FALL 2021, VOL. 34 NO. 3; ISSN 1051-0796

RMT

RASCH MEASUREMENT

► Overview of this Issue of RMT – Stefanie A. Wind & Leigh M. Harrell-Williams
► Precision of measurement, the unit, and Fisher's information function in the Rasch and Gauss distributions - David Andrich
► Review of *Applying the Rasch Model: Fundamental Measurement in the Human Sciences,* Fourth Edition – Stefanie A. Wind
► Call for Nominations for Rasch Measurement SIG Officers and Benjamin Drake Wright Senior Scholar Award – Jue Wang, Eli Jones, Dandan Liao (Rasch Measurement SIG Officers)
► Women in Measurement Announcement - Susan Lyons

Transactions of the Rasch Measurement SIG
American Educational Research Association

Overview of The Issue

In this issue of RMT, we have included one research note, a book review, and several announcements that may be of interest to the Rasch community.

The issue begins with a research note from David Andrich on the Precision of measurement, the unit, and Fisher's information function in the Rasch and Gauss distributions. Following the research note is a review of the fourth edition of *Applying the Rasch Model.*

Then, we provide two announcements that may be of interest to our readers. The first is a call for nominations for Rasch Measurement SIG Officers and (due NOW) and a call for the Rasch Measurement SIG Benjamin Drake Wright Senior Scholar Award. The second announcement provides information about the non-profit organization Women in Measurement and a link to register for a mentoring circle for women of color.

As always, we welcome your contributions to the next issue for RMT. We would appreciate receiving your research note, conference or workshop announcement, etc. by January 1, 2022. Please contact us at the email addresses below if you wish to submit something for inclusion.

Sincerely,

Your RMT Co-editors, Leigh and Stefanie

Rasch Measurement Transactions
www.rasch.org/rmt

Precision of measurement, the unit, and Fisher's information function in the Rasch and Gauss distributions

This note follows up on Andrich (2020a) in which I elaborated some properties of the Rasch measurement distribution for ordered categories in which the thresholds were equidistant, and therefore, the ordered categories are equivalent to a common unit of a typical measuring instrument. In particular, I drew attention to the equivalences between it and the Gauss (Normal) distribution of errors of replicated measurements.

The Rasch distribution of concern here takes the form:

$$p_{xi} = P\{x_{ni}\} = \exp\{x(m_i - x)\Delta_i / 2 + x(\beta_n - \delta_i)\} / \gamma_{ni}, \quad (1)$$

where:

(i) X_{ni} is a random variable of the measurement as a unit count $x_{ni} = 0, 1, 2, \ldots, m$ of object n with parametric measure β_n when measured with instrument i with origin δ_i and unit $\Delta_i = \tau_{xi} - \tau_{(x-1)i}$,

(ii) τ_{xi} is the threshold at which the conditional probability of x or $x - 1$ is 0.5 and $\sum_{x=1}^{m} \tau_{xi} = 0$,

(iii) m_i is the range of the instrument, and

(iv) γ_{ni} is simply the sum of the numerators ensuring that that $P\{x_{ni}\}$ is a probability distribution. The term *measure* denotes the theoretical or hypothesized magnitude of an object and the term *measurement* denotes the value obtained from a measuring instrument as an estimate of the measure.

It is stressed that the distribution in Eq. (1) is an implied distribution of statistically independent replications of measurement of the same object n with the same instrument i. The empirical distribution of such replications is inferred from estimates from the analysis of a set of data with the distribution. The count of the number of units as a measurement, $x_{ni} = 0, 1, 2, \ldots, m$, is referred to as a *unit count*. The count x_{ni} indicates that the object's magnitude is deemed to have exceeded x thresholds and failed to exceed $m_i - x$ thresholds of the instrument which has m_i thresholds. Clearly, the observed measurement is an approximation of the hypothesized measure of the object.

Because this note deals only the with the Rasch distribution of replicated measurements of a single object, the subscript n is dropped in the following exposition. Recall that in the Rasch distribution, the unit count $x_i = 0, 1, 2, \ldots, m$ is a sufficient statistic for the parametric measurement β, which implies that for every unit count x_i there is a parametric measurement β_{xi}. Specifically, the value of β_{xi} for any value of x_i =

0, 1, 2, …, m is a solution to the first derivative of the logarithm of the likelihood function, the probability of Eq. (1), and set to zero to obtain the value of β_{xi} which maximizes the probability, that is:

$$\frac{\partial \ln \mathrm{P}\{x_i\}}{\partial \beta} = x_i - E[X_i] = 0, \quad x_i = 0,1,2,...m_i - 1, \qquad (2)$$

with finite values for $x_i = 0$, m_i extrapolated—their respective values technically being $\pm\infty$.

Also recall that any measurement using a typical instrument is discrete in the unit of the instrument, for example, a parametric measurement of 37.5 centimeters in length, measured with an instrument which measures to the nearest 0.5 cm, is a unit count of 75 half-centimeters. In this case the operational unit of the instrument is 0.5 cm relative to the standard magnitude of 1cm. The discrete nature of measurements is explicit with modern electronic instruments.

In physical measurement, the discrete unit count x_i in the operational unit and the parametric measurement β_{xi} are generally, and helpfully, blurred. Thus, the unit count of 75 in the above example would be reported immediately as 37.5 cm without reference to the 75 half centimeters. Of course, the actual measure β of the object is taken as a real number referenced to some standard, and is to be estimated. In the Rasch distribution the discrete integer, unit count x_i = 0, 1, 2, …, m (say 75 half centimeters) is distinguished from the parametric measurement β_{xi}. This note follows the consequences of having this distinction from the perspective of precision of measurement as defined by the Fisher information function in relation to the magnitude of the unit.

Three properties of the Rasch distribution

In the derivation of the Gauss distribution, to which we return below, Gauss noted that the symmetric quadratic form only holds provided the object is sufficiently well aligned to the range of the instrument that probabilities of measurement near the extreme counts $x_i = 0$; $x_i = m_i$ vanish (Eisenhart, 1983b). In the relationship explained in this note, the same relative alignment is required. Then if the alignment ensures that the probabilities near the extremes of the instrument's range vanish, the following three properties of Eq. (1) hold (Andrich, 2020a; 2020b).

First, the variance of the unit count, $V[X_i]$, is the inverse of the operational unit Δ_i:

$$V[X_i] = 1/\Delta_i. \qquad (3)$$

Second, there is a range of parametric measurements β_{xi} in which $\beta_{xi} - \beta_{x-1,i} = \Delta_i$. This is consistent with typical measurement and is the basis of the results shown below.

Third, the Rasch distribution is identical in the same range and for the same instrument i as the discrete Gauss distribution of the form

$$p_x = \exp\{-\frac{1}{2}\frac{(x_i - \mu_i)^2}{\sigma_i^2}\} / \gamma_i \qquad (4)$$

where $\mu_i = E[X_i]$ and $\sigma_i^2 = \mathrm{V}[X_i]$, and γ_i is the sum of the numerators. The values μ_i, σ_i^2 are the same as those of the complementary continuous Gauss distribution.

Distributions of replicated measurements from two instruments with different units

Eq. (3) may give the impression of a paradox - that the greater the unit, Δ_i, the smaller the variance $V[X_i]$, and therefore the smaller the unit, the worse the precision of measurement. However, the paradox is resolved by realizing that the smaller value of the variance is also expressed in the greater unit. This resolution is shown below.

Table 1 shows the equivalent measurements with two instruments (i = 1, 2), in which $\Delta_1 = 1$; $m_1 = 10$ and $\Delta_2 = 0.5\Delta_1$; $m_2 = 20$ over the same range of the variable. Instrument 1 is taken as the standard which provides a frame of reference to study the effects of a different unit from Instrument 2. Table 1 also shows the maximum likelihood values of the measurements β_{x1}, β_{x2} calculated to a convergence criterion of 0.0001 and the probabilities of the unit counts and corresponding parametric measurements when the object's measure is $\beta = 0.5$ on the common metric. The arbitrary origin for the parametric measurements is set to $\delta_i = 0$. The Table shows, in italics, the range of parametric measurements from both instruments where $\beta_{xi} - \beta_{x-1} = \Delta_i$; it shows that outside this range $p_{xi} = 0$. It also shows that within this range the probabilities are symmetrical. Unlike typical tables of measurements in which the unit count and the parametric measurement are blurred, these are distinguished in Table 1. It is emphasized that Table 1 shows the distribution of replicated measurements of the measurement of the same object by each of the two instruments. We now examine the theoretical means and variances, ($E[X_1]$; $V[X_2]$) of the unit counts, followed by the theoretical means and variances of the parametric measurements, ($E[\beta_1]$; $V[\beta_2]$).

Table 1. Distributions of unit counts and parametric measurements from two instruments with units $\Delta_1 = 1$, $\Delta_2 = 0.5\ \Delta_1$, range $m_1 = 10$, $m_2 = 20$ for a measure $\beta = 0.5$ on a common metric.

	(x_1, m_1)	(β_{x1}, Δ_1)	(x_2, m_2)	(β_{x2}, Δ_2)	$p_x(\beta_{x1})$	$p_x(\beta_{x2})$
	0	-5.5	0	-5.9	0.00	0.00
			1	-4.8		0.00
	1	-4.2	2	-4.1	0.00	0.00
			3	*-3.5*		0.00
	2	*-3.0*	4	*-3.0*	0.00	0.00
			5	*-2.5*		0.00
	3	*-2.0*	6	*-2.0*	0.02	0.00
			7	*-1.5*		0.01
	4	*-1.0*	8	*-1.0*	0.13	0.03
			9	*-0.5*		0.10
	5	*0.0*	10	*0.0*	0.35	0.22
			11	*0.5*		0.28
	6	*1.0*	12	*1.0*	0.35	0.22
			13	*1.5*		0.10
	7	*2.0*	14	*2.0*	0.13	0.03
			15	*2.5*		0.01
	8	*3.0*	16	*3.0*	0.02	0.00
			17	*3.5*		0.00
	9	4.2	18	4.1	0.00	0.00
			19	4.8		0.00
	10	5.1	20	5.9	0.00	0.00
$E[.]$	5.5	0.5	11.0	0.5		
$V[.]$	1.0	1.0	2.0	0.5		

Expected values and variances of the unit counts

First, consider the expected values and variances of the unit counts (x_i, m_1) from Instrument 1, unit $\Delta_1 = 1.0$. The $E[X_1; \Delta_1] = 5.5$ is slightly greater than the middle of the range (0, 10), which accounts for $\beta = 0.5$ being slightly greater than the origin of 0.0. The value $\beta = 0.5$ for the measure of the object was taken for convenience of exposition; it could be any real number within the relevant range. In this special case, which we take as the standard instrument, $V[X_1; \Delta_1] = 1.0$, which is also its inverse and the unit.

Second, consider the expected values and variances of the unit counts (x_2, m_2) from Instrument 2, unit $\Delta_2 = 0.5$. The $E[X_2; \Delta_2] = 11.0$ is slightly greater than the middle of the range (0, 20), which again accounts for $\beta = 0.5$ being slightly greater than its origin 0, and is also double the $E[X_1; \Delta_1] = 5.5$ for Instrument 1. This is consistent with the unit of Instrument 2 being half that of Instrument 1. The $V[X_2; \Delta_2] = 2.0$ is exactly the inverse of the unit, $\Delta_2 = 0.5$, that is, The $V[X_2; \Delta_2] = 2.0 = 1/\Delta_2$.

Thus, the variance of the unit count in the smaller unit is greater than the variance of the unit count in the larger unit, which as noted earlier, could give the impression that the error variance is greater with a smaller unit. However, it has to be appreciated that the greater variance is in terms of the smaller unit, and when referenced to the same metric, the smaller unit provides a smaller variance. This is illustrated in Figure 1 where the two distributions are referenced to the same metric of the parametric measurement. It can be seen that the distribution of the unit counts from Instrument 2 with the smaller unit is narrower than the distribution with Instrument 1 with the larger unit.

Note that with the standard, Instrument 1 with $\Delta_1 = 1$, no measurement has the value of the object's measure $\beta = 0.5$. With the measure taken to be a real number, and a measurement discrete, this feature is typical. In contrast, that a measurement can be a value of the measure with Instrument 2 is contrived for convenience.

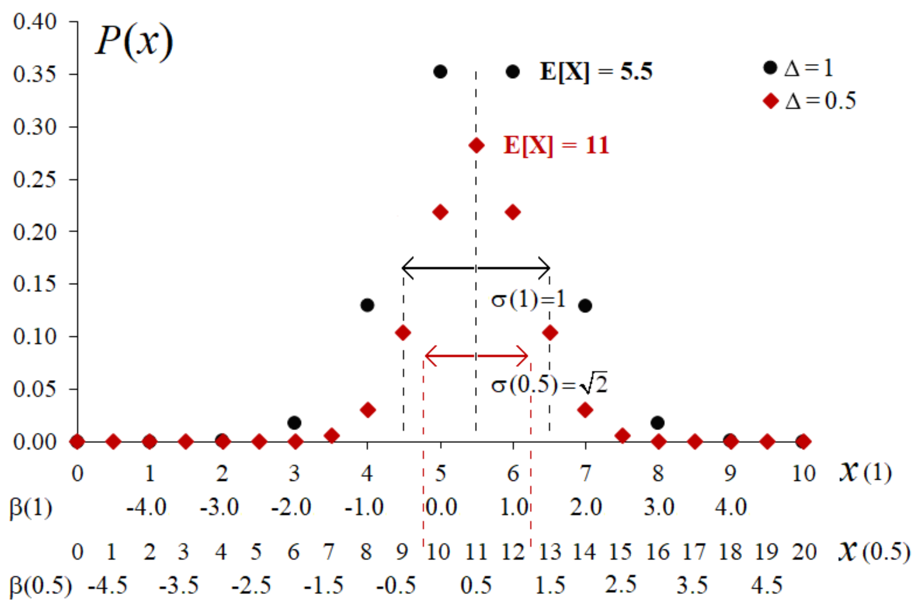

Figure 1. *Distributions of measurements* x *of two instruments with units* $\Delta_1=1$, $\Delta_2=0.5$, *for a measure of* $\beta = 0.5$.

Expected values and variances of the parametric measurements

First consider the expected values and variances of the parametric measurements, (β_1, Δ_1) from Instrument 1, unit $\Delta_1 = 1.0$. $E[\beta_1; \Delta_1] = 0.5$ is exactly the measure of the object, $\beta = 0.5$. This is consistent with maximum likelihood estimate within the range where the Rasch distribution is identical to the Gauss distribution and would be the case no matter the value of β. The $V[\beta_1; \Delta_1]$

= 1.0 is exactly the size of the unit, but unlike the variance of the unit counts, and as seen below, is not a special case.

Thus second, consider the expected values and variances of the parametric measurements (β_2, Δ_2) from the Instrument 2, unit $\Delta_2 = 0.5$. $E[\beta_2; \Delta_2] = 0.5$ again, exactly the measure of the object $\beta = 0.5$. Perhaps of more interest and novelty, and the point of this note, is that the variance $V[\beta_2; \Delta_2] = 0.5 = \Delta_2$, exactly the size of the unit. In this case, the smaller the unit, the smaller the variance, indicating that the smaller unit gives greater precision of the parametric measurements.

Figure 2 shows the distributions of the parametric measurements, (β_{x1}, β_{x2}), from the two instruments referenced to the same metric. It can be seen that the distribution of Instrument 2 with the smaller unit is narrower than that of Instrument 1 with the larger unit.

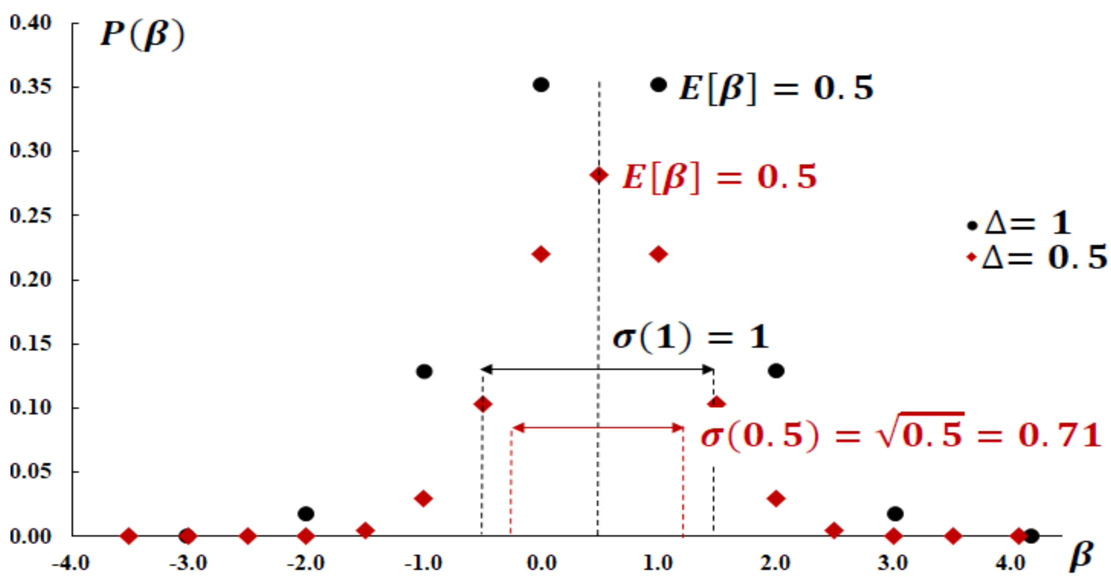

Figure 2. Distributions of parametric measurements β_x of two instruments with units $\Delta_1 = 1$, $\Delta_2 = 0.5$ for a measure of $\beta = 0.5$.

Fisher's information function

The equality between $V[\beta_2; \Delta_2]$ and Δ_2 can be explained by the *Fisher information function,* $I(\beta)$. For the Rasch distribution, derived from an equivalent definition of $I(\beta)$ as that given by Fisher (Samejima, 1969), $I(\beta)$ for instrument i is given by

$$I(\beta_i) = -\frac{\partial^2 \ln \mathrm{P}\{x_i\}}{\partial \beta^2} = V[X_i] \qquad (5)$$

The inverse of the information function, $1/I(\beta)$, , gives the variance of the parametric measurements $\mathrm{V}[\beta_i]$, whose square root provides a standard error of the estimate of β from an analysis of data, that is $V[\beta_i] = 1/I(\beta_i) = 1/V[X_i]$. However, from Eq. (3), $V[X_i] = 1/\Delta_i$. Therefore:

$$V[\beta_i] = \Delta_i \quad (6)$$

Eq. (6) shows a clear relationship between the precision of parametric measurements and the unit of the instrument, one which is elegant and aesthetically pleasing.

Improving precision by increasing the sample size or by reducing the operational unit.

The result in Table 1 is consistent with Eq. (6). In general, relative to a standard unit of Δ_s from Instrument s, if the unit is $\Delta_n = \Delta_s / n$ from Instrument n, then:

$$V[\beta_n;\Delta_n] = \Delta_n = \Delta_s / n = V[\beta_s;\Delta_s] / n \quad (7)$$

Now it is well known that the variance of the mean of a distribution, which is according to a number of criteria the best estimate of the parametric measure of an object and resulted in the Gauss distribution, is inversely proportional to the sample size, that is

$$V[\overline{\beta}_s;\Delta_s] = V[\beta_s;\Delta_s] / n \quad (8)$$

Thus in the above development, to increase the precision of the estimate of a measure relative to a standard unit, taking the mean of a sample size of n measurements with the same instrument is identical to reducing the size of the operational unit of the instrument by $1/n$. This relationship is again elegant and aesthetically pleasing.

Some connections to the development of the Gauss distribution

The development of the Gauss distribution was set in the context of obtaining the best estimate of the measure of an object in the presence of a distribution of measurements.

> *Laws of error, i.e., probability distributions assumed to describe the distribution of the errors arising in repeated measurement of a fixed quantity by the same procedure under constant conditions, were introduced in the latter half of the eighteenth century to demonstrate the utility of taking the arithmetic mean of a number of measurements or observed values of the same quantity as a good choice for the value of the magnitude of this quantity on the basis of the measurements or observations in hand.* (Eisenhart, 1983a, p. 1).

This work exercised the best mathematical minds of the time, including Lagrange and Laplace, and culminated with the Gauss distribution.

> *Nonetheless, Thomas Simpson (1710–1761) wrote to the president of the Royal*

Society in March 1755 that some persons of considerable note maintained that one single observation, taken with due care, was as much to be relied on as the mean of a great number. (Eisenhart, 1983a, p. 1).

From the analysis of the Rasch distribution above, which is identical to the discrete Gauss distribution, and whose mean and variance are the same as the complementary continuous Gauss distribution, we can formalize the equivalence of one measurement *taken with due care* and the mean of measurements. Specifically, if the operational unit is reduced by 1 *nth*, one such measurement is identical in precision to that obtained by the mean of *n* original measurements. This of course is a theoretical relationship obtained from the equivalence of the Gauss and Rasch distributions, both of which were derived theoretically and not to account for any particular data set.

In addition, with some adaptation to more conventional notation, Eisenhart (1983b, p.2) writes regarding Gauss's derivation:

whence

$$f(x-\mu) = c\exp(-1/2)k(x-\mu)^2$$

and k must be positive if the corresponding value of the joint probability density… is to be maximum. … Setting k = h^2 *… the desired law of error is seen to be*

$$f(x-\mu) = h/(\sqrt{\pi})\exp(-h^2(x-\mu)^2,$$
$$-\infty < x < \infty$$

in which, Gauss pointed out [29, art. 178], ''the constant h can be considered as the measure of precision … of the observations.

From the connections with the Rasch distribution above, and Eq. (4), it can be seen that $k = h^2 = (1/2\sigma^2) = (1/2)\Delta$, is a measure of precision, it being half the operational unit of the instrument. Moreover, from the motivation of the Rasch distribution, $k = h^2 = (1/2)\Delta$ must be positive because $\Delta = \tau_k - \tau_{k-1}$ is the directed distance, the unit, between successive thresholds of the instrument. It is emphasized that these results hold when the measure of the object is well within the range of the instrument which in turn ensures that the probabilities of measurements close to the extremes of the range of the instrument are zero, a condition required for the application of the Gauss distribution

Table 2.

Means of the means and variances across each of five instruments, and the estimate of the respective units, for 1000 replicated measurements of each object, $\beta = 0.5$.

Instrument	$\Delta_1 = 1,\ m_1 = 10$	$\Delta_2 = 0.5,\ m_2 = 20$
$E[X_i];\ \bar{X}_i$	5.5; 5.490	11.0; 11.029
$V[X_i];\ S_i^2$	1.0; 0.988	2.0; 1.949
$\Delta_i;\ \hat{\Delta}_i$	1.0; 1.013	0.5; 0.514

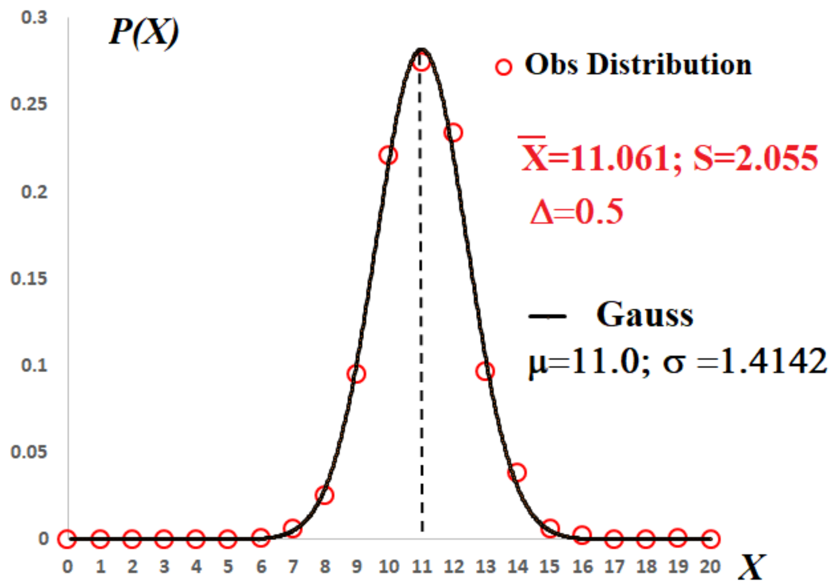

Figure 3. *Distribution of simulated 1000 replicated measurements with one instrument,* $\Delta = 0.5$*,* $m = 20$*, and the continuous Gauss distribution,* $\mu = 11.0$, $\sigma^2 = 1/\Delta = 2.0$; $\sigma = 1.4142$.

An example of an explicit distribution

To show the connection with data based on the theoretical distributions above, data were simulated according to Eq. (1) for 1000 replications of an object with measure $\beta = 0.5$ with an instrument $\Delta_1 = 1$, $m_1 = 10$ and another with $\Delta_2 = 0.5$, $m_2 = 20$ according to the Rasch distribution of Eq. (1). The simulation for each instrument was repeated with four more instruments. Table 2 shows the summary statistics of the means and variances across the five distributions for each kind of instrument. Without studying the sampling distributions, it is evident that the observed values are close to the theoretical values, and that the unit in particular is recovered well.

Finally, to illustrate the relationship of the above properties of the Rasch distribution to the Gauss distribution, Figure 3 shows the empirical distribution of one of the simulations of 1000 replications of an instrument with $\Delta_2 = 0.5$, $m_2 = 20$, and the continuous Gauss distribution with the theoretical values of $\mu_2 = 11.0$, and $\sigma^2_2 = 1/\Delta_2 = 2.0$. and. It is clear that the Gauss distribution overlaps, virtually fully, with the empirical, discrete Rasch distribution.

References

Andrich, D. (2020a) Some History and Recent Understandings of a Rasch Measurement Distribution. *Rasch Measurement Transactions. 33,* 2, 1766-1770.

Andrich, D. (2020b). The Rasch Distribution: a discrete, general form of the Gauss distribution of uncertainty in scientific measurement. *Measurement, 173,* 108672. https://doi.org/10.1016/j.measurement.2020.108672.

Eisenhart, C. (1983a). Law of error I: Development of the concept. In S. Kotz & N.L Johnson (Eds.), Encyclopedia of statistical sciences (Vol. 4, pp. 530-547). Toronto: Wiley.

Eisenhart, C. (1983b) Law of error II: Development of the concept. In S. Kotz & N.L Johnson, (Eds.), Encyclopedia of statistical sciences (Vol. 4, pp. 547 – 562). Toronto: Wiley.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometric Monographs, 34*(2, No.17).

David Andrich

The University of Western Australia

Review of *Applying the Rasch Model: Fundamental Measurement in the Human Sciences,* Fourth Edition

It is my pleasure to review the latest edition of *Applying the Rasch Model* (ARM). With guidance from my graduate school advisor George Engelhard, I learned most of the fundamentals of Rasch measurement theory from the second edition of ARM. Since then, I used the third edition of ARM as one of the required texts in my own graduate seminar on Rasch measurement theory in 2016. The fourth edition of ARM (ARM4) shares the same excellent features as the previous editions. Maintaining the conversational and accessible tone that characterized previous editions, ARM4 provides a gentle yet compelling introduction to Rasch measurement theory that is accessible to beginners while also being thorough enough to satisfy the curiosity and appease potential Rasch-criticisms of advanced researchers from a variety of methodological and theoretical backgrounds. The new features in ARM4 make it an even more comprehensive resource for new researchers and experienced methodologists alike.

In the following paragraphs, I describe my observations and reflections on each chapter in ARM4. I conclude the review with some thoughts about the book as a whole

Reflections on ARM4 Chapters

Chapter 1 (*Why Measurement is Fundamental*) and Chapter 2 (*Important Principles of Measurement Made Explicit*) introduce readers to the fundamental principles that characterize measurement in the physical sciences that are required, but often overlooked, for measurement in the social and behavioral sciences. In this new edition, Chapter 1 emphasizes the definition and use of quantitative scales in psychological sciences that will help readers bridge their understanding of levels of measurement from statistics courses to measurement in a theoretically sound way. The explicit consideration of fundamental principles using concrete, accessible examples provide readers with a solid theoretical foundation on which to build their understanding of additional theoretical and technical aspects of Rasch measurement theory. Reading the new versions of these two chapters ARM4 reminded me that these are chapters that I should read and refer to regularly to reaffirm understanding of the fundamental principles of measurement.

Chapter 3 (*Basic Principles of the Rasch Model*) introduces readers to Rasch measurement theory using examples and step-by-step demonstrations. The pathway analogy and high jump examples are useful and memorable tools for understanding and visualizing key Rasch measurement outcomes (model predictions, estimates, precision, and fit). I especially appreciated the deliberate consideration of misfit and what it means in both a theoretical and practical sense. In addition, the discussion of reliability indices within the Rasch framework is especially important for readers who are familiar with psychometric methods but new to Rasch.

Chapter 4 (*Building a Set of Items for Measurement*) walks the reader through an application of Rasch measurement theory using an example from Bond's Logical Operations Test (BLOT), which is a dichotomously scored test of cognitive development for young adolescents. The demonstration reinforces principles from earlier chapters while demonstrating how to interpret major components of a Rasch analysis of dichotomous data. Among the discussions in this chapter, I particularly appreciated the careful treatment of guessing with examples that demonstrate how it can be detected and understood from a measurement perspective.

Chapter 5 (*Invariance: A Crucial Property of Scientific Measurement*) builds on earlier discussions of invariance, with an emphasis on practical consequences of invariance for practical psychometric tasks such as linking and equating, including vertical scaling and differential item functioning (DIF). I particularly appreciated the connection to Classical Test Theory methods in the discussion of logits and correlations, and the practical nature of the illustrations and examples in this chapter.

Chapter 6 (*Measurement using Likert Scales*) was the first chapter to focus on polytomous (rather than dichotomous) data. This chapter is one that I found particularly interesting given my own research and teaching interests, and it is one that I will recommend to my students and collaborators who work with Likert-type response data. The computer anxiety questions example in the beginning of the chapter made clearly demonstrated the limitations of typical treatment and interpretation of Likert-type responses in "mainstream" survey research. The new example data are from the Instrumental Attitude toward Self-Assessment Questionnaire (IASQ), which includes a four-category rating scale. The four-category scale is likely to be more similar to readers' own Likert-type response data than the original three-category scale example from previous editions. In addition, the step-by-step walkthrough of the polytomous Rasch model analysis of these data helps readers understand how to interpret key outcomes from the model (especially thresholds). Importantly, Chapter 6 considers practical issues such as sample size and dimensionality assessment, while also helping readers understand the theoretical implications of their design and analytic decisions.

Chapter 7 (*The Partial Credit Rasch Model*) extends the introduction to the polytomous Rasch model from Chapter 6 to the more-flexible Partial Credit Model (PCM). The authors highlight important features of the PCM with a theory-driven, purposive example from Piagetian cognitive developmental data. The authors walk through an example analysis step by step to demonstrate the features and benefits of the PCM approach to analyzing the data. Like Chapter 6, the extended understanding section of Chapter 7 included practical considerations for evaluating rating scale functioning (elaborated in Chapter 11) and dimensionality using Rasch Principal Components Analysis (PCA; also discussed in later chapters). The concrete illustration and discussion of Rasch PCA is an excellent reference for this often-misunderstood analysis.

Chapter 8 (*Measuring Facets Beyond Ability and Difficulty*) introduces the Many-Facet Rasch Model (MFRM) using example data from a rater-mediated writing assessment in the United States (originally published by Engelhard in 1992). Like previous chapters, Chapter 8 includes a step-by-step demonstration of the analysis and interpretation of the model results, with attention to both theoretical and practical issues. I appreciated the direct comparison of the MFRM to rater reliability that demonstrated how the MFRM can supplement this popular and limited index to address a variety of concerns in rater-mediated assessment. The extended understanding section of Chapter 8 was also focused on rater-mediated assessments—providing coherence with the earlier parts of the chapter.

Chapter 9 (*Making Measures, Setting Standards, and Rasch Regression*) and Chapter 10 (*The Rasch Model Applied Across the Human Sciences*) include examples of Rasch measurement applications that highlight practical issues across a variety of contexts. The examples in Chapter 9 help readers see how Rasch methods can be applied to solve practical measurement challenges. Although they are important topics that warrant discussion in an introductory Rasch text, I thought the standard setting and Rasch regression examples may be a bit advanced for new researchers. Like Chapter 9, the examples in Chapter 10 serve as useful references for Rasch applications to address practical challenges including creating short versions of existing instruments, revising existing instruments, exploring gains (e.g., growth) in items or persons via “racking” and “stacking” techniques, and applications to classroom assessment. I have personally bookmarked the “Analytic Mill” section of Chapter 10, where the authors demonstrate how Rasch parameters can be used in common analyses, including ANOVA, HLM, and SEM. I have found myself recommending these techniques to students when I serve on dissertation committees, and I am glad to have a reference to share with them for direct guidance.

Chapter 11 (*Rasch Modeling Applied: Rating Scale Design*) is one of my personal favorite chapters in ARM4 because of my own research and teaching emphases on rating scales. The example analyses provide context for explicit guidelines for evaluating rating scale functioning while acknowledging the nuances in this analysis and the need for evidence to support decisions, such as collapsing categories. In addition, the discussion of negatively worded items using data from the IASQ (from Chapter 6) provides direct and clear guidance that helps readers understand the theoretical and practical considerations associated with such items.

Chapter 12 (*Rasch Model Requirements: Model Fit and Unidimensionality*) provides direct and clear instruction on fit and dimensionality—two topics that are often misunderstood in applications of Rasch measurement theory, and that are also sources for criticism of Rasch-based work. In the discussion of fit for individual items and persons, numeric and graphical illustrations of response patterns make the interpretation of infit and outfit statistics concrete. In addition, the discussion of dimensionality with comparisons between Rasch approaches and "traditional" approaches (via factor analysis) helps readers understand how fundamental measurement principles translate to considerations of dimensionality. Chapter 12 also includes a useful consideration of multidimensional Rasch models that helps readers understand the potential use and implications of these models. In addition, the R code to support the examples in this chapter are a nice supplement that allow readers to practice the analyses and interpret the results.

Chapter 13 (*A Synthetic Overview*) is far more than a summary of the previous chapters. This closing chapter reinforces key concepts and themes from earlier in the book with extended discussions and considerations that reflect both classic references and new developments in measurement research and practice. In particular, the thoughtful comparison between CTT and Rasch measurement, the discussion of fit, and considerations of interval-level scales in relation to the requirements for measurement address common misconceptions and criticisms that surround our work with Rasch measurement.

Appendix A (*Getting Started*) is a practical guide that helps readers conduct a Rasch analysis using Winsteps. This Appendix walks readers through the analysis from data entry to variable interpretation in a concise and clear manner that complements the content from earlier in the book. Readers can use Appendix A as a "quick start guide" for proceeding through a Rasch analysis and then supplement their interpretation with guidance from earlier chapters. An annotated resource list at the end of the chapter helps readers identify books, journals, and other sources that can supplement their research using Rasch measurement theory.

Appendix B (*Technical Aspects of the Rasch Model*) makes explicit the details that are essential to understanding and explaining key components of Rasch measurement research. This appendix is an accessible reference where readers can find direct explanations of equations for popular Rasch models and tools for evaluating them (e.g., reliability and information).

Appendix C (*Going All the Way*) is a new addition to ARM4 that includes a discussion of the popular unstandardized Rasch infit and outfit statistics. The direct treatment of weaknesses of these statistics is essential reading that helps readers understand the boundaries of these statistics. Appendix C also demonstrates methods for testing Rasch model requirements using multiple methods, including global model-fit tests, Confirmatory Factor Analysis, graphical approaches, and Structural Equation Modeling. This information is essential for measurement students and scholars who need to understand how Rasch principles align with other popular approaches.

In addition, the tutorials in R using eRm by Tara Valladares are a wonderful resource that align Rasch research with ever-growing interest in R programming.

Conclusions

ARM4 helps readers understand the fundamental principles of measurement and their role in constructing, evaluating, and learning from measurement procedures in the social and behavioral sciences. This new edition brings Rasch methods into the current landscape of quantitative methods so that readers understand how Rasch measurement theory fits within current methodological trends and conversations.

I have whole-heartedly recommended the fourth edition of ARM to my doctoral advisees, and I plan to use it as one of the required texts in my upcoming courses. The book is one I keep handy for quick reference in my own work. My congratulations and sincere thanks to the authors for their contribution to Rasch measurement theory and practice via ARM4.

Stefanie A. Wind
The University of Alabama

Call for Nominations for Benjamin Drake Wright Senior Scholar Award

The Rasch Measurement SIG is currently accepting nominations for the Benjamin Drake Wright Senior Scholar Award. This award is presented to an individual senior scholar for outstanding programmatic research and mentoring in Rasch measurement over the course of a career and who is still active in Rasch measurement research at the time the award is granted. It will be offered in 2022. The award is open to scholars worldwide. Membership in AERA or Rasch Measurement SIG is not required of the nominee.

Eligibility Criteria for the Benjamin Drake Wright Senior Scholar Award

The Rasch Measurement SIG will bestow the Benjamin Drake Wright Senior Scholar Award upon a senior scholar who is active in Rasch measurement research at the time the award is granted (as attested to, for example, by research publications of recent date or current doctoral advisees) and who is nominated by members of the community as an exemplar in regard to the following two basic criteria. Potential nominees will have:

a. designed and carried out programmatic research that originates in Rasch measurement and helps understand crucial phenomena in model definition, parameter estimation, fit assessment, construct specification, novel applications, the place of Rasch measurement in the history and philosophy of science, etc., as represented in a corpus of writings and research projects that have contributed to the theoretical development of the field as well as having been grounded empirically; AND

b. developed the research capacity of the field, as attested to by the existence of a "school of thought" or intellectual heritage associated with the scholar's name, a heritage that includes other individuals whom the scholar has had a direct influence in encouraging and helping become productive in Rasch measurement research or an identifiable domain of Rasch measurement research within which the nominee's constructs and results are used regularly by other researchers.

The Rasch Measurement SIG recognizes that other features of a person's work might add to the criteria above, strengthening a nomination. Among the criteria that could add to the basic ones is one or more of the following. The nominee may also have made:

a. major contributions to broader fields of research in education, psychology, health care, or the social sciences, as represented by his or her participation (as author, speaker, or consultant) in research forums from fields other than Rasch measurement or by the recognition of his or her scholarship in other fields of inquiry (inclusive of all of educational research and the social sciences); OR

b. major impact on the practice of Rasch measurement, as represented by the existence of policy documents, curriculum materials, professional development programs, or a corpus of practitioner- or public-oriented literature to which the nominee has significantly contributed as an author.

The award includes a plaque and an invited address for the 2023 Rasch SIG business meeting at AERA conference. An honorarium is included, and some travel reimbursement may be available.

Nominations should include the following: Individuals will be nominated via a letter of nomination emailed to the Convener of the Awards Committee proposing the name of the nominee and describing the grounds on which the nominee meets the requirements for the award. Three criteria should be addressed in the letter:

1. A brief (no more than 250-word) description of the program of research carried out by the nominee.
2. A list of significant publications representing the contributions described, and a list of scholars who have been significantly affected by the work of the nominee. The list of scholars may include, but need not be limited to, doctoral students who worked with the nominee. Current

contact information for the list of scholars should also be included in the nomination.

3. The nominee's CV.

Self-nominations will not be accepted. The deadline for nominations is December 31, 2021. Nominations are submitted by sending an email to the SIG Chair, Jue Wang, at jue.wang@miami.edu.

Jue Wang, *Chair of Rasch SIG*
Eli Jones, *Secretary of Rasch SIG*
Dandan Liao, *Treasurer of Rasch SIG*

Call for Nominations for Rasch Measurement SIG Officers

The Rasch Measurement SIG is seeking nominations for the next slate of officers. The following positions on the Rasch Measurement SIG are open for election. The length of the term is indicated in parentheses, and all terms start at the end of the 2022 AERA Annual Meeting.

- Chair (2 years)
- Secretary (2 years)
- Treasurer (2 years)

AERA policy requires that all elections be competitive; that is, there must be two or more candidates for each elected office. Only Regular Members of AERA can run for office. Candidates must also be current members of the SIG and of AERA to serve as an officer. Self-nominations are welcome. Members who wish to nominate a candidate for consideration, or self-nominate themselves, should send a one-page biography including the qualifications of the nominee to Dr. Dandan Liao (Rasch SIG Treasure) at dandan.liao@cambiumassessment.com.

The initial call for nominations were due on September 30, 2021. However, you can still contact Dr. Liao before October 30 submission deadline if you are interested.

Duties and Responsibilities of The Officers

Chair: The Chair shall be responsible for the general administration of the SIG, for ensuring that the SIG Bylaws are followed, and shall act as liaison between the SIG and AERA and the SIG and the SIG Executive Committee. The Chair shall preside at all meetings of the SIG's Executive Committee and at the Annual Business Meeting. The Chair shall act as parliamentarian or shall appoint a SIG member to serve in that role for each meeting. The Chair shall appoint ad hoc committees as needed and shall appoint persons to assist officers, to chair committees, or to carry out other SIG work.

Secretary: The Secretary shall be responsible for managing any official correspondence and meeting minutes of the Rasch SIG. This person will also be responsible for maintaining the Rasch SIG website, or appointing an appropriate representative as need.

Treasurer: The Treasurer shall be responsible for managing and reporting on the financial accounts of the Rasch SIG and the safe keeping of all financial documents of the Rasch SIG.

For more information about Rasch Measurement SIG, please check out our website and follow us on Facebook.

Website:
https://www.aera.net/SIG083/Rasch-Measurement
Facebook:
https://www.facebook.com/groups/raschmeasurement

Sincerely,
Jue Wang, *Chair of Rasch SIG*
Eli Jones, *Secretary of Rasch SIG*
Dandan Liao, *Treasurer of Rasch SIG*

Women in Measurement

Women in Measurement is a non-profit organization dedicated to the advancement of gender and racial equity in educational measurement. We work to amplify the voices of women—and minoritized women in particular—and to provide structures of support for women throughout their careers in academia, industry, non-profits, and the public sector. Women in Measurement seeks to understand and dismantle systems of oppression that have led to disproportionate access and representation at the highest levels of our field. Its programs and events include mentoring sessions, research fellowships, networking events, and more. Our next event will be a mentoring circle for women of color in the field, held on December 8th at 12:00-1:30pm EST. Click here to register for that event. To learn more about our organization and sign up for the monthly newsletter visit www.womeninmeasurement.org.

Susan Lyons
Executive Director, Women in Measurement, Inc.