

Things Aren't as Bad as They Seem: A Comment on Storms et al. (2003)

Agnes S. Chan and Yim-chi Ho
The Chinese University of Hong Kong

G. Storms, T. Dirikx, J. Saerens, S. Verstraeten, and P. P. De Deyn (2003) criticized the use of scaling techniques, in proposing "semantic storage deficits" in Alzheimer's disease and schizophrenic patients in some studies, arguing that most reported multidimensional scaling (MDS) models for patients were not adequately fit and did not differ from models generated by random data. The studies cited by G. Storms et al. were reexamined and all available data relevant to their claim were compared. A more complete review revealed somewhat different conclusions; it showed that many of the MDS models seem to meet the criteria of adequate fit, and it does not seem to support the notion that patients' performance is close to random. Suggestions are made to improve the validity of scaling analysis in neuropsychological studies.

The article by Storms, Dirikx, Saerens, Verstraeten, and De Deyn (2003) criticizes a number of studies that had used scaling and clustering techniques to examine the semantic network of groups of patients with semantic network deficits (e.g., Alzheimer's disease, schizophrenia). They questioned the validity of multidimensional scaling (MDS) solutions for patient groups on the grounds of (a) nonsignificant difference between the stress values reported in five MDS studies (namely, Aloia, Gourovitch, Weinberger, & Goldberg, 1996; Bonilla & Johnson, 1995; Chan, Butters, Paulsen, et al., 1993; Paulsen et al., 1996; Rossell, Rabe-Hesketh, Shapleske, & David, 1999) and those obtained by computer-simulated random data; (b) the low inter- and intraindividual consistency in patients compared with normal control participants; and (c) their own empirical data. We discuss some limitations in Storms et al.'s data presentation that might have placed some limits on the conclusiveness of their interpretation.

Inadequate Model Fit?

Having compared the stress values of the five MDS studies with those of simulated random data, Storms et al. (2003) concluded that the MDS models of the patient data in all five studies failed to have adequate fit. Contrary to a systematic breakdown of semantic memory proposed by those studies, they showed that none of the patients' data differed significantly from random data and that the semantic deficits of the patients seemed to be unsystematic and could be attributed to attention deficit or access disorders. They therefore argued that MDS is not suitable for analyzing

patients' data. However, the data presented in their Table 1 did not represent all of the published data (i.e., only part of the results regarding model fitness of the reviewed studies were reported), and some data were not accurately reported. When we evaluated and reviewed all of the model fit indices as presented in those studies and other relevant studies from the literature, we drew a somewhat different conclusion. It seemed that many of the MDS models were able to meet the criteria for adequate fit and that patients' data did not seem to be close to random.

Storms et al. (2003) emphasized that the adequacy of model fit for the MDS solutions had been neglected in the previous publications, and they presented the stress values of the published studies to support their argument. Although Storms and colleagues presented the worst stress values in the studies to highlight their points, the examination of all of the reported stress values revealed a somewhat different picture (Table 1). For instance, the stress values for MDS models for the body parts category (0.23) were much better than those for the animals category (0.33) in nondeluded schizophrenic patients in Rossell et al.'s (1999) study that was reported by Storms and colleagues (note that the stress value for the model for the body parts category for the deluded schizophrenic patients was even better). Moreover, Storms and colleagues emphasized that the model fit for patient data was invariably worse than the fit for normal control individuals; however, our Table 1 shows that in three out of the five studies, similar or even better stress values for patients' data than for normal control participants had been reported (Bonilla & Johnson, 1995; Paulsen et al., 1996; Rossell et al., 1999).

Also note that in all of the reviewed studies, one other model fit index, the R^2 (which measured the percentage of variance accounted for by the solution), was reported in addition to stress level (which measured the accuracy of the MDS model in terms of error; Kruskal & Wish, 1978). From Table 1, it can be seen that except in Aloia et al.'s (1996) study and part of Rossell et al.'s (1999) results, the MDS models for the patient groups generally accounted for 65%

Agnes S. Chan and Yim-chi Ho, Department of Psychology, The Chinese University of Hong Kong, Hong Kong, China.

Correspondence concerning this article should be addressed to Agnes S. Chan, Department of Psychology, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, China. E-mail: aschan@psy.cuhk.edu.hk

Table 1
Task Category, Number of Stimuli, Stress Values, and Their Differences From Random Data, and R² Values for Various Patient Groups and Normal Control Participants in the Five Classical MDS Studies Reviewed in Storms et al. (2003)

Participant group	Category	No. of stimuli	Stress	Difference (SD)	R ²
Aloia et al. (1996)					
Schizophrenia	Animals	11	.37	8.10	.35
Control	Animals	11	.14	-3.40	.89
Bonilla & Johnson (1995)					
Mild to moderate AD	Animals	12	.13	-5.83	.94
	Occupations	12	.22	-0.83	.75
Mild AD	Animals	12	.14	-5.28	.93
	Occupations	12	.15	-4.72	.90
Moderate AD	Animals	12	.18	-3.06	.79
	Occupations	12	.25	0.83	.65
Control	Animals	12	.03	-11.39	1.00
	Occupations	12	.18	-3.06	.89
Chan et al. (1993)					
AD	Animals	12	.22	-0.83	.69
HD	Animals	12	.15	-4.72	.86
Control	Animals	12	.08	-8.61	.97
Paulsen et al. (1996)					
Schizophrenia	Animals	16	.11	-16.30	.87
Control	Animals	16	.20	-7.30	.81
Rossell et al. (1999)					
Nondeluded schizophrenia	Animals	17	.33	4.55	.46
	Body parts	17	.23	-4.55	.76
Deluded schizophrenia	Animals	17	.32	3.64	.46
	Body parts	17	.19	-8.18	.81
Control	Animals	17	.34	5.45	.36
	Body parts	17	.22	-5.45	.73

Note. MDS = multidimensional scaling; AD = Alzheimer's disease; HD = Huntington's disease.

or more of the variance in the data; some patient models even explained more than 80% of the variance. Hence, the results of overall stress values and the R^2 seem to support the idea that the MDS solutions are adequately fit, which is different from the conclusion drawn by Storms et al. (2003).

MDS Models for Patients Did Not Differ From Those for Random Data?

Storms et al. (2003) suggested that the MDS models for patients as evidence for semantic breakdown could alternatively be explained by access disorders or attention deficits. Adopting Spence and Ogilvie's (1973) criteria that reject the null hypothesis of random data only when the observed stress values are at least 3 SDs away from the corresponding values of random data, Storms and colleagues showed that all of the patients' data in the five studies did not differ significantly from simulated random data, and concluded that those MDS solutions were unreliable and might simply represent unpatterned errors. Because Storms and col-

leagues have reviewed only part of the reported data, we redid the same analyses and also included those data that were missed in Storms and colleagues' analysis. We also adopted Spence and Ogilvie's criteria proposed by Storm and colleagues and set the alpha level at .0013. However, note that this criteria is quite conservative and may seriously inflate Type II error. The reevaluation of the difference between the observed stress values in the five reviewed studies and the simulated random values generated by Storms and colleagues yielded different results. The difference in terms of the number of standard deviations away from the mean stress values of Storms and colleagues' random data are presented in Table 1. Because the authors did not provide simulated random data for 16 stimuli as reported in Paulsen et al.'s (1996) study, the data from Spence and Ogilvie were adopted as a reference for evaluating the reported stress value in this case. Results showed that 7 out of 12 reported stress values for the patient groups (not including the combined data of mild-to-moderate Alz-

heimer groups in Bonilla & Johnson's, 1995, study) were qualified as nonrandom data according to Spence and Ogilvie's criteria. Thus, it seems that much of the patients' data are significantly different from random data, and this is not entirely consistent with the conclusion drawn by Storms and colleagues. In addition, whereas Storm and colleagues did not seem to provide empirical evidence to support their claim regarding the alternative hypothesis of access disorders and attention deficits, there are several empirical studies (e.g., Chan, Butters, Paulsen, et al., 1993; Chan, Butters, et al., 1995) that had provided evidence against such alternative explanation of access disorders. Whether the breakdown of semantic structure or accessing deficit should be used to explain the results of scaling analyses needs further debate.

Alternative Interpretation of Storms and Colleagues' Own Empirical Data

Storms et al. (2003) interpreted their own data from Alzheimer's disease patients as similar to random data on the basis of Spence and Ogilvie's (1973) criteria. As stated, setting the p value at .0013 may be regarded as too conservative. Storms and colleagues' patient data have final stress values of 0.125 and 0.054 for two and three dimensions, respectively. These corresponded to 3.00 and 2.55 SDs below their simulated random values. That is, the probability of the obtained stress values being lower than the simulated random values would be only .0013 to .0054. These results seem to suggest that the patient data are significantly different from random proximity data. Furthermore, 6 out of 7 patients in the Storms et al. empirical study had significant moderate-to-strong intraindividual test-retest correlations that ranged from .46 to .98, $p < .05$, which again did not support the notion that the patients' performance is close to random.

Some Possible Factors Affecting Stress Values

Although it has been shown that many of the MDS models in the previous studies were in fact adequately fit, it remains unclear why the model fits varied from study to study, or from one part of a study to another. After reviewing the studies, we have identified two factors that may affect the model fit: first, the population under investigation. The solution obtained from mild Alzheimer's patients seemed to have lower stress values than that obtained from moderate Alzheimer's patients (see Bonilla & Johnson, 1995). Similarly, stress values of the MDS solutions for different types of schizophrenic patients seemed to be different when assessing them with certain semantic categories (see Rossell et al., 1999). Specifically, the stress value of the model for deluded schizophrenic patients seemed to be better (lower) than that for nondeluded schizophrenic patients in the semantic task of the body parts category. Second, is the semantic category. The model fits of the MDS solutions for the animals category were generally better than for the occupations category (see Bonilla & Johnson, 1995). The stress values for the body parts cate-

gory also seemed to be lower than those for the animals category, regardless of the type of participants (see Rossell et al., 1999). Thus, to obtain a better semantic model, we recommend that a more homogeneous sample be recruited, especially among patient groups of the same type of disease but with various severity or cognitive impairment. However, this suggestion does not only apply to MDS modeling but to many other commonly used statistical procedures. In addition, we recommend that more than one category be used in assessing the semantic network of patients in order to obtain a better understanding of their structure of semantic knowledge.

Concluding Remarks

It is probably agreed on that MDS is not the only way to study semantic organization. Nevertheless, as a multivariate measurement, MDS has been regarded as one of the more useful and sensitive tools to understand semantic structure than other traditional univariate analyses, inasmuch as this method can reveal the interrelationship among concepts that may not be easily revealed by other statistical analyses. However, the scaling analysis has its own disadvantages, including subjective interpretation and less well-defined criteria. Given that there are advantages and disadvantages of using scaling analyses, many researchers have advocated the necessity to validate the results obtained from MDS findings with various conventional semantic tasks (e.g., sorting task, naming task, typicality ranking task, and concept identification task; Chan, Salmon, & Pena, 2001), and analyzing the data with various methods, including additive similarity trees (Carroll & Chang, 1970; Chan, Butters, Paulsen, et al., 1993), pathfinder analysis (Dearholt & Schvaneveldt, 1990; Chan et al., 2001), discriminant analysis (Chan, Butters, Salmon, & McGuire, 1993), and correlational and univariate analyses of variance (Chan, Butters, & Salmon, 1997). There is the danger of being too subjective in either rejecting the scaling analysis as a useful statistical tool or regarding it as the finest analysis.

References

- Aloia, M. S., Gourovitch, M. L., Weinberger, D. R., & Goldberg, T. E. (1996). An investigation of semantic space in patients with schizophrenia. *Journal of the International Neuropsychological Society*, 2, 267-273.
- Bonilla, J. L., & Johnson, M. K. (1995). Semantic space in Alzheimer's disease patients. *Neuropsychology*, 3, 345-353.
- Carroll, J. D., & Chang, J. (1970). Analysis of individual differences in multidimensional scaling via an n -way generalization of "Eckhart-Young" decomposition. *Psychometrika*, 35, 283-319.
- Chan, A. S., Butters, N., Paulsen, J. S., Salmon, D. P., Swenson, M. R., & Maloney, L. T. (1993). An assessment of the semantic network in patients with Alzheimer's disease. *Journal of Cognitive Neuroscience*, 5, 254-261.
- Chan, A. S., Butters, N., & Salmon, D. P. (1997). The deterioration of semantic networks in patients with Alzheimer's disease: A cross-sectional study. *Neuropsychologia*, 3, 241-248.

- Chan, A. S., Butters, N., Salmon, D. P., Johnson, S. A., Paulsen, J. S., & Swenson, M. R. (1995). Comparison of the semantic networks in patients with dementia and amnesia. *Neuropsychology*, *9*, 177–186.
- Chan, A. S., Butters, N., Salmon, D. P., & McGuire, K. A. (1993). Dimensionality and clustering in the semantic network of patients with Alzheimer's disease. *Psychology and Aging*, *3*, 411–419.
- Chan, A. S., Salmon, D. P., & Pena, J. D. L. (2001). Abnormal semantic network for "animals" but not "tools" in patients with Alzheimer's disease. *Cortex*, *37*, 197–217.
- Dearholt, D. W., & Schvaneveldt, R. W. (1990). Properties of Pathfinder networks. In R. W. Schvaneveldt (Ed.), *Pathfinder associative networks: Studies in knowledge organization* (pp. 1–30). Norwood, NJ: Ablex.
- Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. Beverly Hills, CA: Sage.
- Paulsen, J. S., Romero, R., Chan, A. S., Davis, A. V., Heaton, R. K., & Jeste, D. V. (1996). Impairment of the semantic network in schizophrenia. *Psychiatry Research*, *63*, 109–121.
- Rossell, S. L., Rabe-Hesketh, S., Shapleske, J., & David, A. S. (1999). Is semantic fluency differentially impaired in schizophrenic patients with delusions? *Journal of Clinical and Experimental Neuropsychology*, *21*, 629–642.
- Spence, I., & Ogilvie, J. C. (1973). A table of expected stress values from random rankings in nonmetric multidimensional scaling. *Multivariate Behavioral Research*, *8*, 511–517.
- Storms, G., Dirikx, T., Saerens, J., Verstraeten, S., & De Deyn, P. P. (2003). On the use of scaling and clustering in the study of semantic deficits. *Neuropsychology*, *17*, 289–301.

Received September 24, 2002

Revision received October 17, 2002

Accepted October 17, 2002 ■

Members of Underrepresented Groups: Reviewers for Journal Manuscripts Wanted

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publications process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write to Demarie Jackson at the address below. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.
- To be selected, it is critical to be a regular reader of the five to six empirical journals that are most central to the area or journal for which you would like to review. Current knowledge of recently published research provides a reviewer with the knowledge base to evaluate a new submission within the context of existing research.
- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In your letter, please identify which APA journal(s) you are interested in, and describe your area of expertise. Be as specific as possible. For example, "social psychology" is not sufficient—you would need to specify "social cognition" or "attitude change" as well.
- Reviewing a manuscript takes time (1–4 hours per manuscript reviewed). If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.

Write to Demarie Jackson, Journals Office, American Psychological Association, 750 First Street, NE, Washington, DC 20002-4242.