Journal of Cognitive Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pecp21

Will the most informative object stand? Determining the impact of structural context on informativeness judgements

Ronaldo Vigo [a] & Basawaraj [a]

[a] Center for the Advancement of Cognitive Science, Psychology Department , Ohio University , Athens , OH , USA

Published online: 11 Mar 2013.

Journal of Cognitive Psychology, 2013
Vol. 25, No. 3, 248–266, http://dx.doi.org/10.1080/20445911.2012.755510

Routledge
Taylor & Francis Group

Will the most informative object stand? Determining the impact of structural context on informativeness judgements

Ronaldo Vigo and Basawaraj

Center for the Advancement of Cognitive Science, Psychology Department, Ohio University, Athens, OH, USA

A fundamental unsolved problem in the cognitive sciences concerns why, how, and to what extent humans judge object stimuli as conveying different amounts of information. Central to this problem is how the notion of informativeness is conceptualised by humans in the first place. In this paper, we investigate this question from the standpoint of how the structure of categories of objects influences informativeness judgements about their members. Results from our two experiments show that the structural or relational context surrounding single-object cues from a categorical stimulus largely determines such informativeness judgements. Moreover, we found that object cues elicit absolute magnitude judgements about their associated concept that are not consistent with the prototype interpretation of the concept. We were able to account for over 90% of the variance in the data from our two judgement experiments with a general theory and measure of information referred to as Representational Information.

Keywords: Categorisation; Complexity; Concept cues; Invariance; Prototype; Representational information; Subjective information.

Humans often encounter situations where the need arises to assess the most representative object among a set of objects, or, more generally, to assess how much information small subsets of a category convey about the entire category. We refer to such a subset of category instances as *concept cues* and to the category under consideration as an *online concept* due to the constant availability of its associated instances a priori. For example, in a political setting, humans may wish to know which candidate of a political party among a set of candidates best represents the views of all the candidates. Likewise, in a scientific setting, a researcher may try to ascertain which method or approach among a set of alternatives best captures the nature of all the approaches. Accordingly, biologists may find it useful to ascertain which genus may be most representative of a certain family of organisms. Finally, and more generally, in human concept learning research, the extraction of representative information facilitates the formation of "concept prototypes" or single-instance representations of a categorical stimulus (Reed, 1972; Rosch, 1978; Rosch et al., 1976; Rosch & Mervis, 1975). In short, the amount of information that is conveyed by an object stimulus about its category of origin lies at the very core of key psychological processes that facilitate our understanding of how sets of objects are represented as concepts in the mind and brain.

Given the central role that representative information plays in human concept learning, a

Correspondence should be addressed to Ronaldo Vigo, Center for the Advancement of Cognitive Science, Psychology Department, Ohio University, Athens, OH, USA. E-mail: vigo@ohio.edu

The authors would like to thank Mikayla Barcus, Charles Doan, Phillip Halsey, and Derek Zeigler for their helpful comments.

key question emerges regarding its relationship to prototypicality (Reed, 1972; Rosch, 1978; Rosch & Mervis, 1975): Are information judgements on concept cues compatible with choosing the most prototypical instance as defined by the family resemblance principle (Rosch & Mervis, 1975)? The basic idea underlying the principle is that, within a family, people share features but no two individuals are identical. However, we can form the central family tendency or prototype by simply taking the features that most members share and draw a composite using these features. Indeed, it is plausible that the same cognitive mechanisms that are responsible for the formation and permanent storage of a concept as a prototype constructed in the aforementioned way may play a role in the formation of concept cue judgements. Once the prototype is generated, degrees of prototypicality become the basis of concept cue judgements.

But, we contend that judgements about the degree of informativeness of concept cues depend on more than an estimate of the most prototypical item(s) in a set. Consider an investigation by Chin-Parker and Ross (2004) which examined how the nature of a learning task would emphasise prototypical versus diagnostic features. In particular, subjects were required to perform one of two types of tasks: a classification task, for which subjects had to predict category membership, and an inference task, for which subjects had to predict missing features from category knowledge. This was followed by other tasks to test the role of diagnosticity and prototypicality on these studies. It was observed that classification learners relied primarily on the diagnostic features when making decisions about category membership but inference learners were sensitive to *both* the diagnosticity of the features (albeit significantly less so than the classification learners) and prototypicality (Chin-Parker & Ross, 2004). These results seem to indicate that representational information involves more than simply prototypicality assessments, a point that we will argue for the remainder of this paper.

Several well-known models that use a measure of prototypicality at their core have been developed to account for categorisation behaviour in humans (Estes, 1986, 1994; Nosofsky & Zaki, 2002; Reed, 1972; Rosch & Mervis, 1975; Smith & Minda,1998; Smith, Murray, & Minda, 1997). Because our interest is on informativeness judgements, and not on categorisation performance, we will focus on the methods used to measure degrees of prototypicality in these models. The first of these methods, which we refer to as the "feature frequency" prototypicality measure (FFPM), was suggested by Rosch and associates (Rosch, 1978; Rosch & Mervis, 1975). The FFPM determines the degree of prototypicality of each object by taking the sum of the number of objects in the set with which it shares the same features. For example, with respect to a set of objects defined by the dimensions of colour, shape, and size, the degree of prototypicality of any object in the set is given by summing three quantities: (1) the number of objects in the set that have the same colour as the object in question, (2) the number of objects in the set that have its same shape, and (3) the number of objects in the set that have its same size. Using this approach, degrees of prototypicality corresponding to the six sets of objects appearing in Figure 1 are shown in the fourth column of Table 1.

The second method for determining degrees of prototypicality comes from the multiplicative prototype model (MPM; Estes, 1986, 1994; Nosofsky & Zaki, 2002). The core idea underlying the MPM is that humans classify objects from a category by assessing the similarity of their

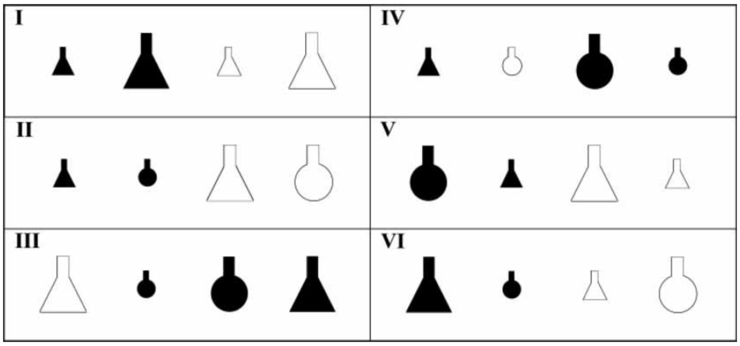
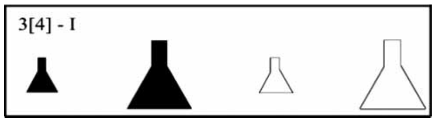
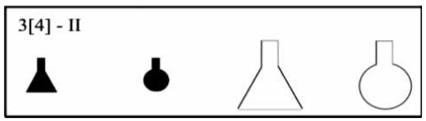
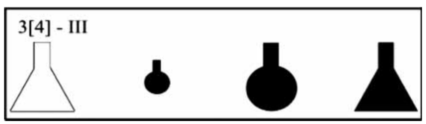
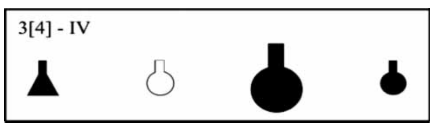
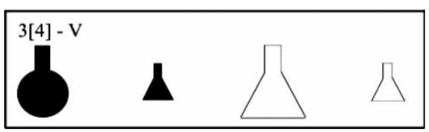
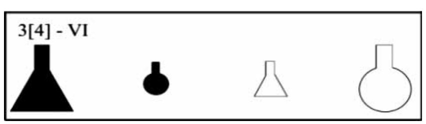


Figure 1. Example of the six types of categorical stimuli used in Experiment 1, the single choice experiment.

TABLE 1
Degree of informativeness values from RIT third column) vs. degree of prototypicality values from FFPM, MPMP-1, and MPMP-2

Stimulus type	*Objects*	*Representational information (Vigo, 2011b)*	*FFPM (Rosch, 1978)*	*MPMP-1 (Estes, 1986, 1994)*	*MPMP-2 (Nosofsky & Zaki, 2002)*
3[4] - I 	{000, 001, 100, 101}	[.20, .20, .20, .20]	[8, 8, 8, 8]	[.37, .37, .37, .37]	[1, .37, .37, .13]
3[4] - II 	{000, 010, 101, 111}	[.05, .05, .05, .05]	[6, 6, 6, 6]	[.22, .22, .22, .22]	[1, .37, .13, .05]
3[4] - III 	{101, 010, 011, 001}	[−.31, −.31, −.08, −.08]	[6, 6, 8, 8]	[.22, .22, .37, .37]	[.37, .13, .37, 1]
3[4] - IV 	{000, 110, 011, 010}	[−.31, −.31, −.31, .78]	[7, 7, 7, 9]	[.29, .29, .29, .47]	[.37, .37, .37, 1]
3[4] - V 	{011, 000, 101, 100}	[−.41, −.22, −.22, .52]	[5, 7, 7, 7]	[.17, .29, .29, .29]	[.13, .37, .37, 1]
3[4] - VI 	{001, 010, 100, 111}	[−.25, −.25, −.25, −.25]	[6, 6, 6, 6]	[.22, .22, .22, .22]	[1, .13, .13, .13]

corresponding exemplars to the category's prototype. Thus, each similarity comparison between an object's exemplar (i.e., mental representation of the object) and the prototype of the set of objects determines the perceived degree of prototypicality of the object. Shepard's similarity measure (1974, 1987) is used in the model to compute these degrees of prototypicality. The similarity between the particular object and the prototype is computed by taking the negative exponent of the sum of the distances (differences) between their features. The Technical Appendix contains a detailed explanation of this procedure. The MPM comes in two versions: In the original MPM, Estes (1986) used the average value across the dimensions of the objects in the category to determine its prototype. The second version (Nosofsky & Zaki 2002) uses the more common feature-frequency notion of prototype (Rosch, 1978; Rosch & Mervis, 1975) explained earlier. Nosofsky & Zaki (2002) argue that this model is the most effective among the well-known prototype models in accounting for classification data. Results from the former MPM prototypicality measure (MPMP-1) are listed on the fifth column of Table 1, and results for the latter (MPMP-2) are listed on the sixth column.

Although there have been multiple studies on prototypicality judgements as the basis for categorisation behaviour (for a brief summary, see Nosofsky, Gluck, Palmeri, McKinley, & Glauthier, 1994), no studies have directly investigated the ‘‘degree of informativeness’’ conveyed by concept cues. By ‘‘degree of informativeness’’, we mean a direct rating on some subjective information scale as to how much information a stimulus object conveys to the observer. One reason for this deficit may be attributed to the fact that the construct of information as proposed by Shannon (1948) does not capture our intuitions about what information means as a cognitive construct. Indeed, the numerous attempts to link this widely used method of measuring information to various psychological phenomena have been effectively thwarted (Devlin, 1991), perhaps by the seemingly daunting task of capturing the structural contextual effects that emerge from the perceived relationships between the instances of a category. Indeed, many experiments in perception science (particularly by Gestalt researchers), and fewer in other areas of cognitive science, have documented the undeniable existence of these contextual effects on response behaviour (for examples, see Medin, Goldstone, & Gentner, 1993; Reicher, 1969). For example, in one study by Medin et al. (1993), subjects were instructed to list common and distinctive properties among two objects from a set of three objects. They found that the properties ascribed to any one object depended on which of the other two objects it was paired with. They concluded that similarity is heavily influenced by context.

This conundrum involving structural context and information was addressed directly by Luce (2003), in his essay titled ‘‘What ever happened to information theory in Psychology?’’ where he writes about the inadequacy of Shannon information theory in the psychological sciences. In particular, Luce concludes (p. 184):

> The question remains: Why is information theory not very applicable to psychological problems despite apparent similarities of concepts? . . . in my opinion, the most important answer lies in the following incompatibility between psychology and information theory. The elements of choice in information theory are absolutely neutral and lack any internal structure; the probabilities are on a pure, unstructured set whose elements are functionally interchangeable.

In this article, we propose a lawful connection between a new general notion of information that overcomes these limitations and subjective informativeness judgements.

Representational information theory (RIT; Vigo, 2011b) offers a possible solution to the context problem. More specifically, RIT is based on five principles: (1) Humans communicate via concepts or, in other words, mental representations of categories of objects (where a category is simply a set of objects that are related in some way); (2) therefore, concepts are the mediators of information; (3) but concepts are relationships between qualitative objects in the environment that are defined dimensionally; (4) the degree of homogeneity of a category (i.e., to what extent its objects are indistinguishable) as well as its cardinality (i.e., size) determine the learnability or complexity of its associated concept; and (5) information is the rate of change of that complexity. The first three principles are frequently adopted by researchers in the field of human concept learning (Bourne, 1966; Estes, 1994; Garner, 1974; Vigo, 2009), and Principles 4 and 5 form the basis of the theories proposed by Vigo (2009, 2011a, 2013b). For example, with respect to the fourth principle, Vigo (2009, 2011a, 2013b) developed a theory and model of degree of concept learning difficulty where the perceived degree of structural complexity of a set of objects (i.e., its learnability) is directly proportional to the cardinality or size of the set and inversely proportional to the exponent of its degree of invariance (where the degree of invariance of a category is proposed as a measure of its degree of homogeneity). The model, referred to as the exponential categorical invariance model (ECIM), accounts extremely well for the variance, $R^2 = .93$, $p < .0001$, in the data from historical data (e.g., Bourne, 1970; Shepard, Hovland, & Jenkins, 1961) and from a recent large-scale experiment involving 84 category structures consisting of two, three, and four dimensions (Vigo, 2011a, 2013b)—a number well beyond the six tested by Shepard et al. (1961).

Last, the fifth principle is the basis of Vigo's (2011b) representational information theory. The basic idea underlying the principle is that what makes objects informative is the way that their inclusion/exclusion impacts the learnability or complexity of sets. The extent of their impact, whether positive (increase in complexity) or negative (decrease in complexity), determines the amount and quality of information they convey. More specifically, the amount of information conveyed by a subset R of a set of objects S about S is the rate of change in the perceived degree of structural complexity in S whenever R is removed from S. In other words, the extent to which an object(s) (in extension) or concept cue (object in intension) is informative about its set of origin is determined by the impact that the removal of the object(s) has on the *relative difference* between the structural complexity of the original set and that of the new set (i.e., the original set minus the object(s)). The higher this rate of change, the more information is conveyed. Thus, the notion of "the degree of surprise of an event" that is often associated with Shannon information is replaced by the notion of "the rate of change in the learnability of a concept" or, equivalently, the rate of change in the perceived degree of structural complexity of the category from which the concept is learned. Moreover, the quality of the information is represented by the direction of the change: Negative rates of change indicate a decrease in perceived complexity (good quality), whereas positive rates indicate an increase in perceived complexity (low quality).

Accordingly, to compute the amount and quality of representational information conveyed by any particular single object about the stimulus set S that contains it, we take the difference between the structural complexity of S′ (S without the object in question) and the structural complexity of S; then, we divide the result by the structural complexity of S to get the percentage rate of change of the structural complexity in S with respect to the particular object. This rate of change, according to RIT, is the amount of information conveyed by the object about S. The structural complexity of both S′ and S is computed by the exponential categorical invariance model (ECIM; Vigo, 2009, 2011a, 2013b), a model intended to account for the degree of learning difficulty of any dimensionally defined stimulus set (i.e., set of objects) in terms of an invariance-based notion of complexity. A sketch of the mathematical details of RIT and ECIM is given in Technical Appendix A and a detailed example of how to compute information values for the objects in the fifth category of objects shown in Figure 1 is given in Technical Appendix B. Representational information values for each of the objects in each of the categories shown in Figure 1 are shown in the third column of Table 1 and in Table 2. Finally, the idea of linking the change in the complexity of a category structure (when some of its elements are removed) to information content addresses the problem of determining the role that context plays on the amount of information humans attribute to

TABLE 2

Amount of information conveyed by each of the four objects of the six 3[4] category structures according to representational information theory (Vigo, 2011b, 2013a).

Stimulus-set type	*Objects*	*Information*
3[4]-I	{000, 001, 100, 101}	[.20, .20, .20, .20]
3[4]-II	{000, 010, 101, 111}	[.05, .05, .05, .05]
3[4]-III	{101, 010, 011, 001}	[−.31, −.31, −.08, −.08]
3[4]-IV	{000, 110, 011, 010}	[−.31, −.31, −.31, .78]
3[4]-V	{011, 000, 101, 100}	[−.41, −.22, −.22, .52]
3[4]-VI	{001, 010, 100, 111}	[−.25, −.25, −.25, −.25]

The third column displays in square brackets the amount of representational information conveyed by each of the four possible single element representations of the stimulus set in the "Objects" column. In this report, the first dimension represents colour, the second shape, and the third size. Negative information values represent a decrease in the complexity of the stimulus set after the object is removed, whereas positive values represent an increase (see Vigo, 2011b, for details).

concept cues. Hence, we will argue that the most informative concept cues will be those whose role in the online set (i.e., in the set presented in its entirety to the subject) is to decrease its perceived complexity. The following two studies show two novel ways of making this key connection empirically.

EXPERIMENT 1

In our first experiment, we asked the participants to judge which single object out of a set of four objects conveyed the most information about the character of the set as a whole (see Figure 1 for examples of six such sets). The sets were structurally different. The ''most informative'' object request was purposely used instead of the ''most typical'' object request because the purpose of our experiment was to elicit responses that are consistent with human intuitions as to the nature of the concept of information. Moreover, we felt that a ''typical object'' request would encourage subjects to indulge in the process of counting features. However, our aim was to elicit Gestalt judgements with respect to the relationships detected in each set of objects. It should be noted that during postexperiment interviews, no subject reported counting features; in addition, during briefing, all subjects reported that they understood what ''most informative'' meant.

Method

Participants. Thirty-six undergraduate students from Ohio University between the ages of 18 and 22 participated in the experiment.

Stimuli. An HP XW4600 workstation with a Dell 1708FP 15-inch flat panel LCD monitor (5 ms response time) was used to display sets of four flasks (flat bottles) defined over three binary dimensions: colour (black or white), shape (triangular or circular), and size (small or large). Each set was displayed on a neutral grey background and conformed to a specific category structure (by ''structure'' we mean the relationships between the dimensional values of the objects of a set). There are only six possible category structures associated with sets consisting of exactly four objects that are defined over three binary dimensions (for a proof of this fact, see Higonnet & Grea, 1958). These are appropriately labelled the 3[4] family of structures, where the number outside the brackets indicates the number of dimensions and the number within the brackets indicates the number of objects (Feldman, 2000; Vigo, 2006). Instances of these six structures are shown in Figure 1. We used these six structures to generate our stimulus sets because of their great significance to categorisation and concept learning research (Nosofsky et al., 1994; Shepard et al., 1961). Note that although there are only six possible structures, there are many instances that correspond to each structure (Feldman, 2000). Generating instances from these six structures is done by systematically remapping dimensional values: For example, round replaces black and triangular replaces white to generate a structurally equivalent stimulus set. The instances were generated using a computer program written in Matlab 6 and Psychophysics Toolbox 3.1. Figure 2 illustrates an example of a stimulus shown to participants during a trial.

Procedure. Participants were tested on four instances of each of the six structures for a total of 24 stimulus sets across 24 trials. The same 24 sets were presented in random order to each participant. Before the start of the experiment participants were told that for each trial, a bottle collector liked the particular set of flasks they would be presented with. Furthermore, they were told that their task during each four-flask presentation was to select the flask that they found most informative about all four flasks that the art collector likes to collect and that the single chosen flask should indicate the kind of flask that the bottle collector likes to collect. Subjects were told to specify their responses within 20 s by clicking on one of four buttons appearing below the set of flasks. The button labelled ''1'' represented the leftmost flask and the button labelled ''4'' represented the rightmost flask; accordingly, buttons labelled ''2'' and ''3'' represented the second and third flasks from left to right as depicted in Figure 2. As soon as a response was entered, a new trial would begin. Accordingly, subjects were told that once a choice was made they would not be able to change it.

Results and discussion

The effects of structural context on the information conveyed by each of the four single-instance

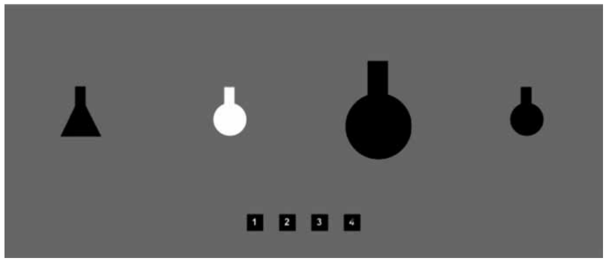


Figure 2. Example of a Type IV stimulus set given in a trial of Experiment 1. Subjects identify which flask they find "most informative" about the entire set by clicking on one of four numbers below the flasks. In this experiment virtually every subject chose the small black bottle (the fourth object in the sequence) as the most informative in this particular stimulus set.

concept cues was tested by examining the percentage of subjects that chose a particular flask on average for each structure type (as determined by the average percentage of selections for all the instances of the structure). A summary of the result appears in Figure 3. Note that, on average, subjects found the concept cues of Type I, II, and VI stimulus sets to be equally informative, and this agrees with our intuitions about these types of structures. A close examination of Figure 1 indicates how this finding conforms to our intuitions as to which individual single-instance concept cues are most informative. Indeed, there was no statistically significant difference between the degrees of information conveyed by the flasks of the instances of these three structure types as seen in Figure 3 and as confirmed by the least favourable pairwise t-test to our hypothesis among all the possible combinations, $t = 1.03$, $p < .05$, $df = 35$; however, Types III, IV, and V structures told a much different story. Subjects judged two of the objects of Type III instances as equally informative and the remaining two objects likewise. Type IV showed a single object

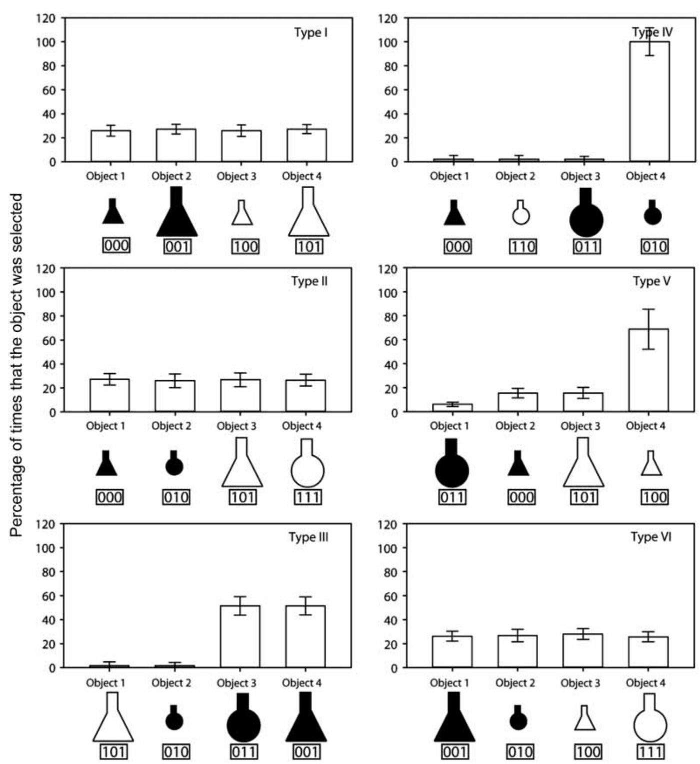


Figure 3. Results of Experiment 1. The frequencies of objects choices are specified on the y-axis of each graph. The error bars represent 95% confidence intervals. Error bars representing standard deviation were nearly identical to those shown. The objects underneath the bars are representative examples only. In fact, each bar in each of the six bar graphs reflects the average perceived degree of informativeness for all of the objects that play the same structural role within the four-object stimulus sets.

as being greatly more informative than the rest, and Type V had the most variance among the degrees of judged informativeness for each object. All of these results were statistically significant as revealed by pairwise *t*-tests on all possible combinations of the magnitude judgements corresponding to the four objects *within* a set.

RIT's information values highly correlate with the observed pattern of results as shown in Figure 4. For example, all four objects in a Type I structure, according to the information measure in RIT, are equally informative with an information value of .20 corresponding to each object. On the other hand, in a Type V structure only two objects convey the same amount of information (i.e., $-.22$). Likewise, the representational information value of each object stimulus for the six category structures is in accord with the idea discussed earlier and proposed in RIT that the objects that convey the greatest amount of information about their categorical stimulus are those objects that, when removed from their category, make the category more complex or harder to learn.

In fact, for Types I, II, and VI, RIT predicts that, due to the fact that they cannot be easily changed by the absence of any single stimulus object, all the objects are equally informative. For Types III, IV, and V, the RIT information measure (without parameters) also predicts the significantly greater variation in information values. In fact, RIT accounts for nearly all of the variance in the data involving each of these three structure types, $R^2 = .99$, $p < .001$, $SE < 0.001$. In Technical Appendix D, we have included more details regarding the regression analysis of RIT. Figure 4 depicts the fits of representational information to the data of Experiment 1.

EXPERIMENT 2

In our second experiment, we used the same types of stimuli as in Experiment 1. However, we asked participants to rank each of the four objects in the stimulus set in terms of how much information it conveyed about the category as a whole. We hypothesised similar results to those obtained under Experiment 1 with respect to single instance cues. Thus, this experiment served as a way of testing the robustness of the first result. However, beyond that, the experiment tested the information conveyed by the remaining three object stimuli in each stimulus set in a more direct fashion, that is, based not on frequency of

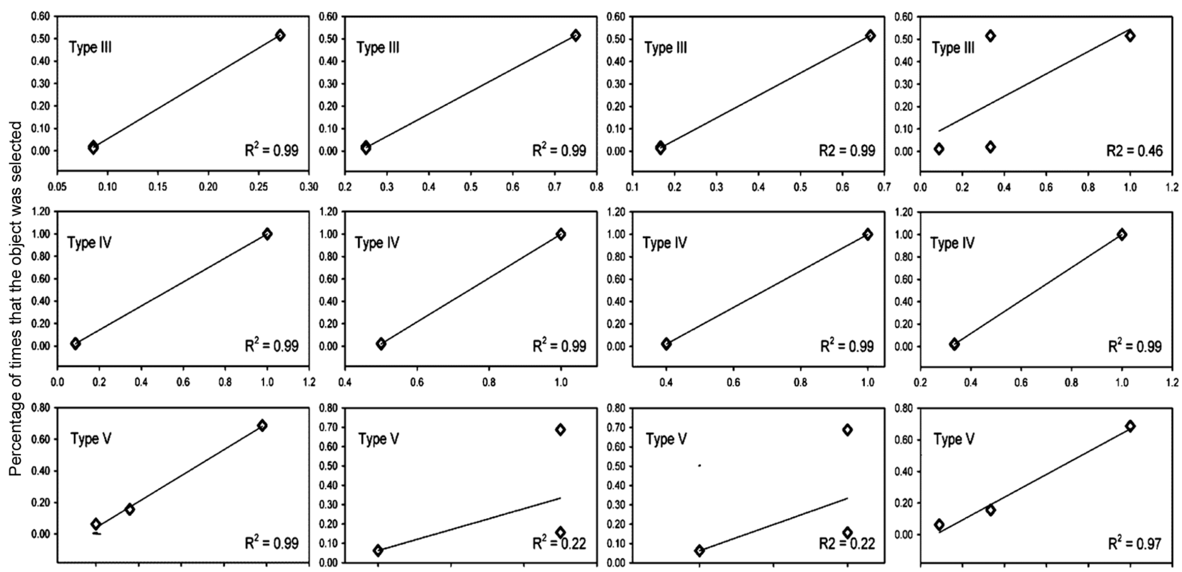


Figure 4. The first column of three graphs features (in the x-axis of each graph) the standardised predictions of informativeness judgements made by RIT for Experiment 1 with respect to stimulus sets of Types III, IV, and V. The four points in each plot stand for the four objects in each stimulus set. Note that only some points are visible because some nearly coincide (see Table 2 for the exact values). The y-axis represents the standardised degree of subjective information as measured by the percentage of times the object was chosen by all of the participants of Experiment 1. The second, third, and fourth columns show the predictions made by the FFPM (Rosch, 1978), MPMP-1 (Estes, 1986), and MPMP-2 (Nosofsky & Zaki, 2002), respectively. As can be seen, these three models failed to predict magnitude judgements on one or two of the three types.

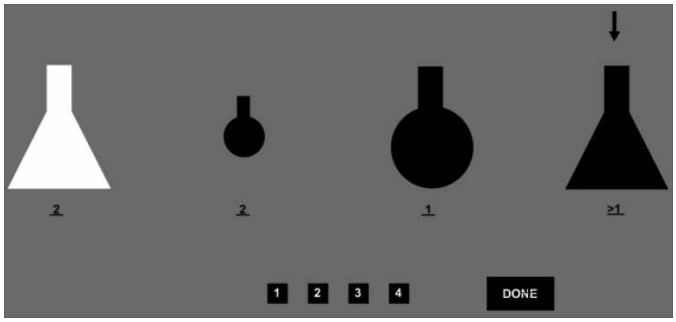


Figure 5. Example of a Type III stimulus set given in a trial of Experiment 2. Subjects identify the degree of informativeness of each flask by first mouse clicking on the flask and then clicking on the number that represents its degree of perceived informativeness.

selection, but direct rankings (see Figure 1 for examples of six sets).

Method

Participants. Thirty-seven undergraduate students from Ohio University between the ages of 18 and 22 participated in the experiment.

Stimuli. The same kinds of stimulus sets used in Experiment 1 were used here but the display per trial looked different. It involved the addition of an arrow to indicate the concept cues of interest, a "DONE" button, and four text entry prompts, one under each stimulus object as pictured in Figure 5.

Procedure. Participants were tested on eight instances of each of the six mentioned structures for a total of 48 stimulus sets across 48 trials. The same 48 sets were presented in random order to each participant. Before the start of the experiment participants were told that a bottle collector liked the particular set of flasks they would be presented with. Furthermore, they were told that their task during each four-object presentation was to assign, on a scale from one to four, a degree of the amount of information conveyed by each of the four flasks that the art collector likes to collect. Moreover, subjects were told to enter their responses for each flask by clicking on the flask first, then clicking on one of the four "buttons" labelled "1", "2", "3", and "4" (where 1 represented the least informative and 4 represented the most informative object). Entry errors could be corrected by simply repeating this process. Upon completion within the 30 s time limit per trial, subjects were instructed to click on the "DONE" button. After exceeding the time limit or after a click on "DONE" the next trial followed immediately.

Results and discussion

Results of this experiment corroborated results from the first. The same relative quantitative and qualitative patterns of informativeness judgements were found as indicated by Figure 6. Again, on average, subjects found the single-instance concept cues of Type I, II, and VI stimulus sets to be equally informative, and this, once again, agrees with our intuitions about these types of structures. And as in Experiment I, there were no statistically significant differences between the degrees of information conveyed by the flasks in these three structure types as seen in Figure 6 and as confirmed by the least favourable pairwise *t*-test to our hypothesis among all possible combinations, $t = -0.90$, $p < .05$, $df = 36$. In Technical Appendix D, we have included a table containing details on the regression analysis (Table A2–A7). Figure 4 depicts the fits of representational information to the data of Experiment 1.

Furthermore, as in Experiment 1, Types III, IV, and V structures conveyed the same story with subjects judging two of the flasks of Type III instances as equally informative. Likewise, as in Experiment 1, Type IV showed a single concept cue as being greatly more informative than the rest, whereas Type V had the most variance among the degrees of judged informativeness of each concept cue. Furthermore, as was demonstrated by Experiment 1, RIT was able to account for virtually all of the variance in the data of Experiment 2 as illustrated in Figure 7: Types III and IV, $R^2 = .99$, $p < .002$, $SE < 0.02$; Type V, $R^2 = .85$, $p = .08$, $SE = 0.14$. In Technical Appen-

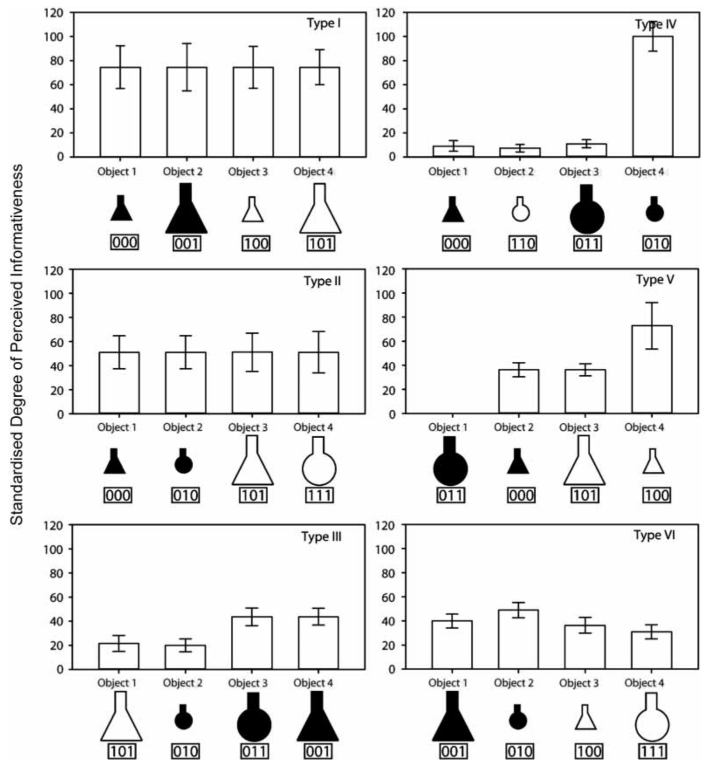


Figure 6. Results of Experiment 2. The judged informativeness of each object is specified on the y-axis of each graph. The error bars represent 95% confidence intervals. Error bars representing standard deviation were nearly identical to those shown. The objects underneath the bars are representative examples only. In fact, each bar in each of the six bar graphs reflects the average perceived degree of informativeness for all of the objects that play the same structural role within the four-object stimulus sets.

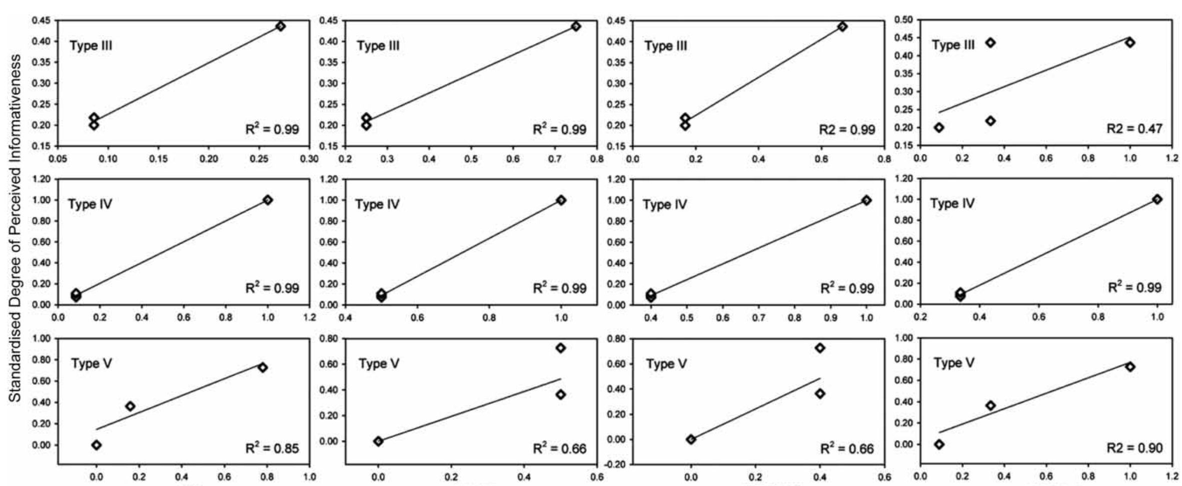


Figure 7. The first column of three graphs features (in the x-axis of each graph) the standardised predictions of informativeness judgements (in the x-axis) made by RIT for Experiment 2 with respect to stimulus sets of Types III, IV, and V. The four points in each plot stand for the four objects in each stimulus set. Note that only some points are visible because some nearly coincide (see Table 2 for the exact values). The y-axis represents the standardised degree of subjective information as measured by the average ratings given by all of the participants of Experiment 2. The second, third, and fourth columns show the predictions made by the FFPM (Rosch, 1978), MPMP-1 (Estes, 1986), and MPMP-2 (Nosofsky & Zaki, 2002), respectively. As can be seen, these three models failed to predict magnitude judgements on one or two of the three types.

dix D, we have included more details regarding the regression analysis of RIT.

Prototypes and concept cue information. We compared the predictions made by RIT to those made by the three well-known models of prototypicality discussed in the introduction. First, we constructed the prototype (in the sense discussed in the introduction) of each stimulus set used in the experiment and determined whether or not it was a good predictor of the objects chosen in Experiment 1 and ranked in Experiment 2. Our goal was to find out if the prototypes and prototypicality ratings generated by the measures discussed in the introduction coincided with those identified by the subjects in both of our experiments. To start with, with respect to the FFPM, the answer in a nutshell is that there was agreement on stimulus sets of Types I, II, III, IV, and VI but not on V: In fact, for Type V, using the generated prototypicality ratings, FFPM could account for only 22% of the variance, $R^2 = .22$, $p = .52$, $SE = 0.31$, in the data of Experiment 1 and only 66% of the variance, $R^2 = .66$, $p = .18$, $SE = 0.21$, in the data of Experiment 2; in contrast, RIT accounted for 99% of the variance and 85% of the variance in the data respectively. These results are illustrated in Figures 4 and 7 along with results from using the more complex MPMP-1 and MPMP-2 approaches to measuring prototypicality. The MPMP-1 and MPMP-2, as did the FFPM, yielded spotty accuracy in their predictions, where the MPMP-1 agreed with the FFPM on all six types of stimulus sets but not with the MPMP-2 which clearly failed to account for the empirical patterns observed with respect to Types I, II, and VI in both experiments (see Table 1 for the models' predictions). With regard to Types V and III, there was also disagreement: specifically, for Type V, using the MPMP-1, we could account for only 22% of the variance, $R^2 = .22$, $p = .52$, $SE = 0.31$, in the data of Experiment 1 and only 66% of the variance, $R^2 = .66$, $p = .18$, $SE = 0.21$, in the data of Experiment 2. On the other hand, the MPMP-2 accounted for only 45% of the variance, $R^2 = .45$, $p = .26$, $SE = 0.32$, in the data of Experiment 1 and for only 47% of the variance, $R^2 = .47$, $p = .12$, $SE = 0.31$, in the data of Experiment 2 with respect to Type III structures. In contrast, RIT (without parameters) accounted for 99% of the variance, $R^2 = .99$, $p < .001$, $SE < .001$, for all the types in Experiment 1 and (see Figures 4 and 7) accounted for 99% of the variance, $R^2 = .99$, $p < .002$, $SE < .02$, for all types in Experiment 2 except Type V. For Type V in Experiment 2, RIT accounted for about 85% of the variance, $R^2 = .85$, $p = .08$, $SE = 0.14$. We have included detailed regression analysis tables for these four models in Technical Appendix D.

Once again, we attributed these considerable differences in model performance to the fact that RIT was designed to measure information by first determining the extent to which the structural homogeneity between the objects in a categorical stimulus affect the perception of the category *as a whole* (i.e., its Gestalt). This is particularly true for Type V stimuli since, according to categorical invariance theory (Vigo, 2009, 2011a, 2013b), these stimuli contain a great deal of variation in the relationship between the features of its objects, which in turn will result in stronger relational or contextual effects that cannot be accounted for by measures of prototypicality as shown.

GENERAL DISCUSSION

Data from Experiments 1 and 2 showed for the first time that, as hypothesised, the perceived degree of informativeness of single-instance concept cues is proportional to the rate of relational complexity reduction of the stimulus set. This is measured in RIT by the proportional change in complexity that takes place when the object corresponding to the single-instance concept cue is removed from the stimulus set. Moreover, we showed that RIT accounts for practically all of the variance in the data of both experiments and provides an explanation by identifying the relational context effects that are often found when humans process multiple object stimuli. This is significant for three reasons. First, this connection between direct informativeness judgements, concept cues, and stimulus context had not been established. In addition, this connection is at the heart of many tangible everyday choice behaviours. One example that immediately springs to mind concerns party elections: A candidate is perceived as conveying more information about his/her party's platform than the rest whenever the absence of that candidate from a debate increases the overall agreeability or homogeneity (i.e., decreases the relational discord or complexity) of ideas among the remaining candidates.

Second, the fact that both experiments, using distinct protocols on the same stimulus sets, produced similar results indicates that the results are robust. Third, the fact that three well-known measures of prototypicality were not able to account for a significant portion of the variance in the data of Type III and V structures (in both experiments) indicates that information judgements are operationally distinct from the interpretation of concepts as prototypes (at least when prototypes are constructed on the basis of either the majority of shared features or on the basis of averages across dimensions). In general, we feel that we have provided strong evidence supporting a fundamental incompatibility between prototype and cue informativeness judgements. More importantly, we have provided evidence that RIT predicts the object that is most representative of a category of objects.

Admittedly, several computational measures of category structure could be used to explore the effect of any particular category member on the overall category structure (see Pothos et al., 2011, for a discussion of several of these measures). However, most of these measures are based on a series of domain-specific and theory-specific goals pertaining to classification behaviour and performance and not directly on complexity notions. RIT, on the other hand, was designed exclusively to capture humans' intuitions about information and informativeness, and to provide a complexity-based and invariance-based alternative to Shannon information theory in situations where the elements of information have structure. Indeed, RIT is founded on the union of several abstract and ubiquitous constructs in universal science, which, we believe, lends the theory generality beyond the scope of psychological research.

Although the basic idea underlying RIT is that information is the rate of change of complexity, the theory assumes that the ECIM is necessary to fully and accurately capture structural complexity. More generally, the assumption is that only measures that conform to the five principles underlying RIT, and a few others discussed in Vigo (2011a, 2013b), would be appropriate. In particular, the fourth of these principles states that the degree of learning difficulty of a concept is a function of the size and homogeneity (i.e., degree of invariance) of the set from which the concept is learned; in other words, that the perceived degree of structural complexity of a category of objects that are dimensionally defined can be accurately accounted for with the appropriate relation between these two variables. Thus, in RIT, any complexity approach that violates this principle does not measure representational information. For example, the Boolean minimisation complexity model (MinC; Feldman, 2000; Vigo, 2006) does not satisfy this principle.

Beyond this technical point, there are other reasons why ECIM plays a central role in RIT (ECIM; Vigo, 2009, 2011a, 2013b). All of these hinge on the generality of the ECIM when compared to alternative complexity-based accounts of concept learning and categorisation. First, as previously stated, ECIM not only predicts the key SHJ ordering (i.e., the 3[4] family of structures; Shepard et al., 1961), but it also predicts the precise ordering of many other structures studied by several researchers (e.g., Bourne, 1970; Feldman, 2000; Nosofsky, 1986; Vigo, 2011a, 2013b). This is accomplished deterministically and without parameters (and, in the most general case, with a single scaling parameter). Second, the ECIM is based on a unification of fundamental constructs such as pattern detection, invariance, similarity, and complexity. Not only are these constructs ubiquitous in universal science, but they offer a plausible, unifying, and general explanatory framework for the degree of learnability of categorical stimuli. Third, the generalised version of the ECIM (named "GISTM"; Vigo, 2013b) can determine the degree of structural complexity of continuous, dichotomous, and binary sets of stimuli, a necessary requirement for a truly general information measure. None of these three conditions is satisfied by MinC. In short, the complexity measure at the heart of RIT has at its core the extraction of relational patterns which must be extracted in a particular way for the information measure to work in the first place.

Interestingly, Pothos and Chater (2002) proposed a computational model named "the simplicity model" that is closer in spirit to the ECIM because in addition to being a nonparametric deterministic model, it attempts to capture the coherence or Gestalt similarity (i.e., homogeneity) of a set of objects as a perceptual organisation principle. However, the model was intended to account for unsupervised categorisation performance with respect to the way that humans will spontaneously divide into groups a set of objects. In contrast, the ECIM was intended to account for the degree of concept learning difficulty of "preformed" categories of objects, which, as we know, is required in RIT. Notwithstanding, in follow-up work, it would be interesting to try the aforementioned alternative models of classification performance within the RIT framework and to compare the results to those obtained by the core measure in RIT.

Although we believe that our findings will help researchers account for the role that representational information, as a subjective and domain-independent psychological construct, plays in influencing a variety of cognitive phenomena, more research needs to be conducted in order to fully understand this role. To start with, experiments on informativeness judgements involving stimulus sets of higher dimensionality and cardinality should be conducted. For example, there is a total of 19 structures associated with sets of four objects defined over four dimensions. Due to their broader range and variability in complexity, testing these structures could further our understanding of the role that changes in the structural complexity of a set, caused by the exclusion/inclusion of objects, play in the process of judging informativeness.

In closing, although we have argued that the information conveyed by concept cues is largely relational in nature, we do not claim that only relational contextual effects are at play when such information is assessed. In fact, other information that is qualitative in nature, such as the kinds of features occurring in the stimulus set and the way that these are encoded and attended to by observers are likely to also play a role; however, our results indicate that these take a backseat to the structural properties of the stimulus set as seen by the very accurate fits of RIT to our data without using free parameters.

Original manuscript received May 2012
Revised manuscript received November 2012
First published online March 2013

REFERENCES

Bourne, L. E. (1966). *Human conceptual behavior*. Boston, MA: Allyn & Bacon.

Bourne, L. E. (1970). Knowing and using concepts. *Psychological Review*, *77*, 546–556.

Chin-Parker, S., & Ross, B. H. (2004). Diagnosticity and prototypicality in category learning: A comparison of inference learning and classification learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 216–226.

Devlin, K. (1991). *Logic and information*. Cambridge, UK: Cambridge University Press.

Estes, W. K. (1986). Array models for category learning. *Cognitive Psychology*, *18*, 500–549.

Estes, W. K. (1994). *Classification and cognition* (Oxford Psychology Series, No. 22). Oxford, UK: Oxford University Press.

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

Garner, W. R. (1974). *The processing of information and structure*. New York, NY: Wiley.

Higonnet, R. A., & Grea, R. A. (1958). *Logical design of electrical circuits*. New York, NY: McGraw-Hill.

Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. New York, NY: Wiley.

Luce, R. D. (2003). What ever happened to information theory in Psychology? *Review of General Psychology*, *7*, 183–188.

Medin, D., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, *100*, 254–278.

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

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.

Nosofsky, R. M., Gluck, M. A., Palmeri, T. J., McKinley, S. C., & Glauthier, P. G. (1994). Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Memory and Cognition*, *22*, 352–369.

Nosofsky, R. M., & Zaki, R. S. (2002). Exemplar and prototype models revisited: Response strategies, selective attention, and stimulus generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 924–940.

Pothos, E. M., & Chater, N. (2002). A simplicity principle in unsupervised human categorization. *Cognitive Science*, *26*, 303–343.

Pothos, E. M., Perlman, A., Bailey, T. M., Kurtz, K., Edwards, D. J., Hines, P., & McDonnell, J. V. (2011). Measuring category intuitiveness in unconstrained categorization tasks. *Cognition*, *121*, 83–100.

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, *3*, 382–407.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*, 274–280.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc. Reprinted in E. Margolis & S. Laurence (Eds.). (1999), *Concepts: Core readings* (pp. 189–206). Cambridge, MA: MIT Press.

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

Rosch, E., Mervis, C. B., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382–439.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 379–423.

Shepard, R. (1974). Representation of structure in similarity data: Problems and prospects. *Psychometrika*, *39*, 373–421.

Shepard, R. N. (1987). Towards a universal law of generalization for psychological science. *Science*, *237*, 1317–1323.

Shepard, R. N., Hovland, C. L., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, *75*, 1–42.

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

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

Vigo, R. (2006). A note on the complexity of Boolean concepts. *Journal of Mathematical Psychology*, *50*, 501–510.

Vigo, R. (2009). Categorical invariance and structural complexity in human concept learning. *Journal of Mathematical Psychology*, *53*, 203–221.

Vigo, R. (2010). A dialogue on concepts. *Think*, *9*, 109–120.

Vigo, R. (2011a). Towards a law of invariance in human conceptual behavior. In L. Carlson, C. Holscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the Cognitive Science Society* (pp. 2580–2585). Austin, TX: Cognitive Science Society.

Vigo, R. (2011b). Representation information: A new general notion and measure of information. *Information Sciences*, *181*, 4847–4859.

Vigo, R. (2013a). Complexity over uncertainty in Generalized Representational Information Theory (GRIT): A structure-sensitive general theory of information. *Information*, *4*, 1–30.

Vigo, R. (2013b). The gist of concepts. *Manuscript under review*.

TECHNICAL APPENDIX

A: Representational information theory

In what follows, a well-defined category is a set containing objects that are defined in terms of a preset finite number of dimensions and their values. In Representational Information Theory (Vigo, 2011b) and in its generalization to continuous domains (GRIT; Vigo, 2013a), the amount and quality of subjective information h_s conveyed by a subset R of a well-defined category $\overset{\frown}{F}$ is defined as the percentage change in the structural complexity of $\overset{\frown}{F}$ whenever R is subtracted from $\overset{\frown}{F}$. This is expressed by the following equation.

$$\hbar_s\left(R \middle| \overset{\frown}{F}\right) = \frac{\psi\left(\overset{\frown}{G}\right) - \psi\left(\overset{\frown}{F}\right)}{\psi\left(\overset{\frown}{F}\right)} = \frac{\left|\overset{\frown}{G}\right| e^{-\Phi\left(G\left(\vec{x}\right)\right)} - \left(\left|\overset{\frown}{F}\right| e^{-\Phi\left(F\left(\vec{x}\right)\right)}\right)}{\left(\left|\overset{\frown}{F}\right| e^{-\Phi\left(F\left(\vec{x}\right)\right)}\right)}$$

In the previous equation, $R \subseteq \overset{\frown}{F}$ or $R \in \wp\left(\overset{\frown}{F}\right)$, $\overset{\frown}{G} = \overset{\frown}{F} - R$, $\left|\overset{\frown}{G}\right|$ and $\left|\overset{\frown}{F}\right|$ are the number of elements in $\overset{\frown}{G}$ and $\overset{\frown}{F}$, respectively, and Φ is the degree of categorical invariance of a set of well-defined objects (see Vigo, 2009, 2011a, 2011b, for a detailed discussion of structural complexity and categorical invariance). As seen in the previous equation, the degree of perceived complexity ψ of a well-defined category of objects is defined as the product between the number of objects in the set $\left|\overset{\frown}{F}\right|$ and the negative exponent of its degree of categorical invariance Φ (Vigo, 2011a):

$$\psi\left(\overset{\frown}{F}\right) = \left|\overset{\frown}{F}\right| e^{-\Phi(F(\vec{x}))}$$

The degree of categorical invariance Φ of $\overset{\frown}{F}$ is defined with the following metric:

$$\Phi(F) = \left[\sum_{i=1}^{D}\left[\left\|\frac{\hat{\partial} F(x_1, \ldots, x_D)}{\hat{\partial}\, x_i}\right\|\right]^2\right]^{1/2}$$

where

$$\left\|\frac{\hat{\partial} F(x_1, \ldots, x_D)}{\hat{\partial} x_i}\right\| = 1 - \left[\frac{1}{p}\sum_{\vec{x}_j \in \overset{\frown}{F}.}\left|\frac{\hat{\partial} F\left(\vec{x}_j\right)}{\hat{\partial} x_i}\right|\right]$$

In the previous definition, $\vec{x}$ stands for an object defined by D dimensional values $(x_1, \ldots, x_D)$. The general summation symbol represents the sum of the partial derivatives evaluated at each object $\vec{x}_j$ from the Boolean category $\overset{\frown}{F}$ (this is the category defined by the concept function F). The partial derivative transforms each object $\vec{x}_j$ in respect to its i-th dimension and evaluates to 0 if, after the transformation, the object is still in $\overset{\frown}{F}$ (it evaluates to 1 otherwise). Thus, to compute the proportion of objects that remain in $\overset{\frown}{F}$ after changing the value of their i-th dimension, we need to divide the sum of the partial derivatives evaluated at each object $\vec{x}_j$ by p (the number of objects in $\overset{\frown}{F}$) and subtract the result from 1. The absolute value symbol is placed around the partial derivative to avoid a value of negative 1.

Using this previous equation, the logical or structural manifold (explained in the following section) of a set of objects is defined as:

$$\Lambda(F) = \left(\left\|\frac{\hat{\partial} F(x_1, \ldots, x_D)}{\hat{\partial} x_1}\right\|, \ldots, \left\|\frac{\hat{\partial} F(x_1, \ldots, x_D)}{\hat{\partial} x_i}\right\|, \ldots, \left\|\frac{\hat{\partial} F(x_1, \ldots, x_D)}{\hat{\partial} x_D}\right\|\right)$$

Thus, the metric Φ computes the distance between the 0 logical manifold (i.e., $0 = (0, \ldots, 0)$ and the logical manifold corresponding to the particular concept function F. This distance in ''ideotype'' space represents the degree of perceived structural complexity of the any categorical stimulus defined by the concept function F (Vigo, 2011a, 2013b).

For a detailed and rigorous characterisation and explanation of these measures and concepts see Vigo (2009, 2011a, 2011b). For a gentle introduction to the major concept learning paradigms see Vigo (2010). Although the measures defined above work only on categories defined over binary dimensions, the theory has been generalised to continuous dimensions using generalised invariance structure theory or GIST (Vigo, 2013b). Finally, please note that there is a parameterised version of the degree of subjective complexity ψ of a category (Vigo, 2011a) that uses a discrimination parameter k and pattern sensitivity weights α_i ($i \in \{1, 2, \ldots, D\}$) as follows:

$$\psi\left(\tilde{F}\right) = \left|\tilde{F}\right| \cdot e^{-k\left[\sum_{i=1}^{D}\left[\alpha_i\left\|\frac{\partial F(x_1, \ldots, x_D)}{\partial x_i}\right\|\right]^2\right]^{1/2}}$$

The previous expression without the pattern sensitivity weights but with the single scaling parameter k is referred to by Vigo (2011a, 2013b) as a candidate ''law'' of invariance for human conceptual behaviour.

B: How to compute representational information

Next, we give an example of to how to compute the information values for each of the objects in the fifth category of Figure 1. The reader will find more details and examples in Vigo (2009, 2011a, 2013a, 2013b). We begin by computing the degree of categorical invariance of the stimulus set. First, we can represent this stimulus set by the following concept function:

$$F(x, y, z) = x'yz + x'y'z' + xy'z + xy'z'$$

Or, alternatively, we can encode the dimensional values of the stimulus set using the digits ''1'' and ''0'' such that $x = 1$ = *white*, $y = 1$ = *round flask,* and $z = 1$ = *large* and get $\tilde{F} = \{011, 000, 101, 100\}$. This instance of the earlier concept function, with object features encoded as described, represent the objects of the fifth stimulus of Figure 1. This set can be transformed along the colour dimension by assigning the opposite colour value to all its objects to obtain the *perturbed* set {111, 100, 001, 000}. Comparing the original set to the perturbed set, we observe that they have two objects in common with respect to the dimension of colour. That is, two out of four objects remain the same. This ratio is a measure of the partial invariance of the category with respect to the dimension of colour. Conducting a similar analysis with respect to the dimensions of shape and size, and arranging the values as an ordered set or vector, we get the logical or structural manifold Λ of the previous concept function F.

$$\Lambda(F) = \left(\frac{2}{4}, \frac{0}{4}, \frac{2}{4}\right)$$

Relative degrees of global invariance can then be measured by taking the Euclidean distance of each structural or logical manifold from the zero logical manifold whose components are all zeros (i.e., $(0, \ldots, 0)$). Thus, the overall degree of invariance Φ of the stimulus set is given by:

$$\Phi(F) = \sqrt{\left(\frac{2}{4}\right)^2 + \left(\frac{0}{4}\right)^2 + \left(\frac{2}{4}\right)^2} \approx 0.707$$

Using the degree of categorical invariance, Vigo (2009, 2011a, 2013b) computes the perceived degree of structural complexity ψ of any well-defined category of objects as being directly proportional to its cardinality (i.e., size) and indirectly proportional to the exponent of its degree of invariance as shown later:

$$\psi(\{011, 000, 101, 100\}) = pe^{-\Phi(F)} = 4.\, e^{-0.707} \approx 1.97$$

Now that we have obtained the degree of perceived structural complexity of the stimulus set using the structural complexity measure, we are now in a position to calculate representational information as shown later. In general, a set of objects is informative about a category whenever the removal of its elements from the category increases or decreases the structural complexity of the category as a whole. That is, the amount of representational information (RI) conveyed by a representation R of a well-defined category $\tilde{F}$ is the rate of change of the structural complexity of $\tilde{F}$. Simply stated, the representational information carried by an object or objects from a well-defined category $\tilde{F}$ is the percentage increase or decrease (i.e., rate of change or growth rate) in structural complexity that the category experiences whenever the object or objects are removed.

TABLE A1

R	$\overleftrightarrow{F}$	$\overleftrightarrow{G} = \overleftrightarrow{F} - R$	$\psi(\overleftrightarrow{G})$	$\hbar_s(R\vert\overleftrightarrow{F})$
{011}	{011, 000, 101, 100}	{000, 101, 100}	1.17	−0.41
{000}	{011, 000, 101, 100}	{011, 101, 100}	1.54	−0.22
{101}	{011, 000, 101, 100}	{011, 000, 100}	1.54	−0.22
{100}	{011, 000, 101, 100}	{011, 000, 101}	3.00	0.52

In Representational Information Theory (Vigo, 2011b) and in its generalization to continuous domains (GRIT; Vigo, 2013a), the amount and quality of subjective information $\hbar_s$ conveyed by a subset R of a well-defined category $\overleftrightarrow{F}$ is defined as the percentage change in the structural complexity of $\overleftrightarrow{F}$ whenever R is subtracted from $\overleftrightarrow{F}$. This is expressed by the following equation:

$$\hbar_s\left(R|\overleftrightarrow{F}.\right) = \frac{\psi\left(\overleftrightarrow{G}\right) - \psi\left(\overleftrightarrow{F}\right)}{\psi\left(\overleftrightarrow{F}\right)} = \frac{\left|\overleftrightarrow{G}\right| e^{-\Phi\left(G\left(\overrightarrow{x}\right)\right)} - \left(\left|\overleftrightarrow{F}\right| e^{-\Phi\left(F\left(\overrightarrow{x}\right)\right)}\right)}{\left(\left|\overleftrightarrow{F}\right| e^{-\Phi\left(F\left(\overrightarrow{x}\right)\right)}\right)}$$

In the previous equation, $R \subseteq \overleftrightarrow{F}$ or $R \in \wp\left(\overleftrightarrow{F}\right)$ $\overleftrightarrow{G} = \overleftrightarrow{F} - R$, $\left|\overleftrightarrow{G}\right|$ and $\left|\overleftrightarrow{F}\right|$ are the number of elements in $\overleftrightarrow{G}$ and $\overleftrightarrow{F}$, respectively, and Φ is the degree of categorical invariance of a set of well-defined objects.

Using this we can calculate the information conveyed by the first element, that is the subset $R = \{011\}$, of the set $\overleftrightarrow{F} = \{011, 000, 101, 100\}$. The set $\overleftrightarrow{G}$ is now defined using set subtraction as

$$\overleftrightarrow{G} = \overleftrightarrow{F} - R = \{011, 000, 101, 100\} - \{011\} = \{000, 101, 100\}$$

The logical manifold of the set $\overleftrightarrow{G}$ is obtained as described earlier. We transform the set along each of its three dimensions and compare the perturbed set with the original to observe the number of objects that are common between the two sets. The logical manifold is:

$$\Lambda(G) = \left(\frac{2}{3}, \frac{0}{3}, \frac{2}{3}\right)$$

Using the logical manifolds we now obtain the degree of categorical invariance as

$$\Phi(G) = \sqrt{\left(\frac{2}{3}\right)^2 + \left(\frac{0}{3}\right)^2 + \left(\frac{2}{3}\right)^2} \approx 0.94$$

and subsequently obtain the degree of perceived structural complexity as

$$\psi\left(\overleftrightarrow{G}\right) = pe^{-\Phi(G)} = 3.\, e^{-0.943} \approx 1.17$$

Substituting the values of $\psi\left(\overleftrightarrow{F}\right)$ and $\psi\left(\overleftrightarrow{G}\right)$ in the previous equation gives us the information conveyed by the object $R = \{011\}$ about its set of origin $\overleftrightarrow{F} = \{011, 000, 101, 100\}$, that is,

$$\hbar_s\left(R|\overleftrightarrow{F}\right) = \frac{\psi\left(\overleftrightarrow{G}\right) - \psi\left(\overleftrightarrow{F}\right)}{\psi\left(\overleftrightarrow{F}\right)} = \frac{1.168 - 1.97}{1.97} = -0.4071 \approx -0.41$$

Repeating the earlier analysis, we can obtain the amount of information conveyed by all possible single element representations of $\overleftrightarrow{F}$. Table A1 shows these information values.

C: Prototypicality measure underlying the multiplicative prototype model

The MPM assumes that categories are represented by stored exemplars in memory. The classification of an item starts by a computation of the weighted distance between the Item i in Category C to be classified and a stored prototype A as shown in Equation 1.

$$d(i, A) = \sum_k w_k \cdot |x_{ik} - P_{Ak}| \quad (1)$$

where x_{ik} and P_{Ak} take the value 0 or 1 depending on the dimension k. The value of w_k ($0 \leq w_k \leq 1$, $\sum_k w_k = 1$) represents an attention weight on the differences between dimensions. The weighted distances are then transformed to similarity judgements by the negative exponent (exponential decay) function (Shepard, 1974, 1987) shown in Equation 2. The scaling parameter indicates overall discriminability between items in the space and yields different overall sensitivity gradients for similarities (i.e., rates at which similarity declines with distance). This parameter is ideally equal to the number of dimensions used to define the set of objects (see Nosofsky, 1984 & 1986, for a discussion).

$$S(i, A) = e^{-v\left[\sum_k w_k \cdot |x_{ik} - P_{Ak}|\right]} \quad (2)$$

For our computations, we set the discriminability parameter v to 3 and, as in the Technical Appendix of Nosofsky (1986), the attention weights are set to .33. The assumption being that attention resources are more or less evenly distributed across the three dimensions; this is a valid assumption considering that the random reassignment of dimensions and dimensional values in our experiments counterbalance attention biases. This is equivalent to cancelling out the influence of parameter values in the model. This allows us to compare Equation 2 to the parameter-free RIT. Finally, note that the MPM uses a response rule (Luce's choice rule; Luce, 1959) to translate degrees of prototypicality into response probabilities as follows (Nososfky & Zaki, 2002):

$$P(A|i) = \frac{S(i, A)}{S(i, A) + S(i, B)}$$

In the previous expression, B is the contrasting or complementary category. Although it is possible to operationalise degrees of prototypicality in terms of response probabilities using this response rule, we advise against it for two reasons: First, the response probability of classifying an item correctly is a significantly different notion from that of degree of prototypicality; second, it is not as direct a measure as ''degrees of perceived similarity to the prototype''. Regardless, operationalising prototypicality in terms of the response probabilities rule yields the same effect sizes reported for the MPMP-1 and MPMP-2 without gain in intuition and validity.

D: Regression analysis of RIT, FFPM, MPMP-1, and MPMP-2

TABLE A2
Experiment 1: Type III

	RIT	*FFPM*	*MPMP-1*	*MPMP-2*
R^2	.99	.99	.99	.45
Std. error	0.0044	0.0044	0.0044	0.2608
p	<.0001	<.0001	<.0001	.3246
F	$F(1, 2) = 13061$	$F(1, 2) = 13061$	$F(1, 2) = 13061$	$F(1, 2) = 1.677$
$y = mx + b$	$m = 2.69$, $b = -0.22$	$m = 1.00$, $b = -0.23$	$m = 1.00$, $b = -0.15$	$m = 0.5$, $b = 0.05$

TABLE A3
Experiment 1: Type IV

	RIT	*FFPM*	*MPMP-1*	*MPMP-2*
R^2	.99	.99	.99	.99
Std. error	0.0024	0.0024	0.0024	0.0024
p	<.0001	<.0001	<.0001	<.0001
F	$F(1, 2) = 122840$	$F(1, 2) = 122840$	$F(1, 2) = 122840$	$F(1, 2) = 122840$
$y = mx + b$	$m = 1.07$, $b = -0.07$	$m = 1.96$, $b = -0.96$	$m = 1.63$, $b = -0.63$	$m = 1.47$, $b = -0.47$

TABLE A4
Experiment 1: Type V

	RIT	*FFPM*	*MPMP-1*	*MPMP-2*
R^2	.99	.22	.22	.97
Std. error	0.0216	0.3067	0.3067	0.0501
p	.0019	.5244	.5244	.0104
F	$F(1, 2) = 519$	$F(1, 2) = 0.5848$	$F(1, 2) = 0.5848$	$F(1, 2) = 94.78$
$y = mx + b$	$m = 0.82$, $b = 0.04$	$m = 0.54$, $b = 0.06$	$m = 0.68$, $b = 0.06$	$m = 0.72$, $b = -0.05$

TABLE A5
Experiment 2: Type III

	RIT	*FFPM*	*MPMP-1*	*MPMP-2*
R^2	.99	.99	.99	.47
Std. error	0.009	0.009	0.009	0.117
p	0.0016	0.0016	0.0016	0.3144
F	$F(1, 2) = 625$	$F(1, 2) = 625$	$F(1, 2) = 625$	$F(1, 2) = 1.77$
$y = mx + b$	$m = 1.22$, $b = 0.10$	$m = 0.45$, $b = 0.10$	$m = 0.45$, $b = 0.13$	$m = 0.23$, $b = 0.23$

TABLE A6
Experiment 2: Type IV

	RIT	*FFPM*	*MPMP-1*	*MPMP-2*
R^2	0.99	0.99	0.99	0.99
Std. error	0.018	0.018	0.018	0.018
p	<.001 (.0005)	<.001 (.0005)	<.001 (.0005)	<.001 (.0005)
F	$F(1, 2) = 1875$	$F(1, 2) = 1875$	$F(1, 2) = 1875$	$F(1, 2) = 1875$
$y = mx + b$	$m = 0.99$, $b = 0.01$	$m = 1.82$, $b = -0.82$	$m = 1.52$, $b = -0.52$	$m = 1.37$, $b = -0.37$

TABLE A7
Experiment 2: Type V

	RIT	*FFPM*	*MPMP-1*	*MPMP-2*
R^2	.85	.66	.66	.90
Std. error	0.14	0.21	0.21	0.113
p	.078	.1835	.1835	.049
F	$F(1, 2) = 11.34$	$F(1, 2) = 4$	$F(1, 2) = 4$	$F(1, 2) = 18.79$
$y = mx + b$	$m = 0.79$, $b = 0.15$	$m = 0.97$, $b = 2.75\text{E}{-17}$	$m = 1.21$, $b = -1.45\text{E}{-16}$	$m = 0.72$, $b = 0.05$

Erratum to "Will the most informative object stand? Determining the impact of structural context on informativeness judgments"

The y-axes of figures 3 and 6 are mislabeled. Instead of percentages, they should display values on the [0, 1] interval. For figure 3, these values were computed on the [0, 1] interval using the min-max normalization procedure (i.e., (z-min)/(max-min)) on the maximum and minimum number of times an object with an equivalent structural role (i.e., with the same relationship to the other objects in its category of origin) was chosen per category structure across all subjects and categories. The error bars are then valid. Regardless of which rescaling is used (percentages or the aforementioned approach), all the reported statistical results, qualitative results, and response patterns remain the same.

Likewise, the y-axis of figure 6 is mislabeled with percentages. For figure 6, these values should also be in the [0, 1] interval using the min-max normalization procedure. The values indicate the degree of perceived informativeness of objects with an equivalent structural role (i.e., with the same relationship to the other objects in the category). This time, the max and min values are the maximum and minimum average informativeness judgment scores across all subjects and categories. Again, all the reported statistical results, qualitative results, and response patterns remain the same.

In either case, we could have scaled within categories instead of using a max and a min across all categories. The latter procedure was selected because it does not change any of the reported statistical results, qualitative results, nor response patterns. Yet, it offers the advantage of making all the graphs uniform, easy to read and comparable to each other.