Discriminant Distance

A New Form of Evaluating the Distance Between Variables

Jose Antonio Martínez García

Department of Management and Marketing, Technical University of Cartagena, Spain

Abstract. In this paper, a new distance measure called the *Discriminant distance* (*Dd*), which overcomes some of the shortcomings of other distance measures and provides an easily understandable geometrical interpretation of distance between two variables, is evaluated. *Dd* summarizes the information provided by Pearson's *r* and Cohens' *d* statistics in a single coefficient. In addition, confidence intervals can also be computed. The Discriminant distance has values enclosed in a [0, 1) interval and it may be mainly used to study convergent and discriminant validity, or scale invariance, when using continuous variables. A web-based computer program is provided to facilitate its computation. Finally, conventions for interpreting *Dd* are discussed along with several limitations.

Keywords: discriminant distance, distance measures, convergent and discriminant validity, scale invariance

The notion of distance is an important and widely used topic in social and behavioral sciences, specifically in marketing research. For example, the estimation of distance between populations is an integral component of several multivariate techniques for data analysis, including multidimensional scaling or discriminant analysis (Bedrick, Lapidus, & Powell, 2000). We can compute distances between vectors, variables, or individuals, to provide a measure of separation or proximity between them.

When the observed data are quantitative, there are different types of distances that are used, depending on the research aims. I agree with Abdi (2007) that some of the most important distance measures are Euclidean, Mahalanobis, Minkowsky, and the Hellinger distances. For the case of studying the distance between two variables, all these distances can be computed by building two vectors of data and certain constraints expressed by a weight matrix conformable with the vectors.

Following the continuous scenario, the absolute value of Pearson's *r* correlation coefficient can also be considered as a measure of distance between variables, anchored in a [0, 1] rank. Recall that two perfectly correlated variables are virtually the same variable from a statistical viewpoint,[1] because their degree of association is maximum (or the distance between them is minimum, i.e., 0), and just the opposite; two uncorrelated variables are not associated, that is, their association is minimum (or the distance is maximum, i.e., 1). Furthermore, the absolute values of the family of the standardized mean difference effect sizes, such as Cohen's *d*, are also distance measures, because when two variables have the same mean, their distance is 0. Conversely, two variables with divergent mean are separated to the extent that means differ (weighted by standard deviation). In this case, the upper bound of the separation is infinite. Therefore, extending the concept of distance to *r* and *d*, we can see how distance is a central topic in hypothesis testing and validation procedures, such as the analysis of convergent and discriminant validity.

The aforementioned distance measures report different information regarding proximity or separation between two variables, but as I will show later, they fail to properly take into account both mean difference and correlation between them. In this paper, I have developed a new distance measure that I have called the *Discriminant distance* (*Dd*), which overcomes some of the shortcomings of other distance measures, and provides an easily understandable geometrical interpretation of distance between two variables. The *Dd* has values enclosed in a [0, 1) interval and it may be mainly used to study convergent and discriminant validity, or scale invariance, when using continuous variables.

Illustration of Shortcomings of Some Distance Measures

We begin by showing some shortcomings of the most commonly used distance measures with the illustration of a specific example. Suppose that a researcher is interested in knowing the proximity or redundancy of two variables in a questionnaire (A and B). These two variables are measured using 5-point Likert-type scales (from 1 to 5).

[1] We will discuss later that two perfectly correlated variables may behave equally for certain statistical analyses, but this does not mean that they are the same variables, because mean values may diverge.

DOI: 10.1027/1614-2241.5.4.137

Table 1. Artificial data to illustrate the shortcomings of some distance measures

	1	2	3	4	5	6	7	8	9	10	$\overline{X}$	σ	Cov(*a*, *b*)
Variable *A*	4	5	4	5	4	4	4	4	3	4	4.1	0.56	.29
Variable *B*	3	4	3	4	3	3	3	3	2	4	3.2	0.63	

Ed	*swEd*	*dwEd*	*Md*	Correlation coefficient *r* (95% CI)	Cohens' *d* (95% CI)
3.00	0.300	4.99	50	.87 (0.52, 0.97)	1.50 (0.99, 2.0)

Let *a* and *b* be two vectors with *J* elements each, indicating the sample responses to items representing both variables. In this example, $J = 10$. Table 1 depicts the data.

Euclidean Distance

The first option would be the computation of *Euclidean distance* (*Ed*), which is defined as

$$Ed(a,b) = \left[(a-b)^T(a-b)\right]^{1/2} = \left[\sum_j (a_j - b_j)^2\right]^{1/2}. \tag{1}$$

This distance measure does not provide much useful information because it is fully dependent on the sample size and the magnitude of the variables. However, we can build two additional relative *Ed*s. The first is the *sample weighted Euclidean distance* (*swEd*), which weights *Ed* by the sample size

$$swEd(a,b) = \left[\frac{\sum_j (a_j - b_j)^2}{J}\right]^{1/2}. \tag{2}$$

The second is the *deviation weighted Euclidean distance* (*dwEd*), which weights *Ed* by the pooled standard deviation (σ_p) of both variables.

$$dwEd(a,b) = \frac{\left[\sum_j (a_j - b_j)^2\right]^{1/2}}{\sigma_p}. \tag{3}$$

However, these modifications of the *Ed* do not take into account the correlation between variables. This consideration is crucial when using paired data.

Minkowski's distance and Hellinger's distance are computed in a similar fashion (see Abdi, 2007); therefore, they do not take into account the correlations either.

Mahalanobis Distance

Mahalanobis distance (*Md*) takes into account the covariance of data. As Abdi (2007) explains, *Md* is defined between rows of a table. The weight matrix *W* is obtained as the inverse of the columns of variance-covariance matrix. Therefore, we can denote by *S* the variance-covariance matrix between the columns of Table 1, being $W = S^{-1}$. In this case, S^{-1} is a 10×10 symmetric matrix.

$$Md(a,b) = \left[(a-b)^T S^{-1}(a-b)\right]^{1/2}. \tag{4}$$

Md weighs raw difference scores between variables by a covariance matrix computed from these pairs of scores. Therefore, it does not consider the covariance between variables, Cov(*a, b*) (it would be the rows instead of the columns), so correlation between A and B is not taken into account.

Pearson's *r* Correlation Coefficient

Correlation coefficient r indicates the strength and direction of a linear relationship between two random variables and it ranges from −1 to 1. Sample correlation is computed with the much known expression

$$r(a,b) = \frac{\mathrm{Cov}(a,b)}{\sigma_a \sigma_b}. \tag{5}$$

As *r* is a sample statistic, we may compute a confidence interval for the estimated population parameter. Several methods have been proposed for that calculus (see Beasley et al., 2007), for example, the Fisher *Z* transformation or bootstrapping.

In our example, both variables are highly correlated (.87), so their association is strong and their distance is small. However, both variables behave divergently if we consider the mean values. Therefore, *r* is an incomplete distance measure because it may suggest that two variables provide the same information or they are redundant when actually their mean values disagree.

Cohen's *d* Effect Size

Cohen's (1988) *d* statistic is defined as the difference between two mean values weighted by a standard deviation. As Cepeda, Pashler, Vul, Wixted, and Rohrer (2006) indicate, the choice of standard deviation is crucial, as it impacts observed effect size. Statisticians differ on the optimal type of standard deviation to use in computing *d*, although the most used criterzia is to average the two standard deviations (σ_{pooled}).

$$d(a,b) = \frac{\overline{X}a - \overline{X}b}{\sigma_{\text{pooled}}}. \quad (6)$$

Like r, d is a sample statistic and confidence intervals can be computed, for example, using bootstrapping procedures (Algina, Keselman, & Penfield, 2005) or approximate methods (Hedges & Olkin, 1985). Reporting confidence intervals for r and d avoid making misleading inferences about the magnitude of the effect due to overestimation (Voelkle, Ackerman, & Wittmann, 2007).

The problem of d is that it does not consider the covariation between variables. Therefore, only by knowing d, we would lose information regarding the proximity or separation of variables; hence, again this is an incomplete distance measure.

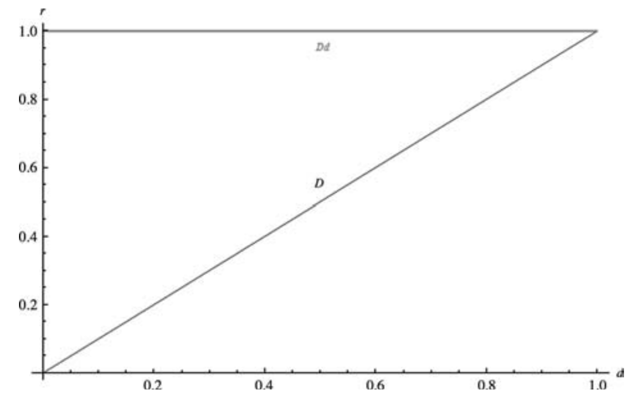

Figure 1. Cartesian two-dimensional coordinate system.

The *Dd*

We propose a new form of evaluating distance between variables considering both mean difference and correlation. The rationale of *Dd* comes from a straightforward premise: *Dd between two variables is minimum (0) if* $|r| = 1$ *and* $d = 0$, *and Dd is maximum (1), if* $|r| = 0$ *and* $d \rightarrow \infty$.

Therefore, two variables with $Dd = 0$ provide exactly the same information, so they are completely redundant.[2] An advantage of *Dd* is that its absolute value ranges from 0 to 1. In addition, *Dd* takes into account the precision of r and d statistics,[3] so confidence intervals are considered. Finally, *Dd* provides an easily understandable geometrical interpretation using the Cartesian two-dimensional coordinate system.

We depict the computation of *Dd* as follows.

Building the Two-Dimensional Plane

The first stage in *Dd* development is to build the Cartesian two-dimensional coordinate system (Figure 1).

Horizontal lower axis represents d and vertical axis represents r. Both variables should be ranged in a $(-1, 1)$ interval. Minimum distance between variables will occur when correlation has the highest value and standardized mean difference has the lowest. However, it would be easier to geometrically interpret the distance if minimum distance occurs when both variables have the lowest value, so $|r|$ has to be transformed. In addition, upper bound of $|d|$ is infinite, so it would also be necessary to transform $|d|$ into a score with an upper bound equal to (or asymptotically equal to) 1.

The diagonal D of the square is the segment that represents the highest distance between both variables. At point (1,1), the segment has the highest length, so *Dd* is maximum; on the contrary, at point (0,0), the segment has the lowest length, so *Dd* is minimum.

Horizontal higher axis represents the orthogonal projection of each point of D on a horizontal segment of length 1. *Dd* is the length of the segment that comes from the origin of the axis to the projection of a point located in D.

Pearson's *r* and Cohen's *d* Transformation

First of all, $|r|$ has to be transformed to change its direction. With the aim to interpret the minimum distance when $|r|$ and $|d|$ have the lowest value, I have to compute r' as

$$r'(a,b) = 1 - |r|. \quad (7)$$

Cohen (1988) made very general recommendations regarding how to evaluate the magnitude of the effect size in social sciences. To evaluate the strength of association between variables, Cohen's conventions indicate that small, medium, and large effects occur when $|r|$ is .1, .3, and .5, respectively. Therefore, the equivalent conventions for r' are .9, .7, and .5, respectively. Then, a small effect corresponds to a high value of r' (around .9), and a large effect corresponds to a low value of r' (around and below .5). Consequently, to the extent that r' is close to 1, the distance between variables increases, and to the extent that r' is close to 0, the distance decreases.

In a second step, I have to transform d to d', d' being a score ranged in a [0, 1) interval. I have chosen the χ distribution with v degrees of freedom, that is the distribution followed by the square root of a χ^2 random variable.

The reason for choosing this transformation is based on two premises.

First, I needed a function whose image was enclosed in a [0, 1) interval. Several cumulative distribution functions pertaining to the family of continuous distributions matched with that requirement, such as half-normal, half-logistic, Rayleigh, χ, Maxwell, and others.

Second, I needed to consider Cohen's conventions for d statistic to relate the magnitude of effects between r' and d'.

[2] Note that variance, skewness, and kurtosis of both variables are the same.
[3] Scale coarseness can be also considered by correcting r (see Aguinis, Pierce, & Culpepper, in press) and measurement error by correcting r and d by reliability (Hunter & Schmidt, 2004).

Table 2. MSE after adjusting some cumulative distribution functions[a]

d	d' (should be)	$\chi(v)$	Half-normal (θ)	Gamma (α, β)	Rayleigh (σ_1)	Exponential (λ)	Maxwell (σ_2)
0.2	0.1	0.104	0.067	0.115	0.0424	0.159	0.0146
0.5	0.3	0.303	0.325	0.26	0.237	0.229	0.18
0.8	0.5	0.500	0.500	0.500	0.500	0.500	0.500
MSE		2.82E−05	0.0017	0.0018	0.0072	0.0085	0.0216

[a]Parameters: $v = 1.21$; $\theta = 1.05$; $\alpha = 0.8$; $\beta = 1.59$; $\sigma_1 = 0.68$; $\lambda = 0.866$; $\sigma_2 = 0.52$.

Cohen's conventions for d indicate that small, medium, and large effects occur when effect size is 0.2, 0.5, and 0.8, respectively. Therefore, it was desirable that a large effect in r and a large effect in d matched at the same point of the Cartesian system. In addition, with the aim to put at the same level the effects on the two Cartesian axes, conventions for effect size in d' should take up the same space over its [0, 1) domain as the conventions for r. Consequently, the required function to use for transforming d to d' should be adjusted to minimize the mean square error (*MSE*) between the transformed effect size conventions for d and r (Table 2).

Functions were calibrated fixing a matched reference point (e.g., (0.8,0.5)) with three decimals of precision, using the quantile procedure in Mathematica 6.0. This process facilitates the identification of parameters needed to define the functions.

χ distribution with 1.21 degrees of freedom yielded the lowest *MSE*, so it was the selected final function. Because I was only interested in searching a function with a required form, there is no problem to admit a noninteger value for degrees of freedom.

Figure 2 shows the form of the selected χ function. Note that this function asymptotically gets closer to the maximum (in this case 1) to the extent that d increases. For example, the final 10% of the distribution of values of the function covers from 1.77 to more. This means that very large effect sizes count in a similar way; they yield values of d' above 0.9, that is, a great distance between variables. In addition, 50% of the distribution of values of d' is above 0.5 (large effect), such as the distribution of r.

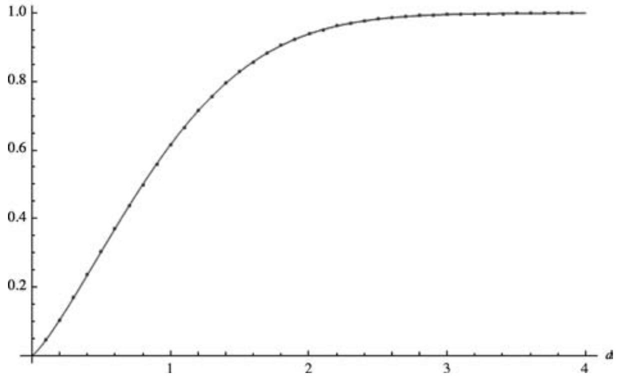

Figure 2. Chi cumulative distribution function with $v = 1.21$.

Point Estimation of *Dd* and Confidence Intervals

Point estimation of Dd comes from point estimates of d' and r' that form the point $P'(d',r')$ on the coordinate system. $P'(d',r')$ is a transformation of the original point $P(d,r)$. Then, I have to obtain the segment S that joins $P'(d',r')$ with origin on (0,0). Dd is obtained from the percentage of D covered by S:

$$Dd(a,b) = S/D. \tag{9}$$

As D is an irrational number (the diagonal of a square), I can approximate its value using two decimals: $D = 1.41$.

Then, $P(d, r) = P(1.497, 0.866)$. Consequently, $P'(d', r') = P'(0.828, 0.134)$. Therefore, $S = 0.835$ and $Dd(a, b) = 0.59$.

I can compute a confidence interval over Dd, using confidence intervals of d' and r', in a similar way as I have computed the point estimate. Using some of the procedures mentioned in the beginning of the paper, I have to convert the lower and upper bounds of d and $r(d_l, d_u, r_l, r_u)$ to the transformed values (d'_l, d'_u, r'_l, r'_u). It is imperative to choose the same level of confidence for the two statistics. In addition, it is also important to note that transformation of the lower bound of $r(r_l)$ corresponds to the upper bound of $r'(r'_u)$, and conversely for the upper bound, because of the opposite direction of r'. Then, I have to situate the points $Pl'(d'_l, r'_l)$ and $Pu'(d'_u, r'_u)$ on the coordinate system. If I join these two points with origin, I obtain S_l and S_u. Finally, the values (Dd_l, Dd_u) that form the confidence interval over Dd are computed as I have shown in Equation 9.

To compute a 95% confidence interval, I calculated 95% confidence intervals for r and d, using Fisher Z transformation and Hedges and Olkin's (1985) approximate method, respectively. Table 1 shows these estimates. By means of the depicted transformation, I obtained the following points: $Pl'(0.61, 0.03)$ and $Pu'(0.94, 0.48)$; therefore, $S_l = 0.61$ and $S_u = 1.05$. In the end, $Dd_l = 0.43$ and $Dd_u = 0.75$.

Figure 3 graphically illustrates the procedure.

I developed a user-friendly computer program that allows researchers to calculate Dd from r and d. The program is available free of charge and can be executed online by visiting the following website www.upct.es/~beside/jose. As an illustration, Figure 4 includes a screen shot of the program. Users input the observed Pearson's correlation and Cohens' effect size together with their respective confidence intervals.

Finally, $Dd(a, b)$ matches with the properties of distance measures (Abdi, 2007):

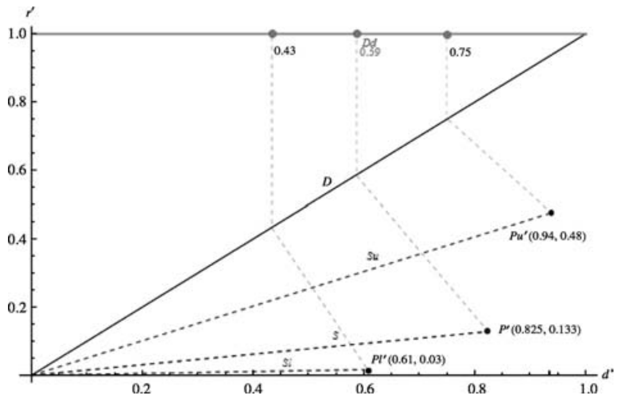

Figure 3. Graphically computation of *Dd;* point estimate and 95% confidence interval.

– $Dd(a, a) = 0$
– $Dd(a, b) = Dd(b, a)$
– $Dd(a, b) \leq Dd(a, c) + Dd(c, b)$, c being another vector with J elements.

Conventions Regarding the Magnitude of *Dd*

We propose general recommendations for evaluating the magnitude of *Dd*, which come from the conventions of *r* and *d*. Point (0.5, 0.5) on *D* represents large effects for *r* and *d* (remember that $d' = 0.5$ is the transformation of $d = 0.8$). This means that there is a large standardized mean difference (high distance) between both variables, but at the same time there is a large association (low distance) between them. Therefore, I can consider its projection of *Dd*(0.50) as a medium distance. To the extent that d' decreases and r' decreases (or *r* increases), the distance between variables also decreases, because the standardized mean difference gets smaller and association gets larger. *Dd*(0.30) represents medium *d* and large *r*, so I can label this distance as a small distance. In addition, *Dd*(0.10) represents small *d* and large *r*, so I may label this combination as a very small distance.

On the contrary, the combination of medium *r* (.3) and large *d* (0.8) (or when $r' = .7$ and d' is above 0.5) means that standardized mean difference is large and association is medium; hence, following the same prior reasoning, I label this distance as large: *Dd*(0.70). Finally, when association is small ($r = .1$ or $r' = .9$) and *d* is large, the distance between variables is very large: *Dd*(0.9). Figure 5 illustrates the procedure.

Note that these recommendations are very general and are only based on Cohen's (1988) conventions. Consequently, these have to be considered as guidelines for evaluating the importance of *Dd*. Researchers are free to interpret the magnitude of *Dd* in a different fashion. However, and in a similar way to Cohen's conventions, these magnitudes can act as a general guide for social science research.

Applications

The main application of *Dd* is in the analysis of convergent and discriminant validity. For example, in the widely used one-dimensional reflective models, a latent variable often underlies a number of observed variables (items). Convergent validity between items is acknowledged if model meets

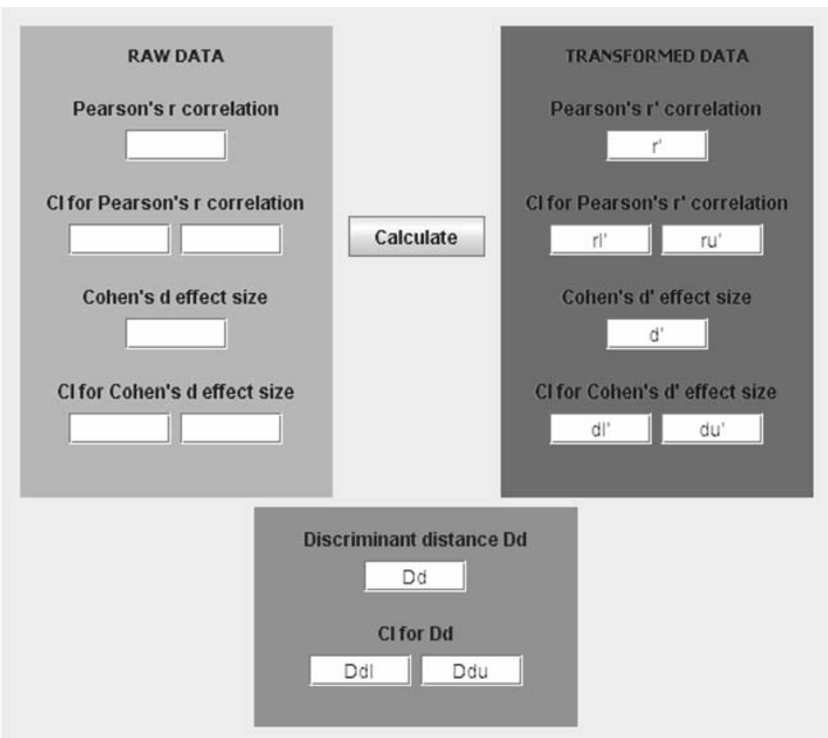

Figure 4. Screen shot of computer program that calculates *Dd.* The program is available at www.upct.es/∼beside/jose.

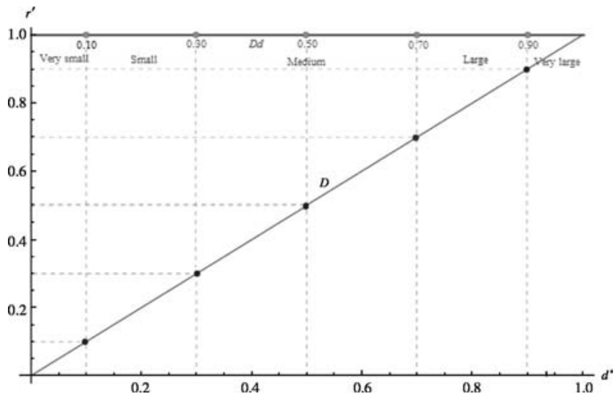

Figure 5. Recommendations for evaluating the magnitude of *Dd*.

the principle of local independence; then conditionalizing on that latent variable will render the observed variables statistically independent. This means that items should be correlated. In addition, equally reliable indicators of a latent variable are interchangeable. Therefore, each single indicator should estimate the mean, that is each indicator should not differ in its mean value from the rest. Consequently, high correlations between indicators are not enough to correctly represent the latent because if indicators are interchangeable, the selection of an indicator with a high score would distort the true mean of the latent if the remaining indicators are lowly scored.

To illustrate how *Dd* can help to analyze convergent validity, I have used simulated data. Let us suppose that I have four indicators underlying a one-dimensional latent variable, and a sample of 400 individuals. Artificial data were generated using STATA 8.0. I drew a sample from a multivariate normal population with a vector of means (3, 3, 3, 4), standard deviations (0.5, 0.5, 0.5, 0.5), all four items being associated with a correlation of .9. All variables were supposed to be measured in the same scale. Standard procedures for evaluating the adequacy of this model include analyses such as covariance structure analysis (LISREL, TETRAD differences), sample pairwise correlations, or standardized mean difference. Table 3 shows the results of these analyses.

It seems clear that Item 4 does not measure as the three remaining items. As we can see, none of the four procedures but for Cohen's *d* has been able to detect the problem. However, Cohen's *d* does not take into account the covariance of items. Recall that an effect size of about 2.1 means that more than 98% of scores of Item 4 are above the mean of the three items (see Cohen, 1988, p. 21). This could lead to think that a very large difference exists between Item 4 and the rest. Nevertheless, *Dd* considers the magnitude of this difference taking into account the correlation between items. Using the proposed conventions for *Dd*, we may say that the magnitude of distance between Item 4 and the rest is medium or medium-large. Therefore, the high correlation between them makes *Dd* not be large. This is very interesting because it indicates that what Item 4 measures is different from the rest, but is not so different as it would seem by only assessing Cohen's *d*. At this point, researchers should evaluate why this item behaves as it does; for example, this behavior can be indicative of a problem of halo effect or common method bias.

Dd can also be used to study discriminant validity between disparate measures. Using covariance structure analysis and pairwise correlations yields the same problem as before, that is, standardized mean difference is not taken into account. In addition, Cohen's *d* does not take into account the correlation. Therefore, computing *Dd* between disparate items in a questionnaire is at the same time a measure of convergent and discriminant validity (proximity or separation between them).

Another potential application of *Dd* (certainly related with discriminant validity) is the study of the proximity or separation between constructs, from individual's viewpoint. Marketing research includes studies that analyze highly related concepts such as perceived quality, satisfaction, disconfirmation, perceived value, and corporate image. Academic researchers distinguish these concepts, but often consumers do not. For example, consumers might judge under certain circumstances that perceived quality and satisfaction are the same concept (Iacobucci, Grayson, & Ostrom, 1994). Recall that all these concepts can be

Table 3. Procedures for evaluating convergent validity

	LISREL	TETRAD differences	Pairwise correlations	Cohen's *d*[a]	*Dd*
	χ^2:0.51	0	*r*(1, 2) = 0.89	*d*(1, 2) = 0.027	*Dd*(1, 2) = 0.078
	Degrees of freedom: 2		*r*(1, 3) = 0.88	*d*(1, 3) = 0.012	*Dd*(1, 3) = 0.084
			r(2, 3) = 0.89	*d*(2, 3) = 0.014	*Dd*(2, 3) = 0.077
	p value: .77		*r*(1, 4) = 0.88	*d*(1, 4) = 2.170	*Dd*(1, 4) = 0.684
	Composite reliability: 1		*r*(2, 4) = 0.89	*d*(2, 4) = 2.116	*Dd*(2, 4) = 0.678
			r(3, 4) = 0.88	*d*(3, 4) = 2.145	*Dd*(3, 4) = 0.682
	AVE: 1				
Model diagnostic	Good model	Good model	Good model	Model problematic	Model problematic
Convergent validity of items	Yes	Yes	Yes	Item 4 very problematic (very large effect size)	Item 4 problematic (medium-large distance)

[a]It is important to note that applied researchers often do not use Cohen's *d* for conducting this type of data analysis; LISREL, TETRAD, and correlations are preferred for that purpose.

Table 4. Measures

	Source
Perceived quality	
I believe this financial institution provides an excellent service	Brady and Cronin (2001)
Satisfaction	
I am satisfied with the service provided by this company	Teas (1993)
Disconfirmation	
Overall, my experience in the race was better/worse than expected	Oliver (1980)
Corporate image	
I like this bank	Andreassen and Lindestad (1998)

considered as consumer attitudes, so it is perfectly plausible that several of them are very proximal.

We have applied *Dd* to a real market research to study the proximity or separation between potentially related concepts. Research was designed to analyze the attitude of consumers toward financial entities in a town. A total sample of 207 consumers were drawn for that purpose. Several concepts were measured: perceived quality, satisfaction, disconfirmation, and corporate image. An item representing each concept was chosen, and it was measured using a 5-point Likert-type scale from strongly disagree to strongly agree (quality and satisfaction) and a semantic-differential scale (disconfirmation and corporate image) (Table 4).

I have computed *Dd* for each pair of variables. Results are shown in Table 5.

In general, the *Dd* between all these variables may be considered small. This means that all variables are very proximal from the consumer's viewpoint. Therefore, distinction in the definition of concepts made by researchers could be questionable, because consumer's responses do not seem to reflect these divergences.

Other applications of *Dd* can also be proposed, such as in the context of the analysis of scale invariance and commensurability between measures that come from different rating scales.

Scale invariance is a characteristic of objects that does not change if the length of the scale is multiplied by a constant factor. For example, given the polynomial function $f(x) = ax^k$, where a and k are constants, $f(cx) = c^k ax^k = c^k f(x)$, where c is a constant, that is, by scaling the argument of the function by a constant factor c, the function is rescaled by a constant factor c^k. If we consider that the measure coming from a Likert scale of rank R_1 is a linear function that represents the punctuation of each individual on the characteristics of interest, then $f(ax) = ax$. In addition, we may suppose that $a = 1$ because the image of x is always $f(x)$, that is, we assign a direct meaning to the scores on the scale considered. Obviously, if we measure the same phenomenon using a different rating scale with R_2, then scale invariance holds $f(cx) = cx$, where c is a function of both ranks. Using a similar reasoning, as the variance S^2 of a sampling distribution of n data is a quadratic function, it can be proved that the function $g(\sum_{i=1}^{n} f(x)) = S^2$, where $f(x) = ax^k$, $x = (x_i - \bar{x})$, $k = 2$, and $a = 1/n$. Therefore, if scale invariance exists, the rescaled variance would be $f(cx) = c^k ax^k$, that is, $g(\sum_{i=1}^{n} f(cx)) = c^2 S^2$, being S^2 the variance of the original scale.

A straightforward way for testing scale invariance between two different rating scales in the same sample of individuals would be using coefficient of correlation r and its confidence interval. In addition, one of these two variables should be rescaled to statistically compare both means. It is quickly viewed that if I use Cohen's d instead of mean difference to generalize the application for comparing any type of rating scale (standard deviation is now considered), scale invariance holds if confidence intervals of r and d include 1 and 0, respectively. Consequently, and acknowledging that *Dd* cannot be negative, scale invariance would exist if *Dd* equals zero, or in the lower bound if its confidence interval is zero. Therefore, *Dd* can be used for evaluating departures of scale invariance, taking into account a single statistic that summarizes the information that comes from r and d.

Limitations

Several issues have to be considered when applying *Dd*.

First, variables susceptible to be compared have to be measured in the same scale, because of the requisites of Cohen's d computation. Nevertheless, when using scales such as Likert-type or semantic, if both variables are measured using disparate scales, they could be transformed in the common unit "POMP" (percent of maximum possible

Table 5. *Dd* in a real market research example

r	r'	d	d'	Dd
$r(1,2) = .67$	$r'(1, 2) = .33$	$d(1, 2) = 0.32$	$d'(1, 2) = 0.18$	$Dd(1, 2) = 0.27$
$r(1,3) = .69$	$r'(1, 3) = .31$	$d(1, 3) = 0.18$	$d'(1, 3) = 0.09$	$Dd(1, 3) = 0.23$
$r(1,4) = .68$	$r'(1, 4) = .32$	$d(1, 4) = 0.27$	$d'(1, 4) = 0.15$	$Dd(1, 4) = 0.25$
$r(2,3) = .70$	$r'(2, 3) = .30$	$d(2, 4) = 0.14$	$d'(2, 4) = 0.07$	$Dd(2, 3) = 0.22$
$r(2,4) = .67$	$r'(2, 4) = .33$	$d(2, 4) = 0.07$	$d'(2, 4) = 0.03$	$Dd(2, 4) = 0.23$
$r(3,4) = .71$	$r'(3, 4) = .29$	$d(3, 4) = 0.07$	$d'(3, 4) = 0.03$	$Dd(3, 4) = 0.21$

Note. 1: Perceived quality; 2: Satisfaction; 3: Disconfirmation; and 4: Corporate image.

score – Cohen, Cohen, Aiken, & West, 1999). This transformation neither distorts r nor distorts d, avoiding misinterpretations of the effect size, yielded by the difference in rank (Höfler, 2008). In addition, when studying scale invariance, as I have commented, one of the variables has to be rescaled in the same metric as the other one.

Second, and very important, the same *Dd* may be obtained from very disparate (r; d) pairs. This means that two variables with the same *Dd* with respect to a third variable do not have to necessarily share the same correlation and standardized mean difference. *Dd* is a form of summarizing r and d magnitudes in a single coefficient, but first of all, researchers must pay attention to the raw r and d scores.

Third, when applying *Dd* to study the proximity or separation between constructs, and convergent and discriminant validity between measures, researchers have to note that *Dd* does not replace content validity assessment. An extreme example would be the computation of *Dd* for incomes and expenses of individuals per month; *Dd* could be very small in some instances, but this would not mean that income and expenses are almost the same variable, or that there would be poor discriminant validity between these two measures; they are very distinct by definition. Therefore, as I have explained, *Dd* may be a helpful coefficient to assess the proximity between concepts when their definition is not clear, and for the analysis of convergent and discriminant validity between items measuring the same concept or measuring a very similar one.

References

Abdi, H. (2007). Distance. In Neil Salkind (Ed.), *Encyclopedia of measurement and statistics* (pp. 1–10). Thousand Oaks, CA: Sage.

Aguinis, H., Pierce, C. A., & Culpepper, S. A. (in press). Scale coarseness as a methodological artifact: Correcting correlation coefficients attenuated from using coarse scales. *Organizational Research Methods*.

Algina, J., Keselman, H. J., & Penfield, R. D. (2005). An alternative to Cohen's standardized mean difference effect size: A robust parameter and confidence interval in the two independent groups case. *Psychological Methods, 10*, 317–328.

Andreassen, T. W., & Lindestad, B. (1998). Customer loyalty and complex services. *International Journal of Service Industry Management, 9*(1), 7–23.

Beasley, W. H., Deshea, L., Toothaker, L. E., Mendoza, J. L., Bard, D. E., & Rodgers, J. L. (2007). Bootstrapping to test for nonzero population correlation coefficients using univariate sampling. *Psychological Methods, 12*(4), 414–433.

Bedrick, E. J., Lapidus, J., & Powell, J. F. (2000). Estimating the Mahalanobis distance from mixed continuous and discrete data. *Biometrics, 56*, 394–401.

Brady, M. K., & Cronin, J. J. (2001). Some new thoughts on conceptualizing perceived service quality: A hierarchical approach. *Journal of Marketing, 5*, 34–49.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*, 354–380.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research, 34*, 315–346.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando: Academic Press.

Höfler, M. (2008). Translocation relative to range: A standardized index for effect intensity. *Methodology, 4*(3), 132–138.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Newbury Park, CA: Sage.

Iacobucci, D., Grayson, K., & Ostrom, A. (1994). The calculus of service quality and customer satisfaction: Theoretical and empirical differentiation and integration. In T. A. Schwartz, D. E. Bowen & S. W. Brown (Eds.), *Advances in services marketing and management* (pp. 1–67). Greenwich, CT: JAI Press.

Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing, 17*(4), 460–469.

Teas, R. (1993). Expectations, performance evaluation, and consumer's perceptions of quality. *Journal of Marketing, 57*(4), 18–34.

Voelkle, M. C., Ackerman, P. L., & Wittmann, W. W. (2007). Effect sizes and F ratios < 1. *Methodology, 3*(1), 35–46.

Jose Antonio Martínez García

Facultad de Ciencias de la Empresa
Universidad Politécnica de Cartagena
Paseo Alfonso XIII, 50
30203 Cartagena
Spain
Tel. +34 968 32 57 76
Fax +34 968 32 70 81
E-mail josean.martinez@upct.es

