Feature Representations and Analytic/Nonanalytic Processing

Samuel Hannah, McMaster University

Abstract Lee Brooks' original formulation of instance theory embedded the notion of an instance within the larger conception of a distinction between analytic and nonanalytic processing. Brooks has recently argued that features can be represented either in terms of their specific feature appearance, or in terms of the abstract information some particular instantiation embodies. This work reviews some recent studies that link reliance on different types of feature representations to different decision-making processes, and to different patterns of categorization behaviour. This in turn complicates the analytic/nonanalytic distinction, suggesting a more precise reformulation.

In his classic chapter on noncomputational approaches to categorization, Lee Brooks (1978) challenged the idea that categorization must depend on an analytic computation of the structure of a stimulus. He argued instead that people could categorize things simply by retrieving from memory the most similar items to the current stimulus. Importantly, nonanalytic processing was seen as relying on the overall or holistic similarity of the current stimulus to items previously encountered. This notion of the priority of a holistic representation was reinforced in work with Regehr (Regehr & Brooks, 1993). This notion of the centrality of a holistic, or global, representation is shared by most instance/exemplar accounts (e.g., Hintzman, 1986; Medin & Schaffer, 1978), and even most prototype accounts treat a prototype as a global or holistic representation (e.g., Hampton, 1995).

However, our recent work (Brooks & Hannah, 2000, 2005; Hannah & Brooks 2005a, b) suggests that similarity to previously encountered features can also have a substantial impact in driving categorization decisions, just as for whole items. This result has led us to suggest that there are at least two levels at which features can be represented. A representation of a feature or property in terms of a previously encountered feature, including the details of its appearance, we call an *instantiated feature*, and the representation of a feature or property considered only in terms of its abstract information we call an *informational feature*.

Acknowledging that there is both an embodied and abstract aspect to feature representations may seem obvious to some. Acknowledging that there are two levels of abstraction for feature representations, however, has significant consequences for understanding behaviour because they offer different affordances for applying concepts. Generic, informational features allow broad application: they have great scope. Particular, instantiated features are more strongly associated with their categories: they have great discriminability. A reliance on one form of feature representation over another allows different strategies or processes to operate, producing different behaviours.

Feature Weighting and Feature Goodness

In Hannah and Brooks (2005a) people classified test items involving four species of imaginary animals after training, and we examined how reliance on different types of feature representations would produce different patterns of decision-making as revealed in error patterns. Participants classified test items, described the two features pointed out by the experimenter – which were consistent with the family-resemblance rule governing the creation of the items – and then classified the items again. We wanted to see how participants would classify items on the second classification step after they had made an error on the first, and after they had been forced to attend to the rule-consistent features. We were interested to see how participants reliant on instantiated features responded to the corrective information in the second step, as compared to how participants reliant on informational representations of features responded to the same.

We reasoned that those reliant on particular instantiations of features might err by deliberately selecting the most recognizable features to base their classification on, and thus would discount less recognizable features, even if they were more numerous. That is, feature recognizability should be a source of feature weighting. Having such participants acknowledge the

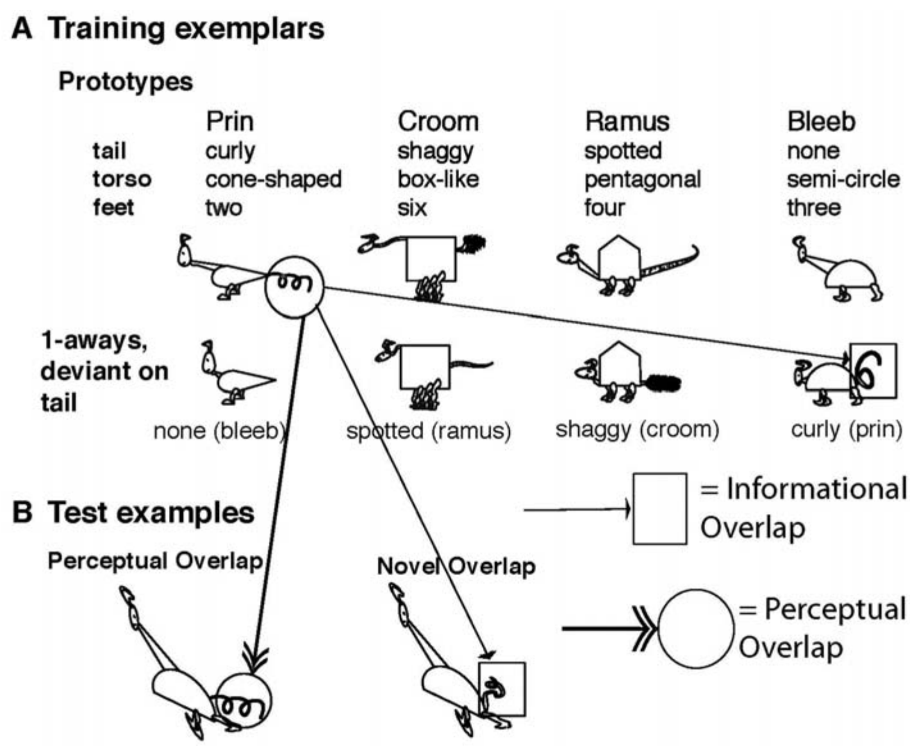

Figure 1. Examples of training stimuli (panel A) and test stimuli (panel B) used in Hannah and Brooks (2005a, b). Training prototypes (A, upper row) had all three characteristic features listed under their name; one-away items (A, bottom row) deviated by one feature from prototypes. The deviant feature overlapped with the characteristic feature of another category at an informational level, but not at an instantiated level. At test (B) the lure feature could be replaced with the actual feature that appeared in the rival category in training (perceptual overlap, right), or with a novel feature that overlaps only at an informational level (informational overlap, left). In Hannah and Brooks (2005b), the novel overlap stimuli were replaced with skewed versions of the perceptual overlap items.

rule-consistent features would have little corrective force after having already discounted these features. Such a strategy is likely to be successful and common in dealing with real-world categories. In the world, as in our training, feature appearance is strongly associated with category identity: Only cats have cat-like paws, only humans have human-like legs. All other things being equal, the features easiest to label usually are the most reliable guides to identity. We call this use of the ease of recognition of features as a source of weighting and classification a *feature-goodness heuristic*.

Those reliant on informational features, however, should pay no heed to the recognizability of the features. What should count is only the informational content. They should have no basis for discounting a feature or two because they are unfamiliar. Although those reliant on instantiated features should persist in their initial error even after acknowledging the rule-consistent features, those reliant predominantly on informational features should incorporate those features into their decision, and revise their initial classification after being forced to attend to the rule-consistent features.

Our earlier research (Brooks & Hannah 2000, 2005) suggested that the feature lists commonly generated by induction-trained participants in categorization experiments function as pointers to instantiated features, and not as implicit family-resemblance rules. We noticed that those few people who generated a counting rule consistent with the family resemblance structure (e.g., “An item is a bleeb if it has a majority of rounded head, rounded torso or two legs”) seemed to attach little weight to particular instantiations, and behaved as if they were reliant largely on the informational content of the features. A counting rule is implicitly acknowledging that no one feature is necessary and sufficient, but rather that a combination of equally weighted characteristic features is necessary for classification. The discovery of such higher-order relations among features and category identity requires that one encode the features at an abstract enough level to notice the overlap of the information across categories within a domain. Thus we took the type of rule or strategy statement produced after test as a proxy measure of the reliance on instantiated or informational representations.

We deviated from the usual induction method of training, and gave direct instruction regarding which features were relevant, and provided labels for these features. By directing participants to which features were relevant, the discovery of the abstract properties was made easier, increasing the proportion of participants who developed counting rules. This procedure ensured that the features our participants extracted mapped onto the features of the test categories. Variability in feature extraction would be unnecessary noise for our purposes. In this study in particular, we required people to explain and describe their decisions, and therefore we needed to ensure they had an adequate feature vocabulary, and one transparent to us.

At the end of the test session, we asked people for their decision-making strategy, so that we could divide them into those using feature lists (e.g., "I based my decisions on the legs and tails, mostly") and those using a counting rule (e.g., "I looked for two features from the same category"). Those few people generating some other strategy than these two were not analyzed. We did not force people to adopt one strategy or the other by giving them a counting rule in training or by relying on induction for reasons related to feature-extraction indicated above. Induction would result in many participants extracting features that did not transfer to the test materials. The behaviour of such participants could differ strikingly from that of the counting participants simply because they did not have feature knowledge adequate to perform well on the test items. Such inappropriate feature extraction would be especially problematic in the biasing study described subsequently in which people may follow the biasing suggestion only because they have no other basis on which to make a decision, having failed to extract features that transfer to the test set.

All test items, and most training items, contained a single lure feature, or, overlap feature. An overlap feature is one that is characteristic of another category. For example, in Figure 1 (Panel A) the second row of animals depicts training items that deviate from the prototypes (Panel A, upper row) by a single feature (*one-away items*). These items, however, deviate by taking on a value that is characteristic of another category, but only at an informational level. Thus, the bleeb one-away has a curly tail, which is characteristic of prins, but the respective curly tails are quite different perceptually. Thus we can distinguish between *informational overlap*, as in the case just mentioned, and *perceptual overlap*, indicating perceptually similar features in members of different categories.

We created our test items by skewing the features of the training one-aways by 20° clockwise or counterclockwise. This skewing produced matched test items that were clearly novel, but that preserved configural relations, and were thus similar to the training pairs at a holistic level, and equally so. Two sets of test items were created by changing the nature of the overlap (lure) feature. In one set, the informational overlap feature was replaced with a perceptual overlap feature. For example, the bleeb tail one-away (Figure 1, Panel B, left) now has the curly tail previously seen only in training prins. The overlap features are now the most familiar looking features in each item. For another group, the overlap features are replaced with novel features that preserve the informational content of the original feature (Figure 1, Panel B, right). These novel informational overlap features (or, "novel overlap features") are the least familiar features in this test set.

We expected that the critical difference in error patterns should arise between users of different strategies. Although those counting participants classifying the perceptual overlap items, for example, may make more errors than counting participants classifying the novel overlap items, their pattern of errors across error type should be fairly consistent. In addition to the persistence and revision errors described above, we also scored errors due to a reinterpretation of the rule-consistent features, and to residual errors involving confusions regarding the category to which a named feature belonged. Within each lure condition (perceptual overlap, novel overlap), users of different strategies differed sharply. If we include confusion responses in comparing error patterns, 53% of feature-listing participants' (n = 26) errors were persistence errors, and 21% were revision errors. For counting participants (n = 42), however, only 28% of their errors were persistence errors, but 45% of errors were revision errors. This difference in the pattern of errors was significant. Excluding confusion errors makes the differences between strategies a little more extreme, but does not change the pattern.

Within strategies, however, the differences are largely no longer significant. For counting participants, the majority of errors were revision errors, regardless of whether confusion errors were included or excluded. For the perceptual overlap participants (n = 19), 46% of all errors were revision errors; for novel overlap participants (n = 23), 44% of all errors were revision errors, a nonsignificant difference. With confusion errors excluded, 58% of errors for both groups consisted of revision errors. Among feature-listing participants, perceptual overlap participants (n = 15) made more persistence errors than any other type of error (60% of all errors), but listing novel overlap participants (n = 11) mainly made confusion errors (45% of all errors). This difference in the distribution of error responses was significant. However, after we exclude confusion responses, significance disappears, and both strategy groups show the same pattern of errors. For listing perceptual overlap participants, 71% of lure-based errors consist of persistence errors, and 57% of lure-based errors for listing novel overlap participants are persistence errors. Reducing the recognizability of lure features for listing novel overlap participants reduces the contribution of a critical source of error, increasing the relative contribution of confusion errors.

Biasing Categorization Decisions

A study in the biasing of categorization decisions provided confirmatory evidence that differences in the reliance of representation type yield differences in categorization behaviour (Hannah & Brooks, 2005b). This study was based on the work of LeBlanc, Norman, and Brooks (2001). They discovered that merely suggesting a tentative hypothesis to doctors in residency or medical students could result in substantial biasing of subsequent diagnoses. We suspected that reliance on instantiated features may play some part in such categorical biasing effects. Medical rules are largely feature lists of the type produced by our instantiated-reliant participants, with some information regarding the relative frequency of different symptoms. In our discussions with practitioners, many have described common diagnostic situations as "pattern matching." If a reliance on instantiated features underlies a susceptibility to biasing suggestions, then, any biasing effect elicited among listing participants should be reduced if we reduce the subjective "goodness" of lure features, but any biasing effect that is elicited among counting participants should be independent of the appearance of overlap features.

We used the same training items and training method as before. The same perceptual overlap test items were used for one group of participants, and for another group the novel overlap items were replaced with a modified overlap condition. This modified overlap condition consisted of the perceptual overlap items with the perceptual overlap feature itself skewed by 20° clockwise or counter-clockwise. The overlap feature was now equally recognizable as its rivals. Early results showed no difference in the biasing effects for novel overlap and modified overlap items, so the former were dropped as the greater similarity of the modified overlap items to the perceptual overlap items was taken as a stronger demonstration of the importance of instantiations. At test, people classified each test item after first rating the plausibility of a suggested classification made by the experimenter. Each item was introduced with the experimenter asking how likely the item was to be a member of either the rule-consistent or the overlap category. For one member of each pair of test items, the experimenter suggested the rule-consistent category, and suggested the overlap category for the other.

We again divided participants after test into feature-listing and counting groups, and analyzed their data separately. As expected, the feature-listing participants showed a significant effect of a suggestion on their classification. Feature-listing participants assigned 28% of items to the overlap category when the overlap category was suggested, but assigned only 18% of items to the overlap category when the correct category was suggested, for a biasing effect of 10 percentage points. More importantly, the size of this bias depended on the similarity of the overlap feature to the features seen in training, as we found a significant Cueing x Lure interaction. For the perceptual overlap participants, suggesting the overlap category increased overlap responses by 13 percentage points, but suggesting the overlap category for modified overlap participants increased overlap responses by only 5 percentage points. Although the counting participants showed a small, but significant, effect of a suggestion, this effect was constant across training items, with the Cueing x Lure interaction proving nonsignificant. A suggestion pointing to the overlap category increased overlap responses by an average of four percentage points compared to suggesting the correct category.

Implications

In two very different paradigms, we found evidence suggesting that reliance on instantiated feature representations produces systematically different behaviour than reliance on informational features when making classification decisions. The difference in the error patterns suggests that people reliant on instantiated features invoke a feature-goodness heuristic to weight features, but those reliant on informational features ignore or override the recognizability of features when making decisions, combining features according to explicit algorithms. The use of a feature-goodness heuristic may explain the interaction between lure familiarity and suggestions found among listing participants in the biasing study. A suggestion may prime the processing of features, and this priming may enhance the subjective goodness or recognizability of features. When lure features are not close to their instantiated representations, however, the saliency of the informational representations may increase, weakening the reliance on a feature-goodness heuristic and weakening the biasing effect of a suggestion.

The application of decision heuristics to instantiated features complicates the analytic and nonanalytic distinction. To break things into features – to analyze – is in many instance-based approaches put in opposition to memory retrieval. Feature-based processing is assumed to be *computational*, in contrast to processing based on some whole (Brooks, 1978; Erickson & Kruschke, 1998; Logan, 1988; Nosofsky, Palmeri, & McKinley, 1994). Yet, in the research summarized earlier, we find noncomputational, memory-driven processing at the feature level.

Instantiated features represent the opposite of the concept of "abstract analogies" that Brooks and Vokey (1991) introduced. Abstract analogies can be thought of as informational representations, but at the level of a whole item, rather than at the feature level. Instantiated features are specific representations retrieved from memory, but at the level of feature rather than whole item. We can, therefore, process abstract representations at either feature or item level, and can retrieve specific representations from episodic memory at either the feature or item level. Does this make the analytic/nonanalytic distinction orthogonal to distinction between abstract and concrete representation, or orthogonal to the distinction between feature and whole item? Or does the distinction need to be clarified with a more precise description of representation and processing?

The nature of this problem may be illuminated by considering the following scenario, which I take to be paradigmatic of analytic processing: adding a sequence of numbers. What is involved is a set of symbols, or representations, that are highly abstract. They are informational features, in the terminology used in this paper. But addition as an analytic process also requires computation: the application of an explicit set of operations that take a specific set of inputs. Analytic operations imply algorithmic operations, as well as abstract representations. Common conceptions of analytic processing conflate the notion of *part* with that of both *abstraction* and *algorithm*.

The research reviewed here suggests that informational features are linked to computational processes (analytic-like), and instantiated features seem linked to heuristic processes (nonanalytic-like). That in turn suggests that the distinction between parts and whole is orthogonal to, not identical with, the distinction between analytic – or perhaps, more properly *computational* – and nonanalytic. Whether level of abstraction and degree of computationality are necessarily joined, or only that one representation type merely favours one particular decision-making mode remains to be explored. Nonetheless, more traction may be made in understanding concept use by redefining processing in terms of computationality and abstraction than in terms of holistic and analytic.

Please address comments about this article to Samuel D. Hannah, Department of Psychology, McMaster University, Hamilton, Ontario L8S 4K1 (E-mail: hannahsd@mcmaster.ca).

References

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

Brooks, L. R., & Hannah, S. D. (2000). *Relation between perceptual and informational learning of family resemblance structures*. Paper delivered at 41st Annual Meeting of the Psychonomics Society, New Orleans, LA.

Brooks, L. R., & Hannah, S. D. (2005). *Instantiated features and the use of "rules"*. Unpublished manuscript.

Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General*, *120*, 316-323.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General*, *127*, 107-140.

Hampton, J. A. (1995). Testing the prototype theory of concepts. *Journal of Memory and Language*, *34*, 686-708.

Hannah, S. D., & Brooks, L. R. (2005a). *Specificity of feature representation influences categorization decision processes*. Manuscript submitted for publication.

Hannah, S. D., & Brooks, L. R. (2005b). *The role of instantiated knowledge in producing categorical biasing*. Manuscript submitted for publication.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, *93*, 411-428.

LeBlanc, V. R., Norman, R. G., & Brooks, L. R. (2001). Effect of a diagnostic suggestion on diagnostic accuracy and identification of clinical features. *Academic Medicine*, *76*, S18-S20.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492-527.

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

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*, 53-79.

Regehr, G., & Brooks, L. R. (1993). Perceptual manifestations of an analytic structure: The priority of holistic individuation. *Journal of Experimental Psychology: General*, *122*, 92-114.

Sommaire

Dans sa formulation initiale de la théorie de l'instance, Lee Brooks intégrait la notion d'instance à la conception large d'une distinction entre traitements analytiques et non analytiques. Brooks a soutenu, il y a peu, que les traits peuvent être représentés soit en fonction de leur aspect particulier soit sous l'angle de l'information abstraite que recèlent certaines instanciations particulières. Le présent article fait état d'études récentes qui établissent un rapport entre différents types de représentation des traits et divers processus décisionnels, de même que diverses tendances du comportement de catégorisation. Cela, en retour, complique la distinction entre traitements analytiques et non analytiques et donne à entendre qu'une reformulation, d'une plus grande précision, est possible.