.

```
SPADE: ... Of course, you
   lied to us about your
   sister and all that, but
   that didn't count, we
   didn't believe you.
```

"Quantitative science begins with identifying conditions and events, qualities, which, when observed, are deemed worth counting." (Wright and Stone, 1999, p. 31)

```
SPADE: ... and because you've
   counted on that with me,
   the same as you counted on
   that with all the others.
```

"To count is to begin to measure." (Wright, 1999, p. 67)

QUOTES 4 & 5: We need to move away from a Guttman like deterministic approach to measurement to a probabilistic / stochastic one.

```
SPADE: ... Those are the
   chances we take.
```

```
SPADE: ... and you'll have to
   take your chances.
```

"In modem physics, beginning at the kinetic theory of gas, and ending up with atomic physics, the deterministic view has been abandoned. No deterministic description for e.g. radioactive emission seems within reach, but for the description of such irregularities the theory of probability has proved an extremely valuable tool." (Rasch 1960/1980, p. 10)

"To return to our attainment tests: We may give a problem to a pupil of whom we know that he could easily solve it, and yet he fails. Or we may give him a task which is much too difficult, and anyhow he solves it. We can never know with certainty how a pupil will react to a problem, but we may say whether he has a good or a poor chance of solving it."

"This way of speaking points to the possibility of mapping upon models of a kind different from those used in classical physics, more like models in modern physics - models which are indeterministic, where chance plays a decisive role: The possible behaviour of a pupil is described by means of a probability that he solves the task." (Rasch 1960/1980, p. 11)

"The Rasch model is a mathematical formulation linking the probability of the outcome when a single person attempts a simple item to the characteristics of the person and the item." (Choppin, 1983, p. 1)

“... item response model which links ability, difficulty and probability.” (Choppin, 1983, p 13)

“A useful way to understand the measurement logic defended by the Rasch model is to ask how we want to think about the relative performance of two persons. Do we want to think that the more able one has a better chance for success no matter what the difficulty of the attempted item? Is that what we intend “more able” to mean?” (Wright, 1977, p. 103)

“Any data with a probabilistic structure which accords with a Rasch model also accords with Campbell concatenation and so supports the estimation of measures which have the same arithmetical measurement properties as length and weight.” (Linacre, 2005, p. 4)

QUOTE 6: Dealing with uncertainty and indeterminism.

```
SPADE: Well, there’s nothing
   certain either way. It
   depends.
```

“Even if we know a person to be very capable, we cannot be sure that he will solve a certain difficult problem, nor even a much easier one. There is always a possibility that he fails - he may be tired or his attention is led astray, or some other excuse may be given. And a person of slight ability may hit upon the correct solution of a difficult problem. Furthermore, if the problem is neither ‘too easy’ nor ‘too difficult’ for a certain person the outcome is quite unpredictable. But we may in any case attempt to describe the situation by ascribing to every person a probability of solving each problem correctly, and this probability will be our indicator of ‘how easily’ the problem is solved. The probability that a very able person solves a very easy problem is near to unity, but not necessarily equal to 1, and the probability that a person of small ability solves a very difficult problem is very near to 0. In the last case mentioned, that of a problem which is neither too easy nor too difficult for the person, the probability may be in the neighbourhood of ½.” (Rasch 1960/1980, p. 73)

“Relativistic thinking begins with the realization that a measure is actually a relationship the student has with the test item. A test result is a probability rather than a fixed value. A relationship takes the place of an event (e.g., an answer to an item) in judging how a test item probably provides certain information about a student.” (Ingebo, 1997, p. 12)

QUOTE 7: Discussing the application of probabilistic measurement, from probable minimums to probable maximums, from zero to infinity.

```
GUTMAN: I will give you
   twenty-five thousand when
   you deliver the Falcon to
   me and another twenty-five
   thousand later on, or, I
   will give you one quarter
   of what I realize on the
   Falcon ...
GUTMAN: ... That would amount
   to a vastly greater sum.
SPADE: How much greater?
GUTMAN: Who knows? Shall we
   say a hundred thousand?
   Will you believe me if I
   name a sum that seems the
   probable minimum? What
   would you say to a quarter
   of a million?
SPADE: Then you think the
   dingus is worth a million,
   huh? Hum, that’s a lot of
   dough. The minimum, huh?
   What’s the maximum?
GUTMAN: The maximum, I refuse
   to guess.
GUTMAN: You’d think me crazy.
   I don’t know. There’s no
   telling how high it could
   go, sir. That is the one
   and only truth about it.
```

“What does this simple model say about the scale on which person ability and item easiness are measured? Odds vary from zero to infinity. Since this model gives the odds in favor of success as the product of person ability and item easiness, the natural scale on which to define ability and easiness is one that also varies between zero and infinity."

"What does that mean? When a person has no ability, his zero ability will give him zero odds in favor of success no matter what item he tries. With no ability he has no chance of succeeding. On the other hand, if an item has no easiness, then it is infinitely hard and no one can solve it. Measurements made on these scales of ability and easiness have a natural zero." (*Wright, 1967, p. 98)*

QUOTE 8: Highlights the nature and problem of guessing.

```
BRYAN: Who killed Thursby?
SPADE: I don't know.
BRYAN: Perhaps you don't, but
   you could make an
   excellent guess.
SPADE: My guess might be
   excellent, or it might be
   crummy, but Mrs Spade
   didn't raise any children
   dippy enough to make
   guesses in front of a
   district attorney, an
   assistant district
   attorney, and a
   stenographer.
```

"Guessing is a problem only when students inclined to guess are also provoked to guess on items that are too difficult for them and then only when those particular students happen to guess correctly ... Problems of guessing are best addressed by targeting test administration so that it does not provoke guessing ..." (Wright & Stone, 1999, p. 117)

From my analysis, there are at least 4 other major lines of dialogue which can be used for teaching purposes. These include references to: the state of modern opinion poll surveys; the power of item anchoring (OECD, 2009); finding good link items (Wilson, 2003; Ingebo, 1997; Claesgens, et al., 2013); and the balance scale metaphor (Massof, 2012).

Now before you start to think that I am getting a little carried away, I will conclude....

```
GUTMAN: Well, sir, the
   shortest farewells are the
   best. Adieu.
```

Nick Marosszeky, Macquarie University

References

Choppin, B. (1983). A Rasch model for item analysis. CSE Report No. 219. Los Angeles: Center for the Study of Evaluation, University of California.

Claesgens, J., Scalise, K., & Stacy, A. (2013). Mapping student understanding in chemistry: The perspectives of chemists. *Emergent Topics on Chemistry Education*, *24*, 407-415.

Engelhard, G., Jr. (2013). *Invariant measurement: Using Rasch models in the social, behavioral, and health sciences.* New York: Routledge.

Ingebo, G. S. (1997). *Probability in the measurement of achievement*. Chicago: MESA Press.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). Studies of illness in the aged - The Index of ADL: A standardized measure of biological and psychosocial function. *JAMA*, *185*, 914-919.

Kempster, P. A., & Lees, A. J. (2013). Neurology and detective writing. *Practical Neurology, 13*, 372-376.

LaPlante, M. P. (2010). The classic measure of disability in Activities of Daily Living is biased by age but an expanded IADL/ADL measure is not. *Journal of Gerontology: Social Sciences*, *65B*, 720-732.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and Instrumental Activities of Daily Living. *The Gerontologist*, *9*, 179-186.

Levine, D. (2012). Revalidating Sherlock Holmes for a role in medical education. *Clinical Medicine*, *12*, 146-149.

Linacre, J. M. (2005). *Measurement, meaning and morality*. Paper presented at the Pacific Rim Objective Measurement Symposium (PROMS) & International Symposium on Measurement & Evaluation (ISME) Kuala Lumpur, Malaysia.

Linacre, J. M. (2012). *WINSTEPS Rasch Tutorial 3 (Point 28)* Retrieved from http://www.winsteps.com/a/winsteps-tutorial-3.pdf

Linacre, J. M., & Wright, B. D. (1989). The 'Length' of a logit. *Rasch Measurement Transactions, 3*(2), 54-55.

Luhr, W. E. (1995). *The Maltese Falcon: John Huston, Director*. New Brunswick, NJ: Rutgers University Press.

Massof, R. W. (2012). Is the Partial Credit model a Rasch model. *Journal of Applied Measurement, 13*, 114-131.

McCrory, P. (2006). "Elementary, my dear Watson" (Editorial). *British Journal of Sports Medicine, 40*, 283-284.

McDowell, I. (2006). *Measuring health: A guide to rating scales and questionnaires (Third Edition)*. Oxford: Oxford University Press.

OECD. (2009). PISA (Programme for International Student Assessment) Data Analysis Manual: SPSS (Second Edition). Paris: OECD.

Rapezzi, C., Ferrari, R., & Branzi, A. (2005). White coats and fingerprints: Diagnostic reasoning in medicine and investigative methods of fictional detectives. *BMJ, 331*, 1491-1494.

Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests*. Chicago: MESA Press.

Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health measurement scales: A practical guide to their development and use (Fifth Edition)*. Oxford: Oxford University Press.

Walker, L. K. (2013). House and Holmes: Clinical reasoning comes full circle. *The Journal of Physician Assistant Education, 24*, 30-31.

Wilson, M. (2003). On choosing a model for measuring. *Methods of Psychological Research Online, 8*, 1-22.

Wright, B. D. (1967). *Sample-free test calibration and person measurement*. Paper presented at the Invitational Conference on Testing Problems, Educational Testing Service ETS, New York.

Wright, B. D. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement, 14*, 97-116.

Wright, B. D. (1999). Fundamental measurement for psychology. In S. E. Embretson & S. L. Hershberger (Eds.), *The new rules of measurement: What every psychologist and educator should know*. Mahway, NJ: Lawrence Erlbaum Associates.

Wright, B. D., & Stone, M. H. (1999). *Measurement essentials (Second Edition)*. Wilmington, DE: Wide Range Inc.

Rasch-related Coming Events

Dec. 6-8, 2017, Wed.-Fri. In-person workshop: Introductory Rasch Analysis using RUMM2030, Leeds, UK (M. Horton), http://medhealth.leeds.ac.uk/info/732/psychometric_laboratory/1489/rasch_courses.

Jan. 5-Feb. 2, 2018, Fri.-Fri. On-line workshop: Practical Rasch Measurement - Core Topics (E. Smith, Winsteps), www.statistics.com

Jan. 10-16, 2018, Wed.-Tues. In-person workshop: Advanced Course in Rasch Measurement Theory and the application of RUMM2030, Perth, Australia (D. Andrich), http://www.education.uwa.edu.au/ppl/raschconference/course.

Jan. 17-19, 2018, Wed.-Fri. Rasch Conference: Seventh International Conference on Probabilistic Models for Measurement, Matilda Bay Club, Perth, Australia, http://www.education.uwa.edu.au/ppl/raschconference.

Apr. 13-17, 2018, Fri.-Tues. AERA, New York, NY, www.aera.net.

Profiles in Rasch Measurement

Greetings! My name is R. Trent Haines, and I am an Associate Professor in the Graduate Program in Psychometrics at Morgan State University. One of our points of pride is that we are currently the only Historically Black University with such a program.

I want to thank Ken Royal for inviting me to provide a personal profile in this issue of *Rasch Measurement Transactions*. I first met Ken when we were doctoral students in the College of Education at the University of Kentucky. Although we were in separate programs, we both worked closely with Kelly Bradley, a faculty member who taught two of my courses and the person I consider to be my Rasch mentor.

I was first introduced to the Rasch family of models when I worked with Kelly on a research project. As part of my training as an Educational Psychologist, Kelly recognized that I had a growing interest in quantitative methods and sent me to a multi-day workshop with Mike Linacre to learn more about Rasch. It was there, with a sense of confusion and intrigue, that my journey with Rasch models began.

After graduation and a stint in a research position, I came to Morgan State as an Assistant Professor and immediately began to teach in our graduate program. The first graduate course I taught was Item Response Theory, where I focused heavily on Rasch measurement models. I have continued to use this course and other opportunities to spread the word on Rasch. In fact, one of my favorite things about being a professor is teaching graduate students about the Rasch model.

For at least two semesters before they get to my class, they have learned about measurement from a perspective that draws heavily on Classical Test Theory. Then, when they get to my Rasch course, I do my best to challenge that perspective as I introduce them to Rasch, the rules of measurement, and how many CTT methods violate these rules. In many ways, my students' journey in that class reminds me of my first introduction to Rasch and how I had to overcome some deeply-ingrained ways of thinking about measurement.

In our class, we typically have very robust debate and, in the end, some students "convert", some do not, but most continue to seek the most appropriate measurement techniques for their lines of research. In any event, we all learn something new about measuring the human condition, sometimes in surprising ways.

In addition to working with my students as a professor and mentor, I have become interested in applying the Rasch model to the development of culturally-responsive measurement techniques. Although this direction for research is somewhat new to me, it comes from my long-held interests in Rasch and in culturally-responsive teaching. Some of my current projects include the validation of tests and measures for use in different cultures, constructing culturally-appropriate rating scales, and using technology to better capture data in culturally-relevant ways.

As I close, I would welcome any reader to contact me about my research or our graduate program. I enjoy getting to know new colleagues, helping my students network with other scholars, and spreading the word about the exciting things we are doing in the Morgan State University Graduate Program in Psychometrics. I can be reached at trent.haines@morgan.edu.