A PSYCHOPHYSICAL INTERPRETATION OF RASCH'S PSYCHOMETRIC PRINCIPLE OF SPECIFIC OBJECTIVITY

R. John Irwin
Department of Psychology, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand
rj.irwin@auckland.ac.nz

Abstract

Rasch's principle of specific objectivity asserts that the comparison of two objects should be independent of everything but the two objects and their observed reactions. A consequence of the principle for psychometrics is that, for homogeneous tests conforming to the Rasch model, the estimated difference in ability between two people is independent of the difficulty of any particular test items used to compare them. The principle is exemplified by the Receiver Operating Characteristic of Luce's psychophysical choice theory. In this application of choice theory, the operating characteristic depicts how the success-rates of two people who attempt the same task co-vary with the difficulty of the task. Because this operating characteristic depends only on the difference in ability between two people and is independent of task difficulty, it satisfies the principle of specific objectivity.

The Rasch model of psychometrics and the psychophysical theory of signal detection rest on different foundations and address different areas of inquiry. Yet, as DeCarlo (2005) has demonstrated, they share many characteristics. He has applied a latent-class generalization of detection theory to problems in psychometrics and shown, for example, how the detection-theory parameter for response bias can elucidate the psychological processes involved in assigning grades to essays.

Here my intention is to relate the Receiver Operating Characteristic (ROC) of detection theory with Rasch's principle of specific objectivity and thereby to extend the compass of both concepts. I first consider how the Rasch model measures the difference in ability between two people, and then describe the principle of specific objectivity derived from that model. Finally, I offer a detection-theoretic interpretation of specific objectivity by means of Luce's choice theory and its Receiver Operating Characteristic.

The Rasch Model

Consider two people, A and B, who independently attempt the same task: for example, two high-jumpers who attempt to clear a bar at the same height, or two students who answer the same test question. For activities of this nature, a person may at any attempt succeed (outcome $X = 1$) or fail (outcome $X = 0$). The Danish mathematician, Georg Rasch, developed a theory – now known as the Rasch model – about how to measure the difference in ability between people engaged in such tasks.

The Rasch model states that the probability of a person succeeding on such a task is a function only of their ability, θ, and the task's difficulty, δ. The standard psychometric design requires a person to attempt a series of related tasks – test items, for example. In this way the probability of success, P, is estimated from the number of successes and failures. According to the Rasch model, the probability of Person A succeeding can be expressed by the logistic distribution function:

$$P_A(X=1\,|\,\theta_A,\delta)=\frac{e^{\theta_A-\delta}}{1+e^{\theta_A-\delta}}, \quad (1)$$

where P_A is the probability of A succeeding, θ_A is the estimate of A's ability ($-\infty < \theta_A < \infty$) and δ the task's difficulty ($-\infty < \delta < \infty$).

Table 1, which is based on an exposition by Masters (2001), shows one method for measuring the difference in mean ability, $\theta_A - \theta_B$, between two people, A and B, when they both attempt the same task. The joint outcomes of two people independently attempting the same task are observed and counted. For this design, outcome $X = 1{,}0$ means A succeeded and B failed, and outcome $Y = 0{,}1$ means A failed and B succeeded.

The information available from this design is the number of times that A succeeded when B failed, denoted N_{10}, and the number of times that B succeeded when A failed, N_{01}. Corresponding probabilities can be estimated from these counts. Let P_{10} be the probability that A succeeds and B fails, a joint probability that equals $P_A\cdot(1-P_B)$, and let P_{01} be the probability that A fails and B succeeds, which equals $(1-P_A)\cdot P_B$. If the logistic function of Equation (1) is substituted into each of the expressions for these joint probabilities, P_{10} and P_{01}, then simplification yields:

$$\ln(P_{10}/P_{01})=\theta_A-\theta_B. \quad (2)$$

So the ratio of the log odds of this comparison between two people, A and B, is constant and completely independent of the task difficulty, δ, which does not appear in Equation (2). Rasch (1966, 1977) asserted that this result demonstrated "specific objectivity".

The Principle of Specific Objectivity

Rasch wanted to eschew a philosophical debate about what he meant by objectivity and stated (Rasch, 1966, p. 104) that specific objectivity obtained whenever "*the comparison of any two subjects* can be carried out in such a way that *no other parameters are involved than those of the two subjects* – neither the parameter of any other subject nor any of the stimulus parameters" (italics in original). He later added (Rasch 1977, p. 77): "The term 'objectivity' refers to the fact that the result of any comparison of two objects within some specified frame of reference is…*independent of everything else within the frame of reference other than the two objects which are to be compared and their observed reactions"* (italics in original). The adjective "specific" qualified the concept of objectivity by restricting it to comparisons made "within some specified frame of reference," which meant that the objects should belong to the same class and be compared by means of the same instrument.

Table 1. The observable outcomes for comparing the joint performance of two people, A and B, who independently attempt the same task. Outcome X = 1,0 means A succeeded and B failed, and outcome Y = 0,1 means A failed and B succeeded (after Masters, 2001).

Parameters	Observation Opportunity	Observable Outcome	
		X	Y
$\theta_A - \theta_B$	Persons A and B independently attempt the same task	1,0	0,1

Rasch anticipated that the principle of specific objectivity would be applicable to problems in psychophysics and foresaw that “with the discovery of specific objectivity we have arrived at concepts of such generality that the original limitation is no longer justified. Extensions into other fields of psychology, such as psychological threshold experiments and experiments on perception of values, offer themselves…” (Rasch, 1966, p. 105). I turn now to such an application.

A Choice-Theory Interpretation of Specific Objectivity

Luce (1959, 1963) offered an important analysis of detection theory by applying his choice axiom to the decisions in a detection experiment. This analysis is known as choice theory (Luce, 1963; Macmillan & Creelman, 2005). In Luce’s parametization of detection theory, logistic distributions replace the Gaussian distributions of the standard version and thereby formally relate choice theory to the Rasch model. Luce’s theory has two parameters: α, a parameter for response strength ($\alpha > 1$), and υ, a parameter for response bias ($\upsilon > 0$). DeCarlo (2005, Equations 1 and 5) has shown how response strength can be identified with psychometric ability (for example, the ability to discriminate between latent classes of essays), and response bias with rated difficulty (interpreted, for example, as the tendency to be strict or lenient in grading essays). Here I show how those parameters can be identified with the difference in ability between two people who attempt the same task. My intention is then to show how the parameters specify a Receiver Operating Characteristic, an operating characteristic which in turn exemplifies the principle of specific objectivity.

A Hypothetical Example

Table 2 presents the hypothetical results of an experiment in which two people attempted the same task. In this example, Person A succeeded 80 times on 100 attempts ($P_A = 0.8$), and B succeeded 30 times on 100 attempts ($P_B = 0.3$). But the number of joint occasions when A succeeded and B failed was 56, that is, $N_{10} = 56$ or $P_A \cdot (1 - P_B) = P_{10} = 0.56$. Similarly, the number of occasions when A failed and B succeeded was 6, that is, $N_{01} = 6$ or $P_B \cdot (1 - P_A) = P_{01} = 0.06$. The table also shows that the number of occasions when both A and B succeeded ($N_{11} = 24$) and when they both failed ($N_{00} = 14$).

Choice-Theory Analysis

A choice-theory interpretation of this table is shown in Figure 1. The probabilistic ability of each person is represented by logistic density functions, labelled A and B with means of θ_A and θ_B respectively. The *x*-axis represents the ability of a person, or the difficulty of a task,

Table 2. Tally of successes and failures of two people attempting the same task 100 times.

		PERSON B		
		Succeeds	Fails	Total
	Succeeds	24	56	80
PERSON A	Fails	6	14	20
	Total	30	70	100

measured in units of log odds ratio. The vertical line labelled δ locates the difficulty of a task that both people attempted. The area under curve A marked by vertical hatching is the proportion of occasions that Person A succeeded on the task (P_A = 0.8), and similarly the shaded area under curve B is the proportion of occasions that Person B succeeded on the same task (P_B = 0.3). The inset shows how P_A and P_B co-vary as a function of δ.

From Equation (2), the difference between the mean abilities of A and B, i. e., $\theta_A - \theta_B$, is equal to $\ln(P_{10}/P_{01})$ = 2.23. Or equivalently, from choice theory, $\alpha = \sqrt{(P_{10}/P_{01})}$ = 3.06, and $2\ln(\alpha)$ = 2.23. Just as the difference in ability of two people can be represented by the detection parameter $2\ln(\alpha)$, so the difficulty of a task can be represented by the response-bias parameter $\ln(\upsilon)$, where $\upsilon = \sqrt{(P_{00}/P_{11})}$, or $\ln(\upsilon) = ½\ln(P_{00}/P_{11})$, which is -0.27 units below the intersection of curves A and B in Figure 1.

Specific Objectivity and the ROC

In this psychometric analysis, the operating characteristic represents the difference in ability between two people, A and B. Moreover, the location of the curve is independent of the difficulty of the task attempted, or indeed independent of the difficulty of any task attempted. In this way the operating characteristic of choice theory satisfies and exemplifies Rasch's principle of specific objectivity. Figure 1 illustrates the difficulty of just one task, but the operating characteristic shows the difficulty of all possible tasks. A task different in difficulty to that illustrated in Figure 1 would lie at a different operating point on the same ROC curve. Very difficult tasks – a high bar – would give rise to points near the bottom left corner of the ROC square, and very easy tasks – a low bar – to points near the top right corner.

The principle of specific objectivity means that the difference in ability between two people can be evaluated without regard to the difficulty of the task. It follows that for a psychometric test consisting of a number of homogeneous items, the estimated difference in

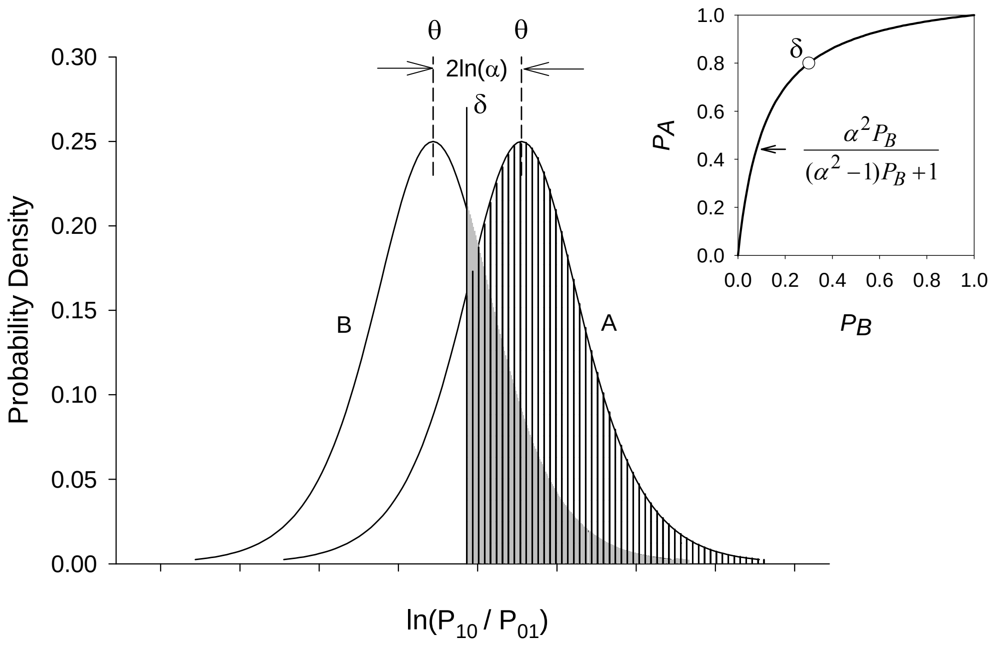

Figure 1. The performance shown in Table 2 for the difference between Persons A and B. The operating characteristic (inset) depicts how the success rates, P_A and P_B, co-vary as a function of task difficulty.

ability between two people who attempt the same test does not depend on which items are selected for the comparison, whether all of the items, or any subset of them. The same conclusion follows from the ROC interpretation of specific objectivity.

An Empirical Example

An analysis of some psychometric data collected for another purpose by K. C. Irwin and Britt (2005) provides an empirical example. In order to evaluate the effectiveness of a project designed to improve the numeracy of students, these authors administered a 21-item test of algebraic thinking to 431 12-year-old students who participated in the project and to 461 students who did not participate. Figure 2 compares the performance of each group of students. The open circles show the proportion of students in each group who passed each of the 21 items. The smooth curve is the best-fitting choice-theory operating characteristic to the data points ($2\ln(\alpha) = 0.77$). Points near the bottom left corner of the square represent difficult items and points near the top right corner represent easy items. The students found the test challenging and so there is a cluster of items near the bottom left corner of the ROC square. In contrast, few items lie near the top right corner of the square.

In so far as the operating characteristic in Figure 2 provides a satisfactory fit to the data, the test of algebraic thinking satisfies Rasch's principle of specific objectivity. This is because the estimated difference in achievement between the two groups ($2\ln(\alpha) = 0.77$) is independent of the difficulty of any particular item or any subset of items of the test. However, the fit is not entirely satisfactory (the data are significantly different from the model according to Hosmer and Lemeshow's test: $\chi^2 = 28.59$, $df = 8$, $p < .01$), and to that extent one may conclude that the test does not exhibit specific objectivity.

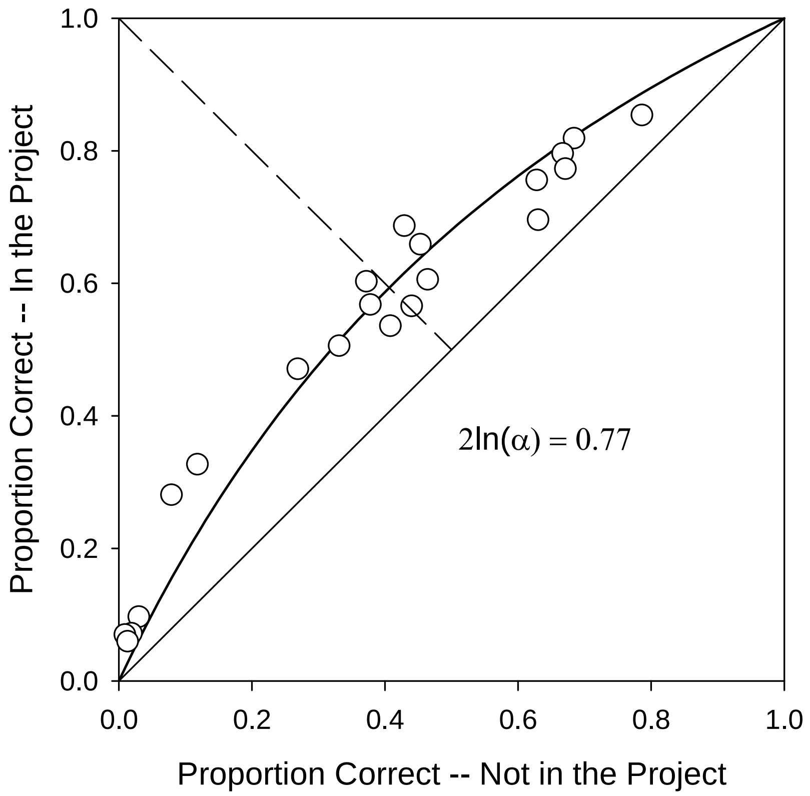

Figure 2. Proportion correct on each of 21 items of a test of algebraic thinking by students who participated in a numeracy project and by those who did not (data from Irwin & Britt, 2005). The smooth curve is the best-fitting choice-theory ROC.

Conclusions

Rasch's principle of specific objectivity can be interpreted by means of the psychophysical ROC of choice theory. According to this interpretation, the estimated difference between the ability of two people can be identified with the choice-theory index of discriminability, $2\ln(\alpha)$, which determines the ROC's location, and the difficulty of a task can be identified with the choice-theory parameter for response bias, $\ln(\upsilon)$, which determines an operating point. Furthermore this ROC represents the difference in ability of two people independently of any particular difficulty of a task because it represents all possible difficulties. The ROC thus exemplifies specific objectivity.

Showing a connection between Rasch's principle of specific objectivity and choice-theory's ROC may contribute to integrating psychometrics and psychology. It further demonstrates the importance of the ROC — one of the most widely used concepts in psychology and medicine. It may also help fulfil Rasch's aspiration that the principle of specific objectivity would apply to other branches of psychology.

References

DeCarlo, L. T. (2005). A model of rater behavior in essay grading based on signal detection theory. *Journal of Educational Measurement*, *42*, 53-76.

Irwin, K. C., & Britt, M. S. (2005). The algebraic nature of students' numerical manipulation in the New Zealand Numeracy Project. *Educational Studies in Mathematics*, *58*, 169-188.

Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (vol. 1, pp. 103-189). New York: Wiley.

Macmillan, N. A. & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, N. J.: Erlbaum.

Masters, G. N. (2001). The key to objective measurement. IPO Research Memorandum, No. 70.

Rasch, G. (1966). An individualistic approach to item analysis. In P. F. Lazarsfeld & N. W. Henry (Eds.), *Readings in mathematical social science* (pp. 89-107.) Cambridge, MA: MIT Press.

Rasch, G. (1977). On specific objectivity: an attempt at formalising the request for generality and validity of scientific statements. *Danish Yearbook of Philosophy*, *14*, 58-94.