DOCUMENT RESUME

ED 214 465 HE 014 933

AUTHOR Bordage, Georges
TITLE The Organization of Medical Disorders in the Memories of Medical Students and General Practitioners.
PUB DATE Mar 82
NOTE 16p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, March 23, 1982).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Classification; *Clinical Diagnosis; Clinical Experience; Comparative Analysis; Diseases; Evaluation Criteria; Foreign Countries; Higher Education; Medical Evaluation; *Medical Students; *Memory; *Physicians; Problem Solving; Recall (Psychology)
IDENTIFIERS Quebec

ABSTRACT

The applicability of Rosch and coworkers' concept of prototypes to the mental categorization of medical disorders, and the influence of clinical experience on those memory structures were studied with 100 preclinical medical students and 77 experienced physicians from Quebec, Canada. The third-year medical students were French-speaking and read English, and the physicians had 2 to 28 years of experience and practiced in predominantly French-speaking communities. Four experiments based on the Rosch (1975) research strategy were conducted to form a logical and converging sequence of evidence about the internal structure of 14 categories of medical disorders. The experiments investigated category norms, prototypicality ratings, family resemblance, and response time. It was found that the mental representation of the categories was better described by the overlapping feature of the prototype view of categorization than the criterial features of the deterministic view. Overall, the disorders with the highest prototypicality ratings were recalled earlier in a free-recall task than were the less prototypical ones; the higher the mean prototypicality rating of an item, the higher the mean family resemblance score. It took about one-half less time for students and one-third less time for physicians to judge category membership of central disorders as opposed to peripheral ones. Additionally, clinical experience did affect the internal structure of the categories: physicians used a narrower range of points on the prototypicality rating scales. It is suggested that prototypes are valuable in describing the internal structure of medical categories. (SW)

THE ORGANIZATION OF MEDICAL DISORDERS IN THE MEMORIES OF MEDICAL STUDENTS AND GENERAL PRACTITIONERS *

GEORGES BORDAGE, MD, MS.

Bureau de pédagogie médicale, Faculté de Médecine
Université Laval, Québec, Canada, G1K 7P4

Investigators of the reasoning process in general (for example, De Groot, 1965) and of medical reasoning in particular (for example, Barrows et al., 1978; Elstein, Shulman and Sprafka, 1980) have reached common conclusions about the critical role of memory in problem solving. Thus, Elstein, Shulman and Sprafka (1980) conclude: "the differences between experts and weaker problem solvers are more to be found in the repertory of their experiences, organized in long-term memory, than in differences in the planning and problem solving heuristics employed" (p. 276). Barrows and Tamblyn (1980) also note that "generations of students in conventional curricula have expressed the desire to repeat basic science courses when they enter their clinical years, testimony to their frustration over their inability to recall subject-based information from earlier years" (p. 12). Hence, How is medical knowledge organized in the memories of medical students and physicians? and Are different structures related to different learning outcomes or clinical performances?

Because of theoretical and methodological shortcomings, there are still no decisive answers to these significant questions. However, most investigators in the field agree on the importance of categorization as a basis for the understanding of memory organization. Most often, categories of objects or events are viewed as rule-defined entities derived from a criterion and expressed within definite boundaries inside which all instances sharing the criterial feature(s) have a full and equal degree of memberships. While this traditional deterministic view is most common, recent investigations stemming from Wittgenstein (1953) in particular have shown that the internal structure of categories in a person's memory may be of different nature. Rosch and co-workers in particular

* Paper presented at the 1982 Annual Meeting of the American Educational Research Association, New York, N.Y., March 23, 1982.

have shown for color and object categories that they are represented in memory in terms of "prototypes", the clearest cases or best examples of the category, surrounded by other instances of decreasing similarity to the prototypes and of decreasing degree of membership (1975, 1978; Rosch and Mervis, 1975).

PURPOSE:

The purpose of the present study was twofold; (1) to determine whether the concept of prototypes as developed by Rosch and co-workers is applicable to the mental categorization of medical disorders, and (2) to describe the influence of clinical experience on those memory structures. Four experiments were conducted with 100 pre-clinical medical students (n_s) and 77 experienced physicians (n_p).

SAMPLES:

The two samples were chosen to represent contrasting levels of diagnostic skills. It was hypothesized that prototypes become more clearly defined as a result of experience and that the categories contain a richer, more closely woven network of information. It was also expected that the experts would have a quicker access to the information in the categories than the novices.

A pre-clinical medical student, representing the novice diagnostician, was defined within the context of a traditional medical curriculum as a medical student who has completed the basic science part of the curriculum, including the various system pathologies, and who is at the beginning or about to begin the in-hospital training or clerkships. Because of limited resources, one such student subpopulation was preselected for the study: the 152 third-year medical students at Laval University in Quebec City. They were sollicited by letter and participation was voluntary. Two thirds accepted to participate and were subsequently randomly assigned to the four experiments. Based on their respective GPA scores (a score believed to be correlated substantially with the dependant variables in the study since both were measures of medical knowledge), the participants ($\bar{x}$ = 4.22, s = .38) did not vary significantly (α = .1; t (150) = .78, p = .22) from the non-participants ($\bar{x}$ = 4.27, s = .36). The average age of the Laval participant is 23.8 years (21-32). Two thirds are

male and all are French speaking. Most are bilingual (all read English) and about half of the textbooks are in English. The courses are organized according to a traditional discipline-oriented curriculum. The students participated in the study some two to three months prior to their clerkships during their three months of introduction to clinical methods.

An experienced physician can be defined in a number of ways. Because of the prevailing trend of many medical schools to emphasize family practice in their curriculum, an experienced physician was defined as a general practitioner who (1) does not have any specialty training, except possibly a certification from the College of Family Physicians, (2) has at least one year of active practice, and (3) is seeing patients on a regular weekly basis. The physicians were selected either from the greater Quebec City chapter of the Quebec Federation of General Practitioners for the first and third experiments and from participants attending one of two continuing medical education conferences, one on emergencies and the other on ophthalmology, for the remaining two experiments. Three quarters of the physicians solicitated agreed to participate. Participation was anonymous and voluntary. All of the participants (except one) graduated from one of the three French-speaking medical schools in the Province of Quebec and all practiced in predominantly French-speaking communities. They had practiced for an average of 8.5 years (2-28) and saw an average of 104 patients per week (15-200). Most worked in group practices (78%) in an urban setting (73%). Only 17% were certified family physicians.

EXPERIMENTS:

The distinctiveness of categories can be achieved in two ways: by means of formal, necessary and sufficient criteria for category membership, or by conceiving of each category in terms of its clear cases. The arguments in favor of the latter view come from the operational definition of two concepts - prototypes and family resemblance. The level of prototypicality of an item is derived from people's perception of how good an example various members are of a category (Rosch, 1973, 1978). The degree of agreement among the judgments is an indication of how applicable the concept is. Prototypes, like

categories themselves, develop through the maximization of family resemblance: "members of a category come to be viewed as prototypical of the category as a whole in proportion to the extent to which they bear a family resemblance to (have attributes which overlap those of) other members of the category" (Rosch and Mervis, 1975, p. 575). Rosch and Mervis (1975) have shown for categories of concrete nouns that the more prototypical of a category a member is rated, the more attributes it has in common with other members of that category.

Four experiments, based on Rosch (1975) and her colleagues' research strategy, were designed and conducted to form a logical and converging sequence of evidence about the internal structure of 14 categories of medical disorders (see Table 1). These categories were chosen because, like in Rosch's initial experiments (1973, 1975), they represent broad and widely accepted category names. The eight organ-system categories were drawn from the classification of medical problems of the Royal College of General Practitioners and the remaining six names from standard medical references.

Experiment-1: Category norms. The list of disorders contained in the categories was obtained seperately from the two samples by means of a free-recall task ($n_s = 42$; $n_p = 21$). These lists were used as the taxonomic norms for the following experiments on internal structure. Experiment-2: Prototypicality ratings. Using a 7-point scale, the subjects were asked to rate the degree of exemplariness of 8 disorders taken from each of the categories ($n_s = 13$; $n_p = 19$). Experiment-3: Family resemblance. The lists of attributes (symptoms, signs, pathophysiological characteristics) for six of the disorders rated for prototypicality were obtained by means of a free-recall task ($n_s = 31$; $n_p = 23$). Two hypotheses were tested about these attributes: (1) that each disorder in a category has at least one, if not several, attributes in common with one or more other disorders in the category (the family resemblance view), but none, or few, of the features are common to all disorders (the criterial view); and (2) that the disorders with the highest prototypicality ratings in a category are those with the greatest number of attributes in common with the other disorders in the category. Experiment-4: Response time. To test the effect of prototypicality on the cognitive processing of the categories, response time to category membership statements of the type "Diabetes mellitus is a kind of

endocrine disorder" was obtained for disorders with high prototypicality ratings (central members) as opposed to those with lower ratings (peripheral members) ($n_s = 14$; $n_p = 14$). The hypothesis was tested that the central members would be recalled faster than the peripheral ones.

All four experiments were conducted in French. If the internal structure of the categories in the long-term memories of the subjects was a prototypical one, then the results from all, if not most, experiments should point toward that conclusion.

RESULTS:

There were three basic and converging findings from this series of experiments.

First, the striking structural feature of the categories resided in the presence of prototypical disorders as opposed to some common criterial attribute(s). --Few attributes were listed which were common to all six disorders in a category. Only rarely was an attribute common to all, or even to the majority, of the disorders. The distribution in Figure 1 also illustrates that the number of attributes decreases as the number of disorders to which the attribute is applicable increases. Both the students and the physicians found the rating task meaningful and made reliable judgments about the degree of exemplariness of selected members from each category. The ratings in all 14 categories departed significantly from a chance distribution (Friedman tests: $p < .001$; see Table 2.). Overall, the disorders with the highest prototypicality ratings were recalled earlier in the free-recall task than the less prototypical ones; the median Spearman rank-order correlation coefficients between the ratings and the item outputs were .83 and .85 for the students and the physicians respectively ($p < .025$).

To derive the measure of family resemblance, each attribute received a score ranging from 1 to 6, representing the number of disorders in a category that shared that attribute. The measure of degree of family resemblance for

a given disorder was the sum of the weighted scores of each of the attributes that were listed for that item (attributes with scores of one or two were not included). Overall, the higher the mean prototypicality rating of an item, the higher the mean family resemblance score; the scores for each level of prototypicality were statistically different from one another (Friedman tests: $\chi_r^2 = 17.3$ and 21 for the students and physicians respectively, $p < .001$). In other words, the more prototypically rated disorders in a category were also the ones with the greatest number of attributes in common with the other members (see Tables 3 and 4). The median Spearman rank-order correlation coefficients between the family resemblance scores and the ratings were .62 and .59 for the students and the physicians respectively (see Table 5). These correlations are not as high as those obtained by Rosch and Mervis (1975) for semantic categories of concrete objects: that is, .88 to .94. A possible explanation may be because they used 20 items per category to perform their experimentation as opposed to only six items here. The call for further experimentation is well founded since there is a definite positive trend between the two measures.

Second, prototypicality affected the cognitive processing of the categories. --It took an average of one-half less time for the students and one-third less time for the physicians to make judgments about the category membership of central disorders as opposed to peripheral ones ($F(1,24) = 5.07$, $p < .05$). There was no interaction between subject types (students and physicians) and disorder types (central and peripheral). Furthermore, while the judgments were virtually error-free for the central members, many more classification errors were made for the peripheral ones (some 17 to 20 times more; Mann-Whitney U-test: $U = 0$, $p < .05$; see Table 6).

Third, clinical experience did affect the internal structure of the categories. --Although there was a high degree of similarity in the individual disorders listed by the students and the physicians in the free-recall tasks, the general practitioners had a richer and more tightly woven network of knowledge than the students. The physicians used a narrower range of points on the prototypicality rating scale (Wilcoxon matched-pairs signed-ranks tests formed by the 14 categories: $T = 18$, $p < .05$) and responded faster to category membership statements than the students ($F(1,24) = 9.24$, $p < .01$). Both of

these results suggest that the associative strength between a disorder and its category label is stronger for the physicians than for the students. Also, the physicians had higher family resemblance scores, (Sign tests for the 84 pairs of scores: $z = 3.38$, $p < .001$), thus indicating that they have more explicit ties among the members of their categories than the students. Despite their increased clinical experience, the general practitioners did not tend in general to list any more than the students did the names of disorders that are encountered most often in the family practice setting. This finding is in keeping with Rosch's contention that frequency is not the major factor in determining the formation of prototypes (1973, 1978). Instead, while frequency may play a role in the formation of some prototypes, prototypes most likely arise from the overlapping structure of the attributes of the various members in a category, the family resemblance structure.

DISCUSSION AND CONCLUSION:

Cantor et al. (1980), in a similar study of the categorization of nine psychiatric disorders, also found that the internal structure of the categories was better described by the "correlated features" of the prototype view than the "defining features" of the traditional deterministic view. They also studied psychiatric diagnosis as a "prototype-matching process" and found that the accuracy and confidence of diagnosis increases with typicality. In their analysis of the changes in the latest edition of the APA's diagnostic manual of mental disorders, they note that "Diagnostic criteria are now presented as prototypes - larger sets of correlated features rather than selected defining ones; guidelines for diagnosis also emphasize the potential heterogeneity of the symptoms of like-diagnosed patients... From the perspective of the prototype view, these changes are important because they help to emphasize, rather than obscure, the probabilistic nature of diagnostic categorizations" (p. 190 and 192). And they add, "diagnoses can be made on the basis of degree of fit between the patient's cluster of symptoms and the prototypes for various different categories" (p. 192).

These observations are also consistent with the differences found between experts and novices. For example, Larkin et al. (1980) state that "Although

a sizable body of knowledge is prerequisite to expert skill, that knowledge must be indexed by large numbers of patterns that, on recognition, guide the expert in a fraction of a second to relevant parts of the knowledge store" (p. 1336). Medical textbooks and much classroom teaching abound in presenting detailed lists of disorders, features and actions to the learner but most often fail to provide a categorization scheme that is best suited for their retrieval in a clinical problem solving situation. The prototypes, with their overlaping attributes, "most reflect the redundancy structure of the category as a whole" (Rosch, 1978, p. 37) and thus may serve as an indexing scheme for the clinician's knowledge. Although, both the medical students and the physicians in the present study exhibited a prototypical structure, the experienced practitioners had a richer and more readily accessible store of knowledge than the novices.

Given the results of the present study and those of Cantor and co-workers in psychiatry, there is a growing body of evidence that prototypes may be better suited to describe the internal structure of medical categories than the traditional deterministic view. Furthermore, the notion of prototypes offers a promising means by which the learning and diagnosctic process of becoming and being an expert clinician can be analysed, understood and eventually improved.

ACKNOWLEDGEMENTS:

The present study was the topic of my doctoral dissertation submitted to the Department of Educational Psychology at Michigan State University. I extend my sincere appreciation to the committee members, and, in particular to the dissertation director, Professor Rose Zacks from the Department of Psychology for their able guidance during the course of this research.

REFERENCES:

Barrows, H.S., Feightner, J.W., Neufeld, V.R. and Norman, G.R., Analysis of the clinical methods of medical students and physicians. McMaster University, Hamilton, Ont., 1978.

Barrows, H.S., and Tamblyn, R.M., Problem-based learning: An approach to medical education. NY: Springer Publ. Co., 1980.

Cantor, N., Smith, E., French, R., Mezzich, J., Psychiatric diagnosis as prototype categorization. Journal of abnormal psychology, 1980, 89(2), 181-193.

DeGroot, A.D., Thought and choice in chess. The Hague: Mouton, 1965.

Elstein, A.S., Shulman, L.S., and Sprafka, S.A., Medical problem solving: An analysis of clinical reasoning. Cambridge: Harvard University Press, 1978.

Larkin, J., McDermott, J., Simon, D., Simon, H., Expert and novice performance in solving physics problems. Science, 1980, 208, 1335-1342.

Rosch, E., On the internal structure of perceptual and semantic categories. In T.E. Moore (Ed.), Cognitive development and the acquisition of language. NY: Academic Press, 1973, 111-144.

Rosch, E., Cognitive representations of semantic categories. J. Exp. Psy: general 1975, 104, 192-233.

Rosch, E., Principles of categorization. In E. Rosch and B.B. Lloyd (Eds.), Cognition and categorization. Potomac: Erlbaum Press, 1978, 27-48.

Rosch, E., and Mervis, C.B., Family resemblances: studies in the internal structure of categories. Cognitive psychology, 1975, 7, 573-605.

Wittgenstein, L., Philosophical investigations NY: Macmillan, 1953.

TABLES and FIGURES:

TABLE 1 Stimuli for category membership experiment.

ORGAN SYSTEMS-RELATED DISORDERS:	PATHOPHYSIOLOGY-RELATED DISORDERS:
1. Respiratory disorders	9. Inflammatory disorders
2. Cardiovascular disorders	10. Infectious disorders
3. Genitourinary disorders	11. Neoplastic disorders
4. Gastrointestinal disorders	
5. Endocrine disorders	SYMPTOM-RELATED DISORDERS:
6. Musculoskeletal disorders	12. Dyspnea
7. Neurological disorders	13. Abdominal pain
8. Hematological disorders	14. Joint pain

TABLE 2 Summary results of mean prototypicality ratings.

	Students					Physicians				
	Frequency range from first experiment									
	Top	High	Mod.	Low	Out	Top	High	Mod.	Low	Out
Categories										
1. Respiratory	7	6.35	5.8	5.05	1	6.7	6.8	6.35	4.95	1
2. Cardiovascualr	6.9	6.45	5.9	5.25	1	7	6.5	6.25	5.4	1.05
3. Genitourinary	6.3	6.8	5.85	4.8	1	6.5	6.3	6.15	5.75	1
4. Gastrointestinal	6.9	6.55	5.7	5.35	1.2	6.8	6.25	6.05	5.5	1.05
5. Endocrine	6.8	6.85	5.75	4.15	1	6.7	6.5	6.6	5.3	1
6. Musculoskeletal	6.8	6.3	6.05	5.15	1	6	6.55	5.15	5.8	1.2
7. Neurological	6.7	6.05	5.6	5.65	1	6.8	4.7	6.5	5.6	1
8. Hematological	6.5	6.85	6.2	3.15	1	7	6.75	6.2	3.9	1.2
9. Infectious	6.7	6.7	5.5	4.5	1.2	6.4	6.3	5.85	6.5	1.4
10. Neoplasms	6.8	6	4.8	6.5	1.2	6.4	6.85	6.45	5.85	1.05
11. Inflammations	6.7	6.3	5.05	3.25	1.2	6.5	6.3	5.75	4.5	1.2
12. Dyspnea	6.8	5.3	5.75	4.65	1.2	7	5.55	5.5	4.25	1
13. Abdominal pain	7	6.35	3.75	4.1	1	6.9	6.6	5.55	4.9	1.3
14. Joint pain	6.8	6.2	4.65	4.3	1	6.6	6.45	6.05	4.2	1.05
Mean total:	6.76	6.36	5.45	4.7	1.07	6.66	6.31	6.03	5.17	1.11

TABLE 3. Students' family resemblance scores at each level of prototypicality (average scores in parentheses).

Category	High prototypicality ratings		Moderate prototypicality ratings		Low prototypicality ratings	
1. Respiratory	48	54	33	30	40	27
2. Cardiovascular	30	24	14	18	17	13
3. Genitourinary	18	6	21	15	15	15
4. Gastrointestinal	51	50	38	60	42	30
5. Endocrine	23	8	13	16	16	10
6. Musculoskeletal	38	24	32	28	11	21
7. Neurological	28	36	31	22	22	22
8. Hematological	25	25	25	17	7	30
	(32.6)	(28.4)	(25.9)	(25.8)	(21.3)	(21)
	(30.5)		(25.8)		(21.1)	
9. Infectious	44	37	45	31	43	31
10. Neoplastic	38	45	40	34	31	23
11. Inflammatory	19	13	22	19	6	3
	(33.7)	(31.7)	(35.7)	(28)	(26.7)	(19)
	(32.7)		(31.8)		(22.8)	
12. Dyspnea	20	33	30	27	10	15
13. Abdominal pain	31	31	24	18	13	21
14. Joint pain	35	52	32	45	44	37
	(28.7)	(38.7)	(28.7)	(30)	(22.3)	(24.3)
	(31.6)		(29.3)		(23.3)	
TOTAL:	32	31.3	28.6	27.8	23.6	21.3
	(31.7)		(27.9)		(22)	

TABLE 4 Physicians' family resemblance scores at each level of prototypicality (average scores in parentheses).

Category	High prototypicality ratings		Moderate prototypicality ratings		Low prototypicality ratings	
1. Respiratory	58	55	40	50	51	33
2. Cardiovascular	36	33	30	17	30	12
3. Genitourinary	49	45	40	29	16	12
4. Gastrointestinal	57	46	53	36	57	21
5. Endocrine	19	9	16	7	13	6
6. Musculoskeletal	20	17	6	11	10	13
7. Neurological	17	45	20	39	30	30
8. Hematological	51	44	29	40	16	30
	(38.4)	(36.8)	(29.3)	(28.6)	(27.9)	(19.6)
	(37.6)		(28.9)		(23.8)	
9. Infectious	36	45	46	30	30	53
10. Neoplastic	35	52	44	50	44	31
11. Inflammatory	14	21	14	17	13	14
	(28.3)	(39.3)	(34.7)	(32.3)	(29)	(32.7)
	(33.8)		(33.5)		(30.8)	
12. Dyspnea	40	26	32	33	36	24
13. Abdominal pain	54	50	27	36	55	12
14. Joint pain	86	60.	51	64	58	30
	(60)	(45.3)	(36.7)	(44.3)	(49.7)	(22)
	(52.7)		(40.5)		(35.8)	
TOTAL:	40.9	39.1	32	32.8	32.8	22.9
	(40)		(32.4)		(27.9)	

TABLE 5 Spearman rank-order correlation coefficients between prototypicality ratings and family resemblance scores for each category (n = 6).

	Students	Physicians
Respiratory	.77**	.77**
Cardiovascular	.87**	.87**
Genitourinary	.6 *	1.0 **
Gastrointestinal	.53	.44
Endocrine	.11	.64*
Musculoskeletal	.76*	.49
Neurological	.76*	- .2
Hematological	.04	.71*
Infectious	.44	- .09
Neoplastic	.71*	.33
Inflammatory	.64*	.71*
Dyspnea	.6 *	.43
Abdominal pain	.86*	.37
Joint pain	- .09	.69*

** $p < .025$

* $p < .05$

TABLE 6: Mean response times and error proportions for central and peripheral sets (standard deviations in parenthesis).

	RESPONSE TIMES (in seconds)	ERROR PROPORTIONS	
		False for a true item	True for a false item
STUDENTS			
CENTRAL (n = 7)	222.9 (25)	.016	.000
PERIPHERAL (n = 7)	338.6 (116.6)	.269	.006
PHYSICIANS			
CENTRAL (n = 7)	180.1 (55.7)	.031	.006
PERIPHERAL (n = 7)	246.4 (88.2)	.181	.006

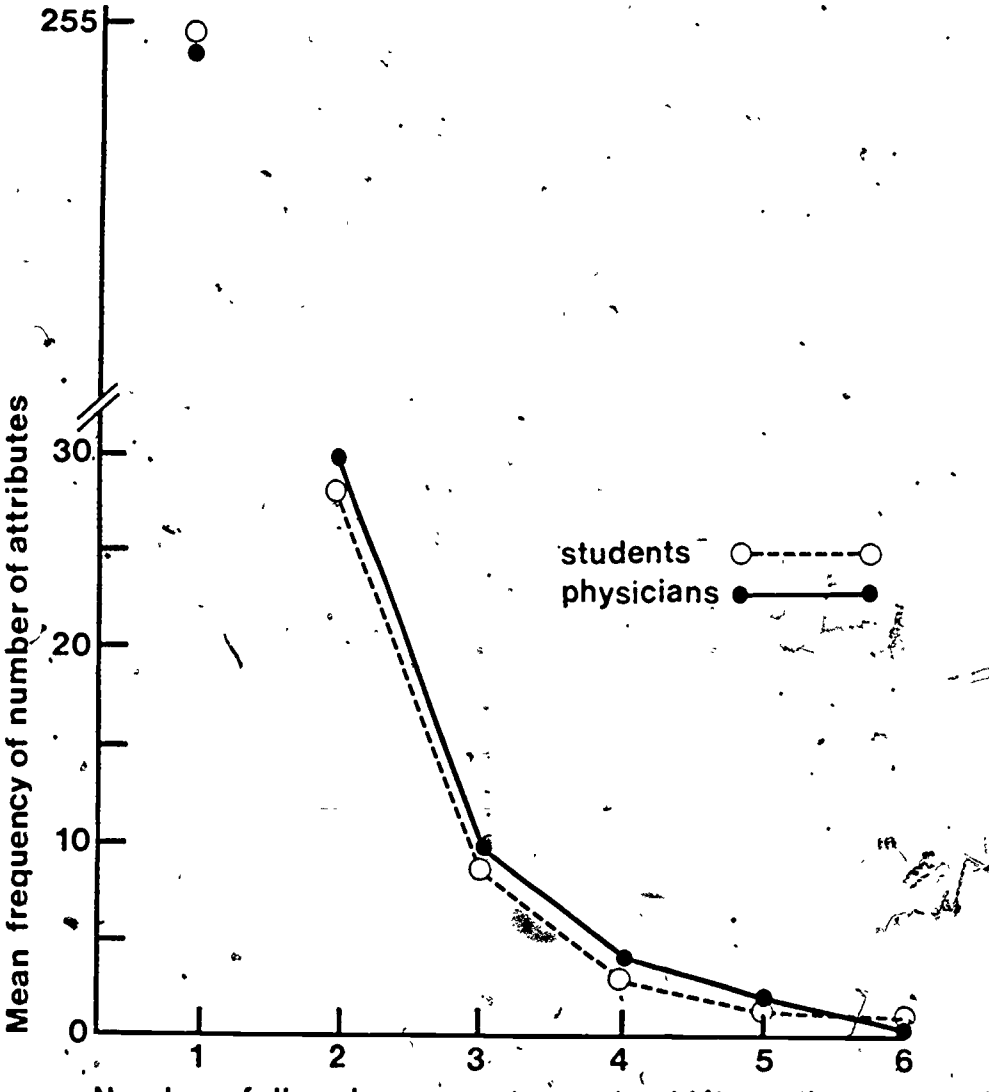

Figure 1. Frequency distribution for the mean number of attributes applied to each number of disorders per category.