An Extension of the Concept of Specific Objectivity*

Hans Irtel
University of Mannheim

Abstract

Comparisons of subjects are specifically objective if they do not depend on the items involved. Such comparisons are not restricted to the 1-parameter logistic latent trait model, but may also be defined within ordinal independence models and even within the 2-parameter logistic model.

KEYWORDS: Latent trait theory, specific objectivity.

The notion of specific objectivity was introduced by Rasch (1960, 1977) to characterize scientific statements about relations between objects in situations, where observations involve instruments which might interfere with an observation's result. Applied to psychological testing, specific objectivity requires that comparisons of subject abilities should in a certain sense (specified below) be independent of the items used to determine the abilities. Specific objectivity within latent trait models is often believed to necessarily enforce the 1-parameter logistic model (Baker, 1992). The purpose of this note is to show that specifically objective comparisons may also be based on substantially different models. We first present a slightly modified version of Rasch's (1977) definition.

Definition 1. Let A be a set of subjects and V be a set of items. Let P be a mapping of $A \times V$ into a set of outcomes . Then $= \langle A, V, P \rangle$ is called a *frame of reference.* Let C be a function mapping all pairs of values of P into a set such that for every quadrupel $(a, v; b, w) \in (A \times V)^2$ the value $C[P(a, v), P(b, w)]$ is defined. Then C is called a *comparing function.* A comparing function is called *specifically objective within the frame of reference* if $C[P(a, v), P(b, v)]$ is independent of v for all a, b in A.

Rasch comments on his definition as follows (the notation has been adapted):

> The term "objectivity" refers to the fact that the results of any comparison of two objects within A is independent of the choice of the agent v within V and also of the other elements in the collection of objects A; in other words: *independent of everything else within the frame of reference than the two objects which are to be compared and their observed reactions.*

*Accepted for publication by *Psychometrika*

> And the qualification "specific" is added because *the objectivity of these comparisons is restricted to the frame of reference* [...]
>
> This also makes clear that the *specific objectivity is not an absolute concept, it is related to the specified frame of reference* (Rasch, 1977, p. 77).

Rasch gives several examples of specifically objective comparing functions. However, they all involve real valued reaction and comparing functions defined on the set of real numbers with "convenient" mathematical properties. These are continuity of the functions and their first derivatives with respect to each argument. Furthermore all examples assume that the elements a in A and v in V are characterized by scalar parameters θ_a and ε_v. He shows that if comparing functions of this type are specifically objective within such a frame of reference then the reaction function is *latently additive* in the sense that there exist strictly monotone Transformations θ', ε', and P' of θ, ε, and P such that $P'(a,v) = \theta'_a+\varepsilon'_v$. A complete proof for this result is given by Fischer (1987) using continuity and strict monotonicity and invertibility with respect to each argument.

The best known example for a specifically objective comparing function is derived from Rasch's 1-parameter logistic model for the response of an examinee to an item. The model states that there exist a subject parameter θ and an item parameter ε such that the probability $P(a,v)$ for a correct response of subject a to item v is given by

$$P(a,v) = \frac{\exp(\theta_a - \varepsilon_v)}{1+\exp(\theta_a - \varepsilon_v)}. \tag{1}$$

It is easy to verify that in this case the comparing function

$$C_R[P(a,v), P(b,v)] := \frac{P(a,v)}{1-P(a,v)} \, \frac{1-P(b,v)}{P(b,v)}$$

is independent of v and thus is specifically objective within the given frame of reference. Fischer (1987) even shows, that given the assumptions for real valued reaction and comparing functions mentioned previously, and also assuming local stochastic independence and that the comparing function only depends on the probability for some response event relating two subjects and a single item, then specific objectivity implies (1).

However, Rasch's (1977) examples and Fischer's (1987) proof must not be interpreted as implying that (1) is the only latent trait model allowing specifically objective comparing functions. Rasch (1977) explicitly states that neither the range of the reaction function P nor the range of the comparing function C have to be sets of numbers. Both may well be based on qualitative comparisons only. Examples for latent trait models based on qualitative comparisons are the ordinal independence models presented by Mokken (1971) and Irtel and Schmalhofer (1982). Assume that instead of (1) the functions θ and ε are such that

$$P(a,v) \geq P(b,v) \quad \text{iff} \quad \theta_a \geq \theta_b,$$

then it is easy to show that the comparing function defined by

$$C_O[P(a,v),P(b,v)] \ := \ \begin{cases} 1 \text{ if } P(a,v) \geq P(b,v) \\ 0 \text{ else} \end{cases}$$

is independent of v and thus is specifically objective within the given frame of reference.

This binary comparing function exactly matches Rasch's definition of specific objectivity given above. There are, however, good reasons to consider this definition too narrow. Consider a statement like "The ratio of the difference between today's temperature T_a and the melting point of ice T_b and the difference between yesterday's temperature T_c and the melting point of ice is 2": $(T_a - T_b)/(T_c - T_b) = 2$. This is a meaningful and objectively verifyable statement for an interval scale of temperature and is independent of any proper device used to measure temperature. It does, however, not fit into Rasch's definition of specific objectivity, because it involves 3 objects not only two. In fact, Rasch's definition excludes all quantitative statements about interval scales since these are invariant under admissible transformations only if at least 3 objects are involved.

So it seems appropriate to extend Definition 1 to comparing functions

$$C[P(a,v),P(b,w),P(c,x),\ldots]$$

of more than two arguments. This also allows non-symmetric situations, where a comparing function is specifically objective for one of the components of P but not for the other. An example for such a comparing function is derivable from the well known 2-parameter logistic model introduced by Birnbaum (1968). In addition to the parameters θ and ε of (1), this model introduces a second item parameter α such that

$$P(a,v) = \frac{\exp[\alpha_v(\theta_a - \varepsilon_v)]}{1 + \exp[\alpha_v(\theta_a - \varepsilon_v)]}.$$

It is easy to verify that with $L(a,v) := \ln\{P(a,v)/[1 - P(a,v)]\}$ the values of the comparing function

$$C_{B_1}[P(a,v),P(b,v),P(c,v)] := \frac{L(a,v) - L(b,v)}{L(c,v) - L(b,v)}$$

are independent of v for all a, b, c in A for which they are defined. Thus C_{B_1} is specifically objective for A within the given frame of reference. It allows item independent comparisons of 3 subjects. Furthermore there also are subject independent comparing functions for items: Whenever they are defined

$$C_{B_2}[P(a,v),P(b,v),P(a,w),P(b,w)] := \frac{L(a,v) - L(b,v)}{L(a,w) - L(b,w)}$$

and

$$C_{B_3}[P(a,v),P(b,v),P(a,w),P(b,w),P(a,x),P(b,x)] :=$$

$$\frac{\dfrac{L(a,v)}{L(a,v)-L(b,v)} - \dfrac{L(a,w)}{L(a,w)-L(b,w)}}{\dfrac{L(a,x)}{L(a,x)-L(b,x)} - \dfrac{L(a,w)}{L(a,w)-L(b,w)}}$$

are independent of a and b in A for all v, w, and x in V.

Specific objectivity as defined in Def. 1 and extended later is based on the response function P which is assumed to be fixed and known. Empirical applications, however, have to estimate P from empirical data. Thus the idea of specific objectivity as independence of certain statements about selected objects from all other objects in the given frame of reference does not necessarily extend to estimates of P. This strongly depends on the methods used to estimate the response functions. From an empirical point of view the parameter estimation problem is solved both for the 1- and the 2-parameter logistic model (Baker, 1992). Until now sufficient theoretical foundations do exist for the 1-parameter model only.

There is a strong connection between specific objectivity of comparisons and the concept of meaningfulness in measurement theory (Irtel, 1987). This is shown by the fact that for all of the examples presented the specifically objective comparing functions correspond to meaningful expressions of scale values. For the 2-parameter logistic model we may define scale values by

$$\begin{aligned} \theta_a &:= \frac{L(a,v)-L(b,v)}{L(c,v)-L(b,v)} \\ \alpha_v &:= L(a,v)-L(b,v) \\ \varepsilon_v &:= -\frac{L(a,v)}{L(a,v)-L(b,v)}. \end{aligned}$$

With these definitions it is easy to verify that the comparing functions C_{B_1}, C_{B_2}, and C_{B_3} are expressions which are invariant under the admissible transformations of the respective scale type. For the interval scales θ and ε these are expressions of the form $(\theta_a-\theta_b)/(\theta_c-\theta_b)$ and for the ratio scale α this is α_v/α_w. Thus given a proper definition of the parameters involved, specifically objective statements are equivalent to meaningful statements in the sense of measurement theory.

References

Baker, F. B. (1992). *Item response theory: parameter estimation techiques.* New York: Marcel Dekker.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F.M. Lord, & M.R. Novick (Eds.) *Statistical theories of mental test scores.* Reading, MA: Addison-Wesley.

Fischer, G. H. (1987). Applying the principles of specific objectivity and generalizability to the measurement of change. *Psychometrika, 52*, 565–587.

Irtel, H. (1987). On specific objectivity as a concept in measurement. In E.E. Roskam und R. Suck (Eds.) *Progress in mathematical psychology* (pp. 35–45). Amsterdam: Elsevier Science Publishers.

Irtel, H., & Schmalhofer, F. (1982). Psychodiagnostik auf Ordinalskalenniveau: Meßtheoretische Grundlagen, Modelltest und Parameterschätzung. *Archiv für Psychologie, 134*, 197–218.

Mokken, R. J. (1971). *A theory and procedure of scale analysis.* Paris: Mouton.

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

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