Debunking the Basic Level

Frédéric Gosselin (GOSSELIF@PSY.GLA.AC.UK)
Philippe G. Schyns (PHILIPPE@PSY.GLA.AC.UK)
Department of Psychology, University of Glasgow
56 Hillhead St., Glasgow G12 8QB UK

Abstract

The goal of this paper is to introduce a new measure of basic-level performance that we will call the "category attentional slip." The idea behind it is very simple: The attentional mechanisms of an ideally rational categorizer are made to "slip" once in a while. We provide a formalization of attentional slip that specifies what an "ideally rational categorizer" is and how its attention "slips." We then compare its predictive capabilities with those of two established basic-level measures: category feature-possession (Jones, 1983) and category utility (Corter & Gluck, 1992). The empirical data used for the comparisons are drawn from eight classical experiments from Murphy and Smith (1982), Murphy (1991), and Tanaka and Taylor (1991).

Real-world "things" may have a number of different names. For example, Scully from the *X-files* television series, is a medical doctor, an FBI agent, a redhead, a female, the partner of agent Mulder, a physician, a creation of Chris Carter, a character portrayed by Gillian Anderson, and so forth. All these names refer to different categories which can share a subset of their members, and so Scully is not the only redheaded FBI agent, and not all FBI agents have red hair (e.g., agent Mulder has brown hair). In this paper, we will not be concerned with all possible categorizations of a single object. Instead, we will concentrate on the idea that categories can be hierarchically organized, so that Scully is a doctor and a human being. Embedded categories are said to denote different *levels of categorization*.

In a seminal paper, Rosch, Mervis, Gray, Johnson, and Boyes-Braem (1976) distinguished three of these levels: the subordinate (*sparrow, BMW*), the basic (*bird, car*) and the superordinate (*animal, vehicle*). They showed that of these levels, the basic was superior in many respects: People tend to designate an object with its basic-level category name; throughout development, basic level names are learned before those of other categorization levels; basic names tend to be shorter and used more frequently than those of other categories; people tend to many more features at the basic level than at the superordinate level, with only a slight increase at the subordinate level; people decide more rapidly that an object belongs to a basic category than to all the other categories of a hierarchy (see also Murphy, 1991; Murphy & Smith, 1982; Tanaka & Taylor, 1991).

Experiments on the basic level have typically probed three embedded categorization levels, using only one or two measures of performance (e.g., response times and feature listings). Obviously, people can often use many more than three levels in their interactions with objects. Berlioz, for instance, was an artist, a human, a mammal, a living organism, a bunch of atoms, and so forth. However, basic-level performance can only exist with respect to the other categorization levels that are probed. Hence, we believe it is more appropriate to speak of the *basic-levelness* (which is a measure of performance) of a level of categorization than to consider the basic level as an absolute level of a categorization hierarchy (Murphy, 1991; Schyns, 1996).

Even though the basic level is important in current theories of object categorization (Murphy, 1991; Murphy & Smith, 1982; Rosch et al., 1976; Tanaka & Taylor, 1991) and recognition (Biederman, 1987), no model of basic-level performance can account for all existing evidence of the basic-level effect. It is the purpose of this research to propose a new and better model of basic-levelness that we call "category attentional slip." The fundamental ideas behind the attentional slip measure are quite simple. We begin with an "ideal" categorizer that performs series of tests on features to decide whether an object belongs to a category. Then, we add noise to the attentional mechanism of this ideal categorizer so that it "slips" once in a while.

We compare the category attentional slip's ability to account for empirical data with the predictions of two established models: category feature-possession (Jones, 1983) and category utility (Corter & Gluck, 1992).

Measures of Basic-Levelness

In this section, we first present a very simple category structure (see Figure 1). We then use this structure to explain category feature possession, category utility, and category attentional slip, respectively.

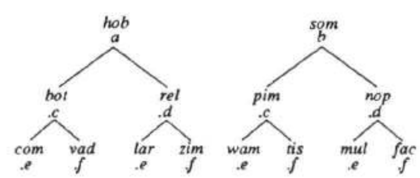

Figure 1: The category structure used to explain category attentional slip, category feature possession, and category utility.

Underneath the category names (e.g., *som, pim, zim*), the letters correspond to the features that define a level of categorization. For example, *a* defines the superior level *hob*, *ac* defines the middle level *bot*, and *ace* define the lower level *com* (the points correspond to the features inherited from the related level[s] above the considered categorization level). Objects are evenly distributed among categories located at a given level of categorization. It is worth mentioning at this stage that the category organization of Figure 1 was chosen for its simplicity, but also to illustrate a general aspect of category organizations that will become particularly important in attentional slip.

Category feature-possession

Jones (1983) proposed that the basic level is the level of categorization where the average *category feature-possession* is maximal. The category feature-possession of a category c_i (B_i) is defined--for a given set of objects composed of n features and for m categories--as the sum of all the b_{ij}, i.e. $B_i = \sum b_{ij}$ over the n features. If $K_{ij} = \max(K_{1j}, K_{2j}, ..., K_{mj})$, b_{ij} is equal to w_j ($w_j = 1$, usually); else b_{ij} is equal to zero. And K_{ij}, the *collocation* of a category c_i and of a feature f_j, corresponds to $K_{ij} = P(c_i|f_j)P(f_j|c_i)$. Finally, $P(c_i|f_j)$ and $P(f_j|c_i)$ are the probability that the object belongs to c_i given that it possesses f_j, and the probability that the object possesses feature f_j given that it belongs to c_i, respectively.

Category feature possession is a four-step process. Consider the category organization of Figure 1 to illustrate the computations. First, we must compute $P(f_j|c_i)$ and $P(c_i|f_j)$ for $i, j \in \{a, b, c, d, e, f\}$. For example, both $P(d|hob)$ and $P(hob|d)$ are equal to .5. Second, we calculate all the collocations. The collocation of category *hob* and feature *d*, for instance, is equal to $P(d|hob)P(hob|d)$, that is .25 (see Table 1 for a listing of all collocations). Third, we locate the largest collocation for every feature in the columns of Table 1. For example, the largest collocations for feature *d*, are equal to .5 (see the underlined figures in Table 1). Finally, a count of the number of underlined figures provide the category feature-possession measure (see the rightmost column of Table 1). For the category organization of Figure 1, feature-possession predicts that reaction times (RT) should be fastest at the higher level of categorization, and that they should be equally slow at the middle and lower levels.

Table 1: Key computations for the numerical simulation of the category feature-possession (Jones, 1983) with the category organization of Figure 1.

Category	Feature *a*	*b*	*c*	*d*	*e*	*f*	Σ
hob	1	0	.25	.25	.25	.25	3
som	0	1	.25	.25	.25	.25	3
bot	.5	0	.5	0	.125	.125	1
rel	.5	0	0	.5	.125	.125	1
pim	0	.5	.5	0	.125	.125	1
nop	0	.5	0	.5	.125	.125	1
com	.25	0	.25	0	.25	0	1
vad	.25	0	.25	0	0	.25	1
lar	.25	0	0	.25	.25	0	1
zim	.25	0	0	.25	0	.25	1
wam	0	.25	.25	0	.25	0	1
tis	0	.25	.25	0	0	.25	1
mul	0	.25	0	.25	.25	0	1
fac	0	.25	0	.25	0	.25	1

Category utility

Corter and Gluck's (1992) category utility measure has a solid logic of construction. For these authors, a category is useful to the extent that it improves the capacity to correctly predict the features of a member of this category. Suppose m features (f_k) describe exemplars. Knowing only $P(f_k)$ (the probability that an object possesses feature f_k) a raw, uninformed *probability-matching* strategy enables to guess that a given object possesses f_k with a probability $P(f_k)$. The probability that this guess is correct is $P(f_k)^2$. However, the prediction might be significantly enhanced if one knew that the object belonged to category c. Category utility measures this gain between an informed and an uninformed prediction of object features. We have already seen what the uninformed guess was, let us now turn to the informed guessing.

Formally, the $P(f_k|c)$ is the prior probability that an object possesses feature f_k given that it belongs to category c. The probability that this guess is correct is $P(f_k|c)^2$. Another prior information, $P(c)$ is the probability that the considered object effectively belongs to c. Thus, the expected increase in predictive power that the object possesses feature f_k from the knowledge that the input belongs to c is given by $P(c)[P(f_k|c)^2 - P(f_k)^2]$. The sum of this expected increase over the m features describing the input object is the *category utility* of c

$$P(c)\sum_{k=1}^{m}\left[P(f_k|c)^2 - P(f_k)^2\right].$$

The average of the category utilities should be maximal at the basic level.

Category utility is also a four-step computation. Consider, again, the category organization of Figure 1 to illustrate the computations. We start with a computation of all $P(c)$s, $P(f_k)$s, and $P(f_k|c)$s. The $P(c)$s are respectively equal to .5, .25, and .125 for the higher, middle, and lower levels of categorization. The $P(f_k)$s are all equal to .5. The $P(f_k|c)$s are easy to compute. For example, $P(a|hob)$ is 1, $P(b|hob)$ is 0, and $P(c|hob)$ is .5. Next, we subtract the squares of the $P(f_k|c)$s from the squares of the $P(f_k)$s. $[P(a|hob)^2 - P(a)^2]$, for instance, is equal to .75 (Table 2 summarizes all these differences).

Table 2: Key computations for the numerical simulation of the category utility measure (Corter & Gluck, 1992) with the category organization of Figure 1.

Category	Feature *a*	*b*	*c*	*d*	*e*	*f*	Σ
hob	.75	-.25	0	0	0	0	0.5
som	-.25	.75	0	0	0	0	0.5
bot	.75	-.25	.75	-.25	0	0	1
rel	.75	-.25	-.25	.75	0	0	1
pim	-.25	.75	.75	-.25	0	0	1
nop	-.25	.75	-.25	.75	0	0	1
com	.75	-.25	.75	-.25	.75	-.25	1.5
vad	.75	-.25	.75	-.25	-.25	.75	1.5
lar	.75	-.25	-.25	.75	.75	-.25	1.5
zim	.75	-.25	-.25	.75	-.25	.75	1.5
wam	-.25	.75	.75	-.25	.75	-.25	1.5
tis	-.25	.75	.75	-.25	-.25	.75	1.5
mul	-.25	.75	-.25	.75	.75	-.25	1.5
fac	-.25	.75	-.25	.75	-.25	.75	1.5

Then, we sum all these differences across categories (i.e., the rows in Table 2). These sums appear in the rightmost column of Table 2. Finally, we obtain the category utilities by weighting each sum by the appropriate $P(c)$. For the higher and the middle level categories, category utility is equal to 0.25; for the lower level categories, it is equal to about 0.188. Thus, category utility predicts that the highest and middle categorization levels are the most basic in the considered organization.

Category attentional slip

To explain what we believe to be the first determinant of basic-levelness (*cardinality*), consider the typical category organization that has elicited a basic-level advantage (see Figure 2 from Murphy & Smith, 1982, Experiment 1). To place the featural description of an unknown object X in a category of the hierarchy, people need to test whether the features defining the category characterize the input. For example, in Figure 2, "does X possess a?", is a test to check that the object X is a *hob*, and "does X possess c?", "does X possess d?", and "does X possess e?" all test that the input is a *bot*. It is important to note, however, that in Figure 2, testing *either* c, d, or e is sufficient to determine the category membership of the object, the other tests are redundant. More generally, for a given categorization level, two tests are *redundant* iff one test can substitute for the other in every possible identification tasks. In Figure 2, the category structure is such that three redundant tests define each middle-level category while a single test defines each higher-level and lower-level categories. Henceforth, we will call s the set of all redundant tests associated with a categorization. For example, s_1 = {"does X possess a?"}, s_2 = {"does X possess c?," "does X possess d?," "does X possess e?"}, and s_3 = {"does X possess o?"} are three sets of redundant tests associated with *hob*, *bot* and *com*, respectively. The *cardinality* of each of these sets is the number of redundant tests it contains (e.g., the cardinality of s_2 is 3).

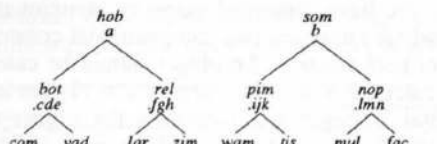

Figure 2: Murphy & Smith's (1982, Experiment 1) categorical structure.

We believe that the second and last determinant of basic-level performance is the *length* of the *optimal strategy* required to reach a categorical decision. This has so far been completly neglected in experiments on the basic level. To illustrate, consider the strategy (S) one could adopt to decide that the input is a *bot* for the category structure of Figure 2. One strategy could be: S = {s_1 = {"does X possess a?"} and s_2 = {"does X possess c?," "does X possess d?," "does X possess e?"}}, where each s_i is a complete set of redundant tests as defined earlier, and where the s_is are performed in a specific order. However, this strategy is far from being *optimal*: a strategy including s_2 alone suffices to determine that the input is a *bot*. Thus, the added features of only one level of the hierarchy could be checked to decide whether the object belongs to this level--in fact, this applies to all three levels of Figure 2. When only the added features of one level of a category structure need to be checked to determine a categorization, the strategies are of length 1.

Most basic-level experiments had categories that requiered only length 1 strategies (see, e.g., Murphy & Smith, 1982; Murphy, 1991; Tanaka & Taylor, 1991). Note, however, that length 1 strategies are too constraining for many real world categorizations. Features do tend to overlap between categories. For example, consider the following cars: a *blue Tercel*, a *blue 911*, and an *orange Tercel*. To identify a *blue Tercel*, one needs to perform two tests: S = {s_1 = {"is the input a *Tercel*?"} and s_2 = {"is the input *blue*?"}}. The hierarchy of Figure 1 illustrates such situation of feature overlap. Although a strategy of length 1 was sufficient to determine that the input was a *bot* in Figure 2, a *bot* categorization needs to test added features of two levels in Figure 1: S = {s_1 = {"does X possess a?"} and s_2 = {"does X possess c?"}}, in a specific order. However, in Figure 1, a strategy of length 1 is still sufficient for the higher level categories (e.g., S = {s = {"does X possess a?"}} to decide that the input is a *hob*), but a strategy examining the added features of all three levels (length 3) is required to decide that the input is a *com*: S = {s_1 = {"does X possess a?"}, s_2 = {"does X possess c?"}, and s_3 = {"does X possess e?"}}.

We now turn to an implementation, *category attentional slip*, that integrates these two determinants of basic-levelness. Suppose an ideal categorizer: a formal model which systematically uses an optimal strategy to decide whether the input belongs to a category. The structure of categories drives its behavior so that it excecutes the smallest series of n sets of redundant tests to arrive at a given categorization. Within each set of the series, only one of the redundant features is tested. Suppose that response time is proportional to the *length of the optimal strategy*. Suppose further that the ideal categorizer has a perfectible attention that "slips" off its rational track with a probability p; it then selects randomly a feature and tests whether the input possesses this feature. This slippage introduces noise that should in principle affect the number of tests required to reach a category decision. Note, however, that the slip does not affect equally all categories. Everything being equal, *low cardinality categories* (those with low feature redundancies such as the low- and high-levels in Figure 2) have fewer chances that attention randomly slips to a relevant feature than *high cardinality categories* (such as the mid-level of Figure 2).

To be more specific, let us first consider cardinality in the simple case of an optimal strategy of length 1, as is needed to decide whether an object is a *bot* in Figure 2. Category attentional slip is related to the number of trials (t) required to complete the strategy. Because the model is stochastic, the measure is t_mean, the average number of trials needed to complete the strategy. So we must derive the probability distribution of t, and compute its mean. To obtain the probability distribution we must address the question: "What is the probability that one test of a single s is performed after t trials?" Recall that p is the probability that attention slips randomly to one feature. When attention slips, it can slip to a relevant feature and perform a relevant test with probability pq, where q is the probability that one relevant test is performed by chance alone--it is the cardinality of the complete set of redundant tests divided by the total number of features of the category organization (for *bot* in Figure 2, e.g., q = 3/22). Hence, (p-pq) is the probability that an irrelevant test is performed. The probability that a relevant test is performed is simply 1 minus the probability that an irrelevant test is performed. Thus, (p-pq)[1-(p-pq)] is the probability that attention slips to an irrelevant feature on the first trial and then performs a relevant test on the second trial, in a length 1 strategy. Generalizing the probability to t trials is now straightforward. We combine the probability that a relevant test has not been performed during the first t-1 trials with the probability that it occurs on trial t: $(p-pq)^{t-1}[1-(p-pq)]$. This defines independent probabilities because two tests are never performed simultaneously. Moreover, between t = 1 and t = +∞ we find all the possible realizations of our critical tests. Thus, we can conceive this formula as the frequency distribution of the trials t. This implies that *t_mean* is equal to

$$\frac{\sum_{t=1}^{+\infty} t(p-pq)^{t-1}[1-(p-pq)]}{\sum_{t=1}^{+\infty}(p-pq)^{t-1}[1-(p-pq)]}.$$

(Note the factor t at the onset of the numerator.) However,

$$\sum_{t=1}^{+\infty}(p-pq)^{t-1}[1-(p-pq)] = 1.$$

Therefore *t_mean* is equal to

$$\sum_{t=1}^{+\infty} t(p-pq)^{t-1}[1-(p-pq)].$$

Let us apply this equation to the higher level category *hob* in Figure 1. From now on, we abitrarly set p = .5. q is equal to 1/6, the cardinality 1 of the s required to achieve a decision divided by the total number of features, that is 6. Thus, (p-pq) is equal to about .417, and applying the equation with these parameters yields a *t_mean* of about 1.714 (i.e., $[1 * .417^0 * (1 - .417)] + [2 * .417^1 * (1 - .417)] + [3 * .417^2 * (1 \quad .417)] + ... = 1.714$). This signifies that the average number of trials needed to decide that a stimulus is a *hob* is 1.714--in fact, this is true of all higher level categories because they share the same q.

As explained earlier, many real-world categorizations will involve optimal strategies of lengths longer than 1. We now turn to the formal expression of *t_mean* for optimal strategies of length 2 (e.g., *hob*, in Figure 1). We start again with the question: "What is the probability that a s_1 test and a s_2 test are performed in that order after t trials (with $t \geq 2$)?" By definition of a strategy (an ordered series of sets of redundant tests), one s_1 test must occur at least once in the first t-1 trials if one s_2 test occurs on trial t. During the interval between the completion of a critical s_1 test on trial i and the critical s_2 test on trial t, no s_2 test occurs. Thus, the probability that a length 2 strategy is completed after t trials is

$$\sum_{i=1}^{t-1}(p-pq_1)^{i-1}[1-(p-pq_1)]$$
$$(p-pq_2)^{t-i-1}[1-(p-pq_2)]$$

where q_j is the probability that one of the tests of s_j is performed by chance alone, in a successfull slip. By the same reasoning as earlier, we can compute *t_mean*

$$\sum_{t=2}^{+\infty} t\sum_{i=1}^{t-1}(p-pq_1)^{i-1}[1-(p-pq_1)]$$
$$(p-pq_2)^{t-i-1}[1-(p-pq_2)].$$

Let us apply this equation to the middle level category *bot* of the category structure of Figure 1. Again, the q_js are equal to 1/6. Applying the last equation yields a *t_mean* of about 3.429, which is also the average of all mid-level categories in Figure 1.

This previous formula can be generalized to any optimal strategy of length n, but for the present purpose, we need to generalize it only to length 3 strategies (which apply to the lower-level categories of Figure 1):

$$\sum_{t=3}^{+\infty} t\sum_{j=1}^{t-i-1}\sum_{i=1}^{t-2}(p-pq_1)^{i-1}[1-(p-pq_1)]$$
$$(p-pq_2)^{j-1}[1-(p-pq_2)]$$
$$(p-pq_3)^{t-1-i-j}[1-(p-pq_3)]$$

To illustrate, consider the application of this equation to the lower level category *com* of the category structure of Figure 1. Once more, all the q_ks are equal to 1/6. It follows that *t_mean* is equal to about 5.142. Because all the lower level categories of the category structure of Figure 1 share the same q the mean *t_means* for lower level categories is also equal to 5.142.

In sum, we have presented category attentional slip, a measure which integrates two computational constraints on basic-level performance: An object should be categorized faster in category X than in category Y if (1) the length of the optimal strategy that identifies the object as X is smaller--all other things being equal--than the length of the optimal strategy that identifies the same object as Y, and if (2) the cardinalities of the sets of redundant tests (or some of them) is larger--all other things being equal--for category X than for category Y.

Comparison of the Basic-Levelness Measures

We now compare the performance of category attentional slip, category feature-possession, and category utility using as benchmarks the results of eight categorization experiments drawn from Murphy and Smith (1982), Murphy (1991), and Tanaka and Taylor (1991). The results of these experiments were gathered on minor variations of Murphy and Smith's generic procedure: Subjects were initially taught the names of objects at three levels of categorization. In a later testing phase, they were shown a picture of a stimulus together with a category name. Subjects' task was to verify that the stimulus was a member of the named category.

Murphy & Smith (1982)

In their Experiment 1, Murphy and Smith (1982) used 16 artificial tools. Their tools were either pounders, or cutters (higher level); they had non-overlapping handles, shafts, and heads (which defined the middle level); and they had one more non-overlapping feature like big or small head, and one or two-parts handle (at the lower level). This category structure mirrors the one of Figure 2. As shown in Table 3, mid-level categories were the fastest with the high-level categories being the slowest.

The category feature-possession and the attentional slip measures correctly predicted the basic-levelness order of two out of three levels of categorization. The category utility measure did slightly worse: it correctly predicted only the basic level (see underlined figures in Table 3). (Recall that both the category feature-possession and the category utility scores should be inversely proportional to the RTs, and that the attentional slip scores should be directly proportional to the RTs.)

In their Experiment 3, Murphy and Smith (1982) used eight of the artificial tools of their Experiment 1, and added eight new tools to produce a total of sixteen. Their artificial tools were either large, or small (higher level); they were either pounders, cutters, scraper, or stirrer (middle level); and they had non-overlapping handles, shafts, and heads (lower level). As shown in Table 3, the lower level categories were the fastest with the middle level categories being the slowest.

The category feature-possession and the attentional slip measures correctly predicted the speed order of two out of three levels of categorization. The category utility measure only predicted the rank of the level of categorization with the second highest basic-levelness measure of performance.

Table 3: Mean values of category utility, feature-possession, and attentional slip score measures (with mean "true" trial identification reaction times) for various categorical structures.

Source	Model	Level		
		Lower	Middle	Higher
Murphy & Smith, Exp. 1	Observation	723 ms	678 ms	879 ms
	Feature-possession	1	3	1
	Category utility	0.453	0.781	0.688
	Attentional slip	1.913	1.76	1.913
Murphy & Smith, Exp. 3, Size	Observation	574 ms	882 ms	666 ms
	Feature-possession	3	1	1
	Category utility	0.483	0.591	0.561
	Attentional slip	1.818	1.935	1.935
Murphy, Exp. 3	Observation	776 ms	688 ms	779 ms
	Feature-possession	1	3	1
	Category utility	0.453	0.781	0.688
	Attentional slip	1.913	1.76	1.913
Murphy, Exp. 4, Simple	Observation	862 ms	811 ms	980 ms
	Feature-possession	1	3	1
	Category utility	0.453	0.781	0.688
	Attentional slip	1.913	1.76	1.913
Murphy, Exp. 4, Enhanced	Observation	1,132 ms	854 ms	955 ms
	Feature-possession	1	5	1
	Category utility	0.640	1.156	0.938
	Attentional slip	1.935	1.714	1.935
Murphy, Exp. 5	Observation	1,072 ms	881 ms	854 ms
	Feature-possession	1	3	4
	Category utility	0.641	1.156	1.438
	Attentional slip	1.931	1.806	1.75
Tanaka & Taylor, Novice	Observation	778 ms	678 ms	746 ms
	Feature-possession	7	12	8
	Category utility	2.387	3.898	3.934
	Attentional slip	1.890	1.818	1.875
Tanaka & Taylor, Expert	Observation	622 ms	623 ms	729 ms
	Feature-possession	10	10	8
	Category utility	2.526	3.803	3.870
	Attentional slip	1.863	1.863	1.889

Murphy (1991)

In Experiment 3, Murphy (1991) used 16 artificial abstract objects of various colors, textures, types of edge, and sizes. In fact, Murphy's categorical structure was identical to Murphy and Smith's (1982, Experiment 1). As shown in Table 3, the middle level categories were the fastest with the two other being about equally slow.

The category feature-possession and the attentional slip measures correctly predicted the whole basic-levelness sequence. The category utility measure did a bit worse, predicting two out of three level of categorization basic-levelness order.

In Experiment 4, Simple Condition, Murphy (1991) replicated Murphy and Smith (1982, Experiment 1). Again, the middle level categories were the fastest with the higher level categories being the slowest (see Table 3). Needless to say, the respective merits of the three measures were the same as before.

Murphy's (1991) Experiment 4, Enhanced condition, and Experiment 5 both used enhanced versions of the 16 artificial tools of Murphy and Smith (1982, Experiment 1). In Experiment 4, Enhanced condition, eight non-overlapping features (i.e., either red dots, yellow circles, green stripes, or blue solid color) were evenly added to the categories at the middle level of categorization. The middle level categories were the fastest with the lower level of categorization being the slowest (see Table 3). In Experiment 5, 16 non-overlapping features (colors and texture cues) were evenly added to categories at the lower level of categorization. The higher level categories were the fastest with the lower level categories being the slowest (see Table 3).

In Experiment 4, Enhanced Condition, the category utility measures scored a perfect three (it is the only case where the category utility measure does better than the two other measures). Whereas the category feature-possession and the attentional slip measures correctly predicted the order of two of the three levels of categorization basic-levelness. The three measures correctly predicted the whole sequence of basic-levelness in Experiment 5.

Tanaka & Taylor (1991)

So far we have used the construction features of objects to build the category structures. Tanaka and Taylor (1991) used natural objects in their experiments; no one knows the "true" construction features of natural objects, but subjects might very well have used the features they listed for the categories. There is some empirical evidence for this. In Experiment 1A, Murphy (1991) asked some subjects to list features for the artificial objects used in Murphy, Experiment 3. He found a mean of 1 feature at the higher level of categorization, a mean of 5.75 features added at the middle level of categorization, and a mean of 0.87 feature added at the lower level of categorization (most of them did not overlap with listed features of contrasting categories) (cf.

the "true" numbers of added features were 1, 3, and 1, at the higher, middle, and lower levels of categorization, respectively). It, thus, seems that we could use such estimates for our simulations.

Tanaka and Taylor's (1991) subjects were taught the names of 16 natural animals at three levels of categorization (e.g., *animal, dog, Beagle*). The subjects were either bird experts and dog novices, or dog experts and bird novices. In Experiment 1, Tanaka and Taylor found that novices listed approximately 8, 12, and 7 new features for the higher, middle, and lower levels of categorization, respectively; and that experts listed approximately 8, 10, and 10 new features for the higher, middle, and lower levels of categorization, respectively (we have rounded these figures). Most of the listed features did not overlap with listed features of contrasting categories. As a simplifying assumption, we take it that no feature was listed in two contrasting categories. In Experiment 3, Tanaka and Taylor found, for the Novice condition, that mid-level categories were the fastest, and that the lower level categories were the slowest, and, for the Expert condition, that middle and lower level categories were about equally fast and that the higher level categories were the slowest (in Table 1 we put the mean RTs of bird novices and of dog novices, and the mean RTs of bird experts and of dog experts).

The category feature-possession and the category attentional slip measures scored a perfect three for both conditions. Whereas the category utility measure correctly identified the basic-levelness order of only one categorization level out of six!

Discussion

This paper presented category attentional slip, a measure of basic-level performance in which two constraints interact: (1) the cardinality of the sets of tests that determine category decision at each categorization level, and (2) the number of different sets of tests necessary to reach a category decision. We then compared the performance of category attentional slip with those of two established models of basic-level performance: category feature-possession (Jones, 1983) and category utility (Corter & Gluck, 1992). The empirical data was drawn from eight classical experiments from Murphy and Smith (1982), Murphy (1991), and Tanaka and Taylor (1991).

Category utility appeared as the worst predictor of basic-levelness with a 12/24 hit rate. In particular, category utility did not predict human performance very well for the category structures of Murphy and Smith (1982, Experiment 3) and of Tanaka and Taylor (1991, Experiment 3). In these experiments, the most basic categories are those at the lowest categorization level. We think this reveals a fundamental problem with category utility: It is strongly biased against lower levels categories. For example, to properly model Tanaka and Taylor's results (1991, Experiment 3, Experts), category utility would require the addition of no less than 29 non-overlapping features to each lower level category!

The category feature-possession and the attentional slip measures are tie with a 20/24 hit rate. In the considered experiments, the category organizations were all composed of non-overlapping features which means that a single set of redundant tests was in each case sufficient to reach a category decision. Thus, performance was in each experiment critically dependent on the first determinant of attentional slip: the redundancy of the sets of tests to determine category membership at all categorization levels. In fact, we can easily demonstrate that, in this case, the cardinality (or the level of redundancy) of a class of redundant test is strictly equal to feature-possession.

However, if we allow category organizations to be composed of partially overlapping features, optimal strategies of length greater than 1 will be required (see section Category attentional slip). And, the second determinant of attentional slip could influence category feature-possession and attentional slip in different ways. The hierarchy of Figure 1 illustrates such situation of feature overlap. Preliminary results using computer synthesized 3D objects indicated that the higher categorization levels of Figure 1 (which require optimal strategies of length 1) were the fastest, and lower level categorizations (which requires optimal strategies of length 3) were the slowest. Category attentional slip was then the only measure to score a perfect three with this category structure; both the category feature-possession and the category utility scored a two (see the simulations for every measure in section Measures of Basic-Levelness).

In sum, we believe that attentional slip is a new powerful experimental and formal platform to study recognition and categorization at the basic-level.

Acknowledgments

This research was partially funded by a NSERC grant to the second author. The first author is a FCAR scholar.

Reference

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94* (2), 115-147.

Corter, J. E. & Gluck, M. A. (1992). Explaining basic categories: Features predictability and information. *Psychological Bulletin, 111* (2), 291-303.

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

Murphy, G. L. & Smith, E. E. (1982). Basic level superiority in picture categorization. *Journal of Verbal Learning and Verbal Behavior, 21*, 1-20.

Murphy, G. L. (1991). Parts in objects concepts: Experiments with artificial categories. *Memory & Cognition, 19* (5), 423-438.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M. & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382-352.

Schyns, P. G. (1996). Diagnostic recognition: Task constraints, objects information, and their interactions. *Submitted for publication.*

Tanaka, J. W. & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology, 23*, 457-482.