Proceedings of the 90th Annual International Meeting of the Psychometric Society, Minneapolis, United States (2025).

Reminiscing Controversies over Joint Graphical Display in Quantification

S. Nishisato *

Ontario Institute for Studies in Education, University of Toronto, Toronto, Ontario, Canada
*Corresponding author. Email: shizuhiko.nishisato@utoronto.ca

Abstract

Graphical display of quantification results was popularized mostly by French researchers (see Benzécri *et al*., 1973) under the name of *French plot* or *correspondence plot*. Nishisato (1980) sheds some light on its mathematical foundation, indicating that the correspondence plot was only an approximation to the exact joint plot of rows and columns in half the space. This view supported the warning by Lebart, Morineau & Tabard (1977) that the row-column distance in the correspondence plot is not exact. This problem appeared totally forgotten when Greenacre (1984) popularized the correspondence plot. Carroll, Green & Schaffer (1986) proposed the so-called *CGS scaling* to remedy the inaccuracy of the plot which, however, was critically examined by Greenacre (1989). Both Greenacre and the CGS scaling group, however, missed a fundamental aspect of the contingency table analysis that we must double the quantification space to arrive at the exact mathematical description of row-column relations. The paper examines limitations of the correspondence plot and introduces the canonical plot as a mathematically exact representation of row–column relations *canonical plot*, a new name proposed at this conference (Nishisato, 2025).

Keywords: dual scaling, correspondence analysis, singular vs eigenvalue decomposition

1. Kretschmer's Typology Data

Let us illustrate the main points of the paper, using numerical data collected by German psychiatrist Kretschmer (1925) from his patients on their mental types and body types (Table 1).

Table 1. Kretschmer's Typology Data

	Pyknic	Leptosomatic	Athletic	Dysplastic	Others	Total
Manic	879	261	91	15	114	1360
Schizo	717	2632	884	549	450	5232
Epileptic	83	378	435	444	166	1506
Total	1679	3271	1410	1008	730	8098

Note: Manic = Manic-Depressive; Schizo = Schizophrenic

2. Contingency Table Analysis

When this 3×5 contingency table is subjected to quantification (see Nishisato, 1980; Greenacre, 1984; Gifi, 1990, Beh & Lombardo, 2014), we obtain the two-dimensional singular-value decomposition as in Table 2 and its two-dimensional correspondence plot as in Figure 1.

Table 2. Two-Dimensional Coordinates of Mental and Body Types

Type	Variable	Dimension 1	Dimension 2
Mental (M)	Manic-Depressive	-1.09	0.16
	Schizophrenic	0.14	-0.18
	Epileptic	0.50	0.48
Body (B)	Pyknic	-0.96	0.11
	Leptosomatic	0.16	-0.29
	Athletic	0.33	0.18
	Dysplastic	0.55	0.45
	Others	0.06	0.09
Singular Value	ρ	0.508	0.261
Angle θ between	M-axis and B-axis	12.9°	3.4°

3. Popular Correspondence Plot

Let us examine the correspondence plot of the quantification results, using those coordinates on two axes. Our first glance at it takes us to the common-sense conclusion that it is a reasonable graph. It looks like a kind of configuration of variables that we can easily expect and interpret.

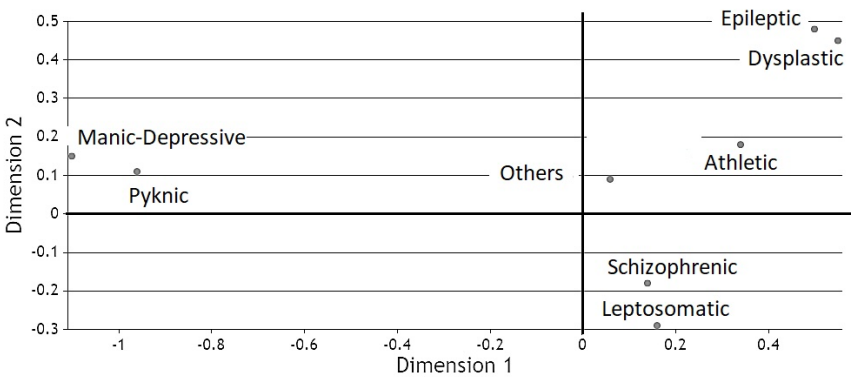

Figure 1. Correspondence Plot of Component 1 Against Component 2

From the correspondence plot, we can roughly summarize the following relations between mental types and physical types.

- "Manic depressive" is close to "pyknic" (chubby).
- "Epileptic" is close to "dysplastic" (unbalanced).
- "Schizophrenic" is close to "leptosomatic" (lean).
- "Others" does not seem to be associated with any mental types.

The above interpretation appears to meet our common-sense expectation. Why do we say then that the correspondence plot is not exact?

4. Problems with Correspondence Plot

The two coordinates are determined so as to maximize the association between mental types and body types. As we know, this task leads us to the singular-value decomposition of the contingency table with two singular values. The singular values, ρ, are equal to the maximized product-moment correlation between the mental and body types. Those correlation coefficients of the two components are, respectively, $\rho_1 = 0.508$ and $\rho_2 = 0.261$, which can be translated into the angles θ between the row vector *mental types* and the column vector *body types* (Nishisato & Clavel, 2003) by

$$\theta_j = \cos^{-1} \rho_j$$

Table 2 shows the angles between the two axes, one for the mental types and the other for body types, and we now know that the angle between the row and the column variables for each component is definitely not zero, as we saw above. This means that the correlation between rows (mental types) and columns (body types) is far from 1.00 for each component, the *condition implied by the correspondence plot!* We should realize here that two non-perfectly correlated variables require two dimensions.

We should thus remember that these non-negligible angles between rows and columns for components are treated as 0 degree (*i.e.*, unidimensional) in the correspondence plot, in other words, as if the row-column correlation is 1 for each component. Nevertheless most researchers ignore this over-simplification depicted by the correspondence plot as logically correct or practical.

To demonstrate this illogical problem of the correspondence plot in a simple way, we can eliminate one row variable from Kretschmer's data as in Table 3 and subject the data to quantification with the results as in Table 4.

Table 3. Kretschmer's Typology Data, with One Row Eliminated

	Pyknic	Leptosomatic	Athletic	Dysplastic	Others	Total
Manic	879	261	91	15	114	1360
Schizo	717	2632	884	549	450	5232

Note: Manic = Manic-Depressive; Schizo = Schizophrenic

Table 4. One-Dimensional Coordinates of Mental and Body Types

Type	Variable	Dimension 1
	Manic Depressive	-0.96
Mental (M)	Schizophrenic	0.25
	Pyknic	-0.85
	Leptosomatic	0.29
Body (B)	Athletic	0.28
	Dysplastic	0.44
	Others	0.01
Singular Value	ρ	0.490
Angle θ between	M-axis and B-axis	60.65°

For this reduced data set, we obtain one component, meaning that rows and columns are located on a straight axis, thus perfectly correlated. Yet, the singular-value is not 1, but 0.490, meaning that we need two-dimensional space for this data set. The correspondence plot depicts this as unidimensional. This result is not consistent with the underlying formulation. The correlation between the rows and the columns is not 1 but only 0.490, thus spanning a two-dimensional space!

To avoid this inaccuracy of the correspondence plot, Nishisato (1980) allocated one full chapter of his 1980 book to some mathematical handling of the information contained in the contingency table by re-writing it into the response-pattern table.

5. Response-Pattern Format

The above contingency table can be represented by the response-pattern table (see Nishisato, 1980), as given in Table 5. Note that the rows and the columns of the contingency table are now occupying only columns. Let us quantify the response-pattern table and look at the results (Table 6). This is the eigenvalue decomposition of the data.

Table 5. Kretschmer's Data in Response-Pattern Format

Pat	MD	Sch	Epi	Pyk	Lept	Ath	Dys	Oth
1	879	0	0	879	0	0	0	0
2	261	0	0	0	261	0	0	0
3	91	0	0	0	0	91	0	0
4	15	0	0	0	0	0	15	0
5	114	0	0	0	0	0	0	114
6	0	717	0	717	0	0	0	0
7	0	2632	0	0	2632	0	0	0
8	0	884	0	0	0	884	0	0
9	0	549	0	0	0	0	549	0
10	0	450	0	0	0	0	0	450
11	0	0	83	83	0	0	0	0
12	0	0	378	0	378	0	0	0
13	0	0	435	0	0	435	0	0
14	0	0	444	0	0	0	444	0
15	0	0	166	0	0	0	0	166

Note: Pat = Pattern; MD = Manic-Depressive; Sch = Schizophrenic; Epi = Epileptic; Pyk = Pyknic ; Lept = Leptosomatic; Ath = Athletic; Dys = dysplastic; Oth = Others.

The quantification of the response-pattern table yields six components as shown in Table 6. When the number of rows and that of columns are different, we see components with the eigenvalues being 0.5. As Nishisato (1980) shows, these components have row-column correlations of zero, thus irrelevant to the interpretation of the data structure. Once we eliminate those components from the outputs, we obtain four components (Table 7). The data require four dimensions to exhaustively explain the correlational structure of the data.

Nishisato (2019) introduced an important concept of *dual subspace*, which spans the two-dimensional space with the sum of the two eigenvalues equal to 1. For the four-component data, we have two such subspaces as we can see in Table 7.

Here we can identify the following sets of dual subspace:

- **Dual subspace 1** = Components 1 and 6: $\rho_1^2 + \rho_6^2 = 0.75 + 0.25 = 1$
- **Dual subspace 2** = Components 2 and 5: $\rho_2^2 + \rho_5^2 = 0.63 + 0.37 = 1$

Other components 3 and 4 are said to belong to *Residual Space*, created by the imbalance of the $m \times n$ contingency table and they do not contribute to the interpretation of the row-column relations. The eigenvalues in residual space are all equal to 0.5. If $m = n$ the dimension of the residual space is nil.

Table 6. Six-Dimensional Coordinates of Mental and Body Types

Type		Com1	Com2	Com3	Com4	Com5	Com6
	MD	1.86	0.48	0.00	0.00	0.37	1.06
Mental	Sch	-0.24	-0.55	0.00	0.00	-0.42	-0.14
	Epi	-0.86	1.47	0.00	0.00	1.12	-0.49
	Pyk	1.63	0.34	-0.28	-0.04	-0.26	-0.93
	Lep	-0.28	-0.89	-0.34	0.18	0.68	0.16
Body	Ath	-0.58	0.54	0.89	-1.92	-0.41	0.33
	Dys	-0.94	1.36	-1.44	1.28	-1.04	0.54
	Oth	-0.10	0.27	2.45	1.24	-0.21	0.06
	ρ^2	0.75	0.63	0.50	0.50	0.37	0.25

Note: Pat = Pattern; MD = Manic-Depressive; Sch = Schizophrenic; Epi = Epileptic; Pyk = Pyknic ; Lept = Leptosomatic; Ath = Athletic; Dys = dysplastic; Oth = Others.

Table 7. Kretschmer's Data: Dual Subspace

Component	Dual 1	Sub 1 6	Dual 2	Sub 2 5
Manic - Depressive	1.86	1.06	0.48	0.37
Schizophrenic	-0.24	-0.14	-0.55	-0.42
Epileptic	-0.86	-0.49	1.47	1.12
Pyknic	1.63	-0.93	0.34	-0.26
Leptosomatic	-0.28	0.16	-0.89	0.68
Athletic	-0.58	0.33	0.54	-0.41
Dysplastic	-0.94	0.54	1.36	-1.04
Others	-0.10	0.06	0.27	-0.21
Eigenvalue ρ^2	0.75	0.25	0.63	0.37

Let us now look at two two-dimensional graphs of the two sets of coordinates in each dual subspace. The rearranged table of coordinates in dual subspace is as in Table 7 and the corresponding graphs are as in Figures 3 and 4 together with the two-dimensional correspondence plot in Figure 2. Note that in each of the two sets of dual subspace, all the row variables lie on one straight line (row-variable axis) and the other on the other straight line (column-variable axis), crossed with the angle associated with the arc cosine of the singular value. Is this not what we learned in statistics, that the two correlated variables can be depicted by two straight lines in 2-dimensional graph? Note that each component in the correspondence plot is not unidimensional, but two-dimensional as we see in each dual subspace plot. In other words, the 2-dimensional correspondence plot is in fact 4-dimensional! Two components in each dual subspace correspond to a single component of the correspondence plot!

6. Discussion

The purpose of the current paper is to overview the history of the currently popular correspondence plot and present a mathematically correct graphical display of correlated variables. To follow our discussion, let us look at our example. The first graph in Figure 2 is the traditional, currently most popular correspondence plot. Notice that, through the use of common sense, the interpretation of

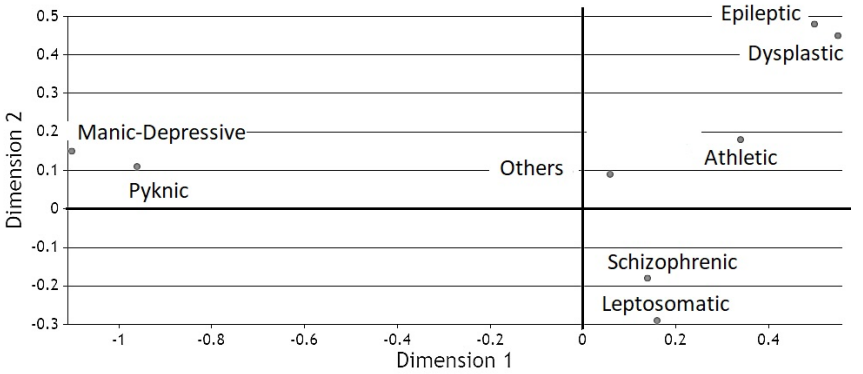

Figure 2. Correspondence Plot of Component 1 Against Component 2

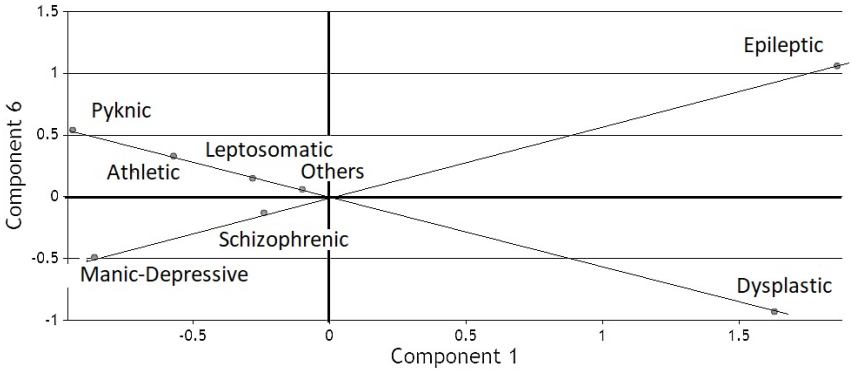

Figure 3. Dual Subspace 1 for Contingency Component 1

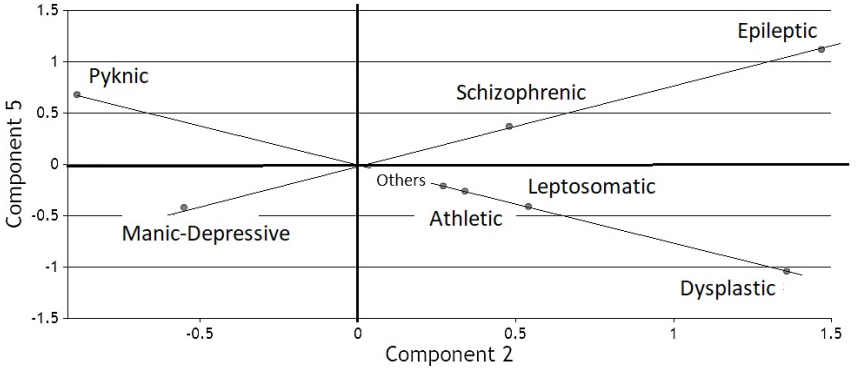

Figure 4. Dual Subspace 2 for Contingency Component 2

this two-dimensional graph appears easy. But, its mathematically correct representation of the entire data are not two-dimensional, but four-dimensional, as represented by the two 2-dimensional graphs in Figures 3 and 4, *each associated with one contingency-table component!*

When we look at the correct two-dimensional graph of each contingency-table component, we notice the following:

- All the row variables lie on one axis and the column variables on the other.
- The angle between these two axes at the origin is equal to the arc cosine of the singular-value (Nishisato & Clavel, 2003).

Let us further examine these graphs. Each component is typically interpreted by those variables which lie far from the origin and close to one another. This common-sense approach is now in jeopardy because, for example, Epileptic and Dysplastic, which are very close to each other in the correspondence plot, are now very far apart in each of the dual subspace plots. Another example is the closeness of Manic-Depressive to Pyknic in the correspondence plot, which are now relatively far apart in the two mathematically correct dual-subspace plots. In dealing with data of many dimensions, we are so used to plot a pair of relatively dominant components, such as component 1 versus component 2, but this approach now poses a problem for its appropriateness.

The above observation comes from the fact that when we have two axes, one for rows and the other for columns, the farther the points from the origin the farther the two points on the two axes move away. Recall our common practice to interpret a configuration of the variables in a graph in terms of those variables which are further away from the origin. Now, as pointed above, this practice leads to different conclusions whether we use our familiar correspondence plot or the exact dual-subspace two-dimensional graph.

On one side, one can say that correspondence plot is much more practical than the dual-subspace plot because the former uses components with relatively large eigenvalues, while the other would say that the dual-subspace plot is the mathematically correct one. This contrast may be called the practical French wisdom versus the impractical but mathematically correct dual-subspace plot. This is a significant contrast between the practical and the exact plots.

This dilemma of the choice between these two views suggests a far more important problem of how to represent a multidimensional configuration of variables than the discussion of preference between these two modes of joint plots. We need a practical way of summarizing a multidimensional, not just two-dimensional, configuration of variables. Such a method may be a high-dimensional graph or cluster analysis. No matter what, we should be able to discuss a multidimensional configuration in a much more convincing way, with accuracy, than what we do now with the correspondence plot.

The current paper is an exposition of the mathematical properties of quantified components, using a simple situation. There is still a long way to go to arrive at satisfactory routine procedures for summarizing multidimensional outputs from analysis.

7. Historical Notes and Moving Forward

Following his critical paper (Greenacre, 1989), Greenacre was invited to give a talk on the same topic at the 1989 Conference of the Federation of Classification Societies (IFCS) at the University of Virginia, Charlottesville. Douglas J. Carroll was in attendance. Immediately after Greenacre's talk, denouncing the CGS scaling, Carroll left the auditorium, whom Nishisato followed and promised him that he would write a paper in support of the CGS scaling. In that year, Nishisato submitted a paper to the Editor of the journal where all the relevant papers by the CGS group and Greenacre were published. However, without reviewing Nishisato's paper, the editor returned the paper to him, stating that "All the relevant matters have already been discussed by those leading researchers. Therefore, no more paper on the CGS scaling and related topics will be published in the journal."

The fact then is that both the CGS group and Greenacre were arguing only in the contingency space, not in the response-pattern space, thus both never discussed graphs for the doubled space, that is, space used by the corresponding response-patterns. This was a problem of differences between singular-value decomposition and eigenvalue decomposition of data. As Nishisato (1980) showed, the scaling outcomes from the contingency table and its response-pattern table are proportional within the contingency-table space, where the arguments were exchanged by the two groups. However, the total association in the contingency table is included in the response-patterns of the contingency table. It is important to know that the substantive difference between singular-value decomposition and eigenvalue decomposition exists, which is not well understood.

Regarding the last point of the data decomposition approaches, Ruben Gabriel, the proposer of biplot (Gabriel, 1971), once suggested to Nishisato to consider his biplot as an approach to the decomposition of data, but Nishisato stayed away from it since he could not imagine about the complications of further generalization of singular-value decomposition. The two often talked about it, but Gabriel could not convince Nishisato to expand the scope of his work from singular-value decomposition to this generalized singular-value decomposition, called biplot.

When Nishisato (1980) was published, Michael Greenacre was writing his doctoral thesis under the supervision of Jean-Paul Benzécri in Paris, and his thesis committee asked Nishisato if he could send Greenacre a copy of his book, which Nishisato agreed to and sent Greenacre a copy of his book. In Nishisato (1980), he devoted one chapter to the analysis of the contingency table in two ways, one analysis of the contingency table by singular-value decomposition and the other the analysis of the response-pattern table by the eigenvalue decomposition of the response-pattern table. The message there was that one must double the quantification space through the analysis of the response-pattern table, generated from the contingency table, to obtain exhaustive analysis of the information. This was exactly what Carroll, Green and Schaffer (1986) intended to do with their CGS-scaling proposal, but the CGS proposal was confined within the contingency-table space, an important limitation. Thus, Greenacre (1989) criticized the CGS scaling totally in vain, or rather, his criticisms was not consistent with the proposed framework to Nishisato.

Nishisato then sent the manuscript to a reputable journal of behaviormetrics, and after four revisions requested by the editor to shorten the paper, his final paper was rejected for the reason that Nishisato's paper was *fundamentally flawed*. The editor was, or reviewers were, not supported by the mathematical formulation. The paper was on a mathematically correct display of rows and columns of the contingency table and there was not even a trace of fundamental flaws. After this challenging review process, Nishisato submitted the same paper, written in Japanese, to the editor Ryozo Yoshino of the Japanese Classification Society Bulletin, and the paper was unconditionally accepted and published (Nishisato, 2019). This paper was published after more than 25 years of struggle against the supporters of the correspondence plot. In the current paper, Nishisato decided to call this plot *canonical plot*. This work also brought Nishisato the honor of the Japanese Classification Society's career achievement award.

Let us repeat: the correspondence plot is not an exact graph. Canonical plot (Nishisato, 2025) is our answer to the correct joint graphical display of the contingency table information.

When canonical plots are examined, it is obvious that those rows and columns which are very close to one another in the correspondence plot are actually widely separated. Then, a question arises: Do you think that graphical display is an appropriate tool to find clusters of variables? Some 60 years ago, Nishisato (1966) wrote a PhD thesis on minimum entropy clustering. Now cluster analysis appears more appropriate than graphical analysis to summarize the quantification outcome. A lesson from the current paper is not to pursue the relations between rows and columns in terms of major components, but we should compute inter-variable distances in the total space (the entire set of components), and the total distance matrix should be subjected to cluster analysis. Some relevant studies began several years ago, and the interested readers are referred to Nishisato (2014), Nishisato,

Beh, Lombardo & Clavel (2021) and Nishisato (2022, 2023a, b, c, 2024, 2025).

Acknowledgement

José Garcia Clavel, my Spanish collaborator for many years, flew from Madrid to Toronto to help my trip to Minneapolis for the conference and also kindly checked this paper in LaTeX. For the harsh criticism of my paper *fundamentally flawed*, Eric Beh offered me his unconditional support for my work. The interaction with him led to the publication of a book (Nishisato, Beh, Lombardo & Clavel, 2021). Furthermore my best allies Eric Beh, Rosaria Lombardo & José G. Clavel published my Festschrift as co-editors (Springer, 2024). Yasumasa Baba organized a special session on my view about graphical display of quantification at a conference. My gratitude also goes to the late Professor Ryozo Yoshino for publishing my paper on the canonical plot. Finally, this paper is dedicated to my wife Lorraine A. M. Nishisato for her 57 years of constant help until she passed away on December 15, 2025.

Funding Statements

None.

Competing Interests

None.

References

Beh, E. J. & Lombardo, R. (2014) *Correspondence Analysis: Theory, Practice and New Strategy*. Wiley.

Beh, E.J., Lombardo, R. & Clavel, J. G. (2023), *Analysis of Categorical Data from Historical Perspectives: Essays in Honour of Shizuhiko Nishisato*. Springer.

Benzécri, J. P. et al. (1973). *L'Analyse des Données: II. L'Analyse des Correspondances.* Paris: Dunod.

Carroll, J. D., Green, P. E. & Schaffer, C. M. (1986). Interpoint distance comparisons in correspondence analysis. *Journal of Marketing Research, 23*,271-280.

Carroll, J. D., Green, P. E. & Schaffer, C. M. (1987). Comparing interpoint distances in correspondence analysis: A clarification. *Journal of Marketing Research, 24*, 445-450

Carroll, J. D., Green, P. E. & Schaffer, C. M. (1989). Reply to Greenacre's commentary on the Carroll-Green-Scaffer scaling of two-way correspondence analysis solutions. *Journal of Marketing Research, 26*, 366-368.

Gabriel, K.R. (1971). The biplot graphical display of matrices with applications to principal component analysis. *Biometrics, 27*, 453-467.

Gifi, A. (1990) *Nonlinear Multivariate Analysis.* New York: Wiley.

Greenacre, M. J. (1984). *Theory and Applications of Correspondence Analysis.* London: Academic Press.

Greenacre, M. J. (1989). The Carroll-Green-Schaffer scaling in correspondence analysis: A theoretical and empirical appraisal. *Journal of Marketing Research, 26*, 358-365

Lebart, L., Morineau, A. & Tabard, N. (1977). *Techniques de la Description Statistique: Méthodes et Logiciels pour l'Analyse des Grands Tableaux.* Paris: Dunod.

Nishisato, S. (1966). The minimum entropy clustering of test items. Ph.D. thesis, the University of North Carolina at Chapel Hill.

Nishisato, S. (1980). *Analysis of Categorical Data: Dual Scaling and Its Applications (Mathematical Expositions No.24.)* Toronto, Canada: The University of Toronto Press.

Nishisato, S. (2014). Structural representation of categorical data and cluster analysis through filters. In Gaul, W., Geyer-Schmidt, A., Baba, Y. & Okada, A. (eds.), *German-Japanese Interchange of Data Analysis Results*. Springer, 81-90.

Nishisato, S. (2019). Reminiscence: Quantification theory and graphs. *Theory and Applications of Data Analysis, 8*, 47-57. (in Japanese).

Nishisato, S. (2020). From joint graphical display to bi-modal clustering. [1] A giant leap in quantification theory. In Imaizumi, T.et al(eds.) *Advanced Studies in Classification and Data Science*, 157-168. Singapore: Springer.

Nishisato, S. (2022). *Optimal Quantification and Symmetry*. Singapore: Springer.

Nishisato, S. (2023a). *Measurement, Mathematics and New Quantification Theory*. Singapore: Springer.

Nishisato, S. (2023b). Gratitude: A life relived. In Beh, E. J., Lombardo, R. & Clavel, J. G. (eds.). *Analysis of Categorical Data from Historical Perspectives: Essays in Honor of Shizuhiko Nishisato*. Singapore: Springer Nature.

Nishisato, S. (2023c). Propositions for quantification theory. In Okada, A., Shigemasu, K., Yoshino, R. & Yokoyama, S. (eds.) *Facets of Behaviormetrics: The 50th Anniversary of the Behaviormetric Society.* Singapore: Springer Nature.

Nishisato, S. (2024). *Optimal Analysis of Qualitative Data with Illustrations*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Nishisato, S. (2025). Reminiscing controversies over joint graphical display in quantification. An invited talk for the Career Award for Lifelong Achievement, the Psychometric Society, Minneapolis.

Nishisato, S., Beh, E. J., Lombardo, R. & Clavel, J. G. (2021). *Modern Quantification Theory: Joint Graphical Display, Biplots, and Alternatives*. Singapore: Springer Nature.