Running head: FEATURE FAMILIARITY AND CATEGORIZATION

Featuring familiarity: How a familiar feature instantiation influences categorization

Samuel D. Hannah

Queen's University

Lee R. Brooks

McMaster University

Send correspondence to:

Sam Hannah

Department of Psychology

Queen's University

Kingston, ON

Canada K7L 3N6

Tel: (613) 533-2876

FAX: (613) 533-2499

email: hannah.sam@gmail.com

Abstract

We demonstrate that a familiar-looking feature can influence categorization through two different routes, depending on whether a person is reliant on abstract feature representations or on concrete feature representations. In two experiments, trained participants categorized new category members in a three-step procedure: Participants made an initial categorization, described the rule-consistent features indicated by the experimenter, and then re-categorized the item. Critical was what happened on the second categorization after participants initially categorized an item based on a familiar, but misleading, feature. Participants reliant on abstract features most commonly reversed themselves after the rule-consistent features were pointed out, suggesting the familiar feature had biased attention. Participants reliant on concrete feature representations, however, most commonly persisted with the initial response as if the familiar feature were more important than its rivals—the familiar feature biased decision-making.

Word count: 134

Featuring familiarity: How a familiar feature instantiation influences categorization

People often rely on familiar-looking features when making a categorization decision, (Brooks & Hannah, 2006). This happens even when one familiar[1] feature is surrounded by more numerous rival features, and the information conveyed by those rival features is understood. In this paper, we seek to explain how the familiarity of a feature's instantiation can influence categorization decisions.

There are at least two routes by which the familiarity of a feature's instantiation could influence categorization—an attention route, and a decision-making route. Perhaps a familiar-looking feature diverts attention from less familiar features during feature search. Alternatively, perhaps people weight a feature by its similarity to known feature instantiations when making a categorization decision.

Chun and colleagues (reviewed in Chun, 2000) have shown that search for objects in scenes is influenced by the familiarity of distractor identity or of surround configuration. The findings regarding novel pop-out in visual search (Johnston, Hawley, Plewe, Elliot & Dewitt, 1990; Johnston, Hawley & Farnham, 1993) also support the notion that visual search is influenced by stimulus familiarity. However, a feature-familiarity weighting heuristic would be consistent with arguments that relying on particular instantiations is optimal under ordinary conditions of categorization (Brooks & Hannah, 2006; Hannah & Brooks, 2006).

Critically for this feature-weighting argument, the instantiation of features is correlated with categorical identity for many natural categories. Cats have paws, dogs have paws, and even monkeys have paws, but *paws* as a feature is instantiated uniquely within each category, and within each category the instantiations of *paws* is similar across exemplars. The paws of one cat, therefore, look very similar to the paws of any other cat, but very different from the paws of a monkey, or even a dog. Surface structure is not mere accidental variation, but is systematically related to deep structure: Instantiation is information.

As illustrated in Figure 1, a feature representation consisting of a particular instantiation (bottom left) is necessarily more selective than a more generic feature representation (denoted *paw*, bottom right). The narrower selectivity of particular instantiations—or, *instantiated features*—relative to more general feature representations provides advantages in categorization for instantiated features. Knowing what cats' paws look like can be enough to recognize a cat from just a flash of paw darting out from under a bed. For many natural categories, instantiated features allow rapid categorization even under impoverished viewing conditions.

Given that there is a nontrivial relationship between feature instantiation and category identity, anomalies in the appearances of features are also nontrivial. People may treat the information that a particular feature instantiates as taking on a concept-specific range of instantiations. Features with instantiations falling outside this range would be suspect as legitimate features. As Schyns' and colleagues (Schyns, Goldstone & Thibaut, 1998; Schyns & Murphy, 1994; Schyns & Rodet, 1997) have pointed out, features are concepts in themselves; an unfamiliar concept should be treated gingerly. The notion that the frequency of an instantiation within a category carries information was leveraged by Barsalou (1985) in his account of the role of frequency of exemplar instantiation as a determinant of typicality. Although there are some important parallels between our treatment of feature instantiation and conceptual flexibility, and Barsalou's treatment of exemplar frequency and conceptual flexibility, there are important distinctions as well, which we will take up the General Discussion.

So far we have phrased the issue of the locus of effect for familiar instantiations in a simple dichotomy: either an attentional, feature search stage or a later decision stage. However, an important aspect of Brooks and Hannah's (2006) findings was that people have access to, and can use either instantiated features or more general feature representations, *informational features*. The locus of effect for instantiation familiarity may vary depending on whether informational or instantiated features are what are most accessible to a person when making a categorization decision.

For persons reliant on informational features, feature instantiations may have no effect at all on categorization decisions. Alternatively, people using informational features to make categorization decisions may ignore feature instantiations when making decisions, but be influenced by particular instantiations during some pre-decision processing stage, such as feature search. People reliant on instantiated features may be influenced by feature instantiation at both attention and decision-making stages. It is possible, then, that how instantiation familiarity influences categorization decisions will vary depending on whether people are reliant on instantiated features or on informational features at different stages. Is there any way of inferring what types of features people are reliant on when making categorization decisions?

Features types and rule types

Brooks and Hannah (2006) made a distinction between two types of rules: strong rules and weak rules. They defined strong rules as rules containing an explicit decision criterion. For example, Brooks and Hannah created a category "bleeb" based on the rule that bleebs have at least three of rounded head, rounded torso, striped pattern and two legs. The stipulation "at least three of …" is the decision criterion, and helps to resolve situations where an item has features consistent with different categories, or *overlap features*. If only one overlap feature is present, the item is still a bleeb, if two it is not a bleeb.

If we eliminated this decision criterion, and simply said a bleeb usually has a rounded head, a rounded torso and so on, then we would have a feature-list rule. Such a rule would tell a person what features were important, and is therefore informative, but it is helpless to resolve what happens when overlap features are present. Weak rules, especially in the form of feature-list rules, are common to everyday concept use (Rosch & Mervis, 1975), and are routine even in areas such as medicine.

Brooks and Hannah (2006) argued that feature-list rules are not simply inarticulate versions of strong rules, but are qualitatively different from strong rules. Weak rules are collections of terms that point to the instantiated features the rule giver has relied on when making categorization decisions. Weak rules are useful to orient a novice to particular features; once the instantiations are acquired, however, it is instantiated features driving the decision-making. The use of weak rules, we suggest, is indicative of a reliance on instantiated features in decision-making.

In contrast, we argue that strong rules are usually grounded in informational features. For most natural categories, it is only some informational content, not appearances, that overlaps across categories. Cats, dogs, and monkeys all share the feature *paw*, but only at an informational level, the *paw* in the bottom right of Figure 1. If strong rules help resolve feature conflict, and feature conflict exists only at an informational level, then the features the rule is operating on are likely to be informational features. The spontaneous use of strong rules, we suggest, is indicative of a reliance on informational features in decision-making.

That different types of rules track the kind of structure the observer encounters is an argument first made by Shephard, Hovland and Jenkins (1962) in their seminal paper on rule use and category structure. Their work demonstrated that the complexity of rule generation tracked the complexity of the underlying category structure. What we are adding is that because people can encode features at different levels of abstraction, the same category structure presented to different learners can result in the emergence of conceptual structures differing in complexity across learners.

This flexibility in conceptual structure is observable, however, only when informational features and instantiated features can be differently distributed across categories. This differential distribution arises when, for example, informational representations of *paw* overlap categories, but paw instantiations do not. This differential distribution of instantiated and informational features is typical of natural categories, and our materials. As will be discussed further in the General Discussion, our work is suggesting that flexibility in feature construction (e.g., Brooks & Hannah, 2006; Schyns, Goldstone & Thibaut, 1998) drives both flexibility in conceptual structure (e.g., Love, Medin & Gureckis, 2004) and flexibility in decision-making (e.g, Erickson & Kruschke, 1998).

Where informational features have a more complex distribution than instantiated features by virtue of informational features overlapping categories, while instantiated features are confined within categories, then it is possible for people to come to rely on either the complex informational structure or the simpler instantiated structure. Given the greater complexity of strong rules, it is reasonable that participants who generate strong rules are relying on the informational structure in decision-making, while those producing simpler feature-list rules are reliant on the simpler, instantiated structure in their decision-making.

If users of strong rules are reliant on informational features to make categorization decisions, then we might expect that they should be immune to the influence of familiar instantiations. Consistent with this, Hannah and Brooks (2006) found participants who produced strong rules at the end of the biasing experiment showed small diagnostic biasing effects that were constant regardless of the familiarity of an overlap feature. However, Thibaut and Gelaes (2006, Experiments 3A & 3B) also gave participants a strong rule, and found that their participants were still influenced by feature similarity. Thus even users of perfectly reliable strong rules are susceptible to the influence of familiar feature instantiations, at some stage of processing.

Familiarity's influence and rule type

The diversion of attention by a familiar instantiation could explain the findings of Brooks and Hannah (2006), Hannah and Brooks (2006) and Thibaut and Gelaes (2006). Strong rules may encourage their users to examine all relevant features to decide if the decision criterion is met, making the search of a cat or of a bleeb more effective than when a weak rule is in use. As a result, strong-rule users may be more likely to encounter the information that would offset either a misleadingly familiar ramus feature or a biasing suggestion reinforced by a misleadingly familiar feature. This would lead to a reduced tendency to follow familiar features, although a particularly salient familiar feature may be so seductive as to undermine proper rule execution.

If this search-alone hypothesis (outlined in Figure 2, top panel) were correct, then the only difference between a strong and a weak rule is the increased effectiveness of feature search, then there is certainly no need to posit a feature-familiarity weighting heuristic. Furthermore, this would also call into question Brooks and Hannah's (2006) contention that there were two levels of feature representation upon which categorization operated. Instead, the two levels of feature representations would be active at different points. Instantiated features would be relied on when searching for features, and informational features relied on when deciding what the features mean, and this would hold for all categorizers. Under this view, any influence a familiar feature instantiation exerts should be traceable to its influence upon attention, regardless of which type of rule a person produces, although its influence on strong-rule users should be reduced.

However, the original treatment of instantiated and informational features (Brooks & Hannah, 2006; Hannah & Brooks, 2006) could also account for an influence of feature familiarity upon users of strong rules, as found in Thibaut and Gelaes (2006). The original argument implied that rule use reflected the types of features that categorization decisions operated on, and did not address attentional processes. That the attentional processes of all categorizers may be sensitive to feature familiarity does not rule out the hypothesis that the decision-making of weak-rule users is also influenced by feature familiarity. This search-versus-decision hypothesis (Figure 2, bottom panel) would require that users of weak rules display a different pattern of responding to familiar features than users of strong rules. It requires that a familiar overlap feature bias attention for users of strong rules, but bias both attention and decision-making for users of weak rules.

Experiment 1A

To determine how a familiar feature exerts its effects on categorization, we need to separate out the effects of a familiar feature instantiation upon attention from its effects on decision-making, confusion or other sources of error. To distinguish between categorization responses that do not follow an experimenter's rule because of memory lapses or confusions from non-rule responses based on the overlap feature (*overlap responses*), we need more than two categories. Hannah and Brooks (2006) used four categories, and their training materials (panel A, Figure 1) are useful here.

To distinguish between overlap responses that arise from an influence upon attention and those overlap responses that reflect an influence upon decision-making, we had participants categorize a test item, and then categorize it again after having been forced to attend to the features consistent with the experimenter's rule. The experimenter forced participants to attend to the rule-consistent features by having them describe the rule-consistent features, as pointed out by the experimenter (e.g., "Please describe the legs and torso"). This allows us to see what effect attending to the rule-consistent features following an initial overlap response has on a participant's second categorization of the same item.

If a familiar overlap feature diverts attention from less familiar rule-consistent features, then participants should revise their second answer to accommodate the more numerous rule-consistent features, producing a "correct"[2] categorization on the second categorization step (*revision response*). However, if the familiar overlap feature was weighted more heavily than the less familiar rule-consistent features, then redirecting attention back to these less-important features should not change their categorization, and participants should persist in their initial response on the second categorization step (*persistence response*). Another effect of a familiar feature could be to change the interpretation of the less familiar features (*interpretation response*). We can assess this by inspecting how participants describe the rule-consistent features.

Hannah and Brooks' (2006) training procedure also aided the acquisition of strong rules by identifying characteristic features of each category at the outset of training, and by having training rounds in which participants identified the characteristic features of exemplars. Over 50% of Hannah and Brooks' participants produced strong rules at the end of test, while fewer than 10% of Brooks and Hannah's (2006) produced strong rules, unless features were identified at the outset of training. Teaching all participants the relevant features also ensures a common vocabulary for describing features. Describing the rule-consistent features at test is essential to distinguishing between attention-related and weighting-related overlap responses.

Feature training also leads to the emergence of a common feature parsing scheme across groups that is consistent with the experimenter's test materials. This is critical to ensure that differences between rule groups reflect reliance on representations of the same features at different levels of abstraction, rather than reliance on different features at the same level of abstraction. This concern precluded giving one group a strong rule, and forcing another to learn the categories by pure induction, without any feature training. Because induction with feature training results in half or more of participants learning strong rules anyways, it was felt that the explicit teaching of strong rules to one group was pointless. Finally, using Hannah and Brooks (2006) training procedure allowed us to maintain contact with formal learning situations, as Hannah and Brooks' procedure was aimed at producing an analogue of a medical diagnostic biasing effect. Domains where decision-making is most likely to involve a reliance on informational features, such as science and medicine, are also reasonably likely to involve some formal learning.

When the overlap feature is the most familiar feature present, we expected that weak-rule producers would make a greater proportion of persistence responses than strong-rule producers after first following the overlap feature. When all features are approximately equally novel, then the differences between rule groups should be at least muted. Some differences between rule groups may persist even when all features are novel, as it may be impossible to ensure all features are equally dissimilar from old features. Further, what we are calling an effect of familiarity may reflect a more general mechanism, such as a fluency of processing heuristic

(Begg, Anas, & Farinacci, 1992; Jacoby & Dallas, 1981), and some novel overlap features may be more easy to process than their rivals.

On the other hand, if all participants relied on informational features when making a categorization decision, and the familiarity of instantiation influences only attention, then two things should emerge in the response patterns. Both weak- and strong-rule users should make mainly revision responses after first following the overlap feature, and the proportion of revision responses should be constant regardless of the familiarity of the overlap feature.

Method

Participants

Eighty-three McMaster University students participated in exchange for course credit in a first-year psychology course. All spoke English as their first language. Two participants were dropped for not following directions, and one for failing to meet the learning criterion (described in the Procedures section below). The data reported here, therefore, come from 80 participants, with 40 participants in each test-item condition (familiar overlap, novel overlap). Participants were run in cohorts ranging from two to six participants per session.

Stimuli and apparatus

Stimuli consisted of line drawings of imaginary animals displayed using an overhead projector, and measured approximately 20 cm X 35 cm on the screen. The drawings comprised four species of imaginary animals, with each category created around a family-resemblance rule based on three features: tail type, torso shape and number of feet/legs. All members of a category had at least two features diagnostic of that category. The respective diagnostic values of these features are listed under each species name in Figure 2.

The informational structure of the training set is defined in Table 1. Examples are shown in the upper panel of Figure 2. For each category, the training set consisted of a prototype exemplar, with all the diagnostic features, and three exemplars that differed from the prototype by one diagnostic feature. For all non-prototype exemplars, the deviant feature had an informational value that was diagnostic of another category. That is, the deviant feature was an overlap feature, making these *overlap exemplars*. For prins, for example, the overlap exemplar shown in the upper panel of Figure 2 has a novel manifestation of a pentagonal torso, which is the diagnostic value for the ramus torso.

Nondiagnostic features (head shape, horn curvature, neck length) varied within each category across three values, but these values were shared across categories. All prototypes shared the same values for these nondiagnostic features (backward-curving horns, oval head and medium neck length), but instantiated them differently. Each overlap exemplar took on one of these values for one nondiagnostic feature, and had a different value on the other two features. For example, one overlap exemplar had backward curving horns, but a triangular head and short neck, another an oval head but forward-curving horns and a differently shaped short neck, and another overlap exemplar had the same neck as the prototype, but a crescent-shaped head and straight horns.

Test items consisted of equivalents of the training overlap exemplars, and were made by skewing the features of training items 20º clockwise and 20º counterclockwise, producing two skewed versions of each feature. Skew was implemented using Adobe™ Photoshop's™ skew filter, rotating items around the horizontal axis only. These skewed features were assembled to yield two test items for each training overlap exemplar. Skewing the training features was intended to hold overall similarity to training items to a moderate level across all test items, while applying a constant distortion across each feature.

For the 24 familiar-overlap test items, the overlap feature (torso, for the prin example in Figure 2) was replaced with the unskewed feature found in the overlap category during training. The prin overlap exemplar shown in the upper panel of Figure 2, for example, gives rise to the familiar-overlap test item shown in the left column of the lower panel of Figure 2 by replacing the original torso with the torso seen in training in ramus exemplars. For the 24 novel-overlap test items, the original informational overlap feature was replaced with a novel feature value that embodied the same diagnostic information. An example of a prin novel-overlap test item is shown in the right column of the lower panel of Figure 2.

Procedure

Training. Participants were told at the start of training that their task was to learn a set of four species of imaginary animals, and to apply that knowledge later to categorizing new exemplars into one of the four categories. Across the training period, participants then saw eight presentations of each item. Training began with the experimenter pointing out the consistent and inconsistent features (e.g., “This prin does not have two feet, but it does have a conical torso and a curly tail.”). This maximizes the chance that participants have a vocabulary to describe those features that is transparent to experimenters. Neither the overlap status of the inconsistent features nor the rules organizing the categories were disclosed to participants. To maintain consistency with Hannah and Brooks (2006), and to replicate formal learning situations, features were explained within the context of an evolutionary cover story. Examples of the cover stories are given in Appendix A.

After this initial training round, participants (a) silently identified the characteristic features of individually presented training items, with the experimenter providing feedback, (b) silently studied the training items with no feedback from the experimenter ("free study"), (c) silently identified the characteristic features of training items presented in pairs, with pair members drawn from different species, with the experimenter providing feedback, (d) silently categorize pair members drawn from different species, with the experimenter providing feedback, (e) provide a written categorization of pair members, with experimenter feedback, (f) free study, (g) written categorization of single items, with experimenter feedback. Training was thus extensive, taking about 40 minutes, and provided substantial support for learning the informational structure of the categories and the informational descriptors for individual features.

Performance on the final presentation of individual training items and a presentation of the same items in a different order after the categorize-describe-categorize task was used to assess learning. Only participants who achieved over 70% accuracy on both assessment rounds were treated as having learned the material sufficiently for analysis. This threshold is 60% of the distance between chance (25%) and perfect performance [.6(1.0 – chance)].

Test. The experimenter presented the 24 test items individually in a randomly generated order held constant across all participants. Each item was displayed twice in a trial, with an initial display of eight seconds. Participants wrote down the identity of the item, and the basis for their categorization. Following this, the experimenter redisplayed the item, indicating the two features consistent with the experimenter's rule, and asked participants to describe them—e.g., "Please describe the torso and feet." Upon completing this step, participants repeated the first step by categorizing the item, and justifying their decision. The second display (probe and second categorization step) was untimed.

Post-test segregation: Feature listing versus feature counting. Participants were asked at the end of the experiment to describe how they made their categorization decisions during test, and from this report they were assigned to one of three strategy groups: counting, listing, or single-feature rule.

Feature-counting participants described their decision-rule involving a threshold based on the number of features present (e.g., "I based my decision on which species had the most features present", or "Every item had two features from one species, so I just looked for two consistent features"). Feature-listing participants simply listed several features as important, without mentioning any critical threshold involving the number of features (e.g., "I based my decisions by which of the three relevant features were present," or, "I mainly relied on torsos and tails. Sometimes I looked at feet, but not much."). Examples of both types are given in Appendix B.

Single-feature participants stated that they exclusively used a single, specific feature (e.g., torso) to make distinctions. Data from the six participants in each test-item condition (familiar overlap, novel overlap) who gave a single-feature rule was not analyzed because it is not clear whether single-feature rules represent an abbreviated counting rule, an abbreviated feature list, or some other approach.

As the distinctions between rule types is obvious (whether only one feature is mentioned or more than one is mentioned, whether an explicit decision rule involving the number of features is present or not), segregation was done by the first author, rather than having multiple independent raters.

Results and discussion

Our central analysis was of the response patterns made by participants producing a feature-listing strategy statement at the end of testing versus those producing a feature-counting strategy given that an overlap response was made on the initial categorization step. We also compared the performance of participants receiving novel-overlap test items to those receiving familiar-overlap test items. We compared their accuracy on the two assessment rounds involving the training items, their accuracy on test items, and their proportion of items eliciting overlap responses at test. We expect there should be no difference in training performance, and that overlap responses become more common when the overlap feature is familiar than when it is novel, replicating the basic findings of Brooks and Hannah (2006). For all analyses, $\alpha = .05$.

Overall performance: Familiar versus novel overlap

Table 2 suggests that training accuracy for both the familiar-overlap and novel-overlap conditions was high, and nearly identical. This is confirmed by A 2 X 2 mixed-design ANOVA on training accuracy, with test-item condition (familiar overlap, novel overlap) as a between-subjects factor and assessment round (end of training, after test); for test-item condition, $F(1,78) < 1$, and no other effects were significant.

Table 2 also suggests that participants in the familiar-overlap condition made more errors at test than did novel-overlap participants, $t(65) = -3.13$ (*df* corrected for heterogeneity of variance[3]), $SE = 0.84$, $p < .005$. Not only were familiar-overlap participants less likely to categorize items according to the experimenter's rule, they made more overlap response than did novel-overlap participants, $t(55) = 4.05$ (*df* corrected for heterogeneity of variance), $SE = 0.67$, $p < .001$. This influence of a familiar-looking feature is found within both strategy groups, including feature-counting participants. Feature-counting participants made an initial overlap response on 8.3% of familiar-overlap test items, but on only 2.2% of novel-overlap items, normal approximation to the binomial: $z = -9.02$, $p < .05$. Feature-listing participants made an initial overlap response on 24.4% of familiar-overlap items, but made overlap responses on only 8.0% of novel-overlap items, normal approximation to the binomial: $z = -11.56$, $p < .05$.

Familiar-overlap participants were also more likely to make an overlap response as a proportion of errors. For familiar-overlap participants, 85.7% of errors were overlap responses; for novel-overlap participants only 66.2% of errors were overlap responses (normal approximation to the binomial: $z = -4.88$, $p < .05$).

Response patterns

We first scored the responses according to whether the initial categorization was consistent with our rule. Non-rule responses were further broken down into persistence, revision, interpretation and confusion responses. Only the first three represent overlap responses. We have restricted our analyses to overlap responses, which are shown in Figure 3for both rule groups within each test-item conditions. A summary of all non-rule response frequencies is given in Table A of Appendix C. Differences in response patterns between groups were tested using the maximum-likelihood chi-square (G^2, or L^2).

The top panel of Figure 3 shows the influence of a familiar-overlap feature on overlap response patterns. Feature-counting participants (black bars) made mainly revision responses, but feature-listing participants (grey bars) made mainly persistence responses. This pattern persists for novel-overlap items as well (bottom panel), although it appears much muted for feature-listing participants. Regardless of overlap feature familiarity, there is a significant difference in the pattern of overlap responses across test item as a function of rule group, $G^2(2) = 15.18$, $p < .0001$. Different strategy statements were therefore linked to different patterns of overlap responses, with weak-rule users producing mainly persistence responses, and strong-rule users producing mainly revision responses.

Comparing the top and bottom panels of Figure 3, the predominance of persistence responses over revision responses declines for feature-listing participants as the familiarity of the overlap feature declines. When the overlap feature is from training, over 70% of overlap responses are persistence responses, and only 24% are revision responses, a difference of over 45 percentage points. When the overlap feature is novel, however, only 57% of overlap responses are persistence responses, and 38% are revision responses, a difference of only 19 percentage points. However, for the feature-counting participants, there is little appreciable change in error pattern with changes in the familiarity of the test item. When the overlap feature is from training, 37% of overlap responses are persistence responses, and 58% are revision responses, a difference of 21 percentage points. When the overlap feature is novel, however, only 33% of overlap responses are persistence responses, and 58% are revision responses, a difference of 25 percentage points. This is the pattern we would expect if feature-listing participants, but not feature-counting participants, were weighting features by familiarity. Partitioning the data by rule group, unfortunately, does not find the pattern for the feature-listing participants to be reliable, $G^2(2) = 1.66$, $p > .40$. The much smaller difference found among feature-counting participants is also not significant, $G^2(2) = 0.62$, $p > .90$.

Feature-listing participants made more persistence responses than revision responses when following a familiar overlap feature. The reverse held for feature-counting participants, who made more revision responses than persistence responses when following a familiar overlap feature. Both weak- and strong-rule users are influenced by the familiarity of feature instantiation, but in different ways. For weak-rule users, a familiar feature provides a source of weighting information, as well as a bias on attention; for strong-rule users, a familiar feature only biases attention.

Although the change in dominance of persistence responses for feature-listing participants did not prove significant, the nonparametric test used is relatively low in power, and the observed pattern is exactly what would be expected under the feature-weighting hypothesis. The persistence of persistence responses with novel-overlap items is not surprising; as we noted earlier, it is not likely that all the features of the novel-overlap test items are identically similar to training items. The feature-listing participants in the all-novel condition made only an average of 1.1 persistence responses across all 24 items. This is such a low absolute rate of responding that it could plausibly arisen due to novel overlap features being randomly more similar to training features than their rivals. Further, if a feature-familiarity weighting heuristic is really a proxy for some more general mechanism, such as fluency of processing (Jacoby & Dallas, 1981), then other sources of fluency differences across features may remain even after eliminating those due to structural similarity to training features.

Interpretation responses were low for all participants, suggesting that the stimuli were not ambiguous to our participants. Members of both strategy groups showed both persistence and revision responses, which is reasonable if we assume that people generate both instantiated and informational feature representations. A varied stock of feature representations contributes to the flexibility with which people can apply knowledge both across and within tasks. It would be simplistic to assume that a reliance on one type of feature representation is absolute, even within a task. We take participants' rule statements as descriptions of what they typically did, not what they always did.

Experiment 1B: Yoked controls

That feature-counting participants made fewer overlap responses than feature-listing participants is to be expected. They were using the kind of rule used by us to generate the category, so their performance should be more consistent with our categorization of the test items than a group of people seemingly not using our rule. However, the difference in base rates for overlap responses raises some problems for our feature-weighting account. It could be that we simply do not have a large enough sample of overlap responses from the feature-counting group to get a stable pattern. Or, it could be the differences reflect a lack of some critical knowledge other than feature representations.

Experiment 1B used yoked participants who were denied any experience with particular training exemplars, being trained solely on verbal descriptions of the four categories. At test, they were given the strategy statements generated by ten participants from each of the counting and listing groups of Experiment 1A, familiar-overlap condition only. These yoked control participants were instructed to follow these rules in making their decisions. We expect that regardless of which set of rules are given to the yoked participants in this experiment, they can only respond on the basis of informational features, which are all they learned in training. We therefore expect that both yoked groups will reproduce the pattern of overlap responses found for the feature-counting rule-giving participants from Experiment 1A, making a larger proportion of revision responses than persistence responses.

Method

Participants

Twenty-four McMaster undergraduate students participated in exchange for course credit in a first-year psychology course. All spoke English as their first language. Four people were dropped for not following directions. Twenty participants supplied data for this experiment, ten in each yoked condition (feature-counting, feature-listing). Participants were run in cohorts ranging in size from one to eight persons.

Stimuli and apparatus

Strategy statements were randomly chosen from those generated by counting and listing strategy groups in the familiar-overlap test condition, subject to the constraint that the rule be coherent and usable. The researcher intuitively decided whether a rule met this constraint. Ten statements were taken from each group. These strategy statements were put on the yoked participants' response sheets, one rule statement per participant. The complete set of rule statements is given in Appendix B.

None of the rules gave descriptions of specific feature values so complete as to be sufficient for a person who had never seen any of the stimuli before. To give the yoked participants knowledge of the informational features, they were given verbal descriptions of the relevant features for each category. The verbal descriptions were put into a table form. The familiar-overlap test items from the previous study were used. Both the table of features and the line drawings were displayed by projecting transparencies on an overhead projector.

Procedure

Participants were told they were to apply the rules they were given to categorizing test items into one of four species, and were told to do this they needed memorize the associations between the categories and the feature labels. Training took place in three stages. First, participants studied the complete feature table for five minutes. Second, participants had two minutes to fill in missing values for each category in a partial table, with feedback after. Participants were then given two minutes to fill in an entirely empty table, with feedback after. To qualify as having learned the categories, participants were required to make no more than three errors on this final task, with no more than one error being made in any category.

The test procedure was identical to that used in Experiment 1A.

Results and discussion

There was no reliable difference between the yoked groups in the accuracy of filling in the category-by-feature table at the end of training. Listing yokes accuracy = 91.7%, counting yokes accuracy = 96.7%, $t(18) = -1.26$, $SE = 0.48$, $p > .20$.

The upper panel of Figure 4 shows equivalent accuracy on test items across all groups. A paired *t*-test found no significant differences in the mean number of overlap responses made between each yoked group and that of their rule-producing partners, $t(9) = 1.03$, $SE = 2.0$, $p > .30$. Feature-listing rule producers made an average of 4.5 overlap responses, and their yoked partners made an average of 6.6 overlap responses. Feature-counting rule producers made an average of 2.7 overlap responses, and their yoked partners made an average of 5.7, $t(9)= 1.83$, $SE = 1.6$, $p > .10$.

Differences between the feature-listing groups (yoked participants and rule producers) are shown in the lower panel of Figure 4, where persistence and revision responses are given as a percentage of total persistence and revision responses. As expected, both yoked groups looked like the feature-counting rule producers, and both looked quite different from the feature-listing rule producers. Both yoked groups made more revisions responses than overlap responses. A maximum-likelihood chi-square analysis on the distribution of persistence and revision responses across strategy groups found a significant difference between the feature-listing rule producers response pattern and that of their yoked partners, $G^2(1) = 4.25$, $p < .05$. No difference in response patterns was found for feature-counting rule producers and their yoked partners, $G^2(1) = 0.38$, $p > .35$.

Feature-listing yoked participants differed in the pattern of persistence and revision responses from their rule-giving counterparts despite equivalent base rates of overlap responses.

The pattern of responses made by the feature-listing group in Experiment 1A, therefore, cannot be attributed to deficiencies in knowledge, or to differences in the base rate of overlap responses between listing and counting participants. This strengthens the argument that the pattern of responding made by the feature-listing participants in Experiment 1A reflects the operation of a feature-weighting heuristic applied to instantiated features.

Both yoked groups produced the same response pattern as the feature-counting participants of Experiment 1A. Because the yoked groups represent categorizers who have no instantiated features available to them, this finding supports the argument that the use of strong rules reflects a reliance on informational features, and is not an artifact of a low base rate of overlap responses from feature-counting participants in Experiment 1A.

General discussion

We have shown that a familiar-looking feature can influence categorization through either of the two routes suggested in the introduction. The familiarity of feature instantiation can bias attention during feature search prior to making a categorization decision. Alternatively, it can be used to weight features when making a categorization decision. The latter finding confirms what Hannah and Brooks' (2006) hypothesized: People can use feature-familiarity to perform the function of a decision criterion in a strong rule, resolving conflicts among features by weighting familiar features more than less familiar features. This dual-route account can explain the pattern of results found not only in Brooks and Hannah (2006), and Hannah and Brooks (2006), it can also explain Thibaut and Gelaes (2006) finding that people given a perfectly reliable strong rule were still influenced by feature familiarity—because of the influence of feature familiarity on attention.

According to the dual-route hypothesis, how a familiar-looking feature influenced our participants depended on what feature representations they relied on when making decisions, as indexed by the types of rules they produced at the end of test. Those producing weak rules (feature lists) seemed mainly susceptible to weighting-related errors (or, persistence responses, in the context of our task), while participants producing strong rules (feature counting rules) made mainly attention-related errors (revision responses) in the presence of a familiar overlap feature (Experiment 1A). Furthermore, when we deprived participants of any chance of acquiring knowledge of feature instantiations prior to test (Experiment 1B), participants produced the same pattern of responses as that generated by strong-rule users in Experiment 1A. The error patterns found in Experiment 1A are not functions of either the base rate of overlap responses, or of level of knowledge.

That these attentional and weighting effects of feature familiarity are correlated with the type of rule statements produced at the end of test provides important confirmation of earlier arguments (Brooks & Hannah, 2006; Hannah & Brooks, 2006). It confirms that people can use either instantiated or informational features to make decisions, and that reliance on a particular level of feature abstraction is typically reflected in different types of rules. If weak rules, like the feature lists given for ordinary categories (Rosch & Mervis, 1975), were merely inarticulate versions of strong rules then we would expect that attention-related responses would dominate over weighting-related responses for feature-listing participants. In this view rules do not work in opposition to similarity-mediated memory processes (Sloman, 1996), nor are they different points along a continuum (Pothos, 2005). Rather, the relation between rule use and similarity varies with the nature of the rules used.

Our feature-listing participants made predominantly persistence responses even in the novel-overlap test condition of Experiment 1A. As noted in the discussion to Experiment 1A, although all features were novel, it is likely that there was still some variance regarding how similar the features of test items were to training features. This is plausible considering the low mean level of persistence response frequency in the novel-overlap condition. A more interesting possibility, however, is that what we have called a feature-familiarity weighting heuristic reflects a more general mechanism than familiarity, such as fluency of processing (Begg, Anas, & Farinacci, 1992; Jacoby & Dallas, 1981). Thus, it may not be the familiarity of the feature per se, that provides feature weights, but the ease of feature *recognition*. In which case, a feature that is a better fit to its learned label would be weighted as more important than its rivals, even if all are equally familiar or equally novel. A novel feature that is a caricature of learned features or of their labels might elicit the same kind of responding from users of weak rules as a familiar instantiation.

Decision-making, attention and memory

For feature-counting participants, revision responses increased as familiarity increased, further evidence that search for objects in displays influenced by familiarity of object, distractors and targets (Chun, 2000; Johnston, Hawley, Plewe, Elliot & Dewitt, 1990; Johnston, Hawley & Farnham, 1993). We are showing a similar episodic influence, but for feature search within an object, rather than for object search within a scene. Further, we are showing how these attentional affects of memory can influence categorization decisions.

In the introduction, we outlined how instantiated features can be thought of as supplying concept-specific bounds to feature representations. The influence of episodic memory upon feature selection can be seen as another, albeit less direct, way of implementing concept-specific bounds upon feature representations, one which could operate on both instantiated and informational features. This can help us understand why people working even perfectly reliable rules can produce errors in response to the presence of a familiar feature as Thibaut and Gelaes' (2006) found, at least if we assume rules tend to be engaged under conditions of uncertainty or information conflict. If a familiar feature diverts attention from conflicting features, then no conflict would be detected and strong rules would fail to be engaged.

People can use rules and conceptual information that take some abstract, generic form such as causal schemas, or propositions (Anderson, 1991; Sloman, Love, & Ahn, 1997). However, not only is there variety in the level of abstraction of feature representation, even decision-making based on abstract representations requires that the relevant inputs be first selected. The evidence here is that input selection is under the influence of episodic memory. The analytic sits atop the nonanalytic.

Conceptual flexibility

The results of our studies imply that decision-making in categorization is at least partially dependent on how features are encoded, and on what level of feature representation seems most reliable in a given context. Other evidence suggests that the perceived reliability of feature representation is partially influenced by the relative predictiveness of categorical identity (Hannah, Jamieson, & Brooks, in preparation). Both the relative predictiveness of abstract and concrete feature representations varies across context, and levels of feature familiarity can change in the course of a task

This work supplements other approaches to conceptual flexibility, a topic Barsalou investigated over 20 years ago (Barsalou, 1985). There are some parallels between Barsalou's (1985) discussion of the frequency of item instantiation and the flexibility of conceptual structure and our treatment of feature instantiations and conceptual flexibility. However, some important distinctions must be noted, besides the obvious difference of Barsalou's analysis operating at the exemplar level and ours at the feature level. More importantly is that he distinguishes between frequency of occurrence within a category, while we are pointing to the frequency of occurrence within a category *relative to the occurrence across categories*. The more frequent a feature occurs within a category relative to its frequency across categories, the more it will behave like an instantiated feature. This puts our analysis of the power of instantiation features closer to Medin and Schaffer's (1978) modeling of exemplar effects on typicality and conceptual structure.

Further, rather than describing this distributional aspect in terms of context-dependent versus context-independent elements, as Barsalou (1985) does, we think the distributional properties are better described as involving a narrow context or a broad context. As we alluded to in the introduction, and have described at greater length elsewhere (Brooks & Hannah, 2006; Hannah & Brooks, 2006), features that have a broad context, as informational features do, also offer critical affordances, especially for reasoning and communication. The role of distributional properties in coordinating the use of instantiated and informational features is the focus of an investigation that we are currently conducting (Hannah, Jamieson & Brooks, manuscript in preparation).

We should also point out that while Barsalou talks about frequent versus infrequent exemplars within a category, he has no distinction that corresponds to that of instantiated versus informational, that is, exemplars do not vary in their degree of abstraction. We are talking less about the role of different patterns of distribution alone, but of patterns of distribution within different feature coding schemes. Most treatments of features assume they are coded in some one fashion. For example, they might be treated as continuous (e.g., Medin & Schaffer, 1978), allowing features to be more or less similar to one another. Or features may be treated as discrete, all-or-none entities (e.g., Tversky, 1977), a coding scheme that readily supports a verbal description, and captures general relations. We are suggesting that in learning about items, people use both such approaches, or perhaps, both approaches represent the ends of a continuum of encoding abstraction. This flexibility in feature encoding supports flexibility in conceptual structure and conceptual processing.

Many modeling approaches to conceptual flexibility acquire flexibility by building in both abstract and concrete levels of representation. Nosofsky, Palmeri and McKinley's (1994) RULEX model tries to learn rules, beginning with simple single-dimension rules, and supplements this by also memorizing whole instances embodying exceptions to its rules, but feature representation is ignored. Further, the model always tries to solve a problem first with its rules, and then using its exemplar exceptions. Erickson and Kruschke's ATRIUM (1998) model takes this rule-versus-instance approach further, allowing a variable reliance on one or the other just as we propose reliance on one level of feature representation can vary.

The SUSTAIN model recently proposed by Love, Medin and Gureckis (2004) is perhaps closest to our approach, being feature-based. Like RULEX, SUSTAIN assumes that people seek the simplest solution first. Unlike RULEX, it is not an explicitly rule-learning model. Love et al.'s model forms clusters of exemplars based on similarity to past exemplars along specific features, and can re-organize and add new clusters as learning progresses. However, as currently formulated, the model treats the feature space in which the clusters form as a given. Flexibility at the level of feature construction is ignored in most theoretical accounts of categorization and decision-making.

Along with Schyns' and colleagues (Schyns, Goldstone & Thibaut, 1998; Schyns & Murphy, 1994; Schyns & Rodet, 1997), this work is pointing to the importance of the construction of feature spaces, with an emphasis on the importance of feature spaces that operate at different levels of abstraction. Further, our findings suggest that how a feature space is used to make decisions depends on the level of abstraction of the features defining that space. Both the structure of a category and how decisions are made regarding that category is shaped in part by what level of feature abstraction a person chooses to work at. Flexibility at the level of feature encoding, or, *encoding variability* (Jamieson & Mewhort, 2005; Martin, 1968), has substantial consequences for the learning and use of concepts, just as it does for memory.

References

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

Barsalou, L. W. (1985). Ideals, central tendency and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: General, 11*, 629-654.

Begg, I., M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, testimonial familiarity, and the illusion of truth. *Journal of Experimental Psychology: General, 121*, 446-458.

Brooks, L. R. & Hannah, S. D. (2006). Instantiated features and the use of "rules". *Journal of Experimental Psychology: General, 135*, 133-151.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Science, 4,* 170-178.

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

Hannah, S. D., & Brooks, L.R. (2006). Producing biased diagnosis with unambiguous stimuli: The importance of feature instantiations. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32*, 1416-1423

Hannah, S. D., Jamieson, R. K., & Brooks, L. R. (2009). *From information to instantiation: Varying people's reliance on different levels of feature abstraction in concept learning.* Manuscript in preparation.

Jacoby, L. L., & Dallas, (1981). On the relation between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*, 306-340.

Jamieson, R. K. & Mewhort, D. J. K. (2005). The influence of grammatical, local, and organizational redundancy on implicit learning: An analysis using information theory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31*, 9-23.

Johnston, W. A., Hawley, K. J., & Farnham, J. M. (1993). Novel popout: Empirical boundaries and tentative theory. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 140-153.

Johnston, W. A., Hawley, K. J., Plewe, S. H., Elliot, J. M., & Dewitt, M. J. (1990). Attention capture by novel stimuli. *Journal of Experimental Psychology: General, 119,* 397-411.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review, 111*, 309-332.

Martin, E. (1968). Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. *Psychological Review, 75*, 421-441.

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

Pothos, E. M. (2005). The rules versus similarity distinction. *Behavioral and Brain Sciences, 28*, 1-49.

Rosch, E. & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*, 573-605.

Schyns, P. G., Goldstone, R. L., & Thibaut, J.-P. (1998). The development of features in object concepts. *Behavioral and Brain Sciences, 21*, 1-54.

Schyns, P. G., & Murphy, G. L. (1994). The ontogeny of part representation in object concepts. In D.L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory, vol. 31,* (305-349). San Diego, CA: Academic Press.

Schyns, P. G., & Rodet, L. (1997). Categorization creates functional features. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23,* 681-696.

Shepard, R. N., Hovland, C. I. & Jenkins H. M. (1961). Learning and memorization of classification. *Psychological Monographs, 75* (13, Whole No. 517).

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3-22.

Sloman, S. A., Love, B. C., Ahn, W. (1997). Feature centrality and conceptual coherence. *Cognitive Science, 22,* 189-228.

Thibaut, J.-P. & Gelaes, S. (2006). Exemplar effects in the context of a categorization rule: Featural and holistic influences. *Journal of Experimental Psychology: Learning, Memory & Cognition, 32*, 1403-1415.

Appendix A: Training Cover Stories

From training protocol. Cover stories below were given while displaying prototypes of each category. Italics indicate category-relevant features

Bleeb: Experimenter explains that Bleebs live on mountains, adding:

> Their *semi-circular torso* resembles the boulders that litter the mountain slopes and meadows, and this camouflage is aided by *their lack of a tail*. Their *3-legged posture* gives them great stability in negotiating the cliff faces and rocks of their home.

Croom: Experimenter explains that Crooms live in warm, swampy lands, that are semi-tropical, adding:

> Their *box-like torsos* maximize surface area, which helps spread their weight out over a wider area, keeping them from sinking in especially wet areas. Their *shaggy tail* helps dissipate heat and helps water evaporate quickly, both of which help cool crooms. *Their 6-legged posture* gives them great traction to power through the mud.

Prin: Experimenter explains that Prins live on flat grasslands, mainly at the edges of lakes, large ponds and large rivers, and are semi-aquatic. Experimenter adds:

> Their *cone-shaped torso* gives them buoyancy in water, while the *curly tail* can be contracted and expanded, to help propel them through the water. *Their two-legged posture* gives them great running speed on the flat prairie.

Ramus: Experimenter explains that Ramuses live on tundra, adding

> Their *pentagonal torso* minimizes surface area, reducing loss of body heat, and their tail is *spotted* with deposits of fat that provide fuel reserves during the cold winters on the mountains. Their *four-legged posture* gives an optimal blend of speed and stability in negotiating the flat, but sometimes boggy land they live in.

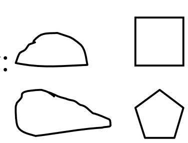
Example of introduction of an item with an overlap feature (bleeb, curly tail—prin feature):
[Experimenter explains]: semi-prehensile tail aids stability for bleebs in *especially steep regions*

Appendix B: Rules From Feature-Counting And Feature Listing Rule Producers, Experiment 1A

Ten of the strategy descriptions from each of the feature-listing and feature-counting strategy groups of Experiment 1A were given to participants in Experiment 1B. The wording of each was changed only to turn it into set of directions. For example, “Memorizing feet, torso and tail for each creature and looking for two of those in the ‘test’ animals to identify it as one of the four originals” was changed to “Memorize feet, torso and tail for each creature and look for two of those in the ‘test’ animals to identify it.”

Feature listing rules (weak rules)

1. Look for the shape of the torso, the type of tail & sometimes how many feet the animals have.
2. Just identify what type of body, then legs and how many were there, then if they have a tail and what kind it was (e.g., bushy or curly).
3. Identify the ‘animal’ by looking at the picture and picking one or two things, i.e., for C, they usually have a bushy tail and for P, they have a very distinct torso, somewhat shorter. By doing so, it will trigger you to remember other characteristics.
4. Remember which characteristics apply to which species.
5. Look at the different shapes of the animals and features, i.e., tail and legs.
6. Identify the animals by their main features. Use a tool where Prin = propeller for the tail and pair for the legs. blEEb= three legs, croom = broom for the tail.
7. Think of the animals in a box: b c. Remember that their legs go 3 6
p r 2 4.

Get to know their body shapes the same way:  Also the same for tails.

8. Look mainly at the body shapes of the animals, as well as the legs and tails. Decide what the animals are based on the torso, tail and legs, and also your first impression of what the animal is.
9. Look at the body types that the animals have as well as the number of feet and the presence of a tail and comparing them to the four that you are taught
10. For the torso use the number of sides or 1/2 circle/cone. For the legs, count how many there are. For the tail, the shape of it (straight, fuzzy, curly or non-existent)

Feature counting rules (strong rules)

1. Memorize feet, torso & tail for each creature and look for 2 of those in the 'test' animals to identify it.
2. Concentrate on two features which link together to a specific creature; any creature with two features belonging to the same initial creature is chosen.
3. Mostly rely on the features: i.e., tail and body shape together, or number of legs and tail together or body shape and legs together.
4. Make decisions by looking at each animal and seeing if they have two features of one of the four categories. The features are: legs, tail, torso.
5. Mostly, link two features of the animal to one of the species. All species have two common features.
6. The animals with the most number of characteristics (of a particular animal) choose as your answer.
7. Match at least two different features together to help recognize which species each animal belongs to.
8. Just look for if an animal has two similar characteristics.

9. To identify the species, look for a dominant number of traits that apply to one species. Although the unknown may contain one trait belonging to another species, two or more traits that fit to one class of species is usually enough to confirm it
10. Match at least two things up. Make our decision on the presence of at least two of the variables (torso, & # of feet or tail & torso, etc.).

Appendix C: Response frequencies for all groups, Experiments 1A & 1B

Table A: *Incorrect responses by type, test conditions within strategy groups, Experiment 1A*.

		Overlap responses			
Strategy	Test	Persistence	Revision	Interpretation	Confusion
Feature counting	Familiar overlap, *n* =19	14	22	2	10
	Novel overlap, *n* =23	4	7	1	4
	Within strategy	*18*	*29*	*3*	*14*
Feature listing	Familiar overlap, *n* =15	62	21	5	15
	Novel overlap, *n* =11	12	8	1	15
	Within strategy	*74*	*29*	*6*	*30*

Table B: *Incorrect responses by type for rule producers and yoked controls, Experiment 1B*.

	Overlap responses			
	Persistence	Revision	Interpretation	Confusion
Feature-counting yokes	19	35	3	10
Feature-counting rule producers	11	13	1	9
Feature-listing yokes	25	28	9	17
Feature-listing rule producers	28	13	3	6

Note. N = 10 for all groups.

Author note

Samuel D. Hannah, Department of Psychology, Queen's University, Kingston, Ontario and Lee R. Brooks, Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Ontario. Funding for the first author has been provided by grants from the National Sciences and Engineering Council for Shepard Siegel and Lorraine G. Allan, and for the second author by the a grant from the National Science And Engineering Research Council to the second author. The research reported here forms part of the first author's Ph.D. thesis. A summary of these results also appeared in a brief review article in the Canadian Journal of Experimental Psychology (Hannah, 2005).

We would like to thank Seth Chin-Parker, Evan Heit, Alan Neville, Geoff Norman, Aimee Skye and Brian Ross for useful suggestions and criticism.

Correspondence concerning the paper can be addressed to Sam Hannah, Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Ontario, Canada, L8S 4K1. The author can be reached via e-mail at hannahsd@mcmaster.ca.

Footnotes

[1] Throughout this paper we use the term *familiar* as a convenient shorthand for *similar to old*, rather than pointing to a subjective experience or specific memory process.

[2] The usage of “correct” or “error”, we should note, reflects another communicative nicety to avoid the cumbersome “consistent-with-the-researcher’s-rule categorization.” This should not be taken to mean that only classifications based on our rule are reasonable or sensible responses.

[3] Degrees of freedom adjusted using the Welch-Satterthwaite solution.

Table 1: *Informational Structure Of Training Items, Experiments 1A And 1B*

	Bleeb			Croom		
	legs	torso	tail	legs	torso	tail
Prototype	1	1	1	3	3	3
Overlap exemplar A	1	1	4	3	3	2
Overlap exemplar B	1	3	1	3	4	3
Overlap exemplar C	2	1	1	1	3	3

	Ramus			Prin		
	legs	torso	tail	legs	torso	tail
Prototype	2	2	2	4	4	4
Overlap exemplar A	2	2	3	4	4	1
Overlap exemplar B	2	1	2	4	2	4
Overlap exemplar C	4	2	2	3	4	4

Note. Legs: 1 = three legs, 2 = four legs, 3 = six-legs, 4 = two legs. *Torso*: 1 = semi-circular, 2 = pentagonal, 3 = box-like, 4 = conical. *Tail*: 1 = none, 2 = spotted, 3 = shaggy, 4 = curly.

Table 2: *Overall Mean "Correct" (And SD) For Familiar-Overlap And Novel-Overlap Groups, Experiment 1A*.

Test group	Assessment round		Test items	
	After training	After test	"Correct"	Overlap responses
Familiar overlap (N = 40)	98.6% (2.6)	96.9% (5.8)	81.0% (18.8)	16.7% (16.1)
Novel overlap (N = 40)	97.8% (6.3)	98.1% (4.9)	92.0% (11.6)	5.3% (1.2)

Note. Assessment rounds involved identifying training items. Overlap responses are given as a percentage of total responses.

Figure captions

Figure 1. Feature representations can involve highly specific representations with narrow selectivity, indicated by the paw at the bottom left, or more generic representations with broad selectivity, indicated by the *paw* at the bottom right.

Figure 2. Schematic of two hypothesis regarding how familiar-looking features influence categorization decisions. The 'search alone' hypothesis (top panel) treats all people as relying on both instantiated (paw picture) and informational features ("paw"), but at different processing stages, so that participants differ only in the quality of feature search. The 'search versus decision' hypothesis (bottom panel) treats strong-rule users and weak-rule users as differing on the type of features used at a decision stage. Not only does this lead to different levels of error[2], but to different patterns of error.

Figure 3. Upper panel: Training stimuli for Hannah and Brooks (2006) and Experiment 1A. Shown above training prototypes (upper row) are the defining features for each category. The overlap feature found in each overlap exemplar (lower row) is identified beneath it; in parentheses is the name of the category the overlap feature is normally found in, or *overlap category*. Lower panel: Two test items used in Experiment 1A. Test items are skewed versions of the training exemplars; both samples shown are versions of the prin overlap exemplar shown above. For familiar-overlap test items (left), the original overlap feature (pentagonal torso) is replaced with a feature seen in training in a rival category. For novel-overlap test items (right), a novel informational equivalent of the original overlap feature is used.

Figure 4. Responses as a proportion of overlap responses for feature-counting (dark bars) and feature-list (grey bars) participants in the familiar-overlap test condition (top panel), and novel-overlap test condition (bottom panel).

Figure 5. Top panel: Mean correct at test for yoked control participants (Experiment 1B) and participants from Experiment 1A who produced the rules used by yoked controls. "F-C" = feature-counting condition; "F-L" = feature-listing condition. Error bars = 1 *SE*. Bottom panel: Persistence responses (hatched bars) and revision responses (grey striped bars) as a proportion of persistence and revision responses for yoked participants and rule-producing partners.

Figure 1

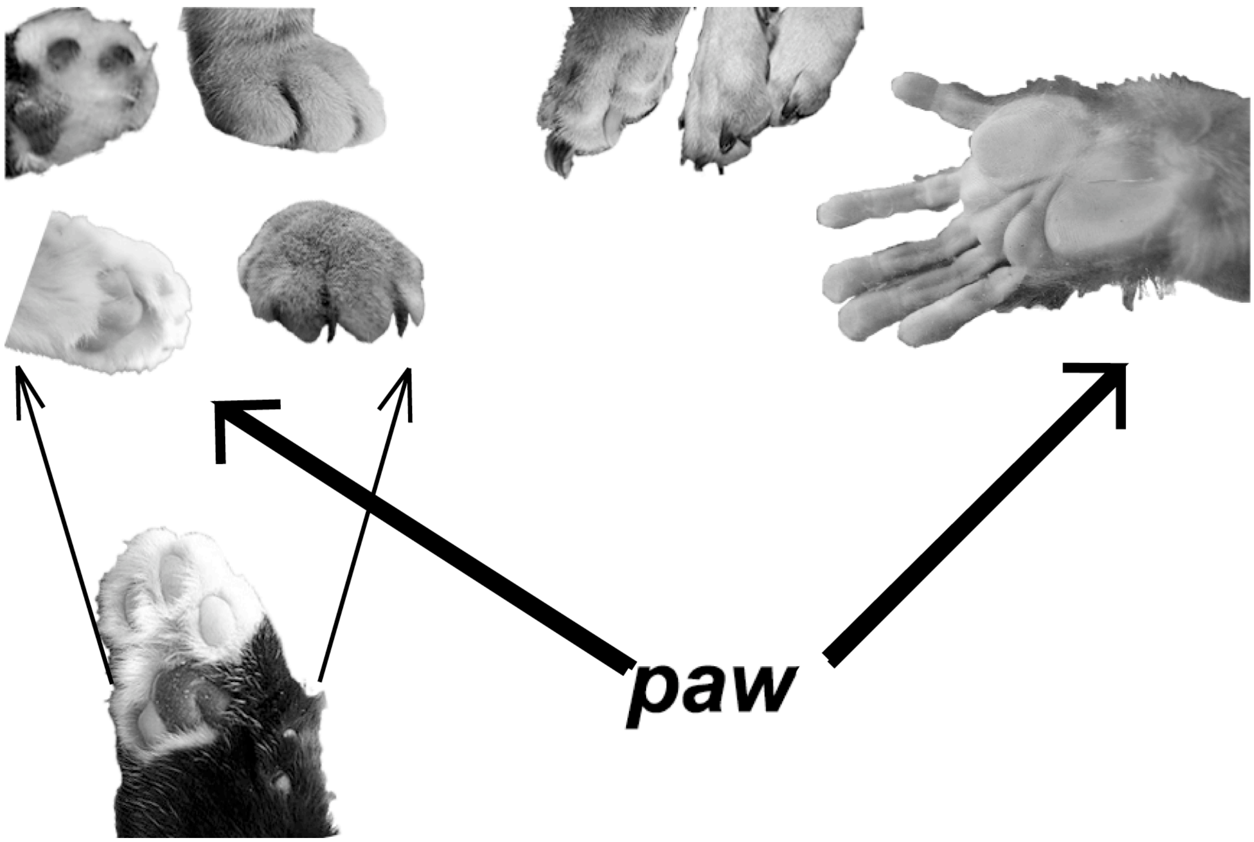


Figure 2

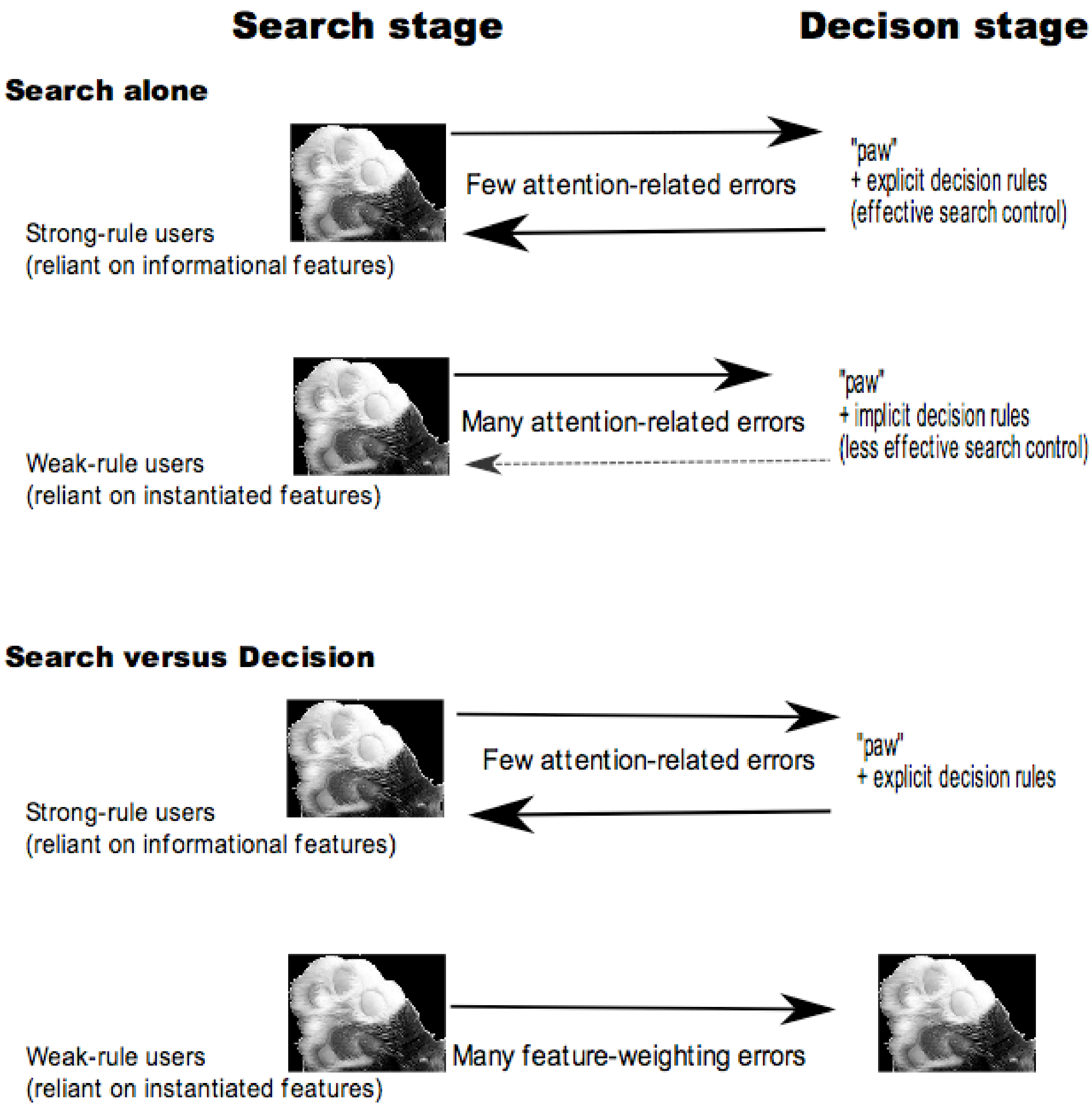

Search stage
Decison stage
Search alone
Few attention-related errors
"paw"
+ explicit decision rules
(effective search control)
Strong-rule users
(reliant on informational features)
Many attention-related errors
"paw"
+ implicit decision rules
(less effective search control)
Weak-rule users
(reliant on instantiated features)
Search versus Decision
Few attention-related errors
"paw"
+ explicit decision rules
Strong-rule users
(reliant on informational features)
Many feature-weighting errors
Weak-rule users
(reliant on instantiated features)

Figure 3

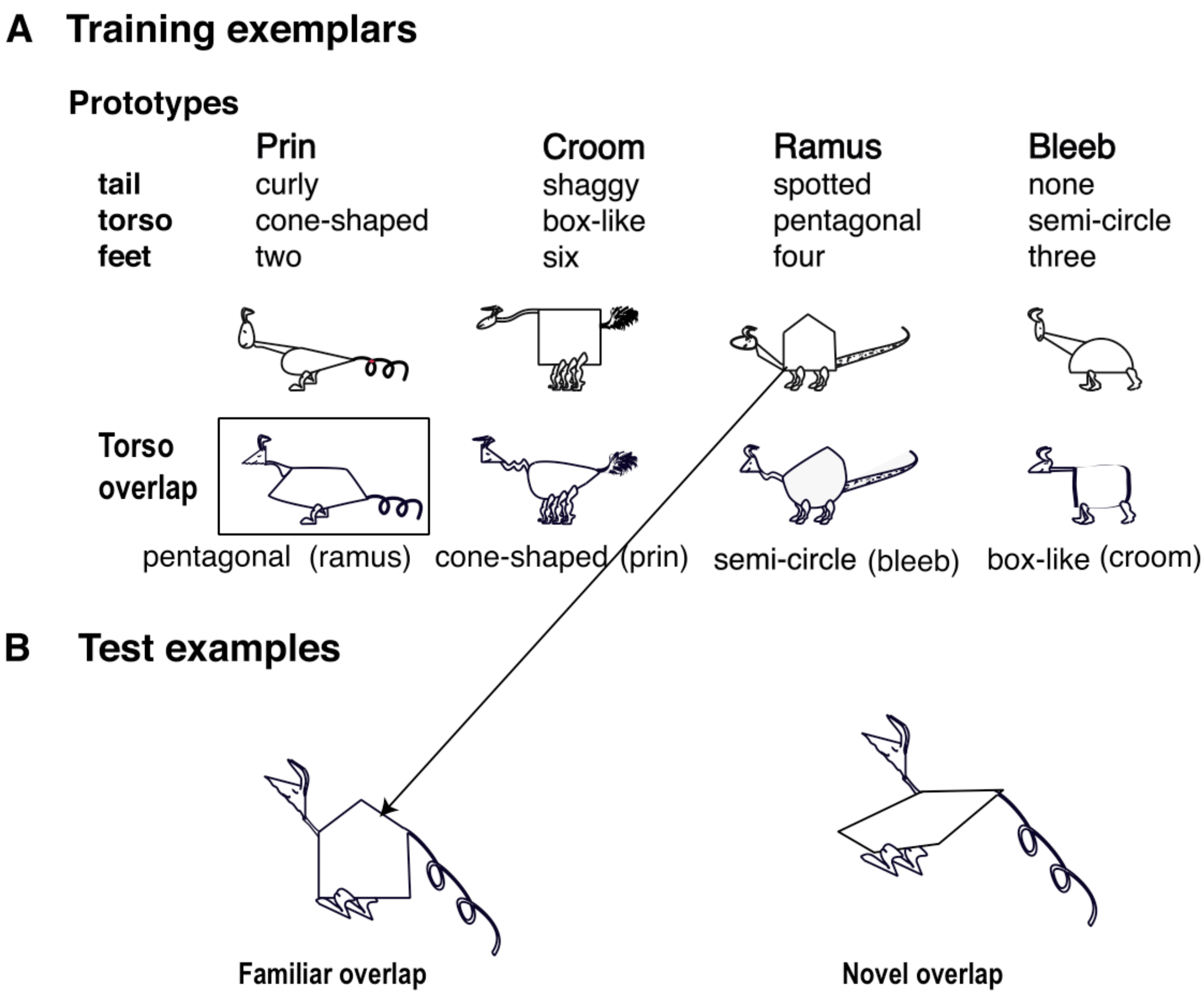

A Training exemplars
Prototypes
Prin
Croom
Ramus
Bleeb
tail
curly
shaggy
spotted
none
torso
cone-shaped
box-like
pentagonal
semi-circle
feet
two
six
four
three
Torso overlap
pentagonal (ramus)
cone-shaped (prin)
semi-circle (bleeb)
box-like (croom)
B Test examples
Familiar overlap
Novel overlap

Figure 4

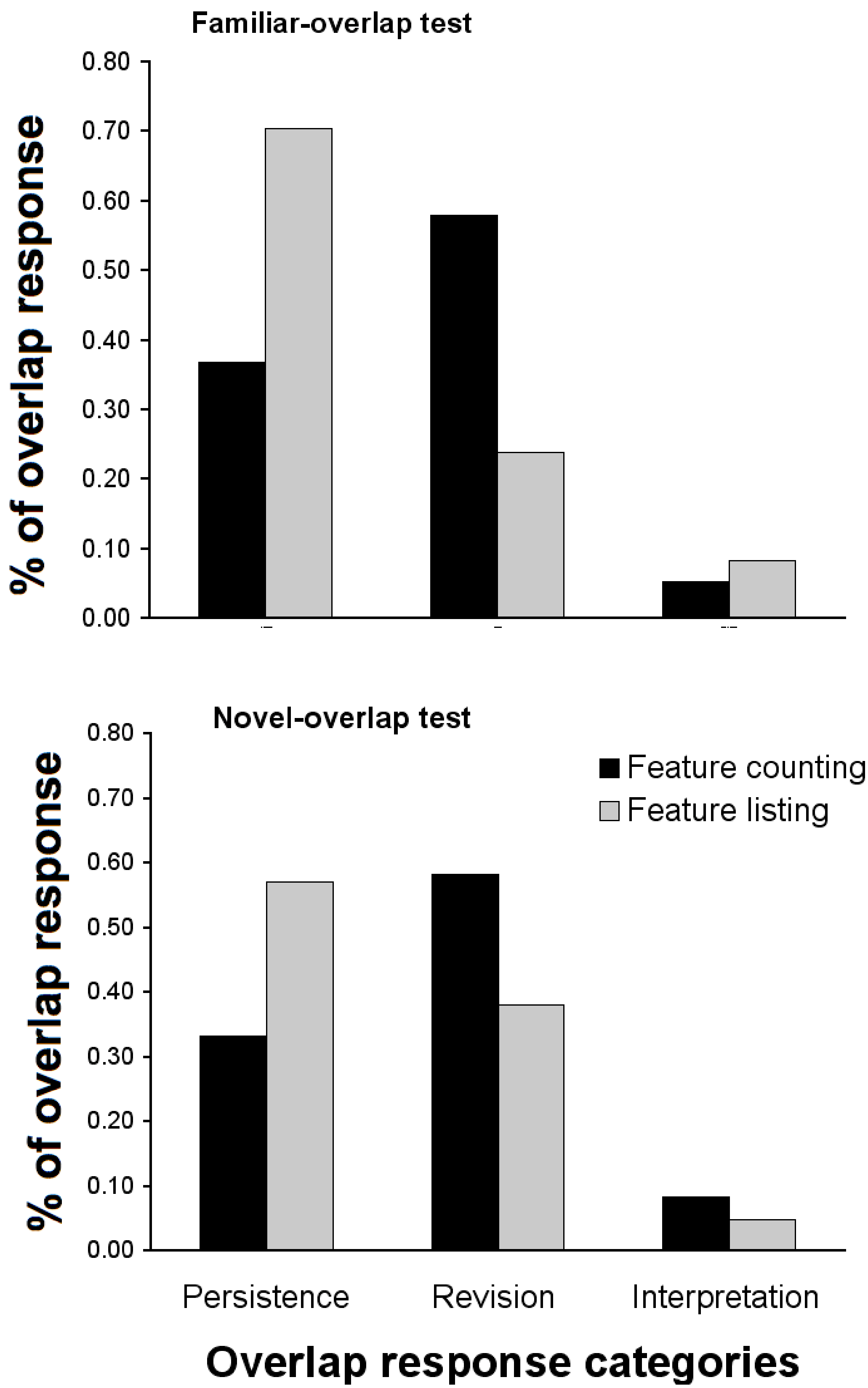
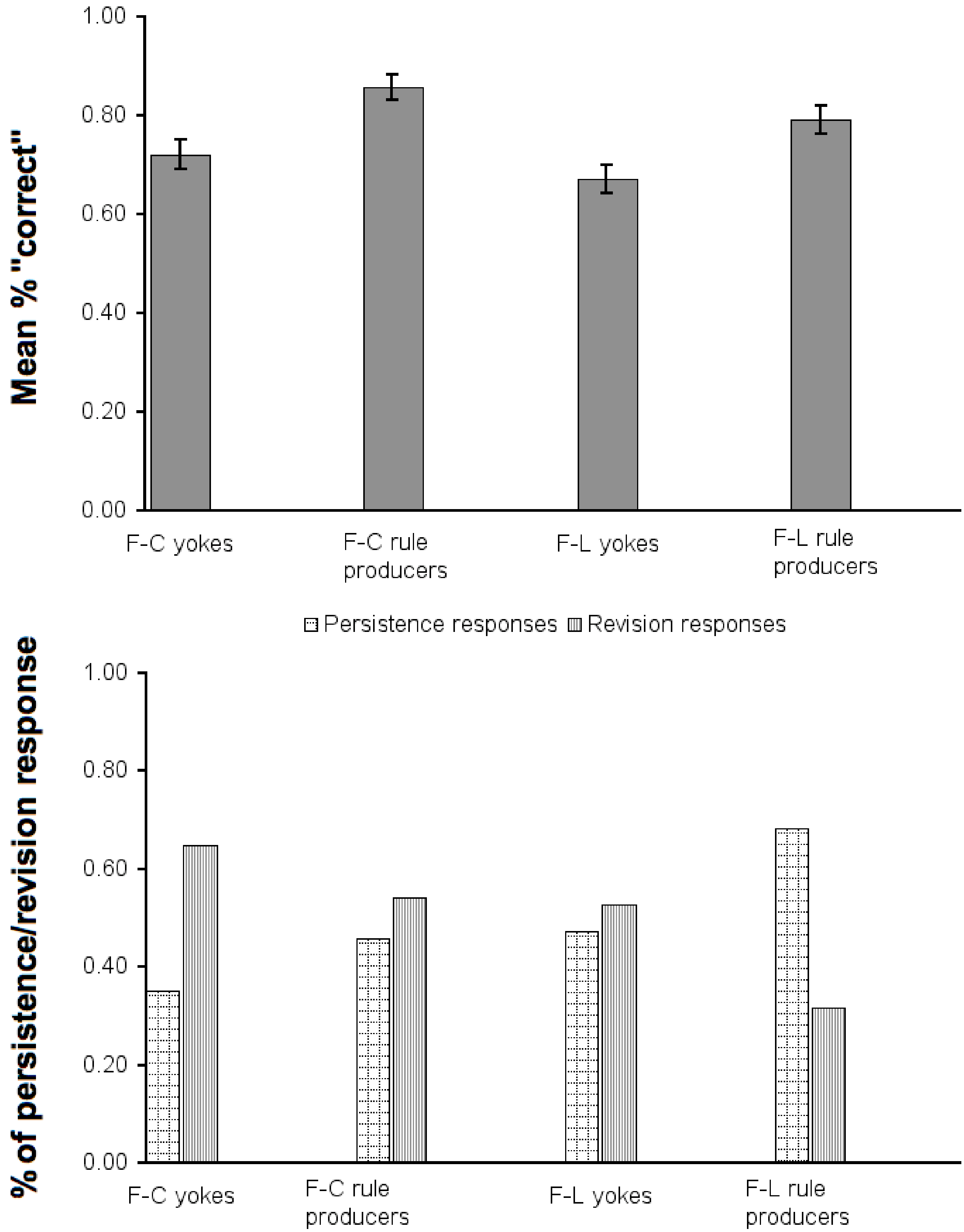

Familiar-overlap test
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
% of overlap response
Novel-overlap test
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
% of overlap response
Feature counting
Feature listing
Persistence
Revision
Interpretation
Overlap response categories

Figure 5


Mean % "correct"
1.00
0.80
0.60
0.40
0.20
0.00
F-C yokes
F-C rule producers
F-L yokes
F-L rule producers
Persistence responses
Revision responses
% of persistence/revision response
1.00
0.80
0.60
0.40
0.20
0.00
F-C yokes
F-C rule producers
F-L yokes
F-L rule producers