Consistent contrast and correlation in free sorting

DORRIT BILLMAN AND JIM DAVIES
Georgia Institute of Technology

Two experiments investigated free sorting, a type of unsupervised learning, with multiattribute drawings of alien animals. In previous research on concept formation, with simpler stimulus structure than ours, participants were insensitive to correlational structure in the stimuli, producing primarily "1D sorts," based on the values of just 1 dimension or attribute. Our experiments showed that participants used many strategies in categorizing but preferred to generate groupings that reflected the correlations in input when this did not violate consistent contrast. The second experiment used hierarchically structured stimuli to show that participants' sort strategies favor consistency within a set of contrasting categories, distinct from any preference for 1D sorting. Finally, both experiments show that correlational sorts are much more likely when the correlation-based sort contrasts consistently. Our data show complexity at work in free sort tasks: People are sensitive to multiple and sometimes conflicting biases for consistency and correlational structure in the category systems they create.

Many researchers have suggested that natural categories capture correlational structure in the world (Rosch, 1978). Some have characterized this structure as one of family resemblance (FR). One approach to investigating the formation of everyday categories is to use free sorting to look at novel category construction (Ahn & Medin, 1992) rather than to reveal the structure of existing categories (Inhelder & Piaget, 1964). In free sort tasks participants are given a collection of items and asked to group them into categories as they see fit. In many situations (Ahn & Medin, 1992; Regehr & Brooks, 1995; Medin, Wattenmaker, & Hampson, 1987; Spalding & Murphy, 1996) adults sort items into groups based on a single dimension, or attribute, a phenomenon called 1D sorting. Participants produce 1D sorts even when there is correlational structure in the input, which could provide an alternative basis for grouping; specifically, researchers have built in FR structure, intended to mimic natural category structure. A reliance on just a single dimension has been surprising to many researchers because participants seem insensitive to just the sort of multiattribute correlational structure that appears in many natural categories.

Medin et al. (1987) had limited success pushing participants away from 1D sorting by varying instructions. Regehr and Brooks (1995) were able

AMERICAN JOURNAL OF PSYCHOLOGY
Fall 2005, Vol. 118, No. 3, pp. 353–383

to increase FR sorting by changing the task so participants were given the prototype, were told to make groups around it, and saw only the last-placed instance. Several variations in stimuli—content, perceptual versus text, qualities of perceptual attributes, details of distributions of attribute values, and numbers of correlated attributes—also had little effect (Ahn & Medin, 1992; Medin et al., 1987; Regehr & Brooks, 1995).

Causal links and the nature of exposure do affect frequency of free sorts. Medin et al. (1987) found that participants produced more FR sorts when correlations were strong and the correlated properties were linked by prior knowledge (e.g., dizziness and earache). FR sorts are more likely when the FR reflects causal links, either by providing scenarios introduced in the experiment or by activating previously known causal relationships (Ahn & Medin, 1992; Ahn, 1999; Medin et al., 1987; Murphy, 2001, on thematic relations; Murphy & Kaplan, 2000; Spalding & Murphy, 1996). FR sorting is also more likely if additional, initial exposure to instances is provided (Kaplan & Murphy, 1999) and if participants predict the value of one attribute given the value of another (Lassaline & Murphy, 1996).

The two experiments reported here looked for situations in which people discover and use correlational structure, thereby relying on more information than one dimension in free sort tasks. We looked for characteristics of the overt stimulus structure that might lead to correlational sorting. We used the traditional free sort task in which all instances are available at the time of sorting rather than sequential exposure.

Correlations, categories, and FR

Correlation has a range of meanings in the concept learning literature: the relationship between quantitative dimensions such as height and weight; a relationship between values of categorical attributes, such as religion and native country; and a predictive relationship or association between two features or predicates such as having a beak and having feathers. We use *correlation* in a broad sense to refer to any of these predictive relationships between properties. We use *property* broadly to refer to any instance characteristic, including a characteristic in which a set of instances varies (attributes of height, religion), particular attribute values (6-foot, Catholic), and features not necessarily thought of as a value of an attribute (magic, has hands). In our experiments, all stimulus variation was categorical rather than continuous. Therefore, *correlation* in our stimuli is the same as *co-occurrence.*

Correlations between properties in a domain might be used as the basis for category formation: Clusters of entities that share multiple, interpredictive properties can be grouped together, based on these co-occurring properties. In such cases, knowing the category of an instance tells you something about the properties of the instance (category validity), and knowing these properties tells you something about category membership (cue validity). Knowing something is a bird gives information about many (but not all) of its properties, and many (but not all) properties are predictive of being a bird. "Family resemblance category" typically refers to category structure in which multiple, correlated properties are likely to be true of category members but there is no property true of all members; we use *FR* to refer to the combination of these two properties.

The correlation structure of a domain involves both the strength of the component associations and their organization. Both are important in understanding when people are sensitive to correlational structure and hence when people might use it in free sorting.

Strength of correlations in category formation

We suspect that in unsupervised learning, correlations must be strong to be noticed. Strong correlation may be particularly important in the initial sample of instances; the stimuli presented in a free sort task provide the initial sample instances for category formation. Furthermore, when participants are learning natural categories through exposure to examples, initial samples may often provide perfect correlations between many relevant properties. Correlation-breaking exceptions (penguins, unripe oranges, or beanbag chairs) may be rarely encountered, and learning them may depend on (later) labeling or other feedback.

In contrast, the correlations typically afforded by stimuli in free sorting are surprisingly low. Table 1 shows the structure used in Ahn and Medin (1992) and in many subsequent experiments. Reading the schema with the instances organized into categories as shown in Table 1, the structure given the grouping appears quite strong. However, both the contingency between attributes and the pairwise association between particular values are low. The correlation between any pair of attributes is only 0.2. The pairwise conditional probability of one attribute value given another is only $p(A2 = 1|A1 = 1) = .6$, close to the base rate of $p(A2 = 1) = .5$. If sample

Table 1. Frequently used stimulus schema grouped by family resemblance

0000	1111
1000	0111
0100	1011
0010	1101
0001	1110

Note. The two columns show instances from two categories. Each instance differs from the prototype (0000 and 1111) in a single dimension or attribute.

sizes are kept small (as needed to make free sorting manageable), it is not possible to introduce any exceptions and maintain strong correlations.

However, opting for small samples with high correlations (indeed, perfect ones) may be at least as representative of naturally occurring unsupervised learning as is providing samples with very low correlations. We used perfect correlations in these experiments.

Organization of correlations in category formation

A set of correlations can be organized differently, largely independently of the strengths of individual correlations. Two aspects of organization are important in our stimulus design: distribution of information that is localized or uniform across attributes and consistency of contrast. In natural domains and natural categories, important information is not uniformly distributed but is localized in a subset of properties. Some properties carry more information than others (i.e., more informative properties correlate more widely or more strongly than less informative ones), and discovering which properties are more informative typically is a key part of learning. To our knowledge, no one disagrees with this idea, but the information in many free sort experiments is uniformly distributed. Thus there is no utility to discovering and preferentially using the most informative properties because all are equal. One purpose of our experiments is to ask whether and how often strong, localized correlations will be used in free sorting. Therefore, properties in our stimuli differ in how predictive they are: The information is localized, not uniformly distributed across attributes.

A second purpose is to investigate how the organizational principle of consistent contrast influences correlational sorting. Given strong correlations localized in a subset of properties, how does the organization of these correlations influence their use in free sort? We expected correlational sorts to be much more likely when this allows consistently rather than inconsistently contrasting categories.

Consistency of contrast reflects the underlying correlation structure in the domain and is defined relative to a set of categories. A system of categories contrasts consistently[1] if the attributes that are important in characterizing and distinguishing one category are the same attributes that are important in characterizing and distinguishing other categories in the contrast set. A contrast set of categories is a set of mutually relevant alternatives, typically subtypes of the same higher-level category. Thus, the categories of "DOG," "BEAR," and "MONKEY" are in contrast with each other. They contrast consistently to the extent that the same attributes (e.g., vocalization, diet, and shelter) are relevant to each: Dogs bark, eat bones and live in doghouses; bears growl, eat honey, and live in caves; and monkeys chatter, eat bananas, and live in trees. (Furthermore, the attributes, perhaps fur length and health, that are irrelevant to membership in one category are irrelevant to others.)

Inconsistent contrast can be found in social role categories. The attributes typical of one occupation often are different from those important for another. Personality and appearance are important for our concept of actors; specifically, actors tend to be extroverted and attractive, whereas education and perhaps prosperity are important for doctors (be they handsome or homely). Furthermore, ranchers usually come from certain regions of the country and have land-holding families.

Consistent contrast is relative to a contrast set of categories. For example, types of birds and types of fish form different contrast sets. If a domain has been organized into multiple contrast sets, as will be the case in hierarchically organized domains, consistency is assessed within and relative to each contrast set. Consistency of contrast overall will be high if the attributes distinguishing kinds of birds are consistent and the attribute set distinguishing kinds of fishes are consistent, even if the attributes important for birds are completely different from the attributes important for fishes.

The influence of consistent contrast has been found in a variety of tasks. Consistently contrasting categories support the formation of overhypotheses (Goodman, 1983); combining an overhypothesis and a single example of a novel category tells one which attribute values of the sample instance should be generalized over the new category. Children (Shipley, 1993) and adults (Nisbett, Krantz, Jepson, & Kunda, 1983) seem to use overhypotheses for organizing and projecting category knowledge. Consistent contrast facilitates concept learning with feedback (Billman & Davila, 2001; Billman, 1996).

Alignability is a related principle that has also been found to influence supervised concept learning (Kaplan, 1999; Sifonis & Ross, 2002; Lassaline & Murphy, 1998; Yamauchi & Markman, 2000). High consistency of contrast (a relationship between categories) depends on fairly high alignability (a relationship between instances). The instances in a domain must be representable in terms of largely the same attributes in order for the categories in the domain to be distinguished by the same attributes.

To our knowledge no prior work has investigated the role of consistent contrast in free sort or any unsupervised learning tasks. We predict that in free sort, as in other category tasks, consistent contrast plays an influential role facilitating use of correlations. Specifically, correlations will be discovered and used much more often when organized to afford consistently contrasting categories.

Explaining prior findings

Ahn and Medin's (1992) two-stage model is the most influential model of free sorting; we consider models focusing on other unsupervised concept formation tasks in the discussion. According to the two-stage model, in the first stage participants select a dimension based on salience of the attribute and the distinctiveness of its values and use it to partition instances. The second stage is entered only if the required number of categories does not match available values of the chosen attribute, as when two categories are required but the chosen attribute has many values. The first stage can produce only 1D sorts, so all other sorting types, such as FR, can be produced only in the second stage. However, the second stage would never be entered for our task and stimuli because we did not specify how many categories the participants should produce. This model was designed to explain the prevalence of 1D sorting. When there is no prespecified number of categories, it does not predict any sensitivity to correlations or relationships between dimensions. It predicts that all participants will generate 1D sorts in our experiments and predicts the high frequency of 1D sorting found in prior studies.

Several different strategies or beliefs might produce 1D sorts in preference to correlational sorting; probably multiple strategies are involved, given the unconstrained nature of the task. Some possibilities are that 1D sorting might be used because it is very easy, because only one attribute need be attended to and little information about the domain is needed; 1D sorting might be used on principle because participants think concepts (at least those in the experiment) should have only one attribute implicated (e.g., they might treat the task as a method for indicating which attribute is most important); 1D sorts might be favored because the structure of 1D sorts often covaries with some related characteristic, such as short description length, and presence of perfectly diagnostic cues (i.e., sufficient features); and 1D sorts might be used as a fallback strategy when the participants do not discover or notice correlational structure.

A background assumption seems to be that the correlations in FR structure are evident, but participants prefer to produce 1D sorting instead. In contrast, we predict that people do care about correlations when they can discover them. Discovery is very difficult when all attribute values are weakly related to all others. We predict that when correlations are strong and hold in a localized and consistent set of attributes, people will discover correlation frequently and then will use it in free sort. Furthermore, we predict that discovery and use of correlations is contingent on consistency. When equally strong contingencies between properties are not organized to produce consistent contrast, those contingencies will be very difficult to notice and hence very rarely used in sorting.

We believe that preference for 1D sorting is not a desire to use just a single dimension but a desire for consistency. If participants discover correlated attributes converging on the same grouping, we believe participants will prefer, not be indifferent to, these correlation-capturing groups. Put differently, correlation between attributes may draw attention to those attributes or prioritize them as important. In addition, categories that can be distinguished by one dimension need not have been generated by a 1D strategy. For example, it is possible that mammals, birds, and fish can be distinguished by their bone structure, but the categories might be motivated by and formed because they capture a rich amount of correlated information.

Hypotheses

The primary goal of both experiments is to test sensitivity to correlated properties in free sorting and to test how consistent contrast modulates sensitivity to correlation. Consider our illustrative comparison between animal and social categories. Broadly, our experiments asked, first, whether people would notice co-occurrences between properties and produce categories based on such associations (e.g., barking and eating bones; growling and eating honey; and chattering and eating bananas). Second, the experiments asked whether people would notice and use correlations to produce consistently contrasting categories more than they use associations motivating inconsistently contrasting categories (e.g., being attractive and extroverted; having a graduate degree and high income; and owning large tracts of land and living in the plains states). The experimental conditions control many aspects of the stimuli, such as the particular attributes and values used, strength of association, and availability of necessary and sufficient features.

We made four predictions: that participants would frequently discover and use correlated properties in free sort when correlations were strong and localized in a subset of attribute values and when the resulting correlation-capturing categories contrasted consistently; that correlation-based sorting would be very unlikely if the resulting categories did not contrast consistently, even when stimuli were equated on many other aspects; for Experiment 2, that correlation-based sorting would persist even when the correlation-based sort could not be generated by any 1D strategy when consistent contrast was preserved; and for Experiment 2, that participants would produce much more complex groupings than the usual two categories. We designed the hierarchical target sort in Experiment 2 to strongly separate sorting based on consistent contrast from 1D sorting, but we were also interested in simply exploring the discovery of correlation in much more complex structures.

In Experiment 1, two types of evidence tested our prediction. First, we compared actual frequency of correlational sorts with the frequency predicted if the sorts were "accidentally" produced by participants independently choosing attributes without regard for correlation. Second, we analyzed the answers participants gave when asked about the basis for their sorts.

Experiment 2 also tested for correlated sorting and its dependence on consistent contrast. It used a more complex, hierarchical stimulus design with multiple contrast sets within the hierarchy. In the consistent contrast condition a consistent set of attributes was used in each contrast set. Critically, different attributes were used for each contrast set. Thus, although the correlations were perfect, as in Experiment 1, there was no 1D sort that could be confused with correlational sorting. Here correlational sorting conflicted with, rather than selected among, 1D-generable sorts.

EXPERIMENT 1

Experiment 1 tested for sensitivity to correlation and hence reliance on more than one dimension in sorting. Our primary reason for using perfect correlations was to make them more easily noticed. In addition, it meant that the categories produced by correlation-based sorts (in both conditions) also had sufficient features and allowed short description length; these are characteristics of 1D sorts. Because our correlational sorts are based on perfect correlations, it is also possible to generate the correlational sort by using a 1D-strategy. Thus, to test whether participants made their sorts based on correlations, we tested whether sorting on correlated dimensions was preferred over the rate predicted if participants were insensitive to correlation.

Experiment 1 also tested whether consistency of contrast is important for discovery and use of correlations by comparing the frequency of correlational sorts in the consistent and inconsistent contrast conditions.

Finally, Experiment 1 provided information about the role of necessary and sufficient features. The correlational sorts in the consistent and inconsistent contrast conditions have equally predictive cues and an equal number of necessary and sufficient features. In both conditions each category in the correlational sort has two perfectly associated attributes. If either sufficiency or necessity were the critical factor for discovery and use of correlational grouping (independent of consistency), then correlation-based sorts should occur in the inconsistent contrast condition with comparable frequency as in the consistent contrast condition.

METHOD

Participants

Forty-nine Georgia Institute of Technology undergraduate students volunteered for extra credit (37 in the consistent and 22 in the inconsistent contrast condition). We ran more participants in the consistent contrast condition because a primary goal was assessing participants' production of correlational sorts in this condition and because three stimulus configurations were needed (between-subject design).

Stimuli

The stimuli were drawings of 18 imaginary alien creatures presented on individual cards (Figure 1). In both consistent and inconsistent conditions, stimuli varied on the six values of each of the six attributes (hair, tail, nose, feet, arms, and environment). In both the consistent and the inconsistent conditions six pairs of these attribute values covaried perfectly. In both conditions these pairs of correlated values provide a seed for six possible correlation-based groups, each based on two perfectly correlated cues.

Table 2 shows the differences between conditions in how these perfectly correlated pairs were assigned. Columns represent attributes, and rows represent category schema. Each schema is instantiated three ways to produce a set of three instances in each target category (details of randomized assignment of uncorrelated values are described later). Thus, both conditions provided the basis for six categories capturing correlational structure, with each category based on two attribute values that always occurred together. For example, pointy hair and wiggly tail always went together for some participants in both conditions, providing a potential grouping of instances in both conditions.

The remaining, uncorrelated attribute values were assigned quasirandomly with two constraints. First, we ensured that no strong spurious correlations were introduced in either condition. Second, we ensured that assignment of "random"

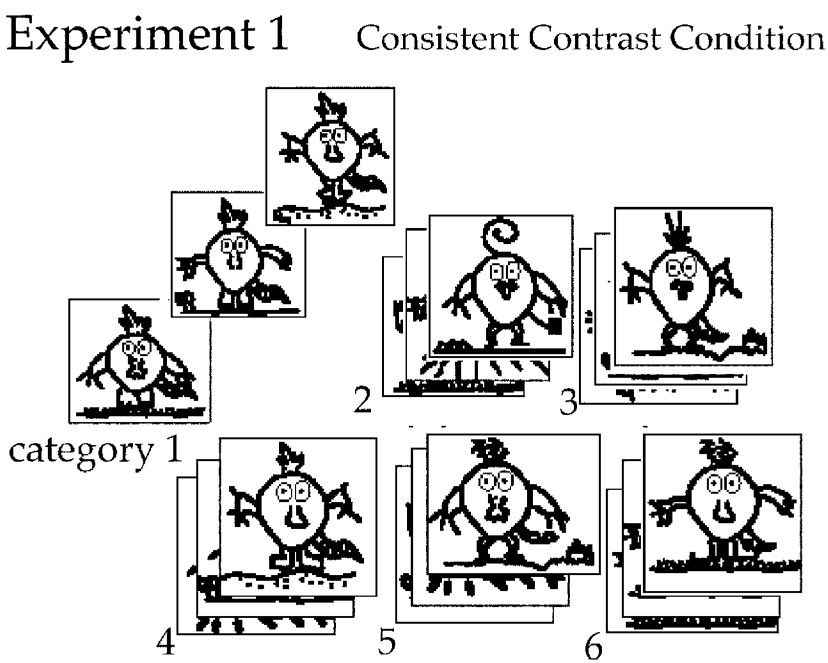

Figure 1. Example stimuli for the consistent contrast conditions of Experiment 1, arranged to illustrate the correlated grouping

Table 2. Stimulus schema for consistent and inconsistent contrast conditions in Experiment 1

Consistent contrast condition			Inconsistent contrast condition		
A1–A2	A3–A4	A5–A6	A1–A2	A3–A4	A5–A6
aa	xx	xx	**aa**	yy	yy
bb	xx	xx	yy	**bb**	yy
cc	xx	xx	yy	yy	**cc**
dd	xx	xx	**dd**	yy	yy
ee	xx	xx	yy	**ee**	yy
ff	xx	xx	yy	yy	**ff**

Note. Columns indicate attributes A1–A6; each attribute has 6 values. Letters *a–f* indicate attribute values that correlated with another attribute value and provided the basis for a possible correlation-based category in free sort. Letters *x* and *y* indicate any of several uncorrelated and quasirandom attribute values. The *x*s in the consistent contrast condition took any of the 6 attribute values. The *y*s in the inconsistent contrast condition had a restricted range of only 4 values, excluding the 2 correlated values (e.g., values *a* and *d* for A1 and A2 were excluded). This restriction was needed to exclude instances such as *aa bb xx* with crossover values, diagnostic of 2 alternative groups (e.g., A1 = *a* and A2 = *a* but also A3 = *b* and A4 = *b*).

attribute values could not undermine interpretation of our results. Because assignments cannot be equated perfectly between conditions, we ensured that assignment of values slightly favored producing the correlational sort in the inconsistent over the consistent contrast. To do this, we restricted assignment of random attribute values in the inconsistent contrast condition to four rather than the full six values (excluding value numbers 1 and 2). This eliminates potential crossover instances, which would disrupt correlations in the inconsistent (but would not in the consistent) condition, and it produces similarity relationships favoring the correlational clustering in the inconsistent versus consistent condition, thus working against our hypothesis (see *Appendix* for additional detail). Our assignment also means that each category in the correlated sort in the inconsistent contrast condition, as in the consistent, has two necessary and sufficient features.

We controlled for effects of attribute salience by using three between-participant configurations of the consistent contrast stimuli, with different pairings of attributes correlating consistently (hair and tail correlated Configuration 1, nose and feet in Configuration 2, or arms and environment in Configuration 3). Because all attributes contributed correlations in the inconsistent contrast condition, just one configuration was needed.

Procedure

Participants were told to imagine they were a zoologist studying alien creatures from the Jovian moon of Callisto. Each was told to study pictures of the animals and then sort them into groups that "make the most sense to you." They were told to make as many groups as they liked. Participants sorted the deck of stimuli given to them in groups on a table. After they were finished, they wrote down a description of how they had sorted (some participants then piloted an induction task; these results are not discussed here). Next, the idea of correlation was explained, and they were asked whether they had noticed any correlations between properties of the creatures, and if so, what these were. Finally, participants were thanked and debriefed.

Scoring

Sorts were scored for whether participants matched the target (correlational) sort, whether they sorted on a single (uncorrelated) attribute, or whether they produced a multiattribute sort (for which no single attribute could reproduce the generated sort.)

We scored participants' sort descriptions in order to capture all sort rationales and to identify any sensitivity to correlations. Descriptions from the consistent contrast condition were coded as one of the following six types: The two correlated attributes were both given as the basis for sorting (or multiple pairs of correlated attribute values), a single correlated attribute was given as the basis, some correlated attribute was mentioned but as part of a larger set of attributes including uncorrelated ones, a single uncorrelated attribute was given, a set of attributes was listed, with no correlated attributes, and no attributes were mentioned. Descriptions from the inconsistent contrast condition were coded as one of the following four types: multiple pairs of correlated attribute values, one pairing of correlated attribute values, multiple attributes given but no correlated pairing, and a single attribute. We scored the final correlation questionnaire both for whether the participant reported noticing the target correlation and for whether the participant reported other properties as co-occurring.

RESULTS

Performance in the consistent contrast condition

Our first question was whether participants in the consistent contrast condition sorted in accord with the correlated attributes. We assessed this from the sorts the participants generated and from what participants said about these sorts. Twenty of 37 participants (54%) produced the exact sort defined by the correlated attributes (Table 3). We compared the actual proportion of correlational sorts to chance to test for sensitivity to correlational structure. If participants had been sorting by a single attribute and unaffected by correlation, they would have a two in six chance of producing the target sort by relying on one of the two correlated attributes independently. They exceeded this, showing an active preference for correlational sorts. Comparing the frequency of correlational sorts to this two in six comparison point shows that participants were reliably sensitive to correlations in their sorts (binomial $p = .003$ where $P(\text{success}) = .33$).

We controlled for effects of initial attribute salience by counterbalancing which attributes were correlated across the three configurations. Neverthe-

Table 3. Experiment 1 numbers (and percentages) of sort types and sort descriptions in consistent and inconsistent conditions

	Correlational sort	Noncorrelational sorts	
	Sort description	Single attribute sufficient	Multiattribute
Consistent contrast ($n = 37$)			
Correlated attributes	13	—	—
Single correlated attribute	7	—	—
Some correlated attribute part of a larger set	—	—	2
Single uncorrelated attribute	—	13	—
Set of uncorrelated attributes	—	1	—
Not attribute-based description	—	—	1
Sort totals	20 (54%)	14 (38%)	3 (8%)
Inconsistent contrast ($n = 22$)			
Correlated attribute values, multiple pairs	—	—	—
Correlated attribute values, 1 value pair	—	2	—
Multiple attributes, no correlated pair	—	—	5
Single attribute	—	14	1
Sort totals	0	16 (73%)	6 (27%)

Note. "—" means no descriptions in the category.

less, it would be useful to know whether discovery or use of correlations was only found in certain configurations. Table 4 shows the distribution of sorts grouped by the attributes used and by stimulus configuration. Sorts on correlated attributes were similarly preferred in all configurations. Furthermore, attributes in each pair were two to six times as likely to be used, given a configuration where they correlated compared with any configuration where they did not.

Sort descriptions provide another measure of the principles used in sorting. The relationship between sort type and sort description is shown in Table 3. When asked to describe their sort, 13 of the 20 participants who produced the correlated sort reported using the pair of correlated attributes to sort (and mentioned no others). Thus, on this intent-based measure 65% of participants who sorted on the correlated attributes clearly intended both attributes to be important.

Table 4. Distribution of correlated sorts in the consistent contrast condition by configuration and attributes used in sorting, Experiment 1

	Attributes used in sort			
Configuration	Hair or tail (13 uses)	Nose or feet (9 uses)	Arms or environment (12 uses)	Other sort types
Hair–tail correlated ($n = 13$)	8	1	3	1
Nose–feet correlated ($n = 12$)	2	6	3	1
Arms–environment correlated ($n = 12$)	3	2	6	1
Correlated use as % of all use	62%	66%	50%	

Note. Bold numbers on the diagonal indicate sorts based on the attributes correlated in that configuration.

Finally, by comparing initial sorts and sort descriptions with answers to the final direct questioning about correlations, we have information about the relationship between discovering the correlations and using them in sorting. Do participants use correlations when they notice them, or do they notice the correlations but decide against using them? Thirteen participants reported the objective correlation and no spurious correlation; these were the 13 who produced the correlational sort and described their sort in terms of the two correlated attributes. Two additional participants mentioned the target attributes as correlated but also mentioned one or more additional, uncorrelated attributes; one of these participants produced a correlational sort but mentioned only one correlated attribute in the initial description, whereas the other produced a multiattribute sort. The remaining 22 participants made no claims about correlations between the objectively correlated attributes either in their initial description or after questioning about correlations; six mentioned other relationships between uncorrelated attributes. In sum, there were no participants with explicit access to the correlation who did not use it to sort; hence there is no evidence of participants noticing the correlation but deciding to sort on another basis. Rather, correlational sorting seems limited by the difficulty of noticing correlations.

Inconsistent to consistent contrast comparison

To assess whether discovery and use of correlations depends on consistency of contrast, we compared the proportion of correlational sorts in the consistent condition with the proportion in the inconsistent condition. As shown in Table 3, no participant in the inconsistent contrast condition produced a correlational sort, significantly fewer than the 54% in the consistent contrast condition, χ^2 likelihood ratio = 24.51, $p << .001$; Goodman & Kruskal's tau = .30, $p << .001$.

On the sort descriptions, only two participants in the inconsistent contrast condition (both producing single attribute sorts) showed any partial knowledge of correlations in their descriptions. These two described their sort as based on a pair of attributes (hair and tail) which had in fact correlated in a subset of the instances, although their sort could be generated by just one of the attributes. The participants producing multiattribute sorts frequently mentioned multiple attributes in their descriptions, but none suggested a knowledge of which attributes correlated.

On the final direct questioning about correlations, five inconsistent contrast participants identified one of the three pairs of objectively correlated attributes, but two of these also mentioned an additional unrelated attribute. Thus only three participants provided clear evidence that they were aware of any actual predictive patterns (one of the three pairs of associated properties). No participants provided evidence of learning much about the correlational structure, and none produced the correlational sort. Participants were unsuccessful in detecting correlation in the inconsistent contrast condition.

Role of 1D sorting. Our stimuli were designed so that correlational sorting in the consistent contrast condition would select from rather than conflict with sorts that could be characterized by a single attribute. In the sorts produced, 8% in the consistent and 27% in the inconsistent condition were incompatible with both correlational sorting and 1D sorting. In the sort rationales, 46% (17 of 37) in the consistent and 32% (7 of 22) in the inconsistent condition reported use of more than one dimension in generating the sort.

Summary. Participants in the consistent contrast condition relied on correlation: They sorted on attributes that were correlated more than the sum of reliance on the same attributes when uncorrelated. Put differently, correlation was the basis for selecting attributes for sorting. This simultaneously argues for sensitivity to correlation and that participants are not just relying on one dimension in sorting. Second, discovery and use of correlational structure depended on consistency of contrast. None of the participants in the inconsistent contrast condition reported or sorted on correlational structure. This was true despite the fact that the attribute values were perfectly correlated, affording two perfectly reliable cues for each group in a correlation-based sort, in the inconsistent as well as consistent condition. Organizing predictive cues at the level of relationships between attributes seems critical for noticing co-occurrence, even when properties are categorical. Third, the reported data support the claim that correlation is valued in free sorting and will be used when discovered.

Finally, across the conditions we found diverse sorting, including sorts that are neither 1D nor correlational.

EXPERIMENT 2

Experiment 2 expands the complexity of the stimulus structure, using stimuli in which the correlational structure invites a hierarchical sort. Our goals are to test for correlational sorting even when such sorts conflict with any 1D sort, to expand the search for correlational sorting to hierarchical organization, and again to ask whether compatibility with consistent contrast is important for people to discover and use correlations.

To do this, the stimuli had correlational structure at two levels. This structure would motivate sorting stimuli into two broad, superordinate categories with three subordinate categories in each. This structure forms three contrast sets: the split into two superordinate categories and the split of each superordinate into its three subordinate categories. In the consistent contrast condition, the different contrast sets would each be organized around a different set of correlated and consistent attributes. Thus, participants in the consistent contrast condition could use correlational structure to produce consistently contrasting categories in their free sorts. However, these sorts could not be produced by use of any 1D-sorting strategy because different attributes are needed at each branch point in the hierarchy. (We consider several extensions of 1D sorting models in the *Discussion.*) In the inconsistent contrast condition, different attributes correlated for each category within a contrast set (producing the inconsistent contrast).

We predicted that participants in both conditions would produce a rich variety of sort types, including many hierarchical ones; that participants in the consistent contrast condition would frequently discover and generate the correlational sort; and that participants in the inconsistent contrast condition would be much less likely to discover and use correlations.

METHOD

Participants

Forty-three Georgia Institute of Technology undergraduate students volunteered to participate for extra credit, 21 in the consistent and 22 in the inconsistent contrast condition.

Stimuli

Commonalities between conditions. The stimuli were drawings of 18 alien creatures similar to those of Experiment 1. However, nine attributes were used (body shape, sound, body color, eyes, mouth, hair, tail, arms, and legs), and all values of these attributes were different from any used in Experiment 1. Table 5 shows the abstract schema for each condition, and Figure 2 shows example instances with perceptual attributes filled in for the schematic ones. Columns A1, A2, and A3 are the two-valued superordinate attributes (body shape, sound, and body color). Columns A4–A9 are the six-valued subordinate attributes (eyes, mouth, hair, tail, arms, and legs). Letters in the table show the correlated attribute values; *x*s represent quasirandomly assigned values of uncorrelated attributes (checked to exclude spurious correlations.) Attributes were paired to carry the correlations (A4–A5, A6–A7, and A8–A9).

Stimulus structure in consistent and inconsistent conditions shared three important commonalties. First, stimuli were designed to embody a hierarchical set of categories, two superordinates with three subordinate categories in each. Piloting confirmed that sound, body shape, and body color were very salient attributes. These three superordinate attributes covaried with each other, assumed either of two values, and provided a strong motivation for splitting the animals into the two broad superordinate target categories. Second, within each of these superordinate categories, subordinate attribute values co-occurred, providing a correlational basis for three subordinate categories. Considered from the perspective of three target subordinate categories shown in Table 5, each subordinate category had two attribute values that occurred only and always in that category, whereas category instances varied on the other four attributes. Third, across the six subordinate categories, different attributes carried the predictive information for different categories, ensuring that the correlational sort could not be produced by a 1D strategy.

For both consistent and inconsistent contrast conditions, there is no single attribute and no 1D sorting strategy that would produce the correlated sort. However, in the consistent contrast condition, sorting based on correlations would produce consistently contrasting categories because the correlated attributes change only when the contrast set changes.

Differences between conditions. Stimulus design in the consistent and inconsistent conditions differed in one critical respect: how the correlated attribute values were organized within each superordinate. For the consistent contrast condition the three pairs of co-occurring values (corresponding to three target categories) within one superordinate were values of the same pair of attributes.

Table 5. Stimulus schema for consistent and inconsistent contrast conditions, Experiment 2

Consistent contrast condition				Inconsistent condition			
A1A2A3	A4–A5	A6–A7	A8–A9	A1A2A3	A4–A5	A6–A7	A8–A9
aaa	**aa**	xx	xx	**aaa**	**aa**	xx	xx
aaa	**bb**	xx	xx	**aaa**	xx	**aa**	xx
aaa	**cc**	xx	xx	**aaa**	xx	xx	**aa**
bbb	xx	**aa**	xx	**bbb**	**bb**	xx	xx
bbb	xx	**bb**	xx	**bbb**	xx	**bb**	xx
bbb	xx	**cc**	xx	**bbb**	xx	xx	**bb**

Experiment 2 Consistent Contrast Condition

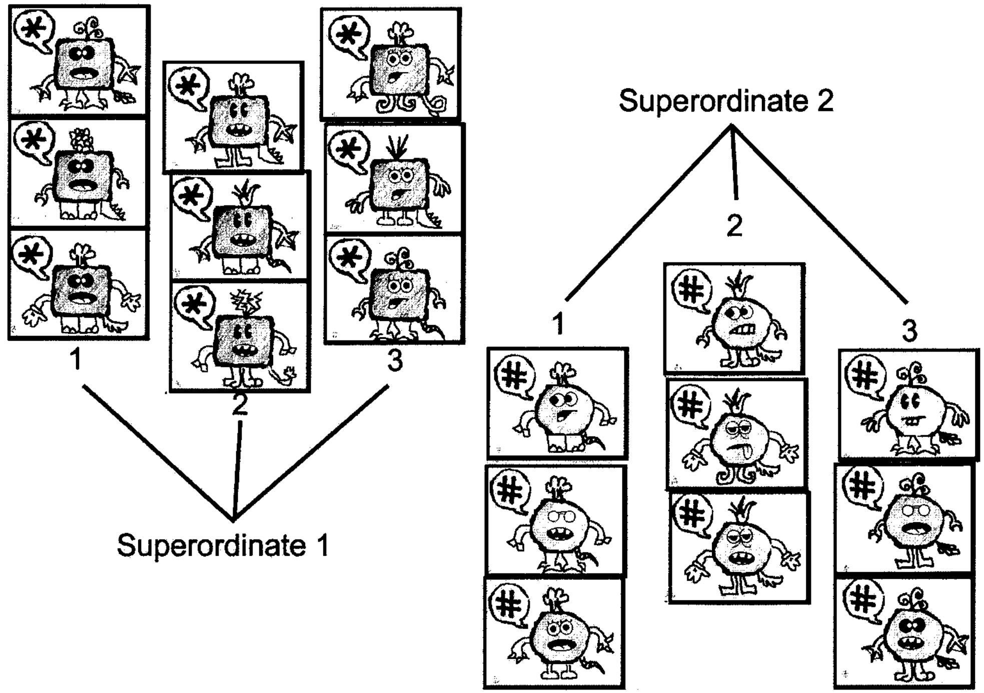

Figure 2. Example stimuli for the consistent contrast conditions of Experiment 1, arranged to illustrate the correlated grouping. In Experiment 2, the six categories are broken up into two superordinate categories, represented with the initial branching line

For example, in the superordinate category of shaggy, square creatures that said "*," the creatures with crossed eyes had an "ah" mouth, high eyes went with teeth mouths, and eyelash eyes went with tongue-out mouths. In the other superordinate the correlated attributes were hair and tail. Thus, knowing the superordinate predicted which attributes were informative at the subordinate level. Categories formed to capture the predictive structure would also contrast consistently, so a correlational sort would not violate consistent contrast.

For the inconsistent contrast condition, within one superordinate each of the three pairs of co-occurring values (for each of the three categories) came from a different pair of attributes. For example, in the superordinate category of shaggy, square creatures that said "*," crossed *eyes* went with "ah" *mouths* (as in the consistent condition), but poof *hair* went with striped *tails,* and three-finger *hands* went with paw *feet.* Thus, in the inconsistent contrast condition superordinates did not tell which attributes carried predictive information, which in turn could be the basis of informative categories. Here, sorting to capture correlational structure would create categories violating consistent contrast.

In the inconsistent contrast condition, "randomly" assigned attribute values could assume any of the six values of each attribute. However pairs of values predictive for another category were excluded. Thus, an *aaa aa xx xx* (category 1) instance might be assigned values such as *aaa aa bd fc* but not *aaa aa bb fc.* This restriction excluded crossover instances that had the pair of values associated with one category and also the value pair associated with another category.

Counterbalancing. To counterbalance the subordinate attributes, we constructed three configurations of the consistent condition stimuli: each of the attribute pairs (A4–A5, A6–A7, and A8–A9) was used twice, in two of the three configurations. In Configuration 1 eyes correlated with mouth in one superordinate, and hair and tail correlated in the other; in Configuration 2 eyes–mouth and arms–legs correlated; in Configuration 3 arms–legs and hair–tail correlated. Only one configuration was used in the inconsistent condition because all three pairs of attributes correlated for an equal number of categories; that is, in both conditions, each attribute pair was used for one third of the correlations.

Procedure

The experiment consisted of an exposure task, sorting task, description of the sorting rationale, and debriefing. Participants were told to imagine they were a zoologist studying pictures of alien animals on the Jovian moon Io. The exposure task was included to give participants more extended contact with these stimuli and the rather complex structure they embody. Participants were handed a deck of cards (same randomized order for all participants) and wrote a description of each picture, such that they could "draw the creature later." They viewed and wrote about the pictures one at a time but could look back over their descriptions. For the sorting task, participants spread out the cards so all were visible and then sorted them. They were asked to classify in the way "that makes the most sense to you as a zoologist on Io," to take their time, and to make as many groups as they saw fit. They were told they would later be asked to provide a rationale for the sort. When participants were finished with the sort, they were asked to write down their rationale such that someone else could "sort the creatures in exactly the same way." Participants could change their sorts when writing the rationale, although this was vary rarely done. The experimenter recorded the sort and debriefed the participant.

Sorts were scored for the structural character of the sort (hierarchical, single partition, matrix, other), whether the superordinate split determined by the three salient attributes was recovered, whether the full correlational sort was recovered, and whether the sort could be produced using a single attribute, within a level or overall. Our scoring looked for any type of structure, such as producing an array of categories with rows and columns grouping on different dimensions, or chaining, not just partitions or hierarchies.

Note that the overall predictive importance of each of the six subordinate attributes was equivalent in both conditions. Thus their global importance, salience, or weight should be the same, and an attentional learning model would predict no difference between conditions because predictive utility of each of the attributes is the same between conditions. Nevertheless, it may be helpful to think of the discovery of correlations as drawing attention to correlated attributes in a very strategic, context-specific way: The correlated attributes are given preference within the contrast set where they are predictive.

RESULTS

Participants produced a variety of sort types, showed sensitivity to correlational structure (more so in the consistent contrast condition), and produced 1D sorts rarely. Figure 3 shows the distributions of numbers of different sort types, with numbers broken down by condition. All participants produced either a single partition or a hierarchy, with the majority sorting hierarchically.

First, we characterized the variety of sorting types produced and assessed the prevalence of 1D sorting. Our structure scoring was designed to capture any type of organization, including a matrix organization laid out with different dimensions for rows and columns; overlapping groups; chaining; and thematic groups. However, only flat partitions or hierarchical groupings were produced.

Overall, 63% (27 of 43) of all participants sorted hierarchically. All the hierarchical sorts used different attributes at different levels in the hierarchy and thus are not 1D sorts. An additional 16% (7 of 43) produced nonhierarchical multiattribute sorts. Only 21% (9 of 43) of sorts could be generated by using only one attribute. Furthermore, four of the nine sorts that could have been generated by a 1D strategy apparently did involve more than one dimension; these sorts consisted of sorting into the two superordinate categories, and in all four cases multiple correlated properties were given in the rationale. Thus only 12% (5 of 43) of sorts were truly 1D, relying on a single attribute, as shown both in the sort itself and the rationale.

Second, we asked whether participants were sensitive to correlation when this does not require violations of consistent contrast. Participants

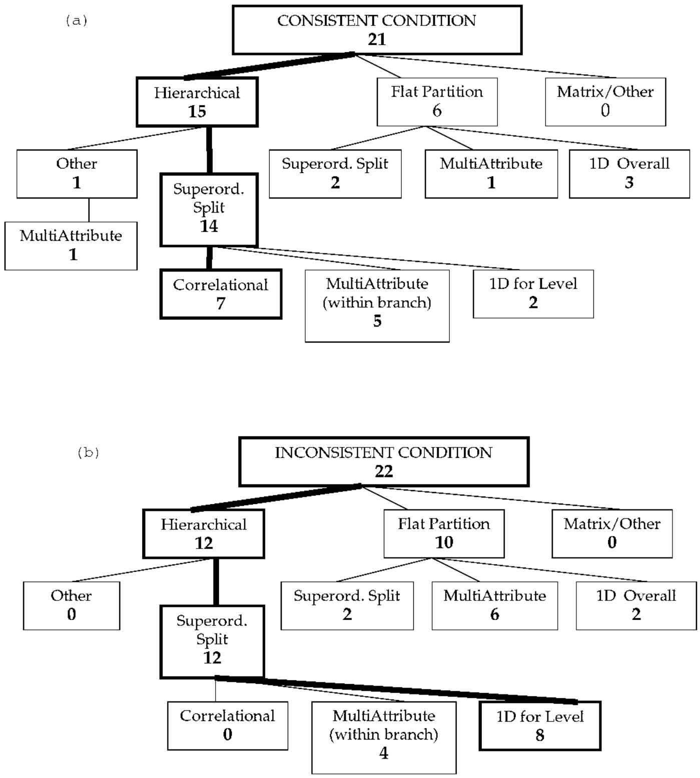

Figure 3. Classification of various sort types used in Experiment 2, by condition

could show sensitivity to correlational structure at the superordinate level and also at the subordinate level. Recall that the consistent contrast condition stimuli were designed so that the sort that captures the most correlational structure contrasts consistently, but this sort cannot be produced by reliance on a single attribute; different attributes correlate at the higher and lower levels, and they differ between the two lower-level partitions. At the superordinate level, this partitioning could be generated by use of any of the three correlated superordinate attributes, and this split was captured by 76% of consistent contrast participants and 64% of inconsistent contrast participants. Seventy percent (21 of 30) of these participants who made the superordinate split gave multiple correlated attributes in their rationale, suggesting that many of these partitionings were not based on a single attribute but were indeed sensitive to correlations.

The clearest evidence for sensitivity to correlations comes from the correlational sort produced at the subordinate level. If participants in the consistent contrast condition produce the full correlational sort, this provides the most direct test of sensitivity to correlational structure when this does not require violating consistent contrast. In the consistent contrast condition, 33% (7 of 21) of the participants recovered this sort. All who produced the sort also mentioned multiple, correlated attributes, differing in each branch of the hierarchy, in their sort description. The correlational sorts were again evenly distributed across the three configurations (2 in Configuration 1, 2 in Configuration 2, and 3 in Configuration 3).

Notice that the correlational sorts cannot be produced at all by a 1D sort strategy or by a more sophisticated procedure that sorts hierarchically, selecting a single attribute for each level of the hierarchy. Thus, these models are insufficient to explain these sorts. Furthermore, 1D sorting models would not predict that participants would report multiple attributes for any of the sort types; in fact, multiattribute reports were frequent, given by 17 of 21 participants in the consistent contrast condition and by 20 of 22 participants in the inconsistent contrast condition, or 86% multiattribute reports overall. In the *Discussion* we consider the probability of producing these sorts under alternative but related models of classification.

Third, comparison between consistent contrast and inconsistent contrast conditions tests whether consistency of contrast does in fact aid discovery of correlations, as in Experiment 1. No one sorted in accord with the correlational structure in the inconsistent contrast condition, a highly significant difference from the third of the participants who sorted by correlations in the consistent contrast condition, χ^2 likelihood ratio = 11.47, $p < .001$; Goodman & Kruskal's tau = .20, $p = .003$. Furthermore, in the consistent contrast condition, no participant produced the six-category lower-level partition without having the superordinate grouping into two sets of three. Thus, establishing the superordinates, and thus separate contrast sets, seemed critical for these participants to recover the correlations between the subordinate attributes in the consistent contrast condition.

Fourth, we can look for the prevalence of consistently contrasting sorts, or all the sorts that use a fixed set of attributes for categories at a given branch point in the hierarchy. Figure 3 shows that in the consistent contrast condition, five of the flat sorts and nine of the hierarchical sorts contrasted consistently and that in the inconsistent contrast condition, four of the flat and eight of the hierarchical sorts contrasted consistently. Thus, 60% (26 of 43) of sorts contrasted consistently, compared with 12% that were 1D sorts.

DISCUSSION

Summary of findings

First, participants favored categories that preserved correlational structure in input. Correlated attributes were preferred, selected more often than chance (more than twice as often as the sum of the same attributes when uncorrelated in Experiment 1). Participants often recovered the full multiattribute, correlational, hierarchical sort in Experiment 2. Participants often reported multiple, correlated attributes as the basis for sorting in both experiments. Finally, no participant reported a correlation but did not use it to sort, suggesting that correlation-based sorting is limited by noticing the correlations, not by prioritizing another sort strategy over sorting by known correlations.

Second, participants favored consistent contrast. The majority of sorts produced consistently contrasting categories, ranging from 95% of sorts in the consistent contrast condition of Experiment 1 to 55% even in the inconsistent contrast condition of Experiment 2. Participants often mentioned multiple, consistent attributes in their rationales.

Third, and the main issue addressed by these experiments, consistent contrast affects frequency of correlational sorts. We hypothesized that people care both about consistency and correlational information and hence that we would be most likely to see correlational sorting when this did not conflict with consistency. In both experiments we predicted and found that correlational sorting was much more likely when the resulting sorts did not conflict with consistent contrast. Indeed, there were no correlational sorts in the inconsistent conditions, where correlational sorting would have produced categories that contrasted very inconsistently. This seems to be primarily because of the greater difficulty of discovering reliable co-occurrences among attribute values when these patterns are distributed over different attributes.

Finally, participants produced a variety of sort types, suggesting that a variety of principles or organizing biases are at work, particularly when the stimuli offer a slightly richer structure than in many previous experiments.

Broadly, participants prefer sorting on correlated attributes over uncorrelated attributes when correlation is discovered. 1D sorting models and our consistent contrast account both emphasize that people prefer to have a dimension matter for a second contrasting category if that dimension matters for a first contrast category. However, our reports suggest that participants do care about multiple correlated attributes rather than simply picking one of the attributes as definitional and the others as unimportant.

Relationships to prior work and 1D sorting

As in prior studies, participants often produced sorts that could have been generated by using a single attribute. However, the strategies used, even for many of these sorts, are richer than a 1D strategy. With the complex stimuli of Experiment 2, only 21% (9 of 43, evenly between conditions) of sorts could have been generated by a 1D strategy, and only about half of these (5 of 43, or 12%) showed a 1D sorting strategy in the rationales. However, 60% (26 of 43) of Experiment 2 sorts contrasted consistently. Thus 39% of the sorts contrasted consistently but could not be produced by a 1D strategy in Experiment 2. Overall, these findings suggest that people may have a preference for consistently contrasting sorts, with the 1D sorting strategy a special case used as a default when no larger set of correlated attributes is noticed.

Indeed, we believe the primary reason participants use 1D sorting is that they have not detected correlational structure, and the primary reason for reduced 1D sorting in our experiments is more frequent detection of correlational structure. Our method may have facilitated discovery of correlation both from the preexposure procedure used in Experiment 2 and from the stimulus designs in both experiments.

Our stimulus structure in the consistent contrast condition differs from traditional FR structures in that our correlations are concentrated in a subset of attributes, rather than having all attributes correlated, and our correlations within this subset of attribute values are completely reliable. The correlation structure in traditional FR structure, where all attributes are equally but only modestly informative, may be very hard to detect, contrary to original expectations. These structures may easily slip through both analytic and implicit detection processes. Implicit processes that might build up knowledge of instances will be working off fairly haphazard representations because there is little task demand to internalize knowledge of instances; the clusters of instances are highly overlapping, and this overlap will be made worse if encoding of instances is degraded and partial. An explicit process that searches for regularities or tests hypotheses will be unlikely to identify any aspects that provide a preferred basis for grouping. Explicit hypothesis testing of association will yield a high rate of disconfirmations, and generating candidate partial groupings is likely to yield overlapping options without a clear basis for preference.

Several of the prior experiments that reduced the amount of 1D sorting are consistent with this failure-to-detect explanation. If an established causal schema links the features, recognizing that the current stimulus exemplifies these associations should be much easier to detect than discovering a novel relationship (Medin et al., 1987; Spalding & Murphy, 1996). The Kaplan and Murphy (1999) study, which provided a preexposure task and found increased correlational sorting, is also consistent with this explanation. In contrast, the shift away from 1D sorts when only the prototype is in view (Regehr & Brooks, 1995) may have a different explanation. Rather than making discovery of correlations easy, it may make applying a 1D strategy harder because the choice of attribute must be maintained in working memory without the constructed layout providing a external representation of what attribute is being used.

The sensitivity to correlations found in our two experiments with free sort fit into a complex of findings showing sensitivity to correlations even in the absence of feedback that is contingent on correlation: in natural categories (Malt & Smith, 1983), in supervised learning (Medin, Altom, Edelson, & Freko, 1982), and in sequential presentation unsupervised concept formation (Billman, 1989; Billman & Knutson, 1996; Clapper & Bower, 1994; Kersten & Billman, 1997). However, discovery of correlation is far from inevitable (Ashby, Queller, & Berretty, 1999), and much remains to be learned about the aspects of stimuli and task that predict successful correlational learning.

Process interpretation

In free sort, participants must do two types of activities, which overlap and interact in unknown ways. Participants must decide on and prioritize the criteria they want their groupings to meet and hence what type of sorting strategy they want to apply. Participants must also consider the instances presented and how the instances would be mapped into groups in light of the criteria or by applying the strategy. However, we do not know much about how the principles in mind and the instances at hand are coupled. Perhaps looking for properties shared between instances leads to discovery of co-occurring values, which prompts search for relationships on other values of the same attribute, which motivates a correlational sort, which in turn rates highly on evaluative criteria including consistent use of at least one attribute. Alternatively, a participant might partially generate a sort based on a single, consistent attribute, unaware of the correlations, make comparisons between instances within a group, notice correlations in reflecting on the sort, and as a result stick with this sort rather than generating others because it rates more highly on evaluative criteria such as high information capture. Both methods can result in the discovery and use of correlational structure in sorting. Probably these and still additional strategies are chosen by different participants, given the very loose demands imposed by the task.

Process models

The two-stage model of Ahn and Medin (1992) was explicitly designed for free sorting to explain the prevalence of 1D sorts. This model cannot explain the preference for sorting on correlated attributes in Experiment 1, and it cannot explain the large number of sorts that cannot be generated by a single attribute. Experiment 1 was designed not to conflict with but to select among sorts that could also be produced by sorting on one dimension. However, even here 15% (9/59) of sorts could not be generated by sorting on a single attribute. This percentage jumps for the more complex stimulus structure of Experiment 2, where 79% (34 of 43) of sorts cannot be generated by a 1D strategy.

The two-stage model originated for sort tasks in which the participant was required to produce only two groupings and not for domains with a hierarchical structure. A sympathetic extension to the model might allow recursive sorting at different levels and an open-ended number of categories but still using only a single attribute for any given level. However, only 10 of the 27 hierarchical sorts were of this type, leaving 63% of hierarchical sorts still unexplained (as well as 7 nonhierarchical sorts).

A further modification would be to suggest that sorting is done recursively but with a single attribute chosen for each branch point rather than at each level (excluding attributes already used at a higher level) and independently of other branch points. This more powerful model can produce the correlational sorts but still shows no preference for them because its fundamental principle is independent selection of one attribute at a time. For the consistent contrast condition of Experiment 2, assume the model makes the correct split at the first level of the hierarchy. Given this split, the model predicts the chance selection of one of the two correlated attributes for the first branch with one of the two correlated attributes for the second branch in 1 in 9 of the hierarchical sorts ($2/6 \times 2/6$). However, the correlational sort was produced for half the hierarchical sorts. Similarly, the extended model could not predict the preference for selecting correlated attributes in the consistent contrast condition of Experiment 1.

In sum, although a 1D sorting principle can be elaborated in increasingly powerful models of free sort, it cannot reflect sensitivity to correlations, nor can it produce the substantial proportion of multiattribute sorts within a contrast set.

A closely related research paradigm is unsupervised concept learning with sequential (i.e., incremental rather than batch) presentation of instances, typically followed or accompanied by a task that assesses knowledge of the structure in input. Whereas research on free sort focuses on invention of groupings, unsupervised concept learning research focuses on discovery of structure in input. However, the tasks are closely related, and models of sequential unsupervised learning can be applied directly to free sort tasks. We consider predictions of these models and focus on two issues they address: correlation learning and attribute-level learning.

Ashby, Queller, and Berretty (1999) proposed a model of unsupervised learning very similar to the two-stage model in that it predicts selection and reliance on one attribute and insensitivity to correlations. It is motivated by their data using much simpler, 2–dimensional, continuously varying perceptual stimuli. They found that with these very different stimuli, participants failed to learn a correlation (here, a quantitative correlation between attributes with continuous values) or to partition instances using both correlated dimensions even though this provided the best-separated clusters. Their model predicts insensitivity to correlation; it does not predict the preference for correlated sorts found with our stimuli in the consistent contrast conditions.

Several models of unsupervised concept learning (Anderson, 1991; Billman & Heit, 1988; Chalnick & Billman, 1988; Clapper & Bower, 1994) do predict sensitivity to correlation and indeed presume that discovery of some type of contingency or correlated features is the basis for unsupervised learning.

Anderson's (1991) and Clapper and Bower's (1994) models assume learning is based only on discovery of co-occurrence at the level of specific features or attribute values. Anderson's rational analysis model explains supervised and unsupervised learning as formation of clusters of interpredictive features, such that if an instance shared many features of the cluster, it is likely to share others. The model forms a flat partition of these categories, and during learning an instance is classified into the best-fitting cluster or used to initiate a new category. Clapper and Bower's model proposes that when similar instances are considered together, common properties are noticed and used to create a category schema. Both models capture the predictive structure in the input domain at the level of specific, co-occurring attribute values, but neither has any mechanism for representing or learning about relationships at the level of attributes. Applied to free sort tasks, both models predict groupings based on correlation, a prediction supported by the findings in our Experiments 1 and 2. However, because there is no role for learning at the level of attributes, their performance is unaffected by organization at the attribute level. Hence, these models predict that, given the equal strength of association between cues (as in our experiments), the discovery of correlated cues will be as easy in the inconsistent as consistent contrast conditions, contrary to our findings.

Billman and Heit's (1988) focused sampling model of unsupervised learning focuses on learning predictive patterns among attribute values and learning which attributes are most predictive. This model predicts faster unsupervised learning when the same attributes are used in multiple predictive relations. Chalnick and Billman (1988) extended this model to build explicit and hierarchical categories based on the correlational patterns discovered and kept the attentional learning component that focuses on the more predictive attributes. When applied to free sort, this approach predicts correlational sorting and more correlational sorting when attributes, not just particular values, are consistently related across the whole learning task. However, these models do not track attribute relevance local to a set of contrasting categories.

A machine learning model of incremental unsupervised learning does include information about relevance local to a contrast set. Martin and Billman's incremental learning model (1991; TWILIX, Martin, 1993) includes information about how homogeneous an attribute is within contrasting categories in determining whether to form a new category. Tests with incremental, trial-by-trial unsupervised learning suggested that it still was less sensitive to consistency than people were and in free sorting probably would show some, but insufficient, advantage for producing correlational sorts in our consistent and inconsistent conditions (Billman & Davila, 1995).

GENERAL DISCUSSION

The primary contributions of the two experiments are the results on how the organization of information in the input domain influences discovery and use of correlations. Although people might notice correlations but decide not to use them in free sorting, we expected that people would use what they noticed. Therefore, we investigated conditions we thought would facilitate learning the predictive structure in input. These included strong and well-organized correlations; stimuli composed of multiple, meaningful, qualitative attributes; and the preexposure task of Experiment 2. We found that people frequently, though far from uniformly, learn and use correlations in free sort. In our circumstances we found that consistent contrast seems to be a critical requirement for noticing correlations. We also found wide variation in sorting, even in our consistently contrasting conditions, particularly with the more complex stimuli of Experiment 2.

How do categories created in free sort relate to natural categories? The situations for acquiring natural categories are much more heterogeneous than in the free sort task. Yet in any one situation, the examples and functions of a new category may be more constrained. In some cases learners may initially establish natural categories by grouping a collection of instances, as in free sort, based on a small number of highly correlated attributes. In other cases categories may be based on similarity to a core instance and brought into contrast with other categories later. In some cases, an initially exceptionless category may introduce partial correlations when exceptional instances are included in the category, by virtue of a common label, for example. Furthermore, as background knowledge develops, learners may have different expectations for categories in different domains; in particular, learners may expect biological kinds to capture rich causal and correlational structure. In short, FR may be an abstraction from a variety of variability structures in natural categories, emerging from a variety of learning conditions. Finally, Western college students may combine an assumption that categories with definitions are desirable with a low-effort strategy of relying on a consistent attribute in forming definitions; preferences for this organizational schema may vary cross-culturally.

Understanding how people organize and learn about an input domain in the absence of explicit feedback or instruction is a tremendously important and complex problem. A small part of the space of tasks and stimuli has been investigated. Yet even over the sampled space it is clear that there is tremendous variation in learning even something as simple as covariation detection. More correlational structure seems to be learned and used when stimuli are meaningful, complex, and categorical than with simple continuously varying perceptual stimuli. More seems to be learned with unsupervised, extended, sequential exposure than in free sort, and learning in free sort seems greater when additional exposure is provided or when tasks that direct the learner to attend to the stimulus structure are included. More is learned when covariation is organized at the level of relationships between attributes and of contingencies between specific values.

Appendix

The Experiment 1 stimulus assigned random values to slightly favor correlational sorting in the inconsistent over the consistent contrast condition. Consider the schema in the inconsistent contrast condition for the first category, *aa yy yy*. If all six values of Attribute 3 were allowed, instances with values such as *aa bb yy* or *aa ab yy* could be included. Such crossover instances could fit into more than one category and also disrupt category-relevant associations, a phenomenon that would not be produced in the consistent contrast condition. Therefore, the random values of each attribute in the inconsistent contrast condition were restricted to only four values, excluding the two values of that attribute (values *a* and *b*) that were correlated as the seed for some other category. This crossover problem does not arise in the consistent contrast condition, and so all six values of random attributes were assigned. This produces greater within-category variability in the consistent contrast condition. As a result, the within-category similarity of the categories in the correlational sort was slightly higher in the inconsistent contrast condition, and the between-category similarity (or confusability) of the categories was slightly lower in the inconsistent contrast condition. Thus both aspects of structure favored correlational sorting in the inconsistent as opposed to the consistent contrast condition, the opposite of our prediction. If we nevertheless find more correlational sorting in the consistent contrast condition, it cannot result from the inadvertent differences in similarity from the random attributes.

In assessing the degree of accidental support for the target groupings, we counted cases in which “random” values occurred consistently with an intentionally correlated attribute pairing, which would provide additional support for that grouping. Again because of the restricted values used in the inconsistent contrast condition, there was more accidental support for the target grouping in the inconsistent contrast than consistent contrast condition, again favoring the inconsistent contrast condition (our second objective). Specifically, there were four “random” attribute values on each of six attributes, or 24 values used in the random assignment for the inconsistent contrast condition. Of these, 10 of 24 values occurred twice in the same category. In comparison, there were six "random" attribute values on each of six attributes, or 36 values used in the random assignment for the consistent contrast condition. Of these, only one value occurred twice in the same category. This confirms the greater homogeneity of the correlational sort categories in the inconsistent over consistent contrast condition.

We also assessed the strength of competitor groupings, those other than the target categories. There was stronger competition in the consistent than the inconsistent condition, again biasing against us. For example, in the consistent contrast condition there were several cross-category pairs of instances that matched on four of six attribute values, making them a very attractive core for a similarity-based category. There were no such pairs in the inconsistent contrast condition. Checking the contingency between all pairs of attribute values consistently showed there were stronger competitor category groupings in the consistent contrast condition.

Notes

This research was supported in part by National Science Foundation grant 9732562. Part of this research was included in Jim Davies's master's thesis and reported at the Psychonomics Society, Chicago, November 1996.

Correspondence about this article should be addressed to Dorrit Billman, Center for the Study of Language and Information, Cordura Hall, Stanford University, Stanford, CA 94305 (e-mail: billman@psych.stanford.edu), or to Jim Davies (e-mail: jim@jimdavies.org). Received for publication April 6, 2004; revision received November 1, 2004.

1. Consistency of contrast is a matter of degree. Consistency of contrast increases when the strength of correlation between attributes and categories increases or when more attributes correlate with the categories across the contrast set. High levels of consistency of contrast imply that the values of informative attributes correlate with one another and with the category. In domains with locally rather than uniformly distributed information, consistent contrast allows focusing on the subset of attributes that are informative across categories. A more formal characterization would specify how these factors combine and normalize the degree of information about an individual category available from the contrast categories relative to information available otherwise.

References

Ahn, W. (1999). Effect of causal structure on category construction. *Memory & Cognition, 27,* 1008–1023.

Ahn, W., & Medin, D. L. (1992). A two-stage model of category construction. *Cognitive Science, 16,* 81–121.

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review, 98,* 409–429.

Ashby, F. G., Queller, S., & Berretty, P. M. (1999). On the dominance of unidimensional rules in unsupervised categorization. *Perception & Psychophysics, 61,* 1178–1199.

Billman, D. (1989). Systems of correlations in rule and category learning: Use of structured input in learning syntactic categories. *Language and Cognitive Processes, 4,* 127–155.

Billman, D. (1996). Structural biases in concept learning: Influences from multiple functions. In D. Medin (Ed.), *The psychology of learning and motivation* (Vol. 35, pp. 283–321). San Diego: Academic Press.

Billman, D., & Davila, D. (1995). Consistency is the hobgoblin of human minds: People care but concept learning models do not. In *Proceedings of the seventeenth annual conference of the Cognitive Science Society* (pp. 188–193). Hillsdale, NJ: Erlbaum.

Billman, D., & Davila, D. (2001). Consistent contrast aids concept learning. *Memory & Cognition, 29,* 1022–1035.

Billman, D., & Heit, E. (1988). Observational learning from internal feedback: A simulation of an adaptive learning method. *Cognitive Science, 12,* 587–625.

Billman, D., & Knutson, J. (1996). Unsupervised concept learning and value systematicity: A complex whole aids learning the parts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 458–475.

Chalnick, A., & Billman, D. (1988). Unsupervised learning of correlational structure. In *Program of the tenth annual conference of the Cognitive Science Society* (pp. 510–516). Hillsdale, NJ: Erlbaum.

Clapper, J. P., & Bower, G. H. (1994). Category invention in unsupervised learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 443–460.

Goodman, N. (1983). *Fact, fiction, and forecast* (4th ed.). Cambridge, MA: Harvard University Press.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child.* New York: W.W. Norton.

Kaplan, A. (1999). *Alignability and prior knowledge in category learning.* Doctoral dissertation, University of Illinois at Urbana–Champaign.

Kaplan, A. S., & Murphy, G. L. (1999). The acquisition of category structure in unsupervised learning. *Memory & Cognition, 27,* 699–712.

Kersten, A., & Billman, D. (1997). Event category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 638–658.

Lassaline, M. E., & Murphy, G. L. (1996). Induction and category coherence. *Psychonomic Bulletin & Review, 3,* 95–99.

Lassaline, M. E., & Murphy, G. L. (1998). Alignment and category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 144–160.

Malt, B. C., & Smith, E. E. (1983). Correlated properties in natural categories. *Journal of Verbal Learning and Verbal Behavior, 8,* 828–835.

Martin, J. D. (1993). Direct and indirect transfer: Explorations in concept formation. (Dissertation, Georgia Institute of Technology, 1993). *Dissertation Abstracts International, 54*(1–B), 330.

Martin, J., & Billman, D. (1991). Variability bias and category learning. In L. A. Birnbaum & G. C. Collins (Eds.), *Proceedings of the Eighth International Workshop on Machine Learning* (pp. 90–94). San Mateo, CA: Morgan Kaufman.

Medin, D. L., Altom, M. W., Edelson, S. M., & Freko, D. (1982). Correlated symptoms and simulated medical classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8,* 37–50.

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology, 19,* 242–279.

Murphy, G. L. (2001). Causes of taxonomic sorting by adults: A test of the thematic-to-taxonomic shift. *Psychonomic Bulletin & Review, 8,* 834–839.

Murphy, G. L., & Kaplan, A. S. (2000). Feature distribution and background knowledge in category learning. *Quarterly Journal of Experimental Psychology, 53A,* 962–982.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review, 90,* 339–363.

Regehr, G., & Brooks, L. R. (1995). Category organization in free classification: The organizing effect of an array of stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 347–363.

Rosch, E. H. (1978). Principles of categorization. In E. H. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Erlbaum.

Shipley, E. F. (1993). Categories, hierarchies, and induction. *Psychology of Learning and Motivation, 30,* 265–301.

Sifonis, C. M., & Ross, B. H. (2002). Alignment effects on learning multiple, use-relevant classification systems. *Memory & Cognition, 30*(7), 1148–1159.

Spalding, T. L., & Murphy, G. L. (1996). Effects of background knowledge on category construction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 525–538.

Yamauchi, T., & Markman, A. B. (2000). Learning categories composed of varying instances: The effect of classification, inference, and structural alignment. *Memory & Cognition, 28,* 64–78.