Title
Linear Separability as a Constraint on Information Integration

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 15(0)

Authors
Wattenmaker, William D.
Schwertz, Stephanie J.

Publication Date
1993

Peer reviewed

Linear Separability as a Constraint on Information Integration[1]

William D. Wattenmaker and
Stephanie J. Schwertz
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, Pennsylvania 15260
wdw@vms.cis.pitt.edu

Abstract

In this paper we examined the extent to which linear separability constrained learning and categorization in different content domains. Linear separability has been a focus of research in many different areas such as categorization, connectionist modeling, machine learning, and social cognition. In relation to categorization, linearly separable (LS) categories are categories that can be perfectly partitioned on the basis of a weighted, additive combination of component information. We examined the importance of linear separability in object and social domains. Across seven experiments that used a wide variety of stimulus materials and classification tasks, LS structures were found to be more compatible with social than object materials. Nonlinearly separable structures, however, were more compatible with object than social materials. This interaction between linear separability and content domain was attributed to differences in the types of knowledge and integration strategies that were activated. It was concluded that the structure of knowledge varies with domain, and consequently it will be difficult to formulate domain general constraints in terms of abstract structural properties such as linear separability.

Introduction

A central issue in cognitive science concerns the domain-generality of learning and processing constraints. Are learning and processing constraints particular to domains or do they generalize across domains? In this article we examined the extent to which linear separability constrained categorization and decision making in different content domains. Linear separability is a principle that is relevant to categorization processes (e.g., Medin & Schwanenflugel, 1981; Waldman & Holyoak, 1990; Wattenmaker, Dewey, Murphy, & Medin, 1986; Wattenmaker, 1993a), connectionist modeling (e.g., Gluck & Bower, 1988), and machine learning (e.g., Nilsson, 1965). In relation to categorization, linearly separable (LS) categories are categories that can be partitioned on the basis of a weighted, additive combination of component information. If it is possible to weight and sum the individual features of members of categories so that there is no ambiguity or overlap in category membership, then the categories are linearly separable. The importance of linear separability as a constraint on categorization was examined in object and social domains.

In the present context, social categorization refers to the categorization of people based on traits or behaviors, whereas object categorization refers to the categorization of concrete entities such as animals, plants, and human artifacts. Object and social domains were selected because there has been extensive investigation of the principles that underlie categorization behavior in both these domains (e.g., Cantor, E. Smith, French, & Mezzich, 1980; Medin & E. Smith, 1984). Linear separability was selected as the principle to investigate because it is central to many categorization models and because it is relevant to categorization in both domains. In both object and social domains it is necessary to weigh and integrate cues to form evaluations and make classifications, and linear separability represents a general constraint on information integration that could, potentially, operate across domains.

Linearly separable structures will be easy to

[1]This research was supported by NIMH grant MH45585 to the first author.

learn if it is natural for people to evaluate each feature individually and then sum these independent evaluations. Linearly separable structures should be difficult to learn, however, if features are evaluated configurally rather than independently. Thus, the strategy that is used to integrate information will be of crucial importance for learning LS structures. Consistent with this idea, Wattenmaker et al. (1986) found that the ease with which LS categories were learned interacted with the knowledge that was brought to bear on the task. Knowledge that promoted a strategy of summing independent features facilitated the acquisition of LS structures, but knowledge that highlighted interactions or relations between features hindered the acquisition of LS structures. These results suggest that if different knowledge structures are activated in object and social domains, then the importance of linear separability as a constraint on categorization might vary in these domains.

Indeed, there appear to be a number of reasons to expect that the knowledge that is activated in the social domain might be more compatible with LS structures. These reasons are outlined below.

(1) Flexibility in incorporating inconsistent information. Social knowledge allows for tremendous flexibility in interpreting features. There is evidence, for example, that when confronted with contradictory social information (e.g., a person is *happy* and *sad*) subjects are able to resolve these conflicts with very little difficulty (e.g., Asch & Zukier, 1984; Kunda, Miller, & Claire, 1990). Indeed, Asch & Zuckier (1984) identified several types of explanations or modes of resolution that people use to resolve contradictory social information. Linearly separable categories often have atypical as well as typical features (i.e., most of the evidence is positive but some negative evidence is also present), thus the ability to resolve inconsistencies is compatible with learning LS structures.

(2) Exemplar versus summary representations. Many researchers have proposed that conceptual representations consist of a mix of abstract and exemplar information (e.g., Smith & Medin, 1981), and it is possible that the availability of these types of information will differ as a function of conceptual domain. Indeed, an intuitive analysis of object categories such as *birds* and *furniture* indicates that several exemplars are readily accessible. Exemplars are also available for social concepts such as *cautious* and *passive*, but these exemplars seem to be less accessible, smaller in number, and less central to the concepts.

If exemplars are more available in object categories, then a natural strategy might be to make decisions by analogy to these instances (e.g., Brooks, 1978; Medin & Schaffer, 1978; Nosofsky, 1984; Wattenmaker, 1993b). Making decisions by analogy to known instances creates the possibility that relational properties will influence decisions. The perceived similarity of instances is often influenced by relations between properties as well as by individual properties (e.g., Goldstone, Medin, & Gentner, 1991), and consequently relational properties can influence the selection and application of analogies. The use of relational properties is inconsistent with summing individual properties and learning LS structures.

Whereas highly structured objects might serve as the primary reference point for object categorization, social categorization would seem to be based on informal theories of personality, and to be influenced by loosely structured collections of abstract, summary features rather than specific examples.

(3) Categorization as an inference process. The features that are used to classify entities such as birds (wings, feathers, flying, etc.) are concrete, perceptual properties, and rather than requiring extensive interpretive and inferential processes, these features are already central components of the concepts. With social categories such as *courageous* or *friendly*, however, the features of the concept seem to be more abstract and to be based on inferential and reasoning processes rather than feature matching.

This reliance on inferential processes with social categorization would seem to allow explanatory systems (e.g., reasoning schemata and modes of resolution) to have a major impact on categorization decisions. This should make it easier to explain inconsistencies, and should also increase the flexibility with which social features can be interpreted. Both of these factors should make it easier to form and learn LS structures.

(4) Flexibility and identifying a basis for summing features. A strategy of summing features is only natural when the features are perceived to possess a common property that can provide a basis for summation. There is so much flexibility in interpreting social features that it should be easier to identify commonalities in social features than in features of objects.

Summary. All of the issues discussed above suggest that there might be important differences in the structures and processes of object and social categorization, and that these domains might be associated with different knowledge, theories, and reasoning schemata. These differences might make different integration strategies salient, and consequently there might be systematic variations in the ease with which LS structures are learned in object and social domains.

Experiments 1-6

Introduction and general method

In these experiments subjects were presented with descriptions of objects or descriptions of people and they were asked to sort the examples into two equal-sized groups. The materials for the object and social tasks were constructed from the same underlying structure (represented in abstract notation in Table 1), but this structure was represented by object or social characteristics. For example, for the social task of Experiment 1 the 1's in Table 1 were represented by features that were rated to be characteristic of extroverts, and the 0's were represented by features that were rated to be not characteristic of extroverts. Thus, for one randomization Exemplar 1 (1110) in Table 1 was *represented as lively, talkative, friendly,* and *timid;* Exemplar 2 (1101) was represented as *lively, talkative, inhibited,* and *entertaining;* etc. For the object materials in Experiment 1, however, the 1's in Table 1 were represented by features that were rated to be good features for a hammer and the 0's were represented by features thatwere rated to be bad features for a hammer. Thus, for one randomization Exemplar 1 was represented as: *is easy to grasp, has a flat surface, is two pounds,* and *is five feet long;* Exemplar 2 was represented as: *is easy to grasp, has a flat surface, is sixty pounds,* and *is one foot in length.* Each exemplar had three characteristic features and one uncharacteristic feature.

For the social task participants were presented with the eight descriptions of people and for the object task participants were presented with the eight descriptions of objects. (As described above, these descriptions were constructed from the notation in Table 1). The category labels were presented in the social conditions (e.g., *extroverted* vs. *non-extroverted*) and the object conditions (e.g., *hammer* vs. *non-hammer*). For the social task participants were asked to place four of the descriptions in the *extroverted* category and four of the descriptions in the *not extroverted* category. For the object tasks participants were asked to place four of the descriptions in the *good hammer* category and four of the descriptions in the *not a good hammer* category. For both tasks, participants were asked to examine the descriptions carefully and place each description in the most appropriate category. (Each subject performed the object and the social sort. The order in which the sorts were performed was counterbalanced across subjects).

There are many possible ways to partition the examples represented in Table 1. One strategy, for example, would be to divide the examples on the basis of the features on one dimension (e.g., if subjects sorted by the first dimension, then Exemplars 1, 2, 3, and 8 would be placed in one category and Exemplars 4, 5, 6, and 7 would be placed in the other category). If a strategy of summing characteristic features is natural, however, then Exemplars 1-4 would be placed in one category and Exemplars 5-8 would be placed in the other category. These categories would be linearly separable. This particular pattern of sorting will be called a Summation sort.

The sorting task was used in all the experiments but the specific features that were used for the social and object sorts varied from experiment to experiment. In Experiments 1 and 2 the social categories were *extroverted* vs. *non-extroverted* and the object categories were *hammer* vs. *non-hammer*. In Experiment 3 the social categories were *active* vs. *passive* and the features were behavior statements rather than trait terms.

Table 1. Abstract Representation of the Exemplars in Experiments 1-6

	DIMENSION			
EXEMPLAR	D1	D2	D3	D4
1	1	1	1	0
2	1	1	0	1
3	1	0	1	1
4	0	1	1	1
5	0	0	0	1
6	0	0	1	0
7	0	1	0	0
8	1	0	0	0

The object categories in Experiment 3 were *animals* vs. *furniture*. In Experiment 4 the social categories were *cautious* vs. *non-cautious* and the object categories were *cars* vs. *airplanes*. In Experiments 5 and 6 the social categories were *cautious* vs. *non-cautious* and the object categories were *bird* vs. *non-bird*. (Thirty-four, 16, 76, 26, 24, and 28 subjects participated in Experiments 1-6, respectively).

If the object and social materials activate different knowledge, and if the knowledge that is activated in the social domain is more compatible with linear separability, then more Summation sorts should be observed with the social materials.

Results and discussion

Many more Summation sorts occurred with the social materials. In Experiment 1, 77% of the social sorts were Summation sorts whereas only 35% of the object sorts were Summation sorts. The corresponding numbers in Experiments 2, 3, 4, 5, and 6, were 69% vs. 19%, 50% vs. 28%, 46% vs. 04%, 79% vs. 50%, and 57% vs. 25%, respectively. Thus, even though very different stimulus materials were used in the different experiments, the results were clear and consistent across the experiments. (Sign tests indicated that all of these differences were significant at the .01 level.) In addition, many subjects in the social conditions reported summing typical features but this strategy was rarely reported in the object conditions. Instead, subjects in object conditions reported making decisions by analogy to known objects, sorting the examples on the basis of a single dimension, or using configural properties involving pairs of features (e.g., "a heavy object is okay as long as it is easy to grasp"). All of these strategies are inconsistent with summing independent features and forming LS categories. Thus, the sorting results and the strategy reports indicate that linearly separability was an important constraint on social but not object classification.

Experiment 7

In this experiment the compatibility between linear separability and object and social domains was examined with a learning rather than a sorting task. Participants attempted to learn to correctly classify examples that had been divided into LS categories. If social knowledge induces summation strategies, then LS structures should be easy to learn in social conditions.

This experiment was also designed to test a key prediction of the knowledge-induced integration hypothesis. Specifically, knowledge that induces a summation strategy should only facilitate the acquisition of categories that are compatible with this strategy. Thus rather than making it easier to learn all structures, social knowledge should primarily facilitate the acquisition of LS structures. Linearly separable and nonlinearly separable (NLS) categories were used to examine this possibility. In terms of the abstract notation, for the LS task the members of one category were 1110, 1011, 1101, and 0111 whereas the members of the contrast category were 0001, 1100, 0110, and 1010. If the typical values (i.e., the 1's) of the first category were summed, then every member of the first category would have more typical values than every member of the second category. If summing features is a natural strategy, then these categories should be easy to learn.

For the NLS task, the members of one category were 1000, 1010, 1111, and 0111, whereas the members of the contrast category were 0001, 0100, 1011, and 0000. In this case there is considerable overlap between the two categories in terms of the numbers of typical features that the examples possess. Thus, summing the typical features would not accurately partition the categories. In the social conditions the 1's and 0's were represented by features of people and in the object conditions the 1's and 0's were represented by features of animals.

If summation is a salient strategy in the social conditions but alternative integration strategies are salient in the object conditions, then the LS structures should be easier to learn in the social conditions. If configural information is salient in the object conditions, however, then the NLS structures should be easier to learn in the object than the social conditions. (The examples in the NLS categories have higher within category similarity and less between category similarity than the examples in the LS categories. Thus, if the overall similarity of individual examples is salient in the object domain, then the NLS categories should be easier to learn than the LS categories.)

Method

Subjects. The subjects were 152 undergraduate students who participated in the experiment in partial fulfillment of course requirements.

Stimuli and procedure. Active vs. passive categories were used as the social stimuli and predator vs. prey categories were used as the object stimuli. In the social conditions the 1's in the above examples were represented by active features (e.g., *talkative*) and the 0's were represented by passive features (e.g., *thoughtful*). In the object conditions the 1's were represented by features that were rated to be characteristic of predators (e.g., *muscular*) and the 0's were represented by features that were rated to be characteristics of prey (e.g., *armored*).

The examples were presented individually during learning and the task was to classify each example as *active* or *passive* in the social conditions, and as *prey* or *predator* in the object conditions. After a response was selected, the subjects were informed of the correct classification. The examples were presented in blocks, and in each block all eight of the examples were presented in a random order. The learning task continued until a subject had two consecutive errorless blocks or until sixteen blocks had been completed.

Results and discussion

The LS structures were much easier to learn in the social than the object condition (an average of 16.24 errors versus an average of 30.79 errors). The opposite result occurred with the NLS structures, however, as fewer errors were made in the object than the social condition (24.03 vs. 28.70 errors). This interaction between stimulus domain and category structure was highly significant, $F(1,148) = 23.06$, $p < .001$, $MSe = 156.56$. The LS structures were significantly easier to learn in the social than the object conditions, $F(1,74) = 20.81$, $p < .01$, $MSe = 193.39$, but the NLS structures were easier to learn in the object than in the social condition, $F(1,74) = 3.88$, $p = .05$, MSe 119.95.

These results indicate that the social materials were highly compatible with the LS categories. The results also indicate that social materials do not produce general facilitation in learning. Instead, ease of learning appears to depend on the compatibility between the form of encoding that is induced and the structure of the to be learned information. The encodings that were induced by social knowledge were compatible with LS but not NLS structures. The encodings that were induced in the object conditions, however, were more compatible with the NLS than the LS structures.

General Discussion

Across several experiments that used a wide variety of stimulus materials and categorization tasks linearly separable categories were more compatible with social than object materials. This result was clear and consistent and held regardless of whether the task involved category construction or category learning.

These results appeared to be attributable to differences in the knowledge structures that were activated in object and social domains. Participants in the object and social conditions seemed to rely on different types of information (e.g., concrete instances vs. abstract representations) and these differences produced different types of integration strategies. In social conditions, summing individual features was a frequent and natural integration strategy. In object conditions, however, clear influences of configural properties, feature conjunctions, specific exemplars, and heavily weighted single dimensions were observed.

A clear implication of these results is that the naturalness or learnability of abstract structures will vary with domain. The structure of knowledge appears to vary with domain, and consequently abstract structures or principles such as linear separability will be more important in some domains than others.

The results indicate that there are a number of important differences between object and social categorization systems. These differences reflect, in part, differences in the structure and nature of the domains. Objects are highly structured entities that can be directly perceived. The structure and organization of our theories of people is less clear, less precise, and more variable, and social concepts are based on interpretations, inferences, and constructions rather than direct perceptions. These basic differences in object and social domains appear to produce many differences in the structure of knowledge and in categorization processes.

The results indicate that there are important differences between conceptual systems, and constraints that operate in one domain might not be important in another domain (see Keil, 1981, 1990; Medin, Wattenmaker, & Mickalski, 1987 for related discussion). However, the results should not be interpreted as supporting the conclusion that there will be no general principles or constraints that extend across systems. Indeed, the way that knowledge, integration strategies, and abstract structures interacted to determine ease of learning appeared to be consistent and systematic for all of the experiments. Thus, regardless of the domain, an awareness of the knowledge and integrations that will be activated can be used to predict ease of learning. The results do suggest, however, that it might be more productive to formulate domain general constraints in terms of processes rather than in terms of products of processes (see Medin et al). It might not be possible to state that linearly separable structures will always be easy to learn, but it might be possible to state constraints in terms of the process by which knowledge, encodings, and abstract structures interact to determine ease of learning.

References

Asch, S. E., & Zukier, H. (1984). Thinking about persons. *Journal of Personality and Social Psychology, 46*, 1230-1240.

Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B.B. Lloyd (Eds.), *Cognition and categorization* (pp. 169-211). Hillsdale NJ: Erlbaum.

Cantor, N., Smith, E. E., French, R., & Mezzich, J. (1980). Psychiatric diagnosis as prototype categorization. *Journal of Abnormal Psychology, 89*, 181-193.

Gluck, M. A., & Bower, G. H. (1988). Evaluating an adaptive network model of human learning. *Journal of Memory and Language, 27*, 166-195.

Goldstone, R. L., Medin, D. L., & Gentner, D. (1991). Relational similarity and the nonindependence of features in similarity judgements. *Cognitive Psychology, 23*, 222-262.

Keil, F. C. (1981). Constraints on knowledge and cognitive development. *Psychological Review, 88*, 197-227.

Keil, F. C. (1990). Constraints on constraints: Surveying the epigenetic landscape. *Cognitive Science, 14*, 135-168.

Kunda, Z., Miller, D. T., & Claire, T. (1990). Combining social concepts: The role of causal reasoning. *Cognitive Science, 14*, 551-577.

Medin, D. L., & Schwanenflugel, P. J. (1981). Linear separability in classification learning. *Journal of Experimental Psychology: Human Learning and Memory, 7*, 355-368.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation., In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual Review of Psychology, 35*, 113-118.

Medin, D. L., Wattenmaker, W. D., & Michalski, R. S. (1987). Constraints and preferences in inductive learning: An experimental study of human and machine performance. *Cognitive Science, 11*, 299-339.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*, 207-238.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 104-114.

Nilsson, N. J. (1965). *Learning machines.* New York: McGraw Hill.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge. MA: Harvard University Press.

Waldman, M. R., & Holyoak, K. J. (1990). Can causal induction be reduced to associative learning? *Proceedings if the Twelfth Annual Conference of the Cognitive Science Society* (pp. 190-197). Hillsdale, NJ: Erlbaum.

Wattenmaker, W. D., Dewey, G. I., Murphy, T. D., & Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties, and concept naturalness. *Cognitive Psychology, 18*, 158-194.

Wattenmaker, W. D. (1993a). Knowledge structures and linear separability: Integrating information in object and social categorization. *Cognitive Psychology. Accepted pending revision.*

Wattenmaker, W. D. (1993b). Incidental concept learning, feature frequency, and correlated properties. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1, 203-222.