# Predicting category intuitiveness with the rational model, the simplicity model, and the Generalized Context Model

**Emmanuel M. Pothos**

Department of Psychology
Swansea University

**Todd M. Bailey**

School of Psychology
Cardiff University

in press: *JEP: LMC*

Please address correspondence to Emmanuel Pothos, Department of Psychology, Swansea University, Swansea SA2 8PP, UK, or to Todd Bailey, School of Psychology, Cardiff University, Cardiff CF10 3AT, UK. Electronic mail may be sent at e.m.pothos@swansea.ac.uk or baileytm1@cardiff.ac.uk

**Running head**: GCM category intuitiveness; **Word count (running text, excluding references)**: 8,721

## Abstract

Naïve observers typically perceive some groupings for a set of stimuli as more intuitive than others. The problem of predicting category intuitiveness has been historically considered the remit of models of unsupervised categorization. In contrast, this paper develops a measure of category intuitiveness from one of the most widely supported models of supervised categorization, the Generalized Context Model (GCM). Considering different category assignments for a set of instances, we ask how well the GCM can predict the classification of each instance on the basis of all the other instances. The category assignment that results in the smallest prediction error is interpreted as the most intuitive for the GCM—we call this way of applying the GCM unsupervised GCM. The paper then systematically compares predictions of category intuitiveness from the unsupervised GCM and two models of unsupervised categorization, the simplicity model and the rational model. We found that the unsupervised GCM compares favorably to the simplicity model and rational model. This success of the unsupervised GCM illustrates that the distinction between supervised and unsupervised categorization may have to be reconsidered. However, no model emerges as clearly superior, indicating that there is more work to be done in understanding and modeling category intuitiveness.


**Keywords**: supervised categorization, unsupervised categorization, exemplar theory, GCM.

# Introduction

The distinction between supervised and unsupervised categorization has been central to the development of categorization theory in cognitive science. Supervised categorization concerns predicting how novel instances will be classified, with respect to a set of existing categories; such predictions can be typically carried out with impressive accuracy. Prominent classes of supervised categorization models include the exemplar theory (Medin & Schaffer, 1978; Nosofsky, 1988; Van Vanpaemel & Storms, 2008), prototype theory (Hampton, 2000; Minda & Smith, 2000; Posner & Keele, 1968), and the general recognition theory (Ashby & Perrin, 1988). Supervised categorization typically involves training procedures with corrective feedback. By contrast, in a typical unsupervised categorization experiment participants are asked to divide some stimuli into categories which are intuitive, without any corrective feedback.

Interest in unsupervised categorization largely originates from the notion of category coherence (Murphy & Medin, 1985). Why do certain groupings of objects form psychologically intuitive categories, but other groupings are nonsensical? For example, most cultures have concepts such as happiness or animal. By contrast, a grouping which includes the Eiffel Tower, children under five, and apples would be considered entirely nonsensical. Murphy and Medin (1985) suggested that a category is coherent if it fits well with our overall understanding of the world; they argued that explanations based on similarity are inadequate (cf. Heit, 1997; Lewandowsky, Roberts, & Yang, 2006; Wisniewski, 1995). Unfortunately, creating categorization models on the basis of general knowledge is extremely difficult (e.g., Fodor, 1983; Pickering & Chater, 1995; but see Griffiths, Steyvers, & Tenenbaum, 2007 or Tenenbaum, Griffiths, & Kemp, 2006). Moreover, there is empirical evidence that people do use similarity in unsupervised categorization, at least in some cases. Accordingly, some researchers have developed unsupervised categorization models which are based on similarity (e.g., Compton & Logan, 1993; Love, Medin, & Gureckis, 2004; Milton & Wills, 2004; Pothos & Chater, 2002).

Unsupervised categorization involves two slightly separate problems: first, identifying the classification for a set of stimuli, which would be preferred by naïve observers. For example, in Figure 1, the preferred classifications for the dots are the ones indicated by the continuous curves (here and elsewhere, points represent objects and the axes are assumed to correspond to dimensions of some putative internal mental space; similarities are inversely related to distances). A second problem in unsupervised categorization is, given a classification for a stimulus set and another for a different stimulus set, deciding which one is more intuitive. In Figure 1, the classification on top should be perceived as more intuitive compared to the classification in the bottom panel (this is because the difference of within- versus between-category similarity in the top panel is higher than in the bottom; Pothos & Chater, 2002). In other words, if real stimuli are created after the Figure 1 points, participants are likely to identify the top classification as preferred more frequently and with more confidence, compared to the bottom classification. In principle, a model of category intuitiveness provides the basis for a model of unsupervised categorization, under the assumption that the most intuitive categorization will also be the preferred one.

--------------------------------------FIGURE 1--------------------------------------

A main objective of this work is to examine predictions of category intuitiveness from computational models of unsupervised categorization, for a range of stimulus sets. As far as we are aware, there has been no systematic investigation of this kind. This is an important shortcoming, given the strong intuitions we can have about which categorizations are more intuitive than others. We consider category intuitiveness predictions from the rational model (Anderson, 1991; Sanborn, Griffiths, & Navarro, 2006) and the simplicity model (Pothos & Chater, 2002). The inclusion of these two models has been partly motivated by the fact that they can readily produce a measure of category intuitiveness (this is not always the case with models of unsupervised categorization; e.g., see Compton & Logan, 1993).

Unsupervised and supervised categorization have typically been assumed to correspond to different psychological processes and the related research traditions have been mostly separate. A model is typically proposed as either a model of supervised categorization or a model of unsupervised categorization. However, this is an assumption which may be inappropriate. The other main objective of this paper is to show that a measure of category intuitiveness can be derived from one of the best known models of supervised categorization, Nosofsky's Generalized Context Model (GCM; Nosofsky, 1988, 1989, 1991, 1992). The version of the GCM which can produce predictions of category intuitiveness will be referred to as unsupervised GCM, to reflect the fact that, in this mode of application, the GCM assesses the intuitiveness of a particular classification instead of classifying new instances with respect to existing categories. Predictions of category intuitiveness from the GCM will be compared to those from the rational model and the simplicity model.

## Unsupervised GCM

The GCM predicts classification probabilities for a set of test stimuli based on their similarity to a set of previously seen training stimuli. The GCM is described by two equations:

$$P(A \mid X) = \frac{\beta_A \eta_{XA}}{\beta_A \eta_{XA} + \beta_B \eta_{XB}} \quad \text{..............................................(1a)}$$

$$\eta_{XA} = \sum_{j \in A} \exp\left\{ -c\left[ \left( \sum_{k=1}^{D} w_k \mid y_{xk} - y_{jk} \mid^r \right)^{1/r} \right]^q \right\} \quad \text{..................................(1b)}$$

P($A$|$X$) is the probability of making a category $A$ response, given instance $X$ (the $\beta$ terms are category biases and $\eta_{XA}$ is the sum of the similarities between $X$ and all the $A$ exemplars). This is Luce's (1963) choice rule; it sometimes involves an exponent to the similarities. In equation (1b), $c$ is a sensitivity parameter, $r$ is a Minkowski distance metric parameter, $q$ determines the shape of the similarity function, $w_k$ are dimensional attention weights, and $y$'s are item coordinates. The input to the GCM consists of the coordinates of a set of training stimuli, information about the assignment of the stimuli to categories, and the coordinates of a set of test stimuli. On the basis of this information, the parameters of the GCM are adjusted so as to predict as closely as possible empirically determined probabilities of how the test items are classified. An error term can be computed as $2 \cdot \sum O_i \ln \frac{O_i}{P_i}$, where $O_i$ are the target probabilities and $P_i$ the predicted probabilities from the model; the summation ranges over all the test items. Target probabilities typically correspond to how participants classify test items into training categories. This equation computes a likelihood ratio chi-square statistic (e.g., see Hahn, Bailey, & Elvin, 2005). We refer to this error term as a log likelihood error term.

How could the GCM compute (relative) category intuitiveness? Suppose that in the top panel of Figure 1 we want to evaluate the intuitiveness of classification {1, 2, 3,4, 5, 6}{7, 8, 9} versus {1, 2, 3}{4, 5, 6, 7, 8, 9}. In evaluating classification {1, 2, 3,4, 5, 6}{7, 8, 9}, we consider the GCM error term in predicting that items 1, 2, 3, 4, 5, 6, are in category {1, 2, 3,4, 5, 6} with 100% probability and likewise for items 7, 8, 9 and category {7, 8, 9}. In other words, exemplars are assigned to categories in accordance with the category structure being evaluated and GCM fits are computed on this basis. A main insight in this paper is that when the GCM self-classifies a set of stimuli in this way, the corresponding error term can be interpreted as a measure of category intuitiveness. We postulate that where the error term is lower, then the corresponding classification is more consistent with the assumptions about categorization ingrained in the GCM and that, therefore, such classifications are considered more psychologically intuitive by the GCM. For example, self-classifying the Figure 1 items relative to the classification {1, 2, 3,4, 5, 6}{7, 8, 9} should be associated with a very small error term, as the two categories are well separated. By contrast, self-classification relative to {1, 2, 3}{4, 5, 6, 7, 8, 9} should lead to a high error term. These results correspond to the obvious impression that, for the stimuli in Figure 1, classification {1, 2, 3,4, 5, 6}{7, 8, 9} is psychologically more intuitive than {1, 2, 3}{4, 5, 6, 7, 8, 9}.

This scheme constitutes a proposal for using the GCM to produce a measure of category intuitiveness (cf. Feldman, 2004; Pothos & Chater, 2002). We believe that given a measure of category intuitiveness one can create a full model of unsupervised categorization, but this is an objective for future work. One can ask what kind of category structures will be predicted as more intuitive by the GCM. For example, Feldman (2004) suggested that the Boolean complexity of concepts defined through logical expressions determines their psychological intuitiveness. Work on basic level categorization has assumed that category structure can be understood in terms of the ratio of within category similarity to between category similarity (Murphy, 1991; Murphy & Smith, 1982), a tradeoff in informativeness vs. specificity (Komatsu, 1992), or a tradeoff between cue and category validities (Jones, 1983). The corresponding claim for the unsupervised GCM is that an intuitive classification will be possible for a set of stimuli if there are groupings which maximize within category similarity, both with respect to the original representation of the stimuli, and the various transformations for this representation allowed by the GCM parameters (suppression of dimensions and stretching/ compression of psychological space). This latter characteristic particularly distinguishes the GCM from other unsupervised categorization models based on similarity. The unsupervised and supervised versions of the GCM are based on the same equations, but applied to answer different questions. In the former case, the computed error term is interpreted as category intuitiveness, in the latter case classification probabilities of novel instances are predicted. Crucially, in the unsupervised GCM parameters are not adjusted to match a particular pattern of empirical results (parameters are determined by item coordinates, so that parameter search is guided by a prerogative to achieve an intuitive classification), while in the supervised GCM parameters are specified so as to achieve particular probabilities for the classification of new instances.

So, our measure of category intuitiveness from the GCM is based on the same equations as the standard GCM (cf. Love, 2002). This is an important point, since it shows how a model which has been considered the hallmark of supervised categorization can be directly applied to unsupervised categorization. The specific details of how we applied the GCM are standard. The city block ($r$=1) and the Euclidean ($r$=2) metrics are the only metrics that have received psychological motivation, and likewise for the exponential ($q$=1) and Gaussian ($q$=2) forms of the similarity function. It has not been possible to motivate more specific values of $r$ and $q$ (Nosofsky, 1992), hence they were included as free parameters within these bounds.

Category biases were allowed to vary freely between zero and one, subject to the constraint that they summed to one, and likewise for the attentional weights. The sensitivity parameter, *c*, determines the extent to which the classification of an instance is influenced by remote exemplars or not. When *c* is very small, all exemplars will have an effect on how a test item is classified. As *c* increases in size, classification of a test item will be influenced primarily by its nearest neighbor amongst the training items, or, in a situation where the training items are the same as the test items, just by itself. This latter situation is pathological, so we required the unsupervised GCM to classify each stimulus on the basis of all the other stimuli in a stimulus set only. Given this requirement, in all our simulations the default approach was to allow *c* to vary freely between zero and infinity. We will later examine directly whether the unsupervised GCM can function adequately with an unrestricted *c*.

It is by no means obvious at the outset that our proposal will necessarily succeed. A common criticism for the GCM (and similar models) is that its parameters allow it too much flexibility in fitting empirical data (Olsson, Wennerholm, & Lyxzen, 2004; Myung, Pitt, Navarro, 2007; Navarro, 2007; Nosofsky, 2000; Nosofsky & Zaki, 2002; Smith, 2007; Smith & Minda, 1998, 2000; 2002; Yang & Lewandowsky, 2004). Accordingly, one can wonder whether our suggestion for the unsupervised GCM might fail because the GCM can perfectly describe any assignment of stimuli into categories (regardless of whether the corresponding classifications are more or less intuitive). The burden is on us to demonstrate that not only is this not the case, but that the unsupervised GCM can perform comparably with established models of unsupervised categorization.

## Rational model

Many models of unsupervised categorization (including the unsupervised GCM and the simplicity model) rely on similarity. It is interesting to include in the comparisons a model that makes no explicit reference to similarity. Anderson's (1991) rational model adopts a category utility approach. In other words, it assumes that categories are formed because they are useful to us, specifically because they allow us to infer unknown information about novel instances (cf. Corter & Gluck, 1992; Gosselin & Schyns, 2001; Jones, 1983; Medin, 1983; Murphy, 1982).

The rational model is an incremental, Bayesian (cf. Tenenbaum & Griffiths, 2001; Tenenbaum, Griffiths, & Kemp, 2006) model of unsupervised categorization. It assigns a new stimulus with feature structure *F* to whichever category *k* makes *F* most probable. For example, a new object with many features of a 'cat', would be assigned to the category of cats, since the feature structure of the object is most probable given this category membership.

We implemented the continuous version of the rational model, which assumes that stimuli are represented in terms of continuous dimensions (for more details see Anderson, 1991; Anderson & Matessa, 1992). The continuous version allows the most direct comparison with the unsupervised GCM, since the latter also assumes continuous dimensions. In the rational model, the probability of classification of a novel instance into category *k* depends on the product $P(k)P(F \mid k)$. P(*k*) is given by equation (3a):

$$P(k) = \frac{cn_k}{(1-c)+cn} \quad \text{(3a)}$$

In equation (3a), $n_k$ is the number of stimuli assigned to category *k* so far, *n* is the total number of classified stimuli, and *c* is a coupling parameter. The coupling parameter determines how likely it is that a new instance will be assigned to a new category.

Thus, $c$ indirectly determines the number of categories that the rational model will produce for a stimulus set. The probability that the new object comes from a new category is given by $P(0) = \frac{1-c}{(1-c)+cn}$. $P(F \mid k)$ is computed as in equation (3b):

$$P(F \mid k) = \prod_i f_i(x \mid k) \quad \text{.........................................................(3b)}$$

where $i$ indexes the different dimensions of variation of the stimuli and $x$ indicates the different values dimension $i$ can take. That is, $f_i(x \mid k)$ is the probability of displaying value $x$ on dimension $i$ in category $k$, and is approximated by $t_{a_i}(\mu_i, \sigma_i \sqrt{1+1/\lambda_i})$, which is the $t$ distribution with $a_i$ degrees of freedom. $\mu_i$ and $\sigma_i^2$ are given by equations (3c) and (3d).

$$\mu_i = \frac{\lambda_0 \mu_0 + n\bar{y}}{\lambda_0 + n} \quad \text{...........................................................(3c)}$$

$$\sigma_i^2 = \frac{\alpha_0 \sigma_0^2 + (n-1)s^2 + \frac{\lambda_0 n}{\lambda_0 + n}(\mu_0 - \bar{y})^2}{\alpha_0 + n} \quad \text{...................................(3d)}$$

*-3d- is the variance for classifying into the i dimension. This tells us how much it 'matters' whether a stimulus has a particular value on dimension i or not, for classification into a particular category.*

In these equations, $\lambda_i = \lambda_0 + n$, $\alpha_i = \alpha_0 + n$, $n$ is the number of observations in category $k$, $\bar{y}$ is their mean along dimension $i$, and $s^2$ is their variance. Finally, $\alpha_0 = 1 = \lambda_0$, $\mu_0$ is the halfway point of the range of all instances and $\sigma_0$ is the square of a quarter of the range (Anderson, personal communication).

It is possible to introduce a dimensional weighting mechanism in the rational model. In equation (3b), assume that the probability of having particular values along each dimension is weighted by $w1$, $w2$ etc., to indicate the relative importance of the dimensions in classifying a new item. In other words, $P(F|k) = P_1^{w1}(j|k) \cdot P_2^{w2}(j'|k)$ ...The question is what kind of weighting scheme is going to be optimal for the rational model. Taking logs in the above equation, we have: $\log(P(F|k)) = w1 \cdot \log(P_1(j|k)) + w2 \cdot \log\left(P_2(j'|k)\right)$ ... Suppose that $P_1 < P_2 \Rightarrow \log(P_1) < \log(P_2)$. Then, clearly, the weight combination which maximizes $P(F|k)$ is w1=0, w2=1. In other words, in the rational model, optimal dimensional weighting corresponds to assigning a weight of 1 to the most useful dimension and a weight of 0 to all the other dimensions (so that, in contrast to the GCM, graded weighting is never optimal for the rational model). Therefore, in the simulations below, where we refer to the 'rational model with dimensional selection', we assess the probabilities for the predicted classifications along all one-dimensional projections.

Note that the standard rational model can compute the probability for a classification, given a particular order of the items. However, all the empirical examples below assume concurrent presentation of the stimuli. Sanborn, Griffiths, and Navarro (2006) provided algorithms for the rational model, which approximate classification probabilities from the rational model, as if all items had been presented concurrently. Sanborn et al.'s examination of their algorithms was shown to both have desirable normative properties and outperform the standard rational model in specific empirical cases. Specifically, we used the Gibbs sampler algorithm to compute the probability for the most probable classification for a stimulus set. Moreover, we adapted the algorithm to compute the probability of any particular classification (not necessarily the most probable one) for a stimulus set. In either case, higher probabilities indicate that the corresponding classifications should be more intuitive.

## Simplicity model

The simplicity model of unsupervised categorization (Pothos & Chater, 2002, 2005; Pothos & Close, 2008) differs from the rational model and the unsupervised GCM in a number of interesting ways. First, the simplicity model is non-metric (a metric space is not assumed), while this is not the case for the other two models. Second, the simplicity model has no free parameters, a characteristic which contrasts most sharply with the unsupervised GCM. Third, the simplicity model aims to maximize within category similarity and minimize between category similarity, but only the former constraint is relevant to the unsupervised GCM. Finally, the simplicity model is currently the only model which has been applied to data from entirely unconstrained categorization procedures; it is therefore interesting to compare it with the rational model and the unsupervised GCM against such data.

According to the simplicity model, more intuitive categories are ones that maximize within category similarity and minimize between category similarity (cf. Rosch and Mervis, 1975). The model is specified within a computational framework based on the simplicity principle (Chater, 1996, 1999). The first step is to compute the information content of the similarity structure of a set of items without categories. This is done by assuming that every pair of stimuli is compared with every other pair. For example, suppose that we have four stimuli, labeled by 1,2,3,4. Then, similarity information would be encoded as similarity(1,2)>similarity(1,3), similarity(1,2)<similarity(1,4), etc., with each comparison requiring one bit of information to specify whether the first pair is more similar or less similar than the second (assuming no exact equalities).

Categories are defined as imposing *constraints* on the similarity relations between pairs of stimuli; similarities within categories are assumed to be greater than all similarities between categories. For example, suppose that we decide to place stimuli 1,2 in one category and stimuli 3,4 in a different category. Then, our definition of categories implies that similarity(1,2)>{similarity(1,3), similarity(1,4)} and that similarity(3,4)>{similarity(1,3), similarity(1,4)}. Thus, the *codelength* for the similarity structure for a set of stimuli can be reduced by using categories, if the constraints specified by the categories are numerous and, generally, correct (note that equalities in similarity relations do not falsify the constraints; Hines, Pothos, & Chater, 2007). If in *u* constraints there are *e* incorrect ones, the number of bits of information required to correct the errors is given by equation (2a).

$$\log_2(u+1)+\log_2\left(\frac{u!}{e!(u-e)!}\right) \quad \text{(2a)}$$

Moreover, we have to take into account the information-theoretic cost of specifying a particular category structure of *n* categories for *r* objects, which is given by $\log_2(Part(r,n))$, where *Part*(*r*,*n*) is given by equation (2b).

$$Part(r,n)=\sum_{v=0}^{n-1}(-1)^v\frac{(n-v)^r}{(n-v)!v!} \quad \text{(2b)}$$

Overall, there is a codelength without categories and a codelength with categories. The ratio of the latter to the former indicates how much codelength *reduction* is afforded by the use of categories; it is typically reported as a percentage and referred to as just ‘codelength’. The lower its value, the more intuitive a particular category structure is predicted to be. The lowest possible value of codelength is about 50%. When trying to identify the most intuitive classification from scratch, the simplicity model employs a search algorithm akin to those in agglomerative clustering procedures (Hines et al., 2007).

## Analyses

Our analyses are divided in three parts. First, all three models are examined with a simple toy stimulus set. For the rational model and the simplicity model, this exercise illustrates the way they are applied and some basic implementational assumptions. Regarding the unsupervised GCM, this exercise is more important, since it corresponds to a preliminary test of whether the model can capture some obvious intuitions about category intuitiveness. Second, we examined a range of classic stimulus sets from the supervised categorization literature, on the assumption that category learnability is related to intuitiveness. Third, we considered data from studies which employed an entirely unsupervised categorization procedure.

### Toy stimulus set/ illustration of the models' operation

Four stimulus sets were created to assess the three models with respect to the straightforward intuition that well-separated, coherent categories should be more intuitive than less-separated ones. Each stimulus set was intended to correspond to a category structure composed of two clusters. The stimulus sets differed on how close the two clusters were to each other, with category prototypes being 2, 3, 4, or 5 units apart. The two most extreme stimulus sets are shown in Figure 2; the other stimulus sets were in between these extremes.

Unsupervised GCM intuitiveness values were obtained as log likelihood error terms, which reflect the deviance between the intended assignment of stimuli into categories and the predicted assignment by the GCM. A lower error term implies that the corresponding classification is considered more intuitive. Constrained optimization of the GCM parameters was done with the fmincon Matlab function (version R2007b). We examined the log likelihood error term for a particular classification at least 50 times using different random initial parameter values (the lowest error term was taken to be the intuitiveness value from the unsupervised GCM). To facilitate comparisons with other models, we normalized the log likelihood values for all category structures onto a 0-1 scale (with 0 corresponding to least predicted intuitiveness and 1 corresponding to the greatest intuitiveness), through the transformation $1-\frac{X-\min}{\max-\min}$, where *X* is the log likelihood error for any of the four category structures, *min* is the least log likelihood error (of these four values), and *max* is the greatest error. By carrying out this (or analogous) transformation for the predictions from all models, we can derive an impression of how the models compare with each other. The normalized scores have been used in all figures, raw model predictions in the tables.

Simplicity model predictions were given as codelength values, so that a lower codelength indicates a more intuitive classification. Codelength values typically range between 50% and 100%. Recall that the input to the simplicity model is not item coordinates, but rather information of which pairs of similarities are greater or smaller than others. In order to derive such information from item coordinates, a distance metric has to be assumed. Consistently with previous examinations of the simplicity model (e.g., Pothos & Chater, 2002, 2005), we opted for the Euclidean metric. The Euclidean metric is an appropriate default choice, since it corresponds better to the way physical distances are perceived psychologically. Of the models considered, the simplicity model was the most straightforward to run, requiring less than a minute per stimulus set. Simplicity values were transformed onto a 0-1 scale as above.

Using Sanborn et al.'s (2006) adaptation of the rational model, it is possible to identify the best classification for a set of items and compute the probability for any particular classification—as noted, these probabilities can be interpreted as predictions for category intuitiveness. The algorithm was run with a different number of iterations for different stimulus sets (at least 10,000), with a view to ensure that not more than 10 hours were required per stimulus set. Sample spacing was set to 20. For the coupling parameter we employed the commonly used value of 0.5. Finally, for the rational model with dimensional selection, intuitiveness values corresponded to the most probable classification regardless of whether all dimensions or a particular dimension were employed.

All models correctly predicted that category structures for which the two categories are closer together should be less intuitive, compared to category structures for which the categories are further apart (Table 1). This is hardly an exciting prediction, but nonetheless an important basic test that the models are consistent with expectations in such an intuitive case. Note that the rational model with dimensional selection correctly predicts that the optimal dimension in all cases is dimension 1 (in other words, the probability of the best possible classification along dimension 1 is greater than the corresponding probability along either dimension 2 or both dimensions). In this straightforward case, there is agreement between the rational model and the rational model with dimensional selection. Regarding the simplicity model, the lowest possible codelength in this case is 51.6; as noted, the exact value will somewhat depend on the particular classification. Also, the worst possible codelength is 117.9, well over 100. This reflects the fact that when the prototypes are only two units apart, the costs associated with correcting errors in the constraints specified by the classification are so high, that we are actually better off describing the similarity information without categories.

The behavior of the models can be seen in Figure 3, where each of the model measures has been transformed on a 0 to 1 scale. While such a transformation involves some arbitrary assumptions, it provides a visually intuitive means of quickly appreciating model similarities and differences (in this case, for example, the fact that the unsupervised GCM and the simplicity model rise quickly to their highest value, while the rational model's rise is more gradual).

-------------------------------------FIGURES 2,3, TABLE 1----------------------------------

**Supervised categorization data**

Supervised categorization data can be used to derive estimates of category intuitiveness in two ways. First, we assume that if classification A is more difficult to learn than classification B, then, in an unsupervised context, classification A will be considered more intuitive compared to B. The empirical evidence supports this assumption. Colreavy and Lewandowsky (in press) found no difference between a supervised categorization condition and a matched unsupervised one, in terms of strategy development and rate of learning (see also Griffiths, Christian, & Kalish, 2008). Love (2002) reached the opposite conclusion, but his supervised and unsupervised stimulus sets were not directly comparable, and the learning task was not entirely equivalent to an unsupervised categorization one.

Second, consider categories A and B and a new instance X. Suppose that participants are more likely to classify X into category A than B. Since participants classified X with category {A} rather than {B}, they must think that the overall grouping {A,X}{B} must be more intuitive than the alternative grouping {A}{B,X}. Therefore, we can assume that classification {A,X}{B} is more intuitive than {A}{B,X}. We can then examine whether the unsupervised categorization models consider{A,X}{B} as more intuitive compared to {A}{B,X}. Note that the application of the unsupervised GCM to such data is very different from the standard application of the GCM, where the objective is to predict classification probabilities for the new instances.

Finally, methodologically, one can ask whether the categories employed in supervised categorization research may be so unstructured that they would never be created in a spontaneous fashion. But, this is not a problem since the difference in relative intuitiveness between two classifications, however unstructured, can always be empirically examined: participants' spontaneous classifications should be more *similar* to the one which is predicted to be more intuitive.

***Shepard, Hovland, and Jenkins (1961)*.** Shepard et al. (1961) considered the difficulty of learning six binary classifications with stimuli made of three binary dimensions (Table 2). Classification 1 is simple to learn because it covaries perfectly with the first dimension of the stimuli. Classification 2 reflects an 'exclusive OR' (non linear) category structure in its first two dimensions, while the third dimension constitutes random noise. Classifications 3, 4, and 5 can be described by one-dimensional rules with exceptions and require attention to all three dimensions. Classification 6 also requires attention to all three dimensions, but in this case there are no obvious regularities. Shepard et al. reported that the cumulative error rate conforms to the following ordering: Classification 1 (easiest) < Classification 2 < {Classifications 3, 4, 5} < Classification 6 (most difficult). This result has become a benchmark for assessing models of supervised categorization (e.g., Love et al., 2004; Kurtz, 2007; Nosofsky & Palmeri, 1996).

------------------------------------TABLES 2, 3, FIGURE 4---------------------------------

As discussed, we assumed that more intuitive classifications should be easier to learn (the results of Griffiths et al., 2008, support this assumption in the case of the Shepard et al. data, with a kind of unsupervised induction task). The unsupervised GCM, the simplicity model, and the rational model were applied by computing the log likelihoods, codelength values, and classification probabilities, respectively. The raw results are shown in Table 3 and normalized predicted intuitiveness scores are shown in Figure 4. The unsupervised GCM performed better than both the simplicity model and the rational model. The unsupervised GCM correctly predicted that Classifications 1, 2 should be the most intuitive, 3,4,5 of intermediate intuitiveness, and, finally, that Classification 6 should be the least intuitive. Note that Classifications 1 and 2 are not distinguished (the former should be more intuitive than the latter). The simplicity model considered all classifications highly unintuitive; the codelength values produced were very close to 100, predicting that participants receiving these stimulus sets would be unlikely to spontaneously produce the Shepard et al. classifications. The model does predict that Classification 1 should be the most intuitive and Classification 6 the least intuitive one. However, the simplicity model was confused by Classification 2, which was predicted to be less intuitive than 3,4,5. The rational model had a similar problem: as with the simplicity model, the rational model correctly predicted Classifications 1 and 6 to be the most and least intuitive, respectively; however, it incorrectly predicted Classification 2 to be less intuitive than Classifications 3,4,5. The same pattern of results was predicted by the rational model with dimensional selection, even though the optimal dimension varied in different cases. To sum up, with the Shepard et al. data, the unsupervised GCM outperformed the models of unsupervised categorization.

***5-4 category structure***. The 5-4 category structure (Medin & Schaffer, 1978) has been extensively explored in the context of the debate between prototype and exemplar theory (e.g., Johansen & Kruschke, 2005; Nosofsky, 2000; Smith & Minda, 2000; but see Homa, Proulx, & Blair, 2008). Medin and Schaffer (1978) reported classification probabilities for seven test items (Table 4) and, as discussed, we can use these probabilities to *infer* the intuitiveness of different classifications. For each categorization model, two computations were made for each test item: One with the test item assigned to the first category and another with the test item assigned to the second category. These two computations corresponded to two intuitiveness values for each item. The difference between these two values should correspond to the classification probabilities reported by Medin and Schaffer (1978).

-------------------------------------TABLES 4, 5, FIGURE 5----------------------------------

The results are shown in Table 5 and Figure 5. To create Figure 5, for the unsupervised GCM we computed the difference in log likelihood error for the classification when the first test item was assigned to the first category minus the log likelihood error for the classification when the first test item was assigned to the second category; and likewise for the other test items. (In other words, we subtracted the values in each of the cells in Table 5.) Subsequently, these differences were converted onto a uniform scale, as in the other examples (in this case, the scale was 0—2; model predictions corresponded to differences between two intuitiveness values, and such differences could be negative). An analogous procedure was adopted for the other models. Correlating classification probabilities and the differences in predicted intuitiveness values, for the unsupervised GCM, the simplicity model, the rational model, and the rational model with dimensional selection respectively, we obtained: -.857, -.960, .068, -.412. Note that a negative correlation is in the predicted direction for the unsupervised GCM and the simplicity model, since for these models *lower* values (lower error or lower codelength) correspond to *more* intuitive classifications and, hence, should be associated with higher classification probabilities in Medin and Schaffer's data. The GCM and the simplicity model competently describe the Medin and Schaffer (1978) data; however, the rational model had difficulty discriminating between the (assumed) less and more obvious classifications.

***Linear separability***. A classification is linearly separable if a straight line (or the equivalent in more than two dimensions), can divide all the items which belong to one category from all items which belong to another. Linear separability is an important consideration in categorization, since exemplar theory is consistent with non-linearly separable categories, but this is not the case for prototype theory. The empirical results have been somewhat ambiguous (Ashby & Maddox, 1992; Kalish, Lewandowsky, & Kruschke, 2004; Kemler Nelson, 1984; Kemler Nelson, 1984; Olsson, Enkvist, & Juslin, 2006; Medin & Schwanenflugel, 1981; Ruts, Storms, & Hampton, 2004; Shepard et al., 1961; Smith, Murray, & Minda, 1997; Wattenmaker, 1995). The latest research with schematic stimuli indicates that linearly separable categories are more intuitive (Blair & Homa, 2001). Of the category structures Blair and Homa used, most relevant are the ones with four categories each, in which each category had nine points (these were the largest stimulus sets). The linearly separable category structure was referred to by Blair and Homa as LS9 and the non-linearly separable one as NLS9. Blair and Homa reported an advantage of the LS9 classification relative to the NLS9 one, in terms of ease of learning.

We modeled the LS9 vs. NLS9 contrast reported by Blair and Homa (2001). Each LS9 category was based around a prototype and nine ‘high distortion’ items from the prototype. Each NLS9 category was based around the same prototypes, six ‘low distortion’ items from the prototype, and one low distortion item from each of the other three prototypes (the items from the other prototypes result in non-linearity). Blair and Homa reported the coordinates of six items from each prototype in three dimensions, which were derived from similarity ratings (Blair, personal communication). To approximate the Blair and Homa stimulus sets, the coordinates of the six items from each prototype were averaged to infer the coordinates of the prototype. Then, we computed the average distance between the prototypes and low/high distortion items, and so created enough low/high distortion items to approximate the LS9 and NLS9 category structures (Appendix). We created two more extreme stimulus sets, referred to as LS9X and NLS9X, in which the prototype coordinates were changed so that the least distance between any two prototypes would be at least 1.5 times the distance between a prototype and a high distortion item. The LS9, NLS9 stimulus sets can only be said to *approximate* the actual Blair and Homa stimulus sets. Therefore, we examined linear separability of each stimulus set with a series of logistic regressions attempting to predict category membership (Ruts et al., 2004). Our re-creation of LS9 is *not* linearly separable (probably because categories are too close to each other), but it is a lot more so compared to NLS9. Moreover, the new stimulus set LS9X is linearly separable.

-------------------------------------TABLE 6, FIGURE 6---------------------------------

The empirical finding we aimed to model was that LS9 was easier to learn compared to NLS9. Although there are no relevant empirical results for LS9X and NLS9X, we tentatively assume that for LS9X increasing the distance between category prototypes would make a category structure more salient (a straightforward assumption), but this would not be the case for NLS9X (a more controversial assumption). The results are shown in Table 6 and Figure 6. The unsupervised GCM and the simplicity model successfully predicted a difference between LS9 and NLS9 and a more pronounced difference between LS9X and NLS9X. Note that using the unsupervised GCM to predict category intuitiveness is a different computation from that corresponding to the standard GCM and, so, our results do not bear on the fact that Blair and Homa (2001) could not identify satisfactory (standard) GCM fits for the classification probabilities of their test items. According to the rational model, all classifications were extremely unlikely, and maybe this obscured any finer differences due to linear separability.

**Unsupervised categorization data**

***Compton & Logan (1999)***. Compton and Logan (1999) reported extensive data on judgments of category intuitiveness, in an entirely unsupervised categorization task. They presented participants with diagrams of dots, as in Figure 1 (but without any curves), and asked participants to classify the dots in a way that seemed immediately natural and intuitive, by drawing curves to indicate their groupings. There were no constraints at all as to how the items should be classified (including no constraints on the number of groups; cf. Murphy, 2004). Compton and Logan measured category intuitiveness in terms of classification variability, that is, the number of unique classifications for each diagram shown to participants, so as to examine whether classification variability changed when the arrangement of dots in a diagram was transformed (e.g., reflected or rotated). Compton and Logan employed 48 unique diagrams, which consisted of 12 examples for each numerosity of dots from 7 to 10, inclusive (each participant classified 144 diagrams, the 48 original ones and various transformations of them). Each unique diagram was created by randomly arranging the appropriate number of dots in a 40x40 grid.

The categorization procedure of Compton and Logan somewhat deviates from the standard procedure in unsupervised categorization experiments: participants drew lines around points, instead of classifying stimuli as separate entities. For example, the classification of dots in a diagram will be affected by the nearest neighbor structure in the diagram; participants would rarely classify in the same cluster dots which are far away from each other. By contrast, in standard unsupervised classification tasks, where participants receive stimuli as separate entities, classification of highly dissimilar items into the same group is sometimes observed. However, Compton and Logan (1999) is currently the most extensive report of unsupervised categorization results. Moreover, perceptual grouping processes in Compton and Logan's experiment is arguably very similar to the grouping by similarity, postulated by models such as simplicity and the GCM: in both cases, the assumption is that participants will prefer groupings which enhance within category similarity. Finally, some researchers have argued that such dot diagrams is a valid way to study unsupervised categorization (Pothos & Chater, 2002).

Compton and Logan only reported the diagrams for which they observed the two highest and two lowest classification variability values in each of their two experiments; we read off the item coordinates from the diagrams (Appendix), so as to examine whether the predictions of the unsupervised categorization models are consistent with the highest/ lowest classification variability results reported by Compton and Logan: there should be less classification variability for stimulus sets for which the models can identify more intuitive classifications.

-------------------------------------TABLE 7, FIGURE 7---------------------------------

The simplicity model and the rational model can identify the best possible classification for a set of items. In the case of the simplicity model, we employed the agglomerative search algorithm described in Pothos and Chater (2002) and for the rational model the Gibbs sampler algorithm in Sanborn et al. (2006), which computes the most probable classification for a set of items, in a way that approximates concurrent presentation of the items.

Regarding the unsupervised GCM, we have no algorithm to identify the preferred classification from scratch. Therefore, we examined the log likelihood error term for the preferred classifications identified by the simplicity model, the rational model with dimensional selection, and K-means two- and three-cluster algorithms (excluding some all-inclusive categories identified by the rational model, since such categories are pathological for the unsupervised GCM); the intuitiveness prediction from the GCM for a stimulus set corresponded to the lowest identified log likelihood error term overall. In this case, we also examined a modification for the unsupervised GCM, for which the sensitivity parameter was fixed to a constant value (we chose $c$=0.5, noting that which value of $c$ is suitable will depend on the coordinate units). Why is this consideration relevant in this case, but not in the case of the stimulus sets from supervised categorization studies? In unsupervised research, the classifications considered are typically chosen to be intuitive to naïve observers. Accordingly, without a restriction on $c$, the GCM can always stretch the representational space in such a way that the corresponding classification is maximally intuitive. Psychologically, by restricting the sensitivity parameter, we suggest that participants spontaneously classify an item not just by considering its single nearest neighbor, but in relation to many of the other items as well (this seems highly plausible in the case of the Compton & Logan results, and also the Pothos & Chater results, considered next).

The results are shown in Table 7 and Figure 7. Note first that the standard unsupervised GCM is unable to discriminate between the intuitive and unintuitive stimulus set, but this is not so when a restriction in the sensitivity parameter is introduced. We next correlated the classification variability results reported by Compton and Logan with the intuitiveness values generated from each model. These (Pearson) correlations were for the unsupervised GCM, .319, the unsupervised GCM with a $c$ restriction, .407, the simplicity model, .430, the rational model, -.032, and the rational model with dimensional selection, -.902 (all in the right direction, noting that higher values from the unsupervised GCM and the simplicity model correspond to *less* intuitive categories, but higher values from the rational model to *more* intuitive categories). Of these correlations, the one involving the rational model with dimensional selection probabilities was the highest, showing that allowing dimensional selection in the rational model is a key modification regarding the model's explanatory power. Moreover, the unsupervised GCM with *c* restricted performs better than the unrestricted GCM. Specifically, it correctly identifies the four low variability stimulus sets as more intuitive than all the four high variability ones.

***Pothos and Chater (2002)***. Pothos and Chater examined four 10-item stimulus sets for which there were differing intuitions about the most intuitive classification (Appendix). In the first one, there were two well-separated clusters of equal size (referred to as 'two clusters'). In the second one, there were also two well-separated clusters, but one was larger than the other (referred to as 'big small'). In the third stimulus set there were three well-separated clusters ('three clusters'). Finally, there was little classification structure in the last stimulus set (referred to as 'little').

We consider Experiment 2 of Pothos and Chater, whereby item coordinates were mapped onto (separable) dimensions of physical variations, to create stimulus pictures that were printed on separate sheets of paper and given to participants to be sorted into groups that were "intuitive and natural". No constraints were imposed on participants' classifications (e.g., participants could use as many clusters as they liked, they could see the stimuli in any order or way they liked, and they could make changes in their classifications). This represents the most naturalistic unsupervised categorization study we found in the published literature. Pothos and Chater employed 28 participants and measured classification performance in terms of three indices, the number of distinct classifications (a higher value implies less participant agreement), the number of robust distinct classifications (these are the distinct classifications with a frequency greater than one), and the frequency with which the best possible classification was produced. Because of the small sample, these three measures did not correlate very well with each other; Pothos and Chater considered the last two as the most valid. We derived two separate rank orderings for the four stimulus sets from these two measures, which we subsequently added together to obtain an overall rank ordering for the observed intuitiveness of different stimulus sets. For the ‘two clusters’, ‘big, small’, ‘three clusters’, and ‘little’ stimulus sets, the summation of the ranks for these two measures produced 3, 2, 5, and 7 respectively, whereby a lower number indicates higher category intuitiveness.

--------------------------------------TABLE 8, FIGURE 8----------------------------

Unsupervised GCM category intuitiveness predictions were computed for the classifications predicted by the simplicity model, the rational model with dimensional selection, and K-means two-cluster and three-cluster algorithms. As before, we explored the version of the unsupervised GCM with and without restricting the sensitivity parameter; for the rational model we employed the Sanborn et al. (2006) algorithms. The simplicity model was applied to the stimulus sets by searching for the best possible classification on the basis of item coordinates. The results are shown in Table 8 and Figure 8. The simplicity model and the restricted unsupervised GCM accurately predicted the ordinal ordering of empirical classification intuitiveness in Table 8, but this was not the case for the unrestricted unsupervised GCM (which failed to discriminate between any of the stimulus sets). Finally, the rational model with dimensional selection was in much closer correspondence to the empirical results than the standard version.

**Conclusions**

Naïve observers can often have very compelling intuitions that a particular grouping for a set of stimuli may be more appropriate than another. Therefore, understanding the computational basis for such intuitions appears an important objective for models of unsupervised categorization. One aim of this paper was to examine predictions about category intuitiveness, from computational models of categorization, against a series of studies from the categorization literature.

Models of unsupervised categorization which can readily produce a measure of category intuitiveness are the rational model and the simplicity model and so these two models were tested in our analyses. Future work could fruitfully include additional models, such as Schyns' (1991) self-organizing neural network, which was used to model category emergence, Compton and Logan's (1993, 1999) perceptual grouping approach to unsupervised categorization, or Love et al.'s (2004; Gureckis & Love, 2002) SUSTAIN model, which assumes two slightly separate mechanisms for supervised and unsupervised categorization (respectively, an explicit error term and surprisingness with a principle of similarity). Finally, there has been an extensive literature on statistical clustering (e.g., Fisher and Langley, 1990; Krzanowski & Marriott, 1995), which looks relevant to studies of unsupervised categorization. Such models could serve as psychological models of categorization, after some additional theoretical elaboration.

Another aim of this paper was to explore the possibility that a measure of category intuitiveness could be derived from a supervised model of categorization, the GCM. In our adaptation of the GCM, a candidate classification for a set of items is examined by considering how well the intended classification of each item can be predicted on the basis of all the other items. A log likelihood error term can thus be computed, which indicates the ability of the GCM to describe the candidate assignment of items to categories. We postulated that when this error term is lower, then the corresponding classification is more consistent with the assumptions of the GCM about categorization, so that such a classification would be predicted (by the GCM) to be more psychologically intuitive. Both the supervised and unsupervised GCM are based on exactly the same equations, but are applied differently. The unsupervised GCM computes a number which can be interpreted as category intuitiveness, while the supervised GCM predicts classification probabilities for novel instances. Crucially, in the unsupervised GCM no parameter fitting is taking place relative to empirical data (parameters are searched so as to identify the best possible classification for a set of stimuli), while in the supervised GCM parameters are specified so as to achieve particular probabilities for the classification of new instances.

The unsupervised GCM favors groupings of items that maximize within category similarity. The crucial difference between the unsupervised GCM and models such as the simplicity one is that similarity groupings are assessed not just against the initial/ unprocessed dimensional representation of the items, but against all possible derivative representations, which would be forthcoming from dimensional weighting, stretching/ compression of psychological space etc. This representational flexibility is, of course, the hallmark of GCM predictions and its characteristic which has allowed it to provide impressive fits to empirical data. It is also a characteristic that has provoked some criticism since, if the GCM is proved to be too flexible, then its explanatory power would be limited. Accordingly, a possible way in which our demonstration could have failed would be if the unsupervised GCM could predict every arbitrary assignment of items into categories to be perfectly intuitive. However, this was not the case, and the accuracy of the predictions from the unsupervised GCM compared favorably from those of the rational model and the simplicity model.

One can ask whether the unsupervised GCM is meant to be understood as a full model of unsupervised categorization. For a full model of unsupervised categorization what is needed is a criterion of category intuitiveness and a search algorithm which can use this criterion to identify the optimal classification for a set of stimuli from scratch. The unsupervised GCM fulfills only the first requirement. It can be used to compute the predicted category intuitiveness for a stimulus set, a prediction which can be compared with the ones from the rational model and the simplicity model. However, the current formulation of the unsupervised GCM fails the second requirement. To appreciate why this is the case, consider first how the simplicity model (equally for the rational model) works. The simplicity model can easily identify the predicted most intuitive classification for a set of stimuli from scratch, with simple search algorithms which take the stimulus configuration and identify the classification which best optimizes the model's criterion for category intuitiveness. For the unsupervised GCM, the problem is that there is no single stimulus configuration, but rather an infinite number of possible ones, defined by stretching/ compressing psychological space or different relative attentional weighting of the item dimensions (in other words, the situation is like having different stimulus sets, the original one and all possible transformations of the original one, as allowed by the GCM parameters). Thus, in the case of the unsupervised GCM the search space is much more extensive, making optimization of its criterion for category intuitiveness intensive and difficult (so that, for example, the straightforward agglomerative algorithm which works for the simplicity model will not work for the unsupervised GCM). With future work, we hope to address this difficulty.

Regarding the results of our simulations, all models performed reasonably well, but no model could be identified as clearly superior when compared to the others. More work needs to be done in order to model category intuitiveness in a satisfactory way. For example, with the Shepard, Hovland, Jenkins (1961) data, the unsupervised GCM performed better than both the simplicity model and the rational model. In the case of the 5-4 category structure (Medin & Schaffer, 1978), the simplicity model came out ahead, with the unsupervised GCM providing the second best fit. In the case of comparing Blair and Homa's (2001) LS/ NLS category structures, the simplicity model and the unsupervised GCM could both provide a perfect account of the empirical findings. Compton and Logan (1999) provided one of the early studies with an entirely unsupervised categorization procedure. The best description for their results was from the rational model with dimensional selection. The unrestricted unsupervised GCM was too powerful for this data. It was necessary to constrain the sensitivity parameter before the unsupervised GCM could accurately predict an intuitiveness difference between the low and high variability stimulus sets (cf. Stewart & Brown, 2005; Olsson et al., 2004). Finally, the restricted unsupervised GCM and the simplicity model could account for Pothos and Chater's (2002) data, and the results from the rational model with dimensional selection were in close correspondence too.

The relative success of the unsupervised GCM calls into question the distinction between supervised and unsupervised categorization which has dominated the literature. We showed that a model of supervised categorization could be straightforwardly adapted to make predictions about category intuitiveness. Also, the converse situation is implied in our demonstration: models of unsupervised categorization were used to describe empirical results from supervised categorization studies, by assuming that when a classification is more difficult to learn then it should be less intuitive (cf. Colreavy and Lewandowsky, in press). It is therefore possible that both supervised and unsupervised categorization could be described within the same mathematical framework (e.g., the GCM), noting, however, that behavioral or neuroscience data may show these to correspond to distinct psychological processes (e.g., cf.  Ashby & Ell, 2002; Ashby & Perrin, 1988; Nomura et al., 2007; Zeithamova & Maddox, 2006; see also Love et al., 2004). Future work will hopefully address these exciting issues, as well as extend formalisms like the rational model and the simplicity model to provide more complete fits to category intuitiveness data.

**Acknowledgments**

This research was partly supported by ESRC grant R000222655 to EMP. We would like to thank Peter Hines, Mark Johansen, Matt Jones, Rob Nosofsky, Amotz Perlman, David Smith, and Safa Zaki. We are grateful to John Anderson for his help with the implementation of the continuous version of the rational model, to Adam Sanborn for helping with the Gibbs sampler, and to Mark Blair for his suggestions for specifying the LS/ NLS stimulus sets, from the Blair & Homa study.

**References**

Anderson, J. R. (1991). The Adaptive Nature of Human Categorization. Psychological Review, 98, 409-429.

Anderson, J. R. & Matessa, M. (1992). Explorations of an incremental, Bayesian algorithm for categorization. Machine Learning, 9, 275-308.

Ashby, G. F. & Perrin, N. A. (1988). Towards a unified theory of similarity and recognition. Psychological Review, 95, 124-150.

Ashby, G. F. & Maddox, W. T. (1992). Complex decision rules in categorization: contrasting novice and experienced performance. Journal of Experimental Psychology: Human Perception and Performance, 18, 50-71.

Ashby, G. F. & Ell, S. W. (2002). Single versus multiple systems of category learning: Reply to Nosofsky and Kruschke (2002). Psychonomic Bulletin & Review, 9, 175-180.

Blair, M. & Homa, D. (2001). Expanding the search for a linear separability constraint on category learning. Memory & Cognition, 29, 1153-1164.

Chater, N. (1996). Reconciling Simplicity and Likelihood Principles in Perceptual Organization. Psychological Review, 103, 566-591.

Chater, N. (1999). The Search for Simplicity: A Fundamental Cognitive Principle? Quarterly Journal of Experimental Psychology, 52A, 273-302.

Colreavy, E. & Lewandowsky, S. (in press). Strategy development and learning differences in supervised and unsupervised categorization. Memory & Cognition.

Compton, B. J. & Logan, G. D. (1993). Evaluating a computational model of perceptual grouping. Perception & Psychophysics, 53, 403-421.

Compton, B. J. & Logan, G. D. (1999). Judgments of perceptual groups: Reliability and sensitivity to stimulus transformation. Perception Psychophysics, 61, 1320-1335.

Corter, J. E. & Gluck, M. A. (1992). Explaining Basic Categories: Feature Predictability and Information. Psychological Bulletin, 2, 291-303.

Feldman, J. (2000). Minimization of Boolean complexity in human concept learning. Nature, 407, 630-633.

Fisher, D., & Langley, P. (1990). The structure and formation of natural categories. In Gordon Bower (Ed.), The Psychology of Learning and Motivation, Vol. 26 (pp. 241-284). San Diego, CA: Academic Press.

Fodor, J. A. (1983). The modularity of mind. Cambridge, MA: The MIT press.

Gosselin, F. & Schyns, P. G. (2001). Why do we SLIP to the basic-level? Computational constraints and their implementation. Psychological Review, 108, 735-758.

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. Psychological Review, 114, 211-244.

Griffiths, T. L., Christian, B. R., & Kalish, M. L. (2008). Using category structures to test iterated learning as a method for identifying inductive biases. Cognitive Science, 32, 68-107.

Gureckis, T.M., Love, B.C. (2002). Who says models can only do what you tell them? Unsupervised category learning data, fits, and predictions. In Proceedings of the 24th Annual Conference of the Cognitive Science Society. Lawrence Erlbaum: Hillsdale, NJ.

Hahn, U., Bailey, T. M., & Elvin, L. B. C. (2005). Effects of category diversity on learning, memory, and generalization. Memory & Cognition, 33, 289-302.

Hampton, J.A. (2000) Concepts and Prototypes. Mind and Language, 15, 299-307.

Heit, E. (1997). Knowledge and Concept Learning. In K. Lamberts & D. Shanks (Eds.), Knowledge, Concepts, and Categories (pp. 7-41). London: Psychology Press.

Hines, P., Pothos, E. M., & Chater, N. (2007). A non-parametric approach to simplicity clustering. Applied Artificial Intelligence, 21, 729-752.

Homa, D., Proulx, M. J., & Blair, M. (2008). The modulating influence of category size on the classification of exception patterns. The Quarterly Journal of Experimental Psychology, 61, 425-443.

Johansen, M. K. & Kruschke, J. K. (2005). Category representation for classification and feature inference. Journal of Experimental Psychology: Learning, Memory, and Cognition, 31, 1433-1458.

Jones, G. V. (1983). Identifying basic categories. Psychological Bulletin, 94, 423-428.

Kalish, M. L., Lewandowsky, S., & Kruschke, J. K. (2004). Population of linear experts: knowledge partitioning and function learning. Psychological Review, 111, 1072-1099.

Kemler Nelson, D. G. (1984). The effect of intention on what concepts are acquired. Journal of Verbal Learning and Verbal Behavior, 23, 734-759.

Krzanowski, W. J. & Marriott, F. H. C. (1995). Multivariate Analysis, Part 2: Classification, Covariance Structures and Repeated Measurements. Arnold: London.

Kurtz, K. J. (2007). The divergent autoencoder (DIVA) model of category learning. Psychonomic Bulletin & Review, 14, 560-576.

Lewandowsky, S., Roberts, L., & Yang, L. (2006). Knowledge partitioning in categorization: boundary conditions. Memory & Cognition, 34, 1676-1688.

Love, B. C. (2002). Comparing supervised and unsupervised category learning. Psychonomic Bulletin & Review, 9, 829-835.

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

Luce, R. D. (1963). Detection and recognition. In R. D .Luce, R. R. Bush, & E. Galanter (Eds.) Handbook of Mathematical Psychology, 1, 103-190. New York: Wiley.

Medin, D. L. (1983). Structural principles of categorization. In B. Shepp & T. Tighe (Eds.), Interaction: Perception, development and cognition (pp. 203-230). Hillsdale, NJ: Erlbaum.

Medin, D. L. & Schaffer, M. M. (1978). Context Theory of Classification Learning. Psychological Review, 85, 207-238.

Medin, D. L. & Schwanenflugel, P. J. (1981). Linear separability in classification learning. Journal of Experimental Psychology: Human Learning and Memory, 75, 355-368.

Milton, F. & Wills, A. J. (2004). The influence of stimulus properties on category construction. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30, 407-415

Milton, F., Longmore, C. A., & Wills, A. J. (in press). Processes of overall similarity sorting in free classification. Journal of Experimental Psychology: Human Perception and Performance.

Minda, J. P., & Smith, J. D. (2000). Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. Journal of Experimental Psychology: Learning, Memory, and Cognition, 27, 775-799.

Murphy, G. L. (1982). Cue validity and levels of categorization. Psychological Bulletin, 91, 174-177.

Murphy, G. L. (1991). Parts in object concepts: Experiments with artificial categories. Memory & Cognition, 19, 423-438.

Murphy, G. L. (2004). The big book of concepts. MIT Press: Cambridge, USA.

Murphy, G. L. & Medin, D. L. (1985). The Role of Theories in Conceptual Coherence. Psychological Review, 92, 289-316.

Murphy, G. L. & Smith, E. E. (1982). Basic-Level Superiority in Picture Categorization. Journal of Verbal Learning and Verbal Behavior, 21, 1-20.

Navarro, D. J. (2007). Similarity, distance, and categorization: a discussion of Smith's (2006). warning about "colliding parameters". Psychonomic Bulletin & Review, 14, 823-833.

Nomura, E. M., Maddox, W. T., Filoteo, J. V., Ing, A. D., Gitelman, D. R., Parrish, T. B., Mesulam, M. M., & Reber, P. J. (2007). Neural correlates of rule-based and information-integration visual category learning. Cerebral Cortex, 17, 37-43.

Nosofsky, R. M. (1988). Similarity, frequency, and category representation. Journal of Experimental Psychology: Learning, Memory and Cognition, 14, 54-65.

Nosofsky, R. M. (1989). Further tests of an exemplar-similarity approach to relating identification and categorization. Journal of Experimental Psychology: Perception and Psychophysics, 45, 279-290.

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. Journal of Experimental Psychology: Human Perception and Performance, 17, 3-27.

Nosofsky, R. M. (1992). Similarity scaling and cognitive process models. Annual Review of Psychology, 43, 25-53.

Nosofsky, R. M. (2000).Exemplar representation without generalization? Comment on Smith and Minda's (2000) Thirty categorization results in search of a model. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 1735-1743.

Nosofsky, R. M. & Palmeri, T. J. (1996). Learning to classify integral-dimension stimuli. Psychonomic Bulletin & Review, 3, 222-226.

Nosofsky, R. M. & Zaki, S. R. (2002). Exemplar and prototype models revisited: response strategies, selective attention, and stimulus generalization. Journal of Experimental Psychology: Learning, Memory, & Cognition, 30, 936-941.

Olsson, H., Wennerholm, P., & Lyxzen, U. (2004). Exemplars, prototypes, and the flexibility of classification models. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30, 936-941.

Olsson, A., Enkvist, T., & Juslin, P. (2006). Go with the flow: How to master a nonlinear multiple-cue judgment task. Journal of Experimental Psychology: Learning, Memory, & Cognition, 32, 1371-1384.

Pickering, M. & Chater, N. (1995). Why cognitive science is not formalized folk psychology. Minds and Machines, 5, 309-337.

Posner, M. I. & Keele, S. W. (1968). On the genesis of abstract ideas. Journal of Experimental Psychology, 77, 353-363.

Pothos, E. M. & Chater, N. (2002). A Simplicity Principle in Unsupervised Human Categorization. Cognitive Science, 26, 303-343.

Pothos, E. M. & Chater, N. (2005). Unsupervised categorization and category learning. Quarterly Journal of Experimental Psychology, 58A, 733-752.

Pothos, E. M. & Close, J. (2008). One or two dimensions in spontaneous classification: A simplicity approach. Cognition, 107, 581-602.

Rosch, E. & Mervis, C. B. (1975). Family Resemblances: Studies in the Internal Structure of Categories. Cognitive Psychology, 7, 573-605.

Ruts, W., Storms, G., & Hampton, J. (2004). Linear separability in superordinate natural language concepts. Memory & Cognition, 32, 83-95.

Sanborn, A. N., Griffiths, T. L., & Navarro, D. (2006). A more rational model of categorization. In R. Sun & N. Miyake (Eds.), Proceedings of the 28th Annual Conference of the Cognitive Science Society.

Schyns, P. G. (1991). A Modular Neural Network Model of Concept Acquisition. Cognitive Science, 15, 461-508.

Shepard, R. N., Hovland, C. L., & Jenkins, H. M. (1961). Learning and memorization of classifications. Psychological Monographs, 75, whole no 517.

Smith, J. D. (2007). When parameters collide: A warning about categorization models. Psychonomic Bulletin & Review, 13, 743-751.

Smith, J. D.& Minda, J. P. (1998). Prototypes in the mist: the early epochs of category learning. Journal of Experimental Psychology: Learning, Memory, & Cognition, 24, 1411-1436.

Smith, J. D. & Minda, J. P. (2000). Thirty categorization results in search of a model. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 3-27.

Smith, J. D.& Minda, J. P. (2002). Distinguishing prototype-based and exemplar-based processes in dot-pattern category learning. Journal of Experimental Psychology: Learning, Memory, & Cognition, 28, 800-811.

Smith, J. D., Murray, M. J. Jr.,& Minda, J. P. (1997). Straight talk about linear separability. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23, 659-680.

Stewart, N. & Brown, G. D. A. (2005). Similarity and dissimilarity as evidence in perceptual categorization. Journal of Mathematical Psychology, 49, 403-409.

Tenenbaum, J. & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. Behavioral and Brain Sciences, 24, 629-641

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. Trends in Cognitive Sciences, 10, 309-318.

Yang, L. & Lewandowsky, S. (2004). Knowledge partitioning in categorization: Constraints on exemplar models. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30, 1045-1064.

Vanpaemel, W. & Storms, G. (2008). In search of abstraction: the varying abstraction model of categorization. Psychonomic Bulletin & Review, 15, 732-749.

Zeithamova, D. & Maddox, W. T. (2006). Dual-task interference in perceptual category learning. Memory & Cognition, 34, 387-398.

Waldron, E. M. & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. Psychonomic Bulletin & Review, 8, 168-176.

Wattenmaker, W. D. (1995). Knowledge Structures and Linear Separability: Integrating Information in Object and Social Categorization. Cognitive Psychology, 28, 274-328.

Wisniewski, E. J. (1995). Prior knowledge and functionally relevant features in concept learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 449-468.

**Appendix. Coordinates of stimulus sets for Blair and Homa (2001), Compton and Logan (2001), and Pothos and Chater (2002).**

The coordinates for the LS9 and NLS9 category structures, after Blair and Homa (2001).

LS9

| dim1 | dim2 | dim3 | category membership |
|---|---|---|---|
| 0.512577 | -0.79309 | 0.599944 | 1 |
| 0.950823 | -0.43238 | -0.64205 | 1 |
| 0.919654 | -0.21622 | 0.84206 | 1 |
| 0.568712 | -0.39062 | 0.806167 | 1 |
| 0.880188 | -1.05977 | 0.25976 | 1 |
| 0.870623 | -0.12556 | 0.828048 | 1 |
| 0.582638 | -0.79996 | -0.43682 | 1 |
| 0.803067 | -0.30819 | -0.65991 | 1 |
| 0.876774 | -0.58663 | -0.61081 | 1 |
| -0.37315 | 0.377606 | -0.70526 | 2 |
| -0.73816 | 0.51154 | 0.740735 | 2 |
| -1.0016 | -0.22806 | -0.23997 | 2 |
| -1.16004 | 0.355098 | 0.491137 | 2 |
| -0.89618 | 0.516986 | 0.680177 | 2 |
| -0.47673 | 1.029432 | -0.24724 | 2 |
| -0.12914 | -0.00409 | -0.49905 | 2 |
| 0.135353 | 0.036625 | 0.084978 | 2 |
| -0.9625 | 0.412114 | 0.664558 | 2 |
| -0.94764 | -0.07662 | -0.57309 | 3 |
| -0.16205 | 0.250033 | -1.35296 | 3 |
| -0.19715 | -0.36733 | -1.42623 | 3 |
| 0.12512 | 0.365261 | -0.24335 | 3 |
| -0.49901 | 0.220019 | -1.30251 | 3 |
| -0.01908 | -0.65789 | -0.18377 | 3 |
| -0.15327 | -0.23203 | -1.45411 | 3 |

| | | | |
|---|---|---|---|
| -0.21237 | -0.32686 | 0.030132 | 3 |
| -0.0272 | 0.535436 | -1.00146 | 3 |
| -0.00022 | 0.108769 | 1.334538 | 4 |
| -0.65028 | 0.589483 | 0.717457 | 4 |
| -0.08156 | 0.110816 | -0.17754 | 4 |
| 0.1981 | -0.19542 | -0.05218 | 4 |
| -0.48844 | -0.4543 | 0.799656 | 4 |
| -0.5698 | 0.011733 | 1.129052 | 4 |
| -0.12106 | 0.094234 | 1.334165 | 4 |
| 0.522919 | 0.424834 | 0.962302 | 4 |
| -0.08454 | 0.180491 | 1.335545 | 4 |

NLS9

| dim1 | dim2 | dim3 | category membership |
|---|---|---|---|
| 0.838829 | -0.36053 | 0.501983 | 1 |
| 0.461945 | -0.17192 | 0.194231 | 1 |
| 0.619334 | 0.015154 | 0.190737 | 1 |
| 0.768688 | -0.26211 | 0.495337 | 1 |
| 0.68267 | -0.60588 | -0.15475 | 1 |
| 0.740189 | -0.04274 | 0.380288 | 1 |
| -0.27036 | 0.610669 | 2.62E-05 | 1 |
| -0.15634 | -0.11304 | -0.30074 | 1 |
| -0.36082 | 0.40637 | 0.609859 | 1 |
| -0.49422 | 0.210742 | -0.35648 | 2 |
| -0.93876 | 0.259376 | -0.10888 | 2 |
| -0.25408 | 0.371623 | 0.289483 | 2 |
| -0.27109 | 0.358076 | 0.309954 | 2 |
| -0.6851 | 0.26592 | -0.35547 | 2 |
| -0.60909 | 0.265358 | 0.425998 | 2 |
| 1.088306 | -0.11084 | 0.327542 | 2 |
| 0.070897 | -0.19185 | -0.99894 | 2 |
| -0.43094 | -0.02517 | 0.646929 | 2 |
| -0.00154 | 0.010836 | -0.38427 | 3 |
| -0.45214 | -0.28763 | -0.98939 | 3 |

| | | | |
|---|---|---|---|
| -0.21596 | -0.50962 | -0.53759 | 3 |
| -0.36785 | -0.23031 | -0.34315 | 3 |
| -0.21239 | 0.267929 | -0.70358 | 3 |
| 0.200385 | -0.10851 | -0.73687 | 3 |
| 0.47956 | -0.19797 | -0.07317 | 3 |
| -0.54098 | 0.321569 | 0.433105 | 3 |
| 0.320622 | 0.068017 | 0.701643 | 3 |
| 0.335456 | 0.08753 | 0.653351 | 4 |
| 0.160945 | -0.04609 | 0.289989 | 4 |
| 0.214626 | -0.04449 | 0.339537 | 4 |
| -0.15039 | 0.004641 | 0.203744 | 4 |
| -0.04018 | 0.27801 | 0.200659 | 4 |
| -0.11774 | 0.162605 | 0.177971 | 4 |
| 0.928562 | 0.017137 | 0.298591 | 4 |
| -0.55534 | 0.339004 | 0.433492 | 4 |
| -0.43699 | 0.087898 | -0.94817 | 4 |

The Compton and Logan (1999) stimulus sets for which highest and lowest classification variability was observed in Compton and Logan's Experiments 1 and 2. Classification variability is expressed in terms of number of unique classifications produced for each dataset, and shown between parentheses for each dataset (this is the way to reference the coordinates in Compton and Logan's diagrams). The horizontal and vertical dimensions in Compton and Logan's diagrams are denoted as 'x' and 'y'.

Experiment 1
Low classification variability (4)

| x | y |
|---|---|
| 0 | 5 |
| 2 | 9 |
| 8 | 8 |
| 9 | 7 |
| 7 | 2 |
| 9 | 0 |
| 11 | 1 |

Experiment 1
Low classification variability (6)

| x | y |
|---|---|
| 0 | 5 |
| 5 | 11 |
| 6 | 10 |
| 5 | 2 |
| 8 | 8 |
| 9 | 9 |
| 10 | 9 |

Experiment 2
Low classification variability (5)

| x | y |
|---|---|
| 0 | 5 |
| 1 | 10 |
| 4 | 11 |
| 4 | 6 |
| 6 | 4 |
| 7 | 11 |
| 9 | 5 |

Experiment 2
Low classification variability (5)

| x | y |
|---|---|
| 1 | 3 |
| 1 | 1 |
| 3 | 10 |
| 4 | 11 |
| 8 | 9 |
| 8 | 7 |
| 9 | 6 |
| 8 | 5 |
| 9 | 4 |

Experiment 1
High classification variability (21)

| x | y |
|---|---|
| 0 | 6 |
| 2 | 6 |
| 3 | 0 |
| 5 | 5 |
| 6 | 2 |
| 8 | 7 |
| 7 | 11 |
| 11 | 1 |

Experiment 1
High classification variability (25)

| x | y |
|---|---|
| 0 | 9 |
| 5 | 8 |
| 4 | 5 |
| 5 | 1 |
| 8 | 4 |
| 9 | 1 |
| 10 | 2 |
| 10 | 4 |
| 11 | 7 |

Experiment 2
High classification variability (20)

| x | y |
|---|---|
| 1 | 5 |
| 2 | 4 |
| 4 | 4 |
| 5 | 10 |
| 7 | 8 |
| 7 | 6 |
| 7 | 4 |
| 10 | 6 |
| 11 | 10 |
| 11 | 1 |

Experiment 2
High classification variability (23)

| x | y |
|---|---|
| 1 | 8 |
| 2 | 5 |
| 5 | 6 |
| 6 | 3 |
| 6 | 8 |
| 9 | 10 |
| 11 | 9 |
| 10 | 8 |
| 11 | 1 |

The coordinates of the four datasets employed by Pothos and Chater (2002). The intended horizontal and vertical dimensions are the first and second dimension respectively.

Two clusters

| | |
|---|---|
| 2 | 2 |
| 2 | 3 |
| 3 | 3 |
| 3 | 2 |
| 3 | 4 |
| 8 | 6 |
| 7 | 7 |
| 8 | 7 |
| 8 | 8 |
| 7 | 9 |

Three clusters

| | |
|---|---|
| 2 | 3 |
| 3 | 3 |
| 3 | 4 |
| 1 | 2 |
| 6 | 6 |
| 7 | 8 |
| 6 | 7 |
| 8 | 0 |
| 9 | 0 |
| 9 | 1 |

Big, small

| | |
|---|---|
| 2 | 5 |
| 3 | 5 |
| 3 | 6 |
| 9 | 4 |
| 7 | 4 |
| 8 | 4 |
| 8 | 5 |
| 7 | 5 |
| 8 | 6 |
| 9 | 5 |

Little

| | |
|---|---|
| 5 | 4 |
| 4 | 5 |
| 2 | 5 |
| 2 | 2 |
| 4 | 1 |
| 6 | 1 |
| 7 | 3 |
| 7 | 6 |
| 5 | 8 |
| 3 | 8 |

## Tables

**Table 1.** The application of the unsupervised GCM, simplicity model, and the rational model to the 'toy' dataset of Figure 2.

| | Distance between prototypes | | | |
|---|---|---|---|---|
| Model | 2 | 3 | 4 | 5 |
| Assumed intuitiveness | least | medium | high | highest |
| GCM[1] | 5.97 | 1.15 | 0 | 0 |
| Simplicity[2] | 118 | 66.9 | 54.7 | 51.6 |
| Rational[3] | | | | |
| all dims | 3.07 | 3.55 | 4.87 | 5.99 |
| best dim[4] | 6.10 | 15.0 | 26.0 | 31.0 |

Notes: [1]Goodness of fit (smaller values predict greater intuitiveness); [2]Codelength % (smaller values predict greater intuitiveness); [3]Classification probability x $10^{-4}$ (larger values predict greater intuitiveness); [4]The best dimension was always dimension 1.

**Table 2**. The Shepard, Hovland, and Jenkins (1961) classifications. The stimuli are specified in terms of three binary features (feature values: 1, 2). Each stimulus is assigned to category A or B, as specified, for each of the six category structures (I-VI).

| | Category structure | | | | | |
|---|---|---|---|---|---|---|
| Stimulus | I | II | III | IV | V | VI |
| 1 1 1 | A | A | B | B | B | B |
| 1 1 2 | A | A | B | B | B | A |
| 1 2 1 | A | B | B | B | B | A |
| 1 2 2 | A | B | A | A | A | B |
| 2 1 1 | B | B | A | B | A | A |
| 2 1 2 | B | B | B | A | A | B |
| 2 2 1 | B | A | A | A | A | B |
| 2 2 2 | B | A | A | A | B | A |

**Table 3**. Predictions of category intuitiveness from the unsupervised GCM, simplicity model, and the rational model for the Shepard et al. (1961) classifications.

| | Classification | | | | | |
|---|---|---|---|---|---|---|
| Model | I | II | III | IV | V | VI |
| Observed | lowest | low | intermediate | intermediate | intermediate | highest |
| GCM[1] | 0 | 0 | 11.1 | 11.1 | 13.1 | 13.6 |
| Simplicity[2] | 93.9 | 107.6 | 103.2 | 101.1 | 104 | 113 |
| Rational[3] | | | | | | |
| all dims | 68 | 38 | 53 | 55 | 51 | 31 |
| best dim[4] | 240 | 214 | 227 | 236 | 229 | 192 |
| | d2 | d1 | d1 | d2 | d2 | d1 |

Notes: [1]Goodness of fit; [2]Codelength %; [3]Classification probability x $10^{-4}$; [4]The best dimension for the rational model with dimensional selection is shown below each probability.

**Table 4**. Medin and Schaffer's (1978) 5-4 dataset and classification probabilities of the test items. Items are represented in terms of four binary dimensions (values 0,1).

| Training items | | Test items | | |
|---|---|---|---|---|
| Category 1 | Category 2 | Label | Coordinates | Probability[1] |
| 1 1 1 0 | 1 1 0 0 | T1 | 1 0 0 1 | 0.59 |
| 1 0 1 0 | 0 1 1 0 | T2 | 1 0 0 0 | 0.31 |
| 1 0 1 1 | 0 0 0 1 | T3 | 1 1 1 1 | 0.94 |
| 1 1 0 1 | 0 0 0 0 | T4 | 0 0 1 0 | 0.34 |
| 0 1 1 1 | | T5 | 0 1 0 1 | 0.50 |
| | | T6 | 0 0 1 1 | 0.62 |
| | | T7 | 0 1 0 0 | 0.16 |

Notes: [1]This is the probability of classification to Category 1.

**Table 5**. The application of the unsupervised GCM, simplicity model, and the rational model to the Medin and Schaffer (1978) data. The first value in each cell corresponds to the intuitiveness of a classification assuming the test item is assigned to the first category, the second number assuming that the test item is assigned to the second category.

| | Test items | | | | | | |
|---|---|---|---|---|---|---|---|
| Model | T3 | T6 | T1 | T5 | T4 | T2 | T7 |
| Empirical probability[1] | .94 | .62 | .59 | .50 | .34 | .31 | .16 |
| GCM[2] | 5.56 – 15.5 | 5.99 – 13.5 | 5.94 – 13.5 | 14.9 – 8.73 | 13.4 – 8.75 | 13.4 – 8.79 | 15.2 – 8.65 |
| Simplicity[3] | 90.3 – 101 | 95.8 – 97.5 | 95.8 – 97.5 | 98.4 – 94.8 | 97.8 – 95 | 97.8 – 95 | 100 – 90.6 |
| Rational[4] | | | | | | | |
| all dims | 9.5 – 7.5 | 4.9 – 18 | 4.9 – 18 | 4.9 – 18 | 8.1 – 12 | 7.8 – 12 | 8.0 – 12 |
| best dim[5] | 46 – 74 | 45 – 76 | 46 – 74 | 46 – 76 | 45 – 74 | 46 – 73 | 47 – 73 |
| | d1/d3 – d4 | d3 – d2 | d1 – d2 | d2 – d4 | d3/d4 – d1 | d1 – d2 | d2 – d1/d3 |

Notes: [1]Empirical probabilities refer to classification into the first category; [2]Goodness of fit; [3]Codelength %; [4]Classification probability x $10^{-4}$; [5]The best dimension for classification into Category 1 – the best dimension for classification into Category 2 is shown below each probability Where two dimensions are shown, this means that the probability of the best classification along one dimension is the same as that of the other.

**Table 6.** Unsupervised GCM, simplicity model, rational model, and rational model with dimensional selection predictions for linearly separable (LS9, LS9X) and nonlinearly separable (NLS9, NLS9X) category structures, created after Blair and Homa (2001).

| | Category structure | | | |
|---|---|---|---|---|
| Model | LS9 | LS9X | NLS9 | NLS9X |
| Predicted intuitiveness | high | highest | lowest | lowest |
| GCM[1] | 49.1 | 1.69 | 100 | 100 |
| Simplicity[2] | 91.4 | 71.3 | 99.5 | 98.6 |
| Rational[3] | | | | |
| all dims | 1.96 | 0.959 | 2.15 | 6.72 |
| best dim[4] | 1.96 | 17.1 | 5.96 | 13.7 |
| | all | d3 | d1 | d2 |

Notes: [1]Goodness of fit; [2]Codelength %; [3]Classification probability x $10^{-24}$; [4]The best dimension for the rational model with dimensional selection is shown below each probability; ‘all’ indicates that no one-dimensional solution was better than the solution with all dimensions.

**Table 7**. Unrestricted and restricted unsupervised GCM, simplicity model, rational model, and rational model with dimensional selection, for the Compton and Logan (1999) data. Stimulus sets for which the lowest and highest classification variabilities were observed are denoted by ‘L’ and ‘H’ respectively (there were two stimulus sets of each kind in Compton and Logan’s study).

| | Category structure | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Model | Exp1 L | Exp1 L | Exp2 L | Exp2 L | Exp1 H | Exp1 H | Exp2 H | Exp2 H |
| Empirical data[1] | 4 | 6 | 5 | 5 | 21 | 25 | 20 | 23 |
| GCM[2] | | | | | | | | |
| unrestricted | 0 | 0 | 0 | 0 | 4.39 | 0 | 0 | 0 |
| restricted ($c$=0.5) | $0.28 \times 10^{-4}$ | $4.52 \times 10^{-4}$ | $10.9 \times 10^{-4}$ | $0.26 \times 10^{-4}$ | 10.2 | 1.24 | 0.64 | 0.07 |
| Simplicity[3] | 69.3 | 53.9 | 84.7 | 60.2 | 87.7 | 74.6 | 76.5 | 71 |
| Rational[4] | | | | | | | | |
| all dims | .128 | .135 | .153 | .091 | .135 | .115 | .117 | .136 |
| best dim[5] | .438 | .287 | .576 | .487 | .135 | .132 | .117 | .136 |
| | d1 | d2 | d2 | d1 | all | d1 | all | all |

Notes: [1] This is the number of distinct classifications produced by Compton and Logan's participants; as there were 30 participants in each of their Experiments 1 and 2, the highest possible value for classification variability is 30 (and the lowest possible value one); [2]Goodness of fit; [3]Codelength %; [4]Classification probability; [5]The best dimension for the rational model is shown below each probability; 'all' indicates that no one-dimensional solution was better than the solution with all dimensions.

**Table 8**. Unrestricted and restricted unsupervised GCM, simplicity model, rational model, and rational model with dimensional selection predictions for the Pothos and Chater (2002) stimulus sets.

| | Stimulus set | | | |
|---|---|---|---|---|
| Model | Two clusters | Big, small cluster | Three clusters | Little |
| GCM[1] | | | | |
| unrestricted | 0 | 0 | 0 | 0 |
| restricted (*c*=0.5) | 0.002 | 0.002 | 0.104 | 4.110 |
| Simplicity[2] | 51.6 | 51.2 | 62.3 | 87.7 |
| Rational[3] | | | | |
| all dims | .640 | .082 | .083 | .144 |
| best dim[4] | .752 | .671 | .360 | .144 |
| | d1 | d1 | d1 | all |

Notes: [1]Goodness of fit; [2]Codelength %; [3]Classification probability; [4]The best dimension for the rational model with dimensional selection is shown below each probability; ‘all’ indicates that no one-dimensional solution was better than the solution with all dimensions.

**Figure captions**

Figure 1. Each point in the diagram represents an item in psychological space. The top panel shows an intuitive category structure, while the bottom one a corresponding less intuitive one.

Figure 2. Shown is the most intuitive (left) and least intuitive (right) category structure, in a set of four category structures which were used to illustrate the function of the models. The distance between the prototypes of the two categories varied between five and two units (in decrements of 1 unit). In the right figure, there was a point that is identical for categories *A* and *B*.

Figure 3. Unsupervised GCM, simplicity, and rational model intuitiveness values for the Figure 2 category structures.

Figure 4. Unsupervised GCM, simplicity, and rational model intuitiveness values for the six classifications of Shepard et al. (1961). The intuitiveness values from each model were converted onto a 0-1 scale.

Figure 5. Unsupervised GCM, simplicity, and rational model intuitiveness values for the classification of the seven test items in the 5-4 dataset of Medin and Schaffer (1978). The intuitiveness values from each model were converted onto a uniform scale (0-2). Also shown are the empirically measured classification probabilities, converted onto a 0-2 scale; the results are ordered in terms of decreasing likelihood

that the test item would be classified in the second category. The horizontal axis refers to the test items (T1-T7).

Figure 6. Unsupervised GCM, simplicity, and rational model intuitiveness values for the LS9 and NLS9 category structures of Blair and Homa (2001), as well as two derivative category structures in which the prototypes were pushed further apart. The intuitiveness values from each model were converted onto a 0-1 scale.

Figure 7. Unsupervised GCM with $c$=0.5, simplicity model, and rational model with dimensional selection intuitiveness values (converted onto a scale between 0 and 1) for the Compton and Logan (1999) datasets. The horizontal axis indexes the datasets, in the same order as they appear in Table 7. In the graph we also show Compton and Logan’s empirical results, converted onto a 0 – 1 scale.

Figure 8. Unsupervised GCM with $c$=0.5, simplicity model, and rational model with dimensional selection intuitiveness values (converted onto a scale between 0 and 1) for the Pothos and Chater (2002) datasets. The aggregate empirical measure of category intuitiveness from Pothos and Chater’s (2002) results is also shown (converted to a 0—1 scale).

**Figures**

Figure 1.

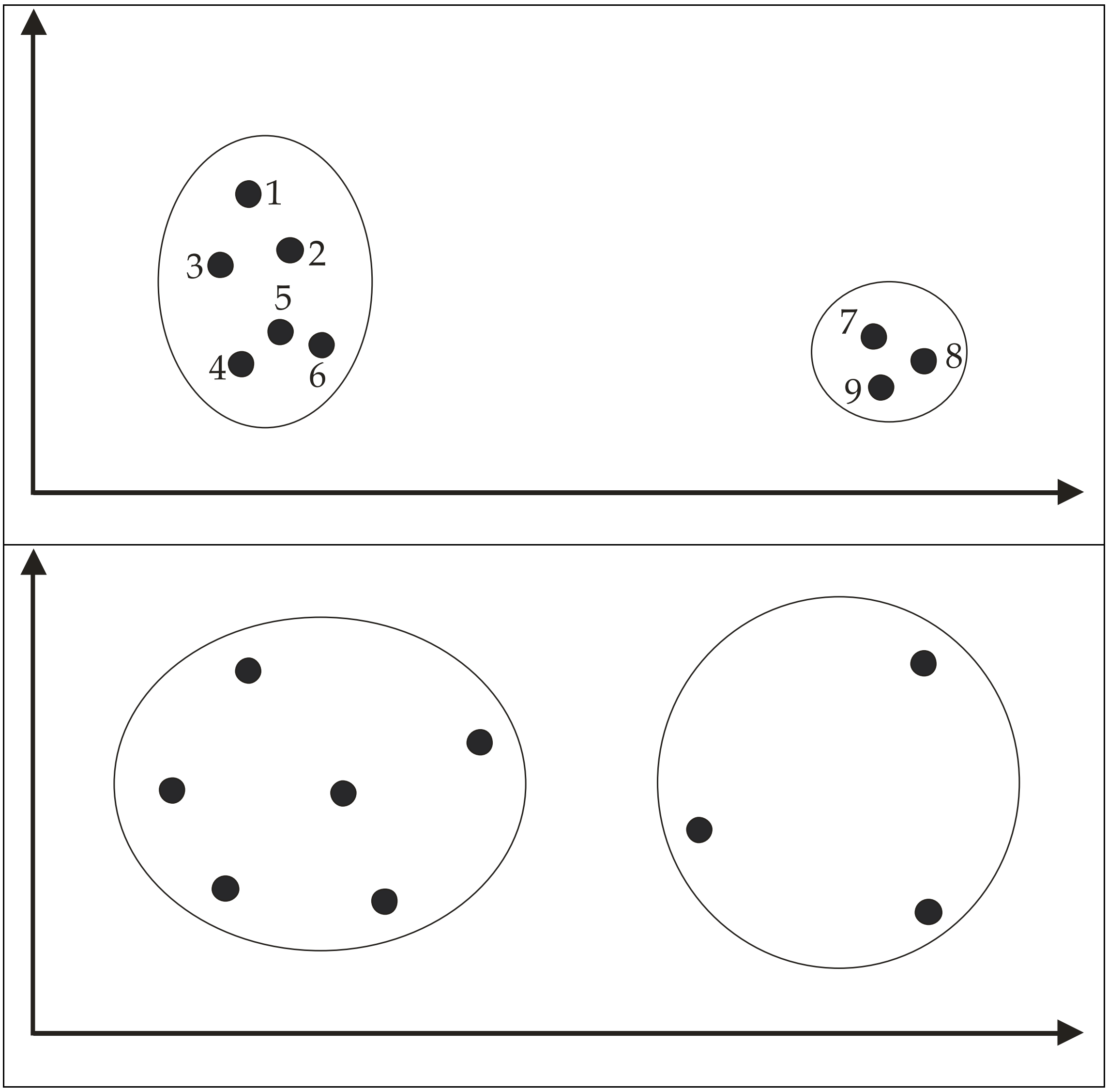


Figure 2.

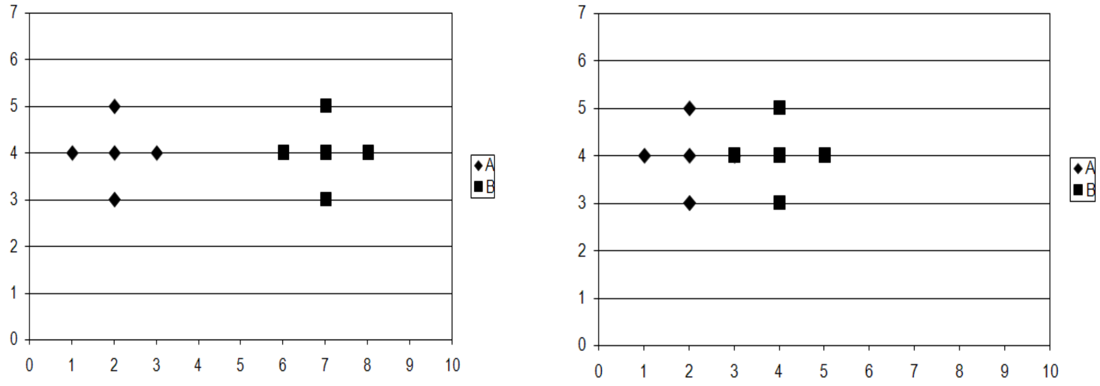


Figure 3.

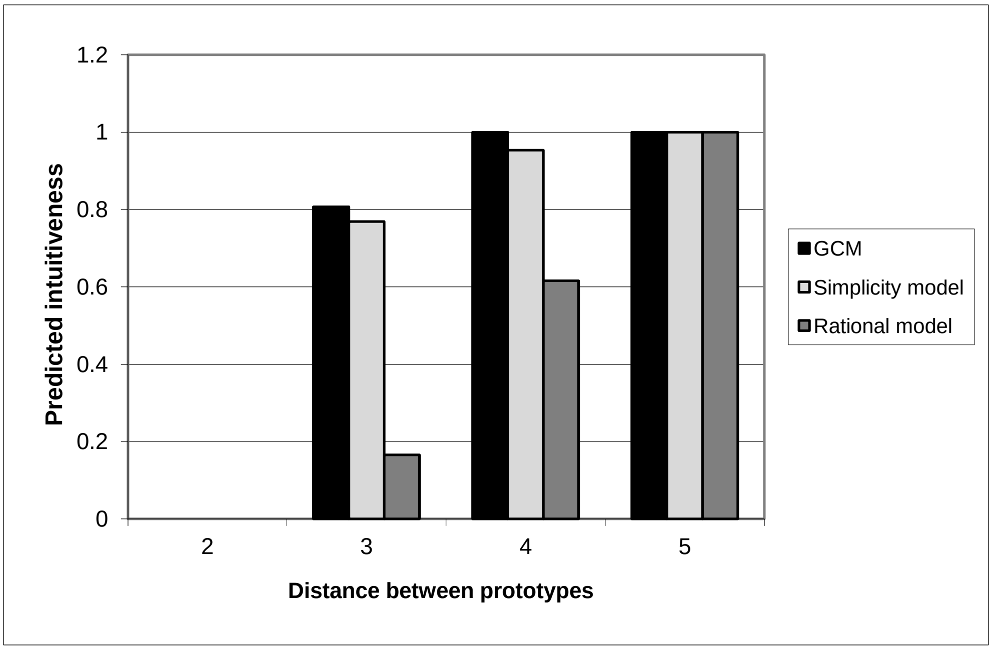

1.2
1
0.8
0.6
0.4
0.2
0
Predicted intuitiveness
2
3
4
5
Distance between prototypes
GCM
Simplicity model
Rational model

Figure 4.

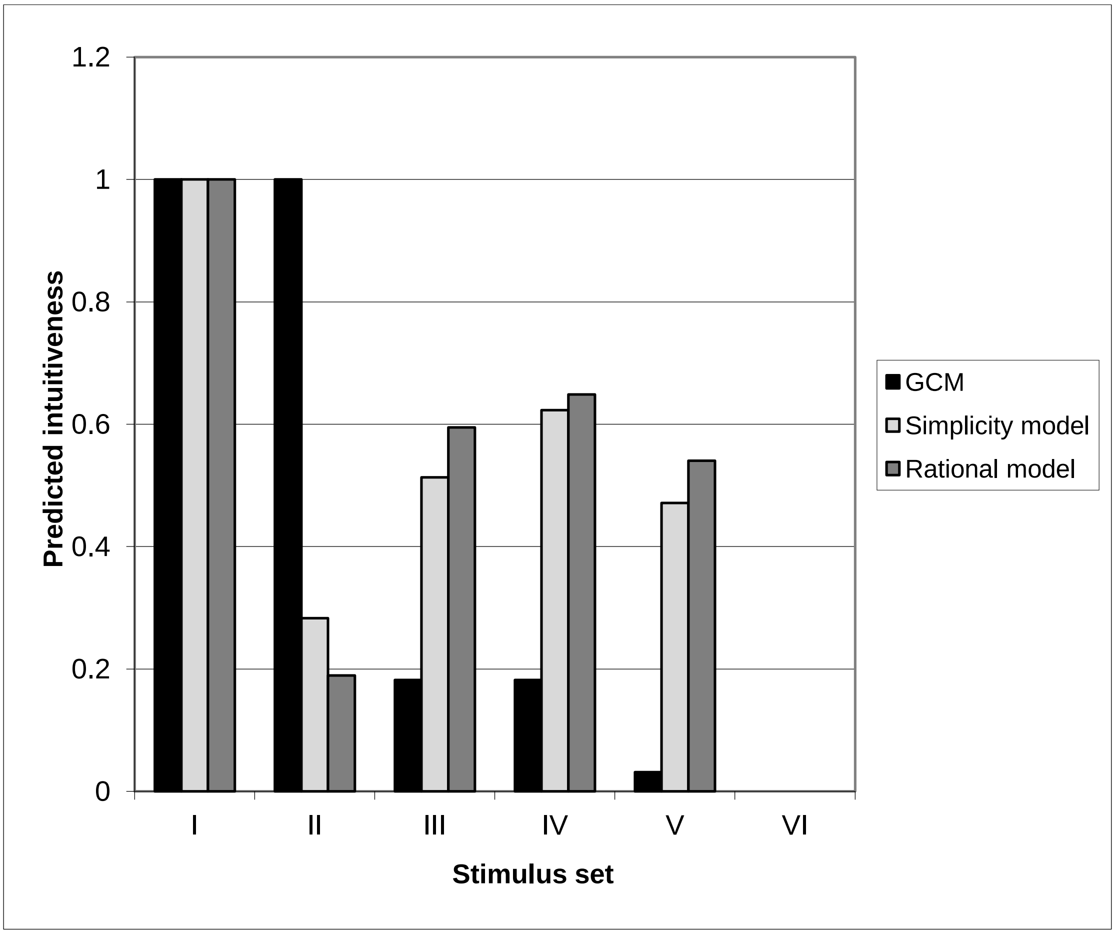

1.2
1
0.8
0.6
0.4
0.2
0
Predicted intuitiveness
I
II
III
IV
V
VI
Stimulus set
GCM
Simplicity model
Rational model

Figure 5.

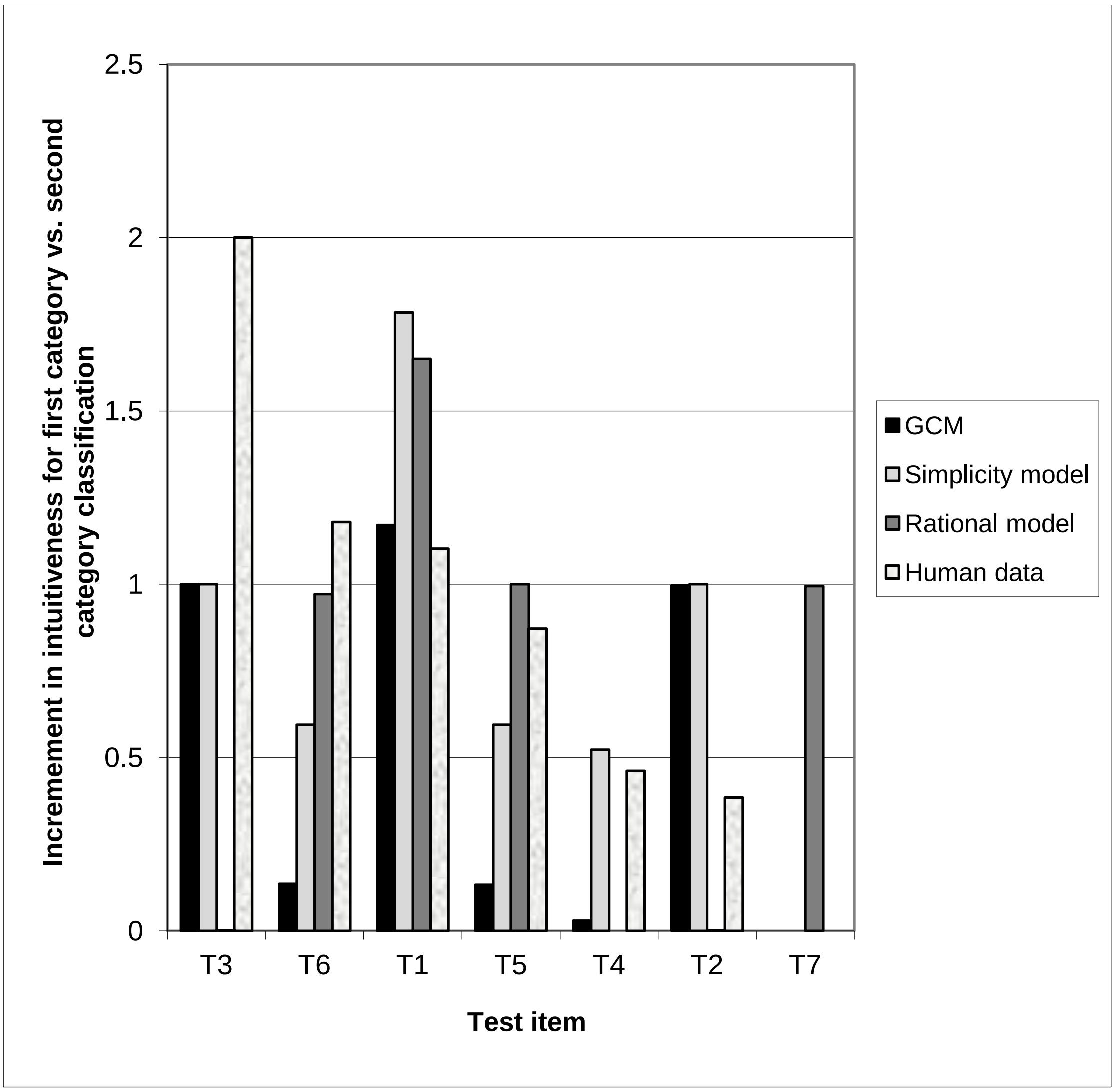

2.5
2
1.5
1
0.5
0
Incremement in intuitiveness for first category vs. second category classification
T3
T6
T1
T5
T4
T2
T7
Test item
GCM
Simplicity model
Rational model
Human data

Figure 6.

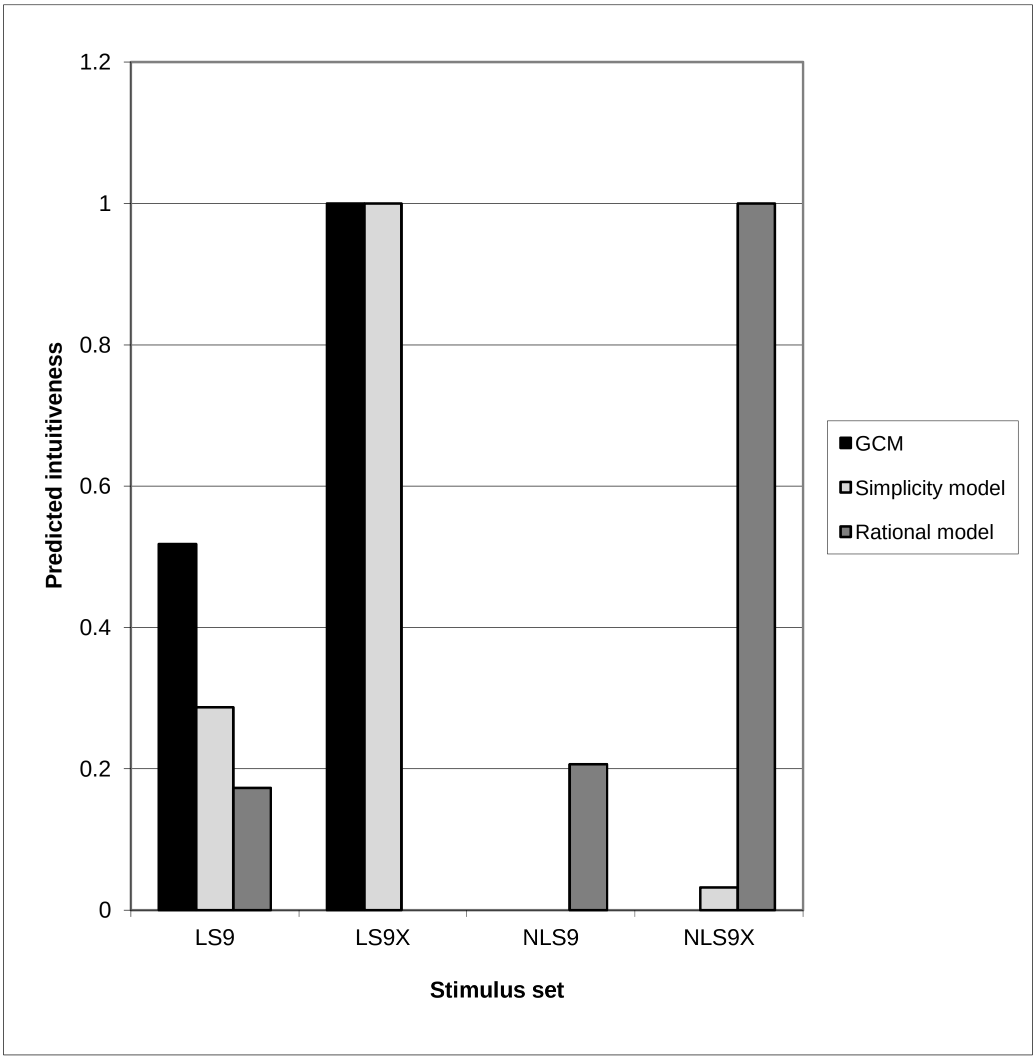

1.2
1
0.8
0.6
0.4
0.2
0
Predicted intuitiveness
LS9
LS9X
NLS9
NLS9X
Stimulus set
GCM
Simplicity model
Rational model

Figure 7.

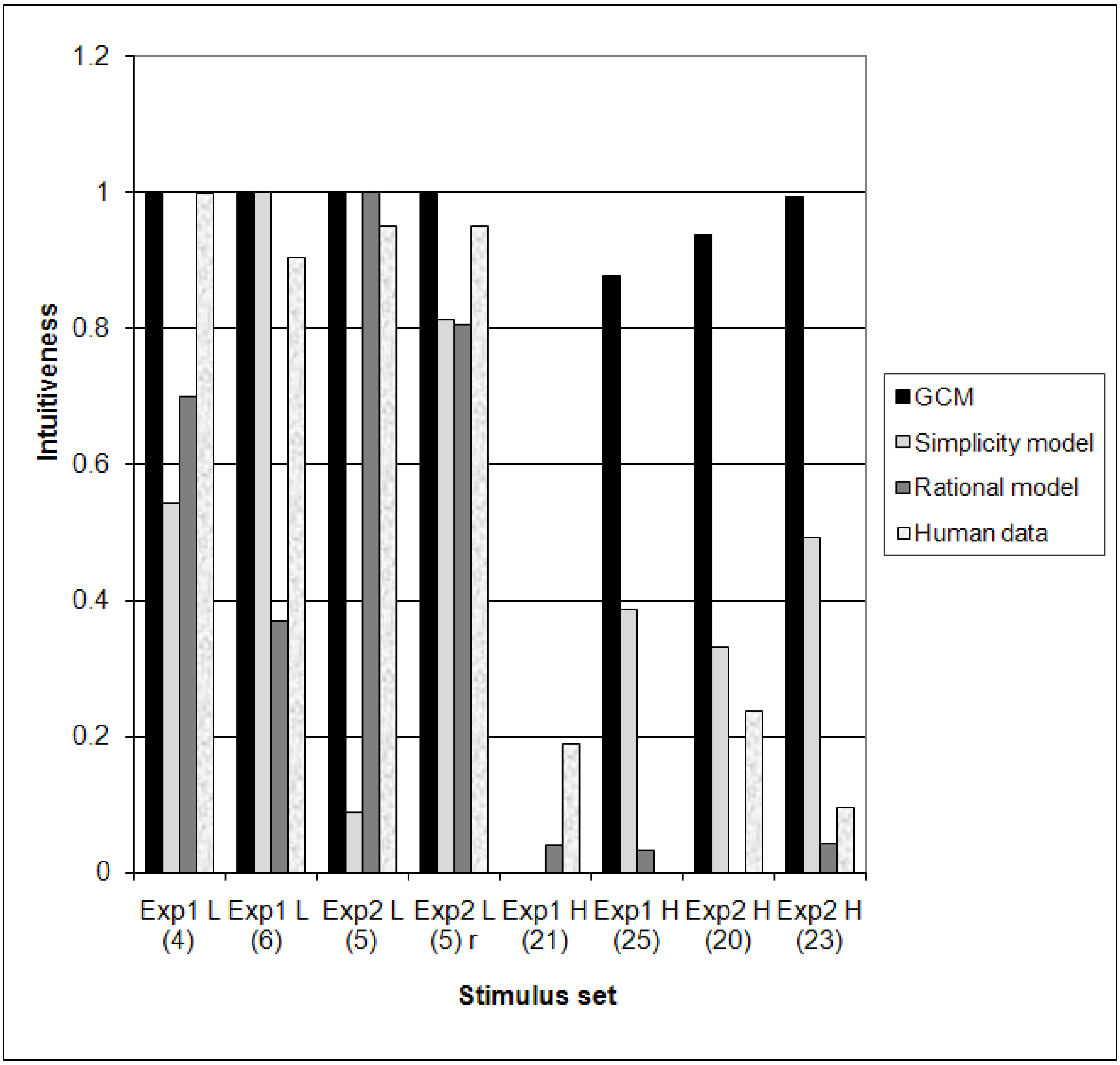
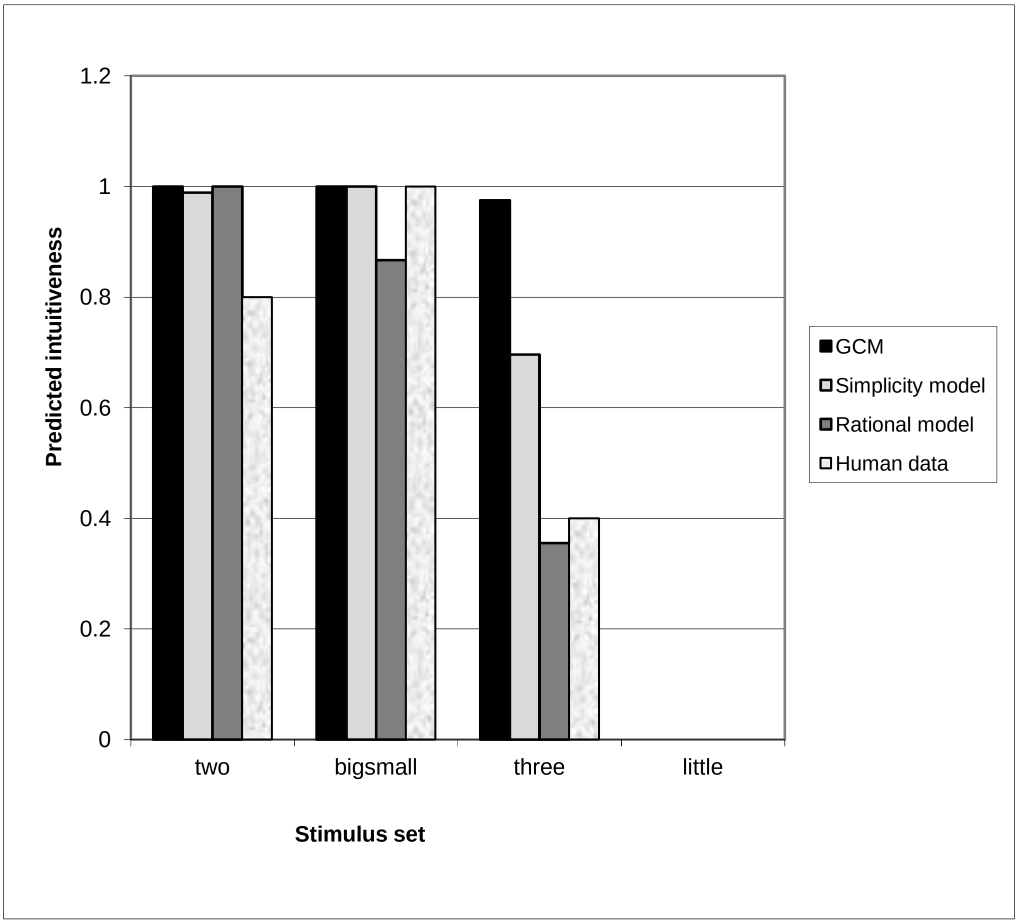

1.2
1
0.8
0.6
0.4
0.2
0
Intuitiveness
Exp1 L (4)
Exp1 L (6)
Exp2 L (5)
Exp2 L (5) r
Exp1 H (21)
Exp1 H (25)
Exp2 H (20)
Exp2 H (23)
Stimulus set
GCM
Simplicity model
Rational model
Human data

Figure 8.


1.2
1
0.8
0.6
0.4
0.2
0
Predicted intuitiveness
two
bigsmall
three
little
Stimulus set
GCM
Simplicity model
Rational model
Human data