Psychological Review
2001, Vol. 108, No. 4, 735–758

0033-295X/01/$5.00 DOI: 10.1037//0033-295X.108.4.735

Why Do We SLIP to the Basic Level? Computational Constraints and Their Implementation

Frédéric Gosselin and Philippe G. Schyns
University of Glasgow

The authors introduce a new measure of basic-level performance (strategy length and internal practicability; SLIP). SLIP implements 2 computational constraints on the organization of categories in a taxonomy: the minimum number of feature tests required to place the input in a category (strategy length) and the ease with which these tests are performed (internal practicability). The predictive power of SLIP is compared with that of 4 other basic-level measures: context model, category feature possession, category utility, and compression measure, drawing data from other empirical work, and 3 new experiments testing the validity of the computational constraints of SLIP using computer-synthesized 3-dimensional artificial objects.

In the 20-question game, one of two players chooses the name of a real-world category. The other player must guess the name after having asked as few yes–no questions as possible. For example, one could ask whether the name begins with a letter prior to *n* in the alphabet, whether it has more than two vowels, whether it sounds like the mating call of a moose, or any other binary query. The usual strategy, however, consists of asking questions that reveal certain relations between the target category and other categories. Suppose that "silver Speedster" was the target. It could be reached after the following sequence of questions (and appropriate answers): "Is it an animate object?" "Is it a vehicle?" "Is it a car?" "A Porsche?" "Is it a Speedster?" and "Is it silver like James Dean's?" If you already knew that the target was a car, you could eliminate a few branches of this search tree and ask more direct questions about the specific features of the target. Your questions would focus on shape (to distinguish Speedster from other cars) and color (to distinguish silver Speedsters from other Speedsters). Strategies of this sort are the most efficient in discovering the name of a category among a set of possibilities.

Humans who must discover the proper categorization of an, as yet, unknown stimulus face a similar situation. Context provides answers to the most general questions, and checking for the presence of specific features is a powerful strategy to categorize the input in a hierarchy. In this article, we present a new model of basic-level performance called *strategy length and internal practicability* (SLIP), in which the categorization strategies predict the time of access to the different categories of a hierarchy. Two factors determine these strategies: *strategy length* (SL) and *internal practicability* (IP); they arise from the organization of category information in a taxonomy.

Simply put, and relating this to the 20-question game, the length of a strategy in SLIP is the number of questions necessary to access a category. In the previous example, one question on shape and one separate question on color were necessary to identify a silver Speedster. We show later that strategy length depends on the overlap of information between the categories of a hierarchy (e.g., many other cars than Speedster are silver). Internal practicability is the ease with which one can obtain information about a category. In the game, more practicable categories are those for which many different questions can provide the same category information (to the enthusiast, a Speedster can be identified from many different shape features, including its silhouette, characteristic hood, windscreen, and so forth). Thus, internal practicability measures the redundancy of information for a categorization.

This article is organized as follows: We first review the evidence that different levels of a taxonomy can be accessed at different speeds. We then examine taxonomies, isolate the two essential principles on the organization of their features, and implement them in SLIP. Next, we compare the predictive power of SLIP with that of four other measures at the basic level of a database based on 22 classic experiments. Finally, three new experiments test the empirical validity of the two computational constraints implemented in SLIP.

Basic-Level Phenomenology

In Rosch, Mervis, Gray, Johnson, and Boyes-Braem (1976, Experiment 7), participants were taught the names of 18 objects at three levels of categorization: the *subordinate* (e.g., *Levis, MacIntosh*), *basic* (e.g., *pants, apple*), and *superordinate* (e.g., *clothes, fruit*).[1] These objects belonged to one of six possible nonanimal taxonomies: *musical instruments, fruits, tools, clothes, vehicles,*

Frédéric Gosselin and Philippe G. Schyns, Department of Psychology, University of Glasgow, Glasgow, Scotland.

This research was supported by Economic and Social Research Council Grant R000237901. We thank Carol A. Fowler for her helpful suggestions. We are also indebted to Simon Garrod and Lizann Bonnar for proofreading drafts of this article.

Correspondence concerning this article should be addressed to Frédéric Gosselin, who is now at Département de Psychologie, Université de Montréal, C. P. 6128, Succ. Centre-ville, Montréal, Quebec H3C 3J7, Canada, or to Philippe G. Schyns, Department of Psychology, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QB, Scotland. Electronic mail may be sent to frederic.gosselin@umontreal.ca or to philippe@psy.gla.ac.uk.

[1] Henceforth, the *basic levelness* of a category denotes a measure of performance. Whenever possible, we refer to the levels of abstraction as the subordinate, basic, and superordinate. Otherwise, we use a set of unambiguous level descriptors (e.g., low, middle, and high).

and furniture. In a verification task, participants were shown a category name followed by a stimulus picture, and had to determine whether they matched. Categories at the basic level were the fastest to verify, and categories at the subordinate level the slowest (see also Hoffmann & Ziessler, 1983; Jolicoeur, Gluck, & Kosslyn, 1984; Murphy, 1991; Murphy & Brownell, 1985; Murphy & Smith, 1982; Tanaka & Taylor, 1991).

The basic level is superior in many other respects: (a) Objects are named more quickly at this level than at any other level of abstraction (Hoffmann & Ziessler, 1983; Jolicoeur et al., 1984; Murphy, 1991; Murphy & Brownell, 1985; Murphy & Smith, 1982; Rosch et al., 1976; Tanaka & Taylor, 1991); (b) objects are described preferentially with their basic-level names (Berlin, 1992; Brown, 1958; Rosch et al., 1976; Tanaka & Taylor, 1991; Wisniewski & Murphy, 1989); (c) more features, especially shapes, are listed at the basic level than at the superordinate level (Rosch et al., 1976; Tversky & Hemenway, 1984); (d) throughout development, basic-level names are learned before those of other categorization levels (Anglin, 1977; Brown, 1958; Horton & Markman, 1980; Markman, 1989; Markman & Hutchinson, 1984; Mervis & Crisafi, 1982; Rosch et al., 1976); and (e) basic names tend to be shorter (Brown, 1958; Rosch et al., 1976). Convergence of all these performance measures is crucial to establish a preferred categorization level, even though verification speed is the most commonly used.

There is considerable evidence that a basic-level superiority holds across cultures for living things (for extensive reviews, see Berlin, 1992; Malt, 1995). Basic-level phenomenology also seems to hold across domains (for a review, see Murphy & Lassaline, 1997). This is true for computer programs (Adelson, 1985), events (Morris & Murphy, 1990; Rifkin, 1985; Rosch, 1978), personality types (Cantor & Mischel, 1979), sign language (Newport & Bellugi, 1978), environmental scenes (Tversky & Hemenway, 1983), clinical diagnosis (Cantor, Smith, French, & Mezzich, 1980), and emotions (Shaver, Schwarz, Kirson, & O'Connor, 1987).

In sharp contrast to this wealth of empirical studies, formal analyses of the basic level have received considerably less attention. In the next sections we develop a new formalism for the basic levelness of categories. We first examine taxonomies and derive two essential principles of the organization of categorical information. We then implement the principles in an ideal model that distinguishes categories at different levels of a taxonomy.

Two Principles for Organizing Information in Hierarchies

To study the computational determinants of basic-level performance, we first consider the typical artificial taxonomies that have elicited the basic-level advantage (see the top taxonomy of Figure 1, from Murphy & Smith, 1982, Experiment 1). Below the category names (e.g., *hob, bot, com*), letters correspond to the features that define the categories. For example, feature *a* defines the higher level category *hob,* three features (*c, d, e*) define the middle level *bot,* and the feature *o* defines the lower level *com.* These features specify the information that must be represented to place any object *X* in the taxonomy (exemplars of such inputs are represented in italics below the taxonomies). For example, if feature *a* represents *X,* then the object is a *hob* at the highest level of the taxonomy. If *o* describes *X,* then it is a *com* at the lowest level. Observe the mid-level *bot* category: *X* is a *bot* if it possesses any of the features *c, d,* or *e.* Technically, features *c, d,* and *e* are redundant for *bot;* any one of them is singly sufficient to access the category—redundant features are therefore interchangeable. Feature redundancy is known to be an important factor of basic-level category selection (Murphy & Smith, 1982; Rosch et al., 1976). It is the first principle for organizing features in taxonomies.

A second important principle emerges from the hierarchical nesting of categories. In the bottom taxonomy of Figure 1 (from Hoffmann & Ziessler, 1983, Hierarchy I), *X* is a *ril* if it possesses feature *a,* a *kas* if it has feature *c,* but it can only be a *lun* if it possesses both feature *a* and feature *g.* This occurs because *a* is present in the definitions of not one, but four different low-level categories; that is, *a* is shared between *lun, fuk, tuz,* and *zut.* This feature overlap is a common property of object taxonomies; think, for example, of the number of objects that have the same color, wheels, or legs, and so forth. Feature overlap is our second principle for organizing features in taxonomies.

Feature redundancy and feature overlap are two unavoidable principles of feature organization in taxonomies, even if the basic-level literature has neglected overlap (but see Hoffman & Ziessler, 1983). The functional role of redundancy and overlap in taxonomies differs markedly: Although several redundant features provide as many singly sufficient, interchangeable ways of accessing one category, one overlapping feature is not singly sufficient to isolate a category (i.e., overlapping features are singly necessary but only jointly sufficient to identify the category). In the next section, we implement feature redundancy and feature overlap in a formal model that predicts the basic levelness of the categories of a taxonomy.

SLIP: An Ideal Categorizer

Imagine SLIP, a formal categorizer that knows perfectly the top taxonomy of Figure 1. Assume further that this knowledge can guide an active search for the features of the as yet uncategorized input object *X.* For example, knowing that feature *a* defines *X* as a *hob,* SLIP will seek this feature in the input to verify that the input is a *hob.* Figure 1 illustrates the mapping between taxonomic knowledge and optimal categorization strategies. For example, strat(*X, hob*) = [{"Does *X* possess *a*?"}]—or strat(*X, hob*) = [{*a*}], for short[2]—is the categorization strategy that searches feature *a* in the input representation of *X* to verify that it is a *hob.* In a slightly more complicated example, strat(*X, bot*) = [{"Does *X* possess *c*?" "Does *X* possess *d*?" "Does *X* possess *e*?"}]—or strat(*X, bot*) = [{*c, d, e*}]—searches for feature *c,* feature *d,* or feature *e,* given that the three features are redundant (i.e., interchangeable) for this categorization.

Although *hob* and *bot* require different categorization strategies of the same input, they share one essential property: Both comprise only one set of redundant features, even though the number of features per set varies (i.e., one for *hob* and *com* and three for *hob*). It is important to draw the reader's attention to the set of redundant features (henceforth called *rf-set*). In SLIP, the rf-set is the basic unit of a categorization strategy. It is essential to remember that the features of a rf-set are always interchangeable (or redundant) for the categorization considered.

We rank categorization strategies according to the number of rf-sets they comprise. For example, Length 1 strategies (those examined so far) make up only one set of redundant features.

[2] In our notation, braces enclose sets of redundant feature tests, whereas brackets enclose strategies.

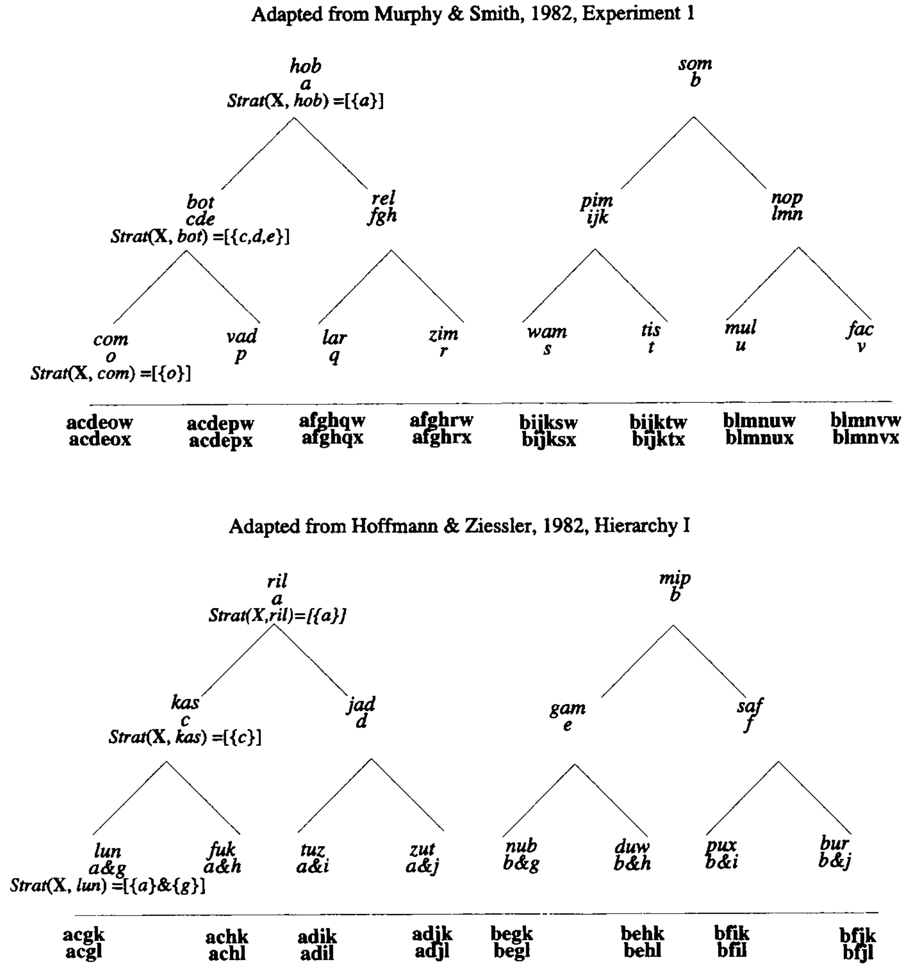


Figure 1. The top taxonomy is that of Murphy and Smith (1982, Experiment 1), and the bottom taxonomy is that of Hoffmann and Ziessler (1983, Hierarchy I). Below the category names are the optimal computational strategies of the strategy length and internal practicability model. The features defining each exemplar are provided at the bottom of the taxonomies. An index for these abstract features is given in Appendix B. Strat = strategy.

Length *n* strategies comprise *n* distinct rf-sets. For ease of presentation, we first formally develop Length 1 categorization strategies and then generalize to length *n* strategies.

In a Length 1 strategy, the successful completion of its single rf-set fully determines the behavior of SLIP. We say that, for any rf-set *j*, it is successfully completed whenever one of its interchangeable features is successfully tested. For example, *j* = ("Does *X* possess *c*?" "Does *X* possess *d*?" and "Does *X* possess *e*?") is completed whenever *c, d,* or *e* are found in the input. When *j* succeeds, Length 1 strat(*X, bot*) is verified and *X* is indeed a *bot.*

Note that a categorization strategy does specify which features to test, but not how to test them. To move from a logical account to a model, we implement the process of testing one rf-set. Arguably, features are optimally searched in parallel. Thus, we implement the optimal completion of one rf-set *j* as a parallel search for its features.[3]

Assume that each rf-set can initiate the parallel search for its own features. For example, in strat(*X, bot*), rf-set *j* = ("Does *X* possess *c*?" "Does *X* possess *d*?" and "Does *X* possess *e*?") will launch in parallel three independent searches for features *c, d,* and *e* in the input *X.* Suppose that *X* = *acdeow.* To search for a feature (e.g., *c*), a dedicated feature agent (i.e., f-agent *c*) will be called on. This agent knows (a) which information channel[4] to scan (here, the second position of the input string) and (b) how to use information from this channel to successfully detect the feature (here, *c*). Thus, the f-agents of a feature search scan their information channels in parallel until they detect their own feature. For example, to apply strat(*X, bot*) to *X* = *acdeow,* f-agent *c,* f-agent *d,* and f-agent *e* scan in parallel the Positions 2, 3, and 4 of the input string, respectively.

[3] Appendix A presents a serial implementation of SLIP. It is important to stress that the parallel and serial implementations make similar ordinal predictions of basic levelness. Their predictions diverge on only one disjunctive case (Lassaline, 1990, Experiment 3, 4D; see *Simple Disjunctions* section for a discussion).

[4] Information channel is here loosely defined. It can be equated with the classical channels of luminance, color, motion and depth, with different spatial frequency channels, or with the dimensions of shape, color and texture, and so forth. Information channels are useful in Experiment 3, where each possible position of features in a string becomes a different information channel.

So far, the parallel implementation is ideal: rf-set j elicits a parallel search for its own features. Specialized f-agents carry out the search in the appropriate information channel by scanning for a specific feature. Uncertainty, however, permeates any psychological task. Here, we assume that f-agents can slip from their dedicated information channel. When they slip, f-agents scan a randomly chosen channel. To illustrate, f-agent c could slip from the channel for c (Position 2 in the string) to scan any position of the string X = *acdeow*. All the ingredients are in place to formalize the completion of Length 1 categorization strategies: Starting from the notion of a slip, we progressively formalize the completion of one rf-set.

When one f-agent slips from its information channel, it can either slip to another channel (and miss the detection of its feature) or, strictly by chance, slip right back to its dedicated channel (and detect its feature). Thus, a feature will be detected whenever the f-agent either does not slip, or slips right back to its dedicated channel. We summarize these two classes of events into φ, the probability of detecting a target feature in one trial—for example, feature c in strat(X, *bot*). To derive φ, we start with the probability Q that a slip does reveal the target feature by chance. It is 1 divided by the number of possible information channels (e.g., 6 in X = *acdeow*, and so $Q = 1/6$). If the f-agent slips with a probability S, SQ combines the probability S of a slip with the probability Q that the target feature (i.e., c) is nevertheless detected. The probability that the slip does not reveal the target feature is therefore $S - SQ$. The complement of this event, $1 - (S - SQ)$, is φ, the probability that one f-agent detects its target feature in one trial.

From φ, the probability of detecting one feature, we can derive ψ_j, the probability that any one of the m_j interchangeable features of rf-set set j is detected, leading to its successful completion. Working a contrario, we start from $(S - SQ)$, the probability that one feature is not detected by its f-agent (which slipped to the wrong channel): $(S - SQ)^{m_j}$ is, therefore, the probability that none of the m_j features of rf-set j is detected. The complement $\psi_j = 1 - (S - SQ)^{m_j}$ is the probability of detecting one of the interchangeable features of rf-set j in one trial of a parallel feature search.

So far, we have derived the probability ψ_j of the successful completion of one rf-set j in one trial. Concretely, the f-agents have been sent to detect their features in their information channels and we have examined what happens after one trial. Some f-agents might have failed (because they slipped to the wrong channel) and some might have succeeded (because they either slipped to the right channel or did not slip at all). However, f-agents are persistent. When they slip to the wrong channel on their first trial (and fail to detect their feature), they keep trying on subsequent trials. We therefore need to develop ψ_j to account for the possibility of several unsuccessful trials, followed by one successful feature test. $(1 - \psi_j)^t$ is the probability of t successive failures over the first t trials. The probability that a Length 1 categorization strategy succeeds in t trials or less is therefore[5]

$$1 - (1 - \psi_j)^t. \tag{1}$$

Note that when ψ_j is large, the probability of a failure $(1 - \psi_j)$ is small, and Equation 1 increases rapidly with increasing t. In other words, a large ψ_j implies that a long series of failed feature tests is unlikely. ψ_j is large whenever rf-set j comprises many redundant features. ψ_j implements the computational constraint of feature redundancy discussed earlier. That is, ψ_j measures the *internal practicability* of a category, the ease with which it can be accessed from its defining features. SLIP predicts that basic-level categories have, on average, a higher internal practicability.[6] Consequently, fewer feature tests should be required, on average, to access *bot* compared with *hob* and *com* in the top taxonomy of Figure 1.

We have so far developed SLIP for strategies composed of only one rf-set j. In these, feature redundancy solely determines the number of feature tests to access the category. In general, SLIP strategies comprise n rf-sets. For example, in the bottom taxonomy of Figure 1, the input is a *lun* whenever the following strategy is completed: strat(X, *lun*) = [{"Does X possess a?"} and {"Does X possess g?"}].

We now generalize Equation 1 to the case of length n strategies. This implements feature overlap, the second computational constraint discussed earlier. In processing terms, the rf-sets of a strategy are handled like the features of one rf-set: A parallel process also drives their testing. Although the completion of one rf-set implies the successful testing of only one of its features, the completion of a strategy requires the successful testing of all of its rf-sets.

The logical dependency of rf-sets in a strategy has performance implications. Although the testing of one rf-set must be as fast as its fastest feature detection, the testing of a strategy is as slow as the completion of its slowest rf-set. Consequently, SLIP predicts that the fastest categorization strategy comprises only one rf-set j (i.e., a Length 1 strategy, no feature overlap) with many redundant features (i.e., a high ψ_j), and the slowest strategy comprises many rf-sets (i.e., length n strategies, high feature overlap), each composed of a single feature (i.e., a low ψ_j).[7]

To generalize to length n strategies, we start with Equation 1, the cumulative probability that one rf-set j is completed in at most t trials. The cumulative probability that the n rf-sets of a length n strategy are completed in trial t or before becomes

$$\prod_{j=1}^{n} [1 - (1 - \psi_j)^t]. \tag{2}$$

To compute the probability of a completion in exactly t feature tests, we subtract two cumulative probabilities: the probability that the length n strategy is completed in most t trials minus the probability that it is completed in at most $t - 1$ trials. Equation 3 expresses the likelihood that length n strategy is completed in exactly t feature tests:

[5] Technically, Equation 1 describes the cumulative geometric distribution.

[6] The internal practicability of a category does not always reduce to the redundancy of its features. Instead, internal practicability can implement sophisticated assumptions about the internal structure of the features themselves (e.g., whether they are independent or form configurations, have graded salience, and so forth). For ease of presentation, we have not included any such assumptions in our examples, so internal practicability reduces here to feature redundancy. However, to model the data of our Experiment 3, we modify SLIP to account for feature configurations.

[7] The advantage of redundant features arises from the fact that the likelihood of the event S = (a is true or b is true) is superior or equal to the likelihood of the single event S = (a is true), whereas nonoverlapping features (i.e., shorter strategies) are advantageous because the event S = (a is true) is more likely than S = (a is true and b is true).

$$\eta_{n,\psi} = \prod_{j=1}^{n} [1 - (1 - \psi_j)^t] - \prod_{j=1}^{n} [1 - (1 - \psi_j)^{t-1}]. \quad (3)$$

To derive the basic levelness of a category in SLIP, we consider each possible number of feature tests (from 1, 2, 3, . . . to infinity), multiply each one of them with the likelihood that it leads to a successful completion of a length n strategy (cf. Equation 3), and add the partial results. That is,

$$\sum_{t=1}^{+\infty} t \times \eta_{n,\psi}. \quad (4)$$

To illustrate, we apply Equation 4 to the two taxonomies of Figure 1. We first derive ψ_j for each category. We start with Q_j, the probability that a slip reveals by chance the target feature. It is one divided by the total number of features in the input; that is, one sixth for the top taxonomy and one fourth for the bottom taxonomy. If the probability of a slip is $S = .50$, $\psi_j = 1 - (S - SQ)^{m_j}$ is $1 - (.50 - .50 \times .25) = .625$, for all categories of the bottom taxonomy, irrespective of level. In the top taxonomy, the computations yield ψ_j values of .583, .927, and .583 for the high, middle, and low levels, respectively.

To compute basic levelness, consider *bot* and *com* in the top taxonomy of Figure 1. Using Equations 2 and 3, we find that the likelihood of a completion of *bot* in one trial is $[1 - (1 - .927)^1] - [1 - (1 - .927)^0] = .927$. Turning to Equation 4, the first term of the weighted sum is $1 \times .927$. The likelihood of a completion in two trials is $[1 - (1 - .927)^2] - [1 - (1 - .927)^1] = .995 - .927 = .068$. In Equation 4, the second term is therefore $2 \times .068$. So far, Equation 4 equals $(1 \times .927) + (2 \times .068) = 1.063$. The likelihood of a completion in three trials is $.9996 - .9946 = .0039$. Equation 4 becomes $(1 \times .927) + (2 \times .068) + (3 \times .0039) = 1.074$. If we pursued this computation for all possible values of t, we would find that the average basic levelness of *bot* is 1.078 feature tests. For *com*, we find that the first three terms of Equation 4 are $(1 \times .583) + (2 \times .243) + (3 \times .102) = 1.375$ (with average basic levelness of 1.714 feature tests). The basic levelness of *bot* is higher than that of *com*. The explanation lies in the terms of Equation 4: A categorization in few feature-test attempts is more likely for *bot* than for *com* (e.g., .927 vs. .583 for one feature test), and it decreases more rapidly with increasing numbers of feature tests (e.g., .0039 vs. .102 for three feature tests). Remember that *bot* is more practicable (i.e., has more interchangeable features) than *com*. The difference in basic levelness is a direct reflection of the effect of ψ_j.

Bot and *com* are both Length 1 strategies; they involve only one set of redundant features. When a strategy comprises n sets of redundant features, the principles just discussed are simply extended. The only difference is the nature of the cumulative probabilities used in Equation 3. Instead of describing the successful resolution of only one set of redundant features (Length 1 strategy, see Equation 1), they involve the resolution of n feature sets (length n strategy, see Equation 2).

In general, to compute the basic levelness of one level of abstraction, we average the basic levelness of its categories. To illustrate, in the top taxonomy of Figure 1, the middle level requires fewer feature tests (1.078) than the top and bottom levels (1.714 for both), isolating the effect of category practicability. Turning to the bottom taxonomy of Figure 1, the high- and mid-level categorizations require, on average, fewer feature tests (1.6) than the low-level categorization (2.03), isolating the effect of strategy length on basic levelness.

Summary

We have developed a computational analysis of basic-level categorization in the spirit of Marr (1982). Starting from two fundamental constraints on the organization of information in a taxonomy (the overlap of features between categories and the redundancy of features within categories), we have seen how the constraints determine different feature testing strategies to categorize the input. Each strategy specifies (a) the features to test and (b) whether they are redundant or overlapped for the categorization considered. In general, greater feature overlap augments the length of a strategy, and higher feature redundancy augments its accessibility.

To implement these constraints, two main philosophies of modeling were available. The first philosophy has been by far the most productive in the categorization literature. It consists of fitting a posteriori the parameters of a model to human data (e.g., Lamberts, 1994; Medin & Schaffer, 1978; Nosofsky, 1986; Tversky, 1977). The main aim is to mimic human performance, and the quality of fit measures the success of the model. In the second philosophy, one constructs a priori an "ideal" model to serve as a benchmark against human performance (see Anderson, 1990, 1991; Kersten, 1990). The ideal model is perfect for the task at hand: It is omniscient and uses all of the information available in the task. The aim is not so much to fit data, but to understand how human performance diverges from the ideal implementation.

Our modeling adheres to the second philosophy. We opted for an ideal implementation of feature redundancy and overlap: SLIP has perfect knowledge of all parameters of the categorization tasks (the possible categorizations and the associated strategies and their features). It also operates in parallel, arguably the better mode for our tasks. SLIP frames categorization as the resolution of a categorization strategy. A categorization strategy comprises n sets of redundant (i.e., interchangeable) features. Equation 1 implements the computational constraint of feature redundancy; Equation 2 implements feature overlap. Equations 3 and 4 coordinate these factors to predict the basic levelness of a category. It is the average number of feature tests that resolves its categorization strategy. We now turn to the competitors of SLIP.

Other Models of Basic Levelness

The competitors of SLIP implement one of two main ideas: *utility* or *similarity*. According to category utility, the most useful level of a taxonomy is the basic level (see Brown, 1958, *level of usual utility*, and Rosch, 1978; Rosch & Mervis, 1975; Rosch et al., 1976). Models that implement a version of category utility are Rosch et al.'s *cue validity*, Jones's (1983) *category feature possession*, Corter and Gluck's (1992) *category utility*, Fisher's (1987, 1988) *COBWEB*, Anderson's (1990) *rational analysis*, and Pothos and Chater's (1998) *compression*. In other models, the basic level maximizes a measure of similarity between exemplars: Rosch et al.'s (1976) *differentiation model* and Medin and Schaffer's (1978; modified by Estes, 1994) *context model*.

We do not review all of SLIP's competitors here. We have left aside Rosch et al.'s (1976) *cue validity* because it cannot predict the classic advantage for an intermediate level (Murphy, 1982), a minimum requirement of any model of the basic level. We have also discarded Fisher's (1987, 1988) COBWEB because it makes predictions similar to Corter and Gluck's (1992) category utility, on which COBWEB is based. Finally, we have excluded Anderson's (1990, 1991) *rational analysis* categorization model because it does not provide a metric of basic levelness. Whenever possible, we examined the likely behavior of each competitor to variations of the two computational constraints of SLIP: internal practicability and strategy length.

Utility and Category Cue Validity

Brown (1958) suggested that "things are first named so as to categorize them in a maximally useful way" (p. 20). For example, a dime is a *dime*, instead of a *metal object* because this is what is mostly relevant about it. Several models built on this idea.

Category Feature Possession

Jones (1983) proposed that the basic level maximizes the average *category feature possession*, a measure of category utility. Category feature possession starts with $P(c_i|f_j)$, the probability of membership to category c_i, given feature f_j. It also considers the probability that the object possesses f_j, given that it belongs to c_i, $P(f_j|c_i)$. The conjunction of these two events, $K_{ij} = P(c_i|f_j)P(f_j|c_i)$, is called the *collocation* of category c_i and feature f_j. Collocations are computed for all categories and features. For each category, the largest feature collocations are counted. This sum—weighed by a constant k is the feature possession of a category, following Corter and Gluck, 1992—we set k to 1 in our simulations. Feature possession reflects the number of strong bidirectional links between a category and its features (i.e., their mutual predictability). The category feature possession of a level of categorization is the average category feature possession of its categories.

In the bottom taxonomy of Figure 1, category feature possession predicts a faster access at the highest level (feature possession = 3) and equally slower accesses to the middle and low levels (feature possession = 1). In the top taxonomy, the model predicts faster access to both the high and middle levels (feature possession = 3) and slower access to the bottom level (feature possession = 1).

How does category feature possession compare with SLIP? Category feature possession is a good measure of internal practicability for the following reasons: A feature that is not shared between categories predicts only one category, and this category, in turn, predicts the feature: $K_{ij} = P(c_i|f_j)P(f_j|c_i) = 1$. Categories with many such features score high in collocation[8] and internal practicability. However, category feature possession does not make specific predictions for variations of strategy length. Numerical simulations reveal that the measure is biased toward higher level categorizations.

Corter and Gluck's (1992) Category Utility Measure

Corter and Gluck (1992; see also Fisher, 1987, 1988) proposed another measure called *category utility*. A useful category better predicts the features of its members. Starting from $P(f_j|c_i)$, the probability of feature f_j given knowledge of category c_i, the probability of a correct guess is $P(f_j|c_i)^2$. When the category is useful, an informed feature guess should outperform an uninformed guess $P(f_j)$. If $P(f_j)^2$ is the probability of guessing the feature correctly, then the category utility of c_i for feature f_j is $P(c_i)[P(f_j|c_i)^2 - P(f_j)^2]$, the subtraction of informed and uninformed feature guesses, given $P(c_i)$, the probability that the object belongs to c_i. Summed across all input features, the category utility (or basic levelness) of c_i becomes

$$P(c_i) \sum_{j=1}^{m} [P(f_j|c_i)^2 - P(f_j)^2]. \quad (5)$$

The basic levelness of a level of abstraction is the average basic levelness of its categories.

In the bottom taxonomy of Figure 1, category utility predicts faster accesses to both the high and mid-levels (utility = .38), and a smaller utility of .25 predicts a slower access to the lower level. In the top taxonomy of Figure 1, the predictions are middle (utility = .78), high (utility = .69), and low level (utility = .45).

Category utility is also biased to the higher levels of categorization. If we distribute the summation over the two terms of category utility, we obtain

$$P(c)\left[\sum_{k=1}^{m} P(f_k|c)^2 - \sum_{k=1}^{m} P(f_k)^2\right], \quad (6)$$

where $\sum_{j=1}^{m} P(f_j)^2$ is constant across levels, and the two remaining variable terms are biased to the higher levels. The probability $P(c_i)$ that an object belongs to category c_i decreases exponentially with increasing category specificity, reducing utility. At the same time, $\sum_{j=1}^{m} P(f_j|c_i)^2$ increases almost linearly with increasing specificity and can only compensate the exponential reduction of $P(c_i)$ with an exponential addition of redundant features at lower levels. Thus, category utility predicts an advantage for the higher levels of taxonomies.

How does category utility compare with SLIP? The models are similar in their sensitivity to redundant features. However, category utility requires an exponential addition of redundant features to compensate for the exponential decrease in the likelihood of lower level categories. When the added features overlap between categories, category utility and SLIP tend to diverge in their predictions.

Compression

Our last reviewed measure of utility is minimum description length (MDL). MDL is a method that partitions data to compress them (Pothos & Chater, 1998). Each level of a taxonomy represents a different partition of the same data set, and MDL measures the amount of compression that each taxonomic level achieves.

[8] This only holds when few nondiagnostic features are present. For nondiagnostic features (those with equal probability of occurring in all low-level categories), collocation is always greater at the higher levels. This is because $P(c_i|f_j)$, equal to $P(c_i)$ in this case, is maximum at the highest level, and because $P(f_j|c_i)$ is constant for nondiagnostic features.

Pothos and Chater (1998) suggested that compression is maximal for the basic-level partition.

In a data set of r objects, there are $s = [r(r - 1)]/2$ possible object pairs and, therefore, s possible pairwise similarities between the objects. $A = [s(s - 1)]/2$ is the number of possible binary relationships (inequalities) between these pairwise similarities. They are encoded on A bits of information (one bit per similarity relationship). A bits describe the data set before any category partitioning. Assume D_i is the number of bits required to describe the same data after one partitioning. $A - D_i$ is then a measure of the compression efficiency of the partition. Pothos and Chater (1998) proposed that this difference is maximal for the basic-level partition. MDL is technical, and the reader can skip to the final paragraph of this section (after Equation 8) without loss of continuity.

To derive D_i, we start from the number of possible partitions of r items into n clusters, part(r, n),

$$\sum_{v=0}^{n} (-1)^v \frac{(n - v)^r}{(n - v)!v!}. \qquad (7)$$

To encode this partitioning, we need $\log_2[\text{part}(r, n)]$ bits of information. We then compute u, the combinatorics of all within-cluster pairwise similarities with all between-cluster pairwise similarities. The scheme assumes that within-cluster pairwise similarities $s(i, j)$ are all greater than any between-cluster pairwise similarities, $s(k, l)$. However, this does not always hold, and e counts the number of times the assumption is violated (e can vary between 0 and u and can thus be encoded on a maximum of $\log_2(u + 1)$ bits). There are $C_e^u = u!/[(u - e)!e!]$ possible ways of selecting e errors among the combinatorics of relationships u. A total of $\log_2(u + 1) + \log_2(C_e^u)$ bits encode the errors.

Remember that A bits specify all possible binary relationships between objects, whereas u specifies those constrained by the clustering. $A - u$ counts the relationships left outside the clustering; $A - u$ bits are necessary to encode these.

The compression of information offered by one partitioning of the data is $A - D_i$, where $D_i = \log_2[\text{part}(r, n)] + [\log_2(u + 1) + \log_2(C_e^u)] + (A - u)$. Hence $A - D_i$ is equal to

$$u - \{\log_2[\text{part}(r, n)] + \log_2(u + 1) + \log_2(C_e^u)\}. \qquad (8)$$

How does compression compare with SLIP? If we compare the compression achieved at all levels, we find the following for the bottom taxonomy of Figure 1: high = 2,759 bits, middle = 2,277 bits, and low = 865 bits, predicting a preference for the high level. In the top taxonomy, compressions are high = 3,569 bits, middle = 2,274 bits, and low = 865 bits, predicting again an advantage for the high level. This bias for the high level occurs because when there is little feature overlap, MDL is only dependent on u, the combinatorics of within and between category similarities. This combinatorics grows with level of generality, and so compression is greater when fewer categories are considered, irrespective of how redundant the features are within the categories. Variations of strategy length create overlap between features and add errors to the MDL description. However, these tend to be insufficient to counterbalance the bias for the high level.

Similarity and the Differentiation Model

The models reviewed so far have equated the basic level with highest usefulness. Another principle is that of *differentiation,* or *dissimilarity.* As Rosch et al. (1976) put it, categories at the basic level "have the most attributes common to members of the category and the least attributes shared with members of other [contrasting] categories" (p. 435). The first component of this definition has been called the *specificity* (Murphy & Brownell, 1985), or the *informativeness* (Murphy, 1991) of a category, and the second component, the *distinctiveness* of a category (Murphy, 1991; Murphy & Brownell, 1985). However, category differentiation is not sufficiently specified to be refuted (Medin, 1983).

The two determinants of SLIP can be loosely mapped onto the two determinants of category differentiation (Schyns, 1998). In general, more specific representations tend to be more informative, but they are also less distinctive from other representations (Murphy, 1991). Subordinate categories tend to score high on informativeness (e.g., two brands of cars convey detailed information), but low on distinctiveness (e.g., two brands of car are similar in overall appearance, at least more so than a brand of car and a type of shoe). In contrast, superordinate categories score low on informativeness, but high on distinctiveness (e.g., *vehicle* and *furniture* have different functions, shapes, parts, colors, textures, and so forth). Thus, strategy length is related to informativeness: It also increases with category specificity. Internal practicability is related to distinctiveness: It also increases with more general categories.

Medin and Schaffer's (1978) influential exemplar model of categorization (see Nosofsky, 1986, and Lamberts, 1994, for further developments) can also be construed as a similarity model of basic level performance. A multiplicative rule computes the similarity $S(a, b)$ between any two exemplars of a category. A match between corresponding attributes is assigned a value of 1, and a mismatch a value of α_D, a dissimilarity parameter—α_D varies between 0 and 1 to assign different weights to the attributes. The matches and mismatches are then multiplied to measure the similarity of the exemplars.

To compute the basic levelness of a category (e.g., *Ford*), Estes (1994) proposed to consider one category exemplar (e.g., a *Ford Mustang*) and compute the ratio between the similarity of this target exemplar with all exemplars from this category (e.g., all Fords) and the similarity of the target with exemplars from its superordinate category (e.g., all cars). The basic levelness of a category is the average ratio of all its exemplars; the basic levelness of a taxonomic level is the average basic levelness of its categories.

In the bottom taxonomy of Figure 1, the index predicts (with $\alpha = .3$, which follows Estes, 1994) a faster access to the top level (.90), the middle level (.88), and finally the lower level (.77). For the top taxonomy of Figure 1, the model predicts equal accessibility to the top and middle levels (.99) and slower access to the lower level (.77).

The context model also has a bias for the higher taxonomic levels. It computes a ratio between two polynomials that approach one with increasing levels of taxonomic generality. The bias for the high level can be overcome by increasing the similarity between the target exemplars and the exemplars of lower level contrast categories. The numerical simulations show that the context model is neither sensitive to internal practicability nor to strategy length, the two factors of SLIP.

Numerical Simulations of Classic Basic-Level Experiments

In the next sections, we compare the predictive performance of SLIP with that of the four models of basic levelness that we presented: Jones's (1983) category feature possession, Corter and Gluck's (1992) category utility, Pothos and Chater's (1998) compression, and Medin and Schaffer's (1978) context model. To compare human performance with the predictions of the models, we compiled a database of 22 classic basic-level experiments. In the following section, we present the experiments according to the factor of SLIP they tested (i.e., internal practicability vs. strategy length). From the outset, it is worth pointing out that SLIP did a good job of predicting the experiments, particularly when these tested taxonomies with overlapping features.

Variations of Internal Practicability Determine Basic Levelness

Faster Access at an Intermediate Level

One of the most influential experiments on the basic level is that of Murphy and Smith (1982, Experiment 1). It is influential because most subsequent basic-level experiments used the same procedure. Their participants were initially taught the artificial taxonomy represented at the top of Figure 1 (see also Figure 2, top taxonomy).[9] In a testing phase, they were shown a category name followed by a stimulus. The participants' task was to verify as quickly as possible whether the name and stimulus matched. As shown in Figure 1, mid-level categories have the highest practicability. Table 1 illustrates that they were also verified faster (see Murphy & Smith, 1982, Experiment 1). Using the same taxonomic organization of category attributes, Murphy (1991, Experiment 4, Simple) replicated these results. In fact, the highest practicability of the middle level is also responsible for its faster access in Mervis and Crisafi (1982) and Murphy (1991, Experiment 4, Enhanced; see Figure 2).

Aside from these experiments with artificial categories, the database also comprises experiments with natural taxonomies. With these, we assumed that the features participants listed reflected their internal representations of categories (see Rosch & Mervis, 1975). In addition, following Tversky and Hemenway (1984) and Tanaka and Taylor (1991), we assumed that the same feature could not be listed for two contrasting categories. Following these principles, five natural taxonomies had a greater feature redundancy at the intermediate level: Rosch et al. (1976, Experiment 7).[10] Tanaka and Taylor (1991, Novice),[11] and Johnson and Mervis (1997, advanced songbird expert, intermediate songbird expert, and novice),[12] and the, or an, intermediate level was accessed faster, see Table 1.

Faster Access at the Lower Level

In Murphy and Smith (1982, Experiment 3), a unique set of attributes was added at the lower level of Murphy and Smith's artificial tools. Figure 3 illustrates that the addition of redundant attributes enhanced the practicability of the lower level categories, and Table 2 reveals that these categories were accessed faster. Tanaka and Taylor's (1991, Expert)[13] experiment is a variation on this theme: In the expertise category, they added new singly diagnostic features at the lower level (i.e., redundant features), and thus sped up its access. In fact, basic and subordinate categories were equally fast, and superordinate categories were slower (see the mean response times [RTs] for *bird* and *dog* experts in Table 2).

Faster Access at the Higher Level

In his Experiment 5, Murphy (1991) added a set of singly sufficient values to the high-level categories of Murphy and Smith's (1982) artificial tools. Figure 4 shows that this level became more practicable, and Table 3 confirms that it was indeed accessed faster.

Variations of Strategy Length Determine Basic Levelness

In the experiments reviewed so far, the length of categorization strategies was held constant, even though the experimenters never mentioned this explicitly. In SLIP terms, categorization strategies comprised only one rf-set. We now turn to the few experiments that tested the effect of strategy length on basic levelness. Hoffmann and Ziessler (1983, Hierarchy I) used artificial objects organized as described in the bottom taxonomy of Figure 5. We explained earlier that strategy length was one at the high and middle levels, but two at the lower level (see the strategies in Figure 1). Table 4 confirms that participants accessed the high- and mid-level categories equally fast and were slower for the low-level categories.

In Gosselin and Schyns (1998), participants learned the taxonomy of Figure 5 applied to artificially textured and colored geometric primitives (similar to Biederman's, 1987, geometric primitives, called *geons*). This taxonomy ascribes strategies of different lengths to the different taxonomic levels: from top to bottom,

[9] To facilitate comparison of taxonomies, we have normalized the notation of information, substituting letters of the alphabet for the actual features. Appendix C lists the mapping between letters and actual features (whenever relevant or possible).

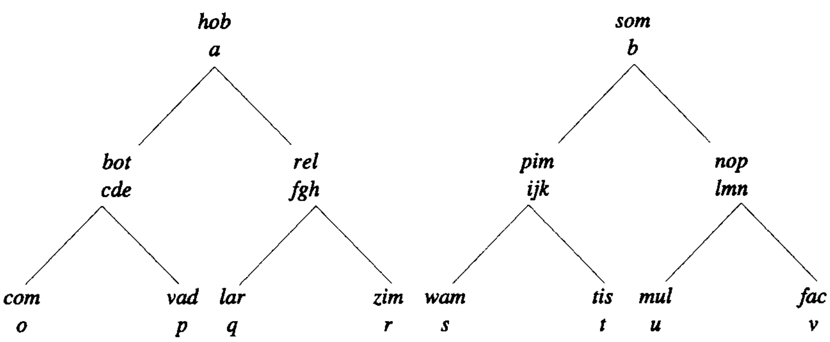
[10] In Rosch et al. (1976, Table 2, nonbiological taxonomies, raw tallies), the mean number of added redundant features was of 1.85, 5.55, and 3.5 for subordinate, basic, and superordinate, respectively.

[11] In Tanaka and Taylor (1991, novice, bird, and dog novices confounded), participants listed approximately 8, 12, and 7 new redundant features for the superordinate, basic, and subordinate levels of categorization, respectively.

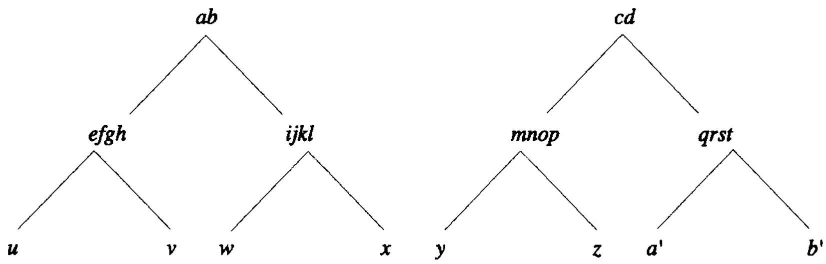
[12] Johnson and Mervis (1997, Experiment 1, songbirds condition) used four-level natural taxonomies in a verification task. Their advanced songbird experts listed 1.75, 5.00, 6.02, and 3.75 for the superordinate, basic, subordinate, and subsubordinate levels, respectively. For the intermediate songbird experts, these numbers were 1.00, 4.87, 4.28, and 2.47 for the same levels. For the novices and the tropical freshwater fish experts, the numbers were 1.08, 2.47, 0.23, and 0.02.

[13] Their experts (bird and dog experts confounded) listed approximately 8, 10, and 10 new features for the superordinate, basic, and subordinate levels of categorization, respectively. Compare this with 8, 12, and 7 for the superordinate, basic, and subordinate levels, respectively, in their condition novice in the previous section.

Adapted from Murphy & Smith, 1982, Experiment 1



Adapted from Mervis & Crisafi, 1982



Adapted from Murphy, 1991, Experiment 4, Enhanced

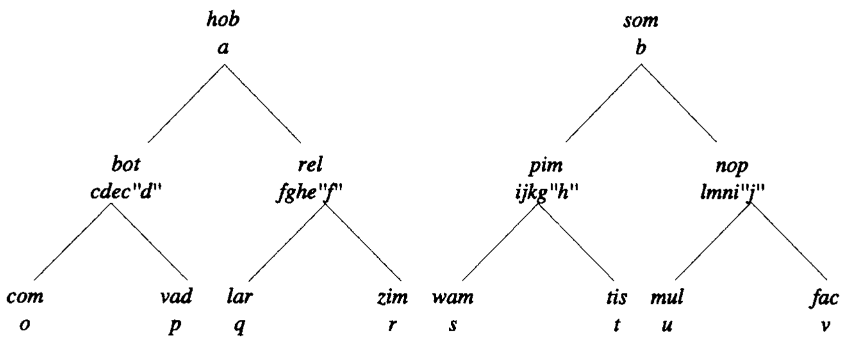


Figure 2. Taxonomies of all experiments that exhibited an advantage for an intermediate level of categorization arising from variations of feature redundancy. From top to bottom: Murphy and Smith (1982, Experiment 1; see also Murphy, 1991, Experiment 4, Simple, for a replication); Mervis and Crisafi (1982); Murphy (1991, Experiment 4, Enhanced). Underneath the category names are the optimal computational strategies of the strategy length and internal practicability model. An index for these abstract features is given in Appendix B.

Table 1
Variations of Internal Practicability Produce Faster Access at an Intermediate Level: Numerical Predictions and Observations

Source/model	Level			
	H_{-3}	H_{-2}	H_{-1}	H
Murphy & Smith (1982, Exp. 1)				
Observation[a]		723	678	879
Context		0.544	0.554	0.516
SLIP[b]		1.714	1.078	1.714
Possession		1	3	3
Utility		0.453	0.725	0.688
Compression[c]		865	2,277	3,569
Murphy (1982, Exp. 4, Simple)				
Observation[a]		862	811	980
Context		0.544	0.554	0.516
SLIP[b]		1.714	1.078	1.714
Possession		1	3	3
Utility		0.453	0.725	0.688
Compression[c]		865	2,277	3,569
Mervis & Crisafi (1982, Exp. 1)				
Observation[a]		3rd	1st	2nd
Possession		1	4	2
SLIP[b]		1.778	1.038	1.237
Utility		0.609	1.094	1.063
Context		0.569	0.572	0.521
Compression[c]		5,992	12,916.59	18,985.00
Murphy (1982, Exp. 4, Enhanced)				
Observation[a]		1,132	854	955
Utility		0.640	1.100	0.938
Possession		1	5	1
SLIP[b]		1.778	1.016	1.778
Compression[c]		865	2,277	3,569
Context		0.561	0.584	0.516
Rosch et al. (1976, Exp. 7)				
Observation[a]		659	535	591
Possession		2	6	4
SLIP[b]		1.266	1.009	1.046
Compression[c]		0	85	185
Utility		1.030	1.701	1.874
Context		0.620	0.607	0.531
Tanaka & Taylor (1991, Novice)				
Observation[a]		777.5	677.5	745.5
Possession		7	12	8
SLIP[b]		1.802	1.333	1.658
Compression[c]		0	85	185
Utility		2.387	3.517	3.934
Context		0.751	0.728	0.607
Johnson & Mervis (1997, Songbird, Novice)				
Observation[a]	~2,100	~1,950	~1,600	~1,900
Possession	2	23	247	108
SLIP[b]	13.889	1.735	1.000	1.018
Utility	16.323	32.519	60.778	59.429
Compression[c]	0	865	2,277	3,569
Context	4th	3rd	2nd	1st
Johnson & Mervis (1997, Songbird, Intermediate)				
Observation[a]	~1,725	~1,600	~1,550	~1,800
Possession	247	428	487	100
SLIP[b]	1.173	1.038	1.023	1.853
Utility	65.551	115.417	137.706	128.219
Compression[c]	0	865	2,277	3,569
Context	2nd	3rd	4th	1st
Johnson & Mervis (1997, Songbird, Advanced)				
Observation[a]	~1,600	~1,625	~1,500	~1,750
Context	2nd	3rd	4th	1st
Possession	375	602	500	175
SLIP[b]	1.063	1.011	1.023	1.363
Compression[c]	0	865	2,277	3,569
Utility	86.434	149.056	165.302	167.574

Note. The cells under Level are ordered from left to right according to increasing generality. The number of levels varies across taxonomies; some taxonomies do not have an H_{-3} level. For each taxonomy, models are ranked from best to worst and ties are ordered alphabetically. H = highest; negative subscript numbers following H = the distance away from the H level of generality; Exp. = experiment; SLIP = strategy length and internal practicability.
[a] Data in milliseconds. [b] Data in attempts. [c] Data in bits.

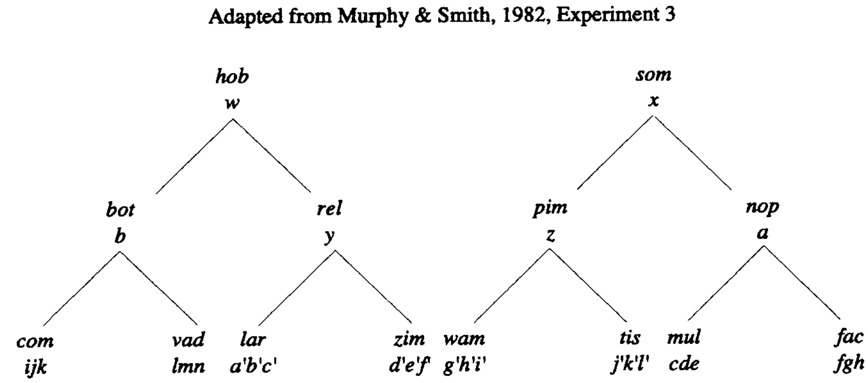


Figure 3. Taxonomy of Murphy and Smith (1982, Experiment 3), the only experiment with varying redundancy that exhibited an advantage for lower level categorizations. Below the category names are the optimal computational strategies of the strategy length and internal practicability model. An index for these abstract features is given in Appendix B.

Lengths 1, 2, and 3, respectively. Faster categorization was found for the high, middle, then low categorization levels (see Table 4).

Disjunctions and Mixtures

In developing SLIP, we have so far assumed that a conjunction of several attributes defines categories. As Smith and Medin (1981) put it: "If we rely on intuitions (our own and those published by semanticists) and restrict ourselves to concepts about naturally occurring objects (flora and fauna) . . . we can think of no obvious disjunctive concepts" (p. 28). Even though we generally share this view, some concepts can be disjunctive. For example, a strike in baseball is either a called or a swinging strike. Researchers in several basic-level experiments examined disjunctive categories. To model these, SLIP must be slightly modified (see Appendix C for details).

Table 2
Variations of Internal Practicability Produce Faster Access at the Lower Level: Numerical Predictions and Observations

Source/model	Level H_{-2}	H_{-1}	H
Murphy & Smith (1982, Exp. 3, Size)			
Observation[a]	574	882	666
Possession	3	1	1
SLIP[b]	1.078	1.714	1.714
Utility	0.483	0.428	0.561
Compression[c]	865	2,277	3,569
Context	0.538	0.544	0.554
Tanaka & Taylor (1991, Expert)			
Observation[a]	621.5	623.0	728.5
Possession	10	10	8
SLIP[b]	1.474	1.474	1.676
Context	0.750	0.726	0.650
Compression[c]	0	85	185
Utility	2.526	3.258	3.870

Note. The cells under Level are ordered from left to right according to increasing generality. For each taxonomy, models are ranked from best to worst and ties are ordered alphabetically. H = highest; negative subscript numbers following H = the distance away from the H level of generality; Exp. = experiment; SLIP = strategy length and internal practicability.
[a] Data in milliseconds. [b] Data in attempts. [c] Data in bits.

Simple Disjunctions

Figure 6 illustrates the Hierarchy II of Hoffmann and Ziessler (1983). At the highest level, feature disjunctions define the categories, whereas feature conjunctions define the lower level. It was found that mid-level categories were accessed fastest, and high- and low-level categories equally slowly (see RT in Table 5; see also Corter, Gluck, & Bower, 1988, for a replication using categories of artificial diseases and conceptual features).

Lassaline (1990; reported in Lassaline, Wisniewski, & Medin, 1992) constructed a disjunctive, two-level taxonomy in which artificial tools similar to those of Murphy and Smith (1982) were used. In two conditions of her Experiment 3 (one dimension and four dimension), two-feature disjunctions defined the high-level categories, and a single feature defined each low-level category. In the four-dimension condition, low-level features were one value of four different stimulus dimensions (e.g., *hammer head, pizza cutter handle, dotted texture,* and a *square internal shape*). In the one-dimension condition, low-level features were four different values of the same dimension (e.g., *hammer, brick, knife,* and *pizza cutter head*). We did not distinguish between features and dimensions in our figures, so the bottom taxonomy of Figure 6 illustrates the taxonomy used in both conditions.[14] In a verification task, an advantage was found for the low-level categories in the one-dimension condition, but the advantage was at the high level in the four-dimensions condition (see Table 5). Note that SLIP is the only model that predicts a difference between these two conditions, even if it does not explain the data completely (see Table 5; Lassaline et al., 1992).

In a taxonomy similar to that of Murphy and Smith (1982, Experiment 1), Murphy (1991, Experiment 3) used 16 artificial

[14] In Appendix C, we explain how SLIP can be made sensitive to the distinction between features and dimensional values. With this distinction, a disjunction of features becomes longer to complete than individual features, but only if these features are different values of the same dimension.

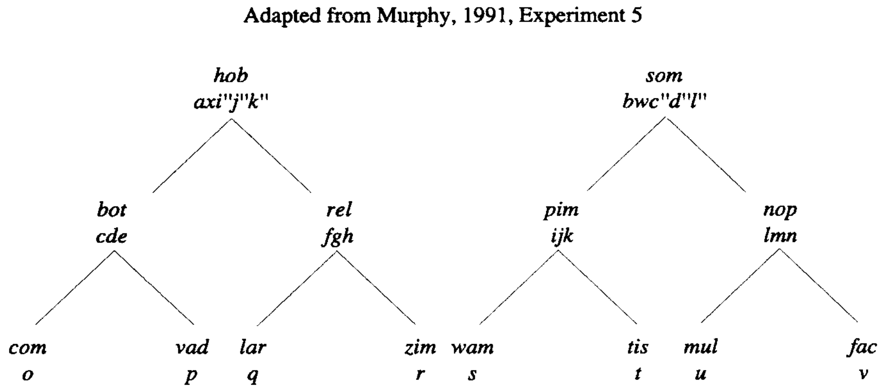


Figure 4. Taxonomy of Murphy (1991, Experiment 5), the only experiment with varying redundancy that exhibited an advantage for higher level categorizations. Below the category names are the optimal computational strategies of the strategy length and internal practicability model. An index for these abstract features is given in Appendix B.

abstract stamps of various colors, textures, types of edge, and sizes, organized as featural disjunctions at the high level (see Figure 6). Access was faster at the middle level, with the other two levels being equally slow (see Table 5).

Mixtures of Disjunctions and Conjunctions

In their Hierarchy III, Hoffmann and Ziessler (1983) rearranged the objects of their Hierarchy I: Conjunctions of two features defined the lowest level, disjunctions of two two-feature conjunctions the middle level, and disjunctions of four two-feature conjunctions the high level (see Figure 7). In this taxonomy, faster access was reported for the high-level categories (see RT estimations in Table 6).

In Lassaline (1990, Experiment 1, see Figure 7), two-feature disjunctions defined categories at the high level, and a conjunction of a feature with a disjunction of features defined the low level. The high level was accessed faster (see Table 6). Figure 7 also illustrates the taxonomy of Lassaline's second experiment. A compound of the form $[\{X\}]$ *AND* $\{Y\}]$ *OR* $[\{X\}$ *AND* $\{Z\}]$ defined low-level categories. A disjunction of two such compounds defined the high level. The low level was faster to verify (see Table 6).

Table 3
Variations of Internal Practicability Produce Faster Access at the Higher Level: Numerical Predictions and Observations in Murphy (1982, Experiment 5)

Model	Level		
	H_{-2}	H_{-1}	H
Observation[a]	1,072	881	854
Compression[b]	0	85	185
Possession	1	3	5
SLIP[c]	1.8	1.096	1.018
Utility	.703	1.281	1.688
Context	0.604	0.554	0.516

Note. The cells under Level are ordered from left to right according to increasing generality. Models are ranked from best to worst and ties are ordered alphabetically. H = highest; negative subscript numbers following H = the distance away from the H level of generality; SLIP = strategy length and internal practicability.
[a] Data in milliseconds. [b] Data in bits. [c] Data in attempts.

Comparison of the Models for the Published Experiments

The presentation of the database of 22 basic-level experiments has been organized according to the factor of SLIP. We now examine the correlation between the database and the predictions of the models (i.e., SLIP, category feature possession, category utility, compression, and context model).

To give the best possible chance to each competitor, whenever possible, we best fitted its free parameters to the database of experimental results. The only free parameter of SLIP, *S*, was set to .5 (that the ordinal predictions of SLIP are invariant for values of *S* comprised between .1 and .9). Jones's (1983) category feature possession comprises a single free parameter *k*. Following Corter and Gluck (1992), we set it to one, but the ordinal predictions of category feature possession are invariant to changes of *k*. Following Estes (1994), we used a single dissimilarity parameter α for Medin and Schaffer's (1978) context model. It was best fitted to $\alpha = .94$. Corter and Gluck's (1992) category utility and Pothos and Chater's (1998) compression have no free parameter.

Overall, SLIP predicted 85% of the experimental results, clearly winning the competition. Second best were Jones's (1983) category feature possession and Corter and Gluck's (1992) category utility with 63%, then came Pothos and Chater's (1998) compression measure with 40%, and, finally, Medin and Schaffer's (1978) context model with 37%. It is instructive to examine the models specifically for their predictions of variations of internal practicability, strategy length, and mixture experiments. Table 7 summarizes this.

When the researchers varied only the internal practicability of categories in the experiments, SLIP accounted for 82% of the available data, feature possession with 77%, category utility with 51%, context model with 38%, and compression with 23%. When the experiments only varied strategy length, SLIP accounted for 100% of the available data, compression and category utility with 83%, feature possession with 50%, and context model with 33%. When experiments comprised mixtures, SLIP was first with a score of 85%, category utility with 80%, compression with 60%, category feature possession with 40%, and context model with 35%.

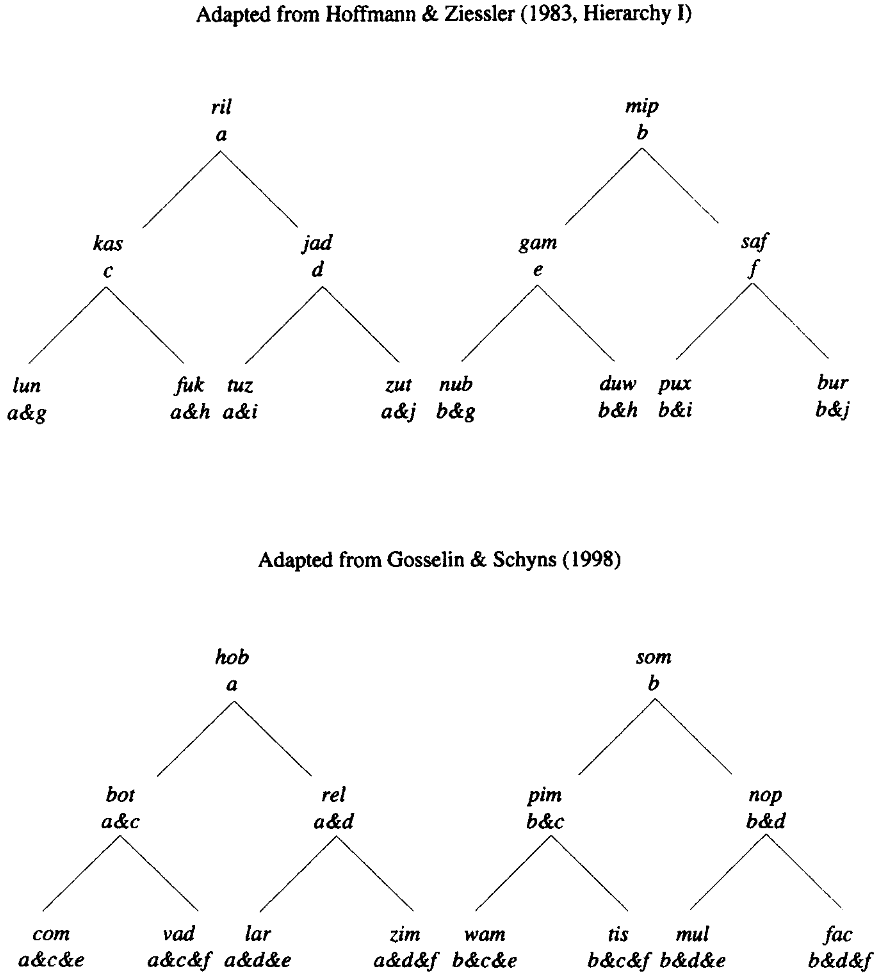


Figure 5. Abstract taxonomies of all experiments with varying strategy length. From top to bottom: Hoffmann and Ziessler (1983, Hierarchy I); Gosselin and Schyns (1998). Below the category names are the optimal computational strategies of the strategy length and internal practicability model. An index for these abstract features is given in Appendix B.

In summary, SLIP is the best predictor with 85% of the data of 22 published basic-level experiments predicted. It performed particularly well on taxonomies with feature overlap. Thus, it would seem that the two computational constraints of SLIP approximate those underlying the basic levelness of categories in humans.

Empirical Testing of SLIP

In the following sections we further examine how the competence of SLIP (feature redundancy and overlap) predicts human categorization at different levels of abstraction. So far, in these experiments on the basic level, we have mainly been motivated by empirical considerations instead of an ideal model like ours. As pointed out earlier, strategy length has never been tested as such. In both Hoffmann and Ziessler (1983, Hierarchy I) and Gosselin and Schyns (1998, Overall), strategy length is confounded with level of abstraction: The most inclusive level always has the shortest categorization strategy, and the least inclusive level has the longest strategy. From an empirical standpoint, the effect of strategy length remains to be explicitly demonstrated. A similar problem affects feature redundancy. Even though we showed earlier that many basic-level experiments modulated feature redundancy, no systematic study of this factor has been carried out so far.

The following experiments were designed to overcome these two shortcomings. In these, researchers all used computer-synthesized artificial three-dimensional (3D) objects, to tightly control feature composition. In Experiment 1 the effect of strategy length on basic levelness was isolated; in Experiment 2 the effect of internal practicability was tested; in Experiment 3, the interactions between the two factors were examined.

Experiment 1

In Experiment 1, we isolated strategy length and examined how a variation of this factor at different levels of a hierarchy influences their basic levelness. As pointed out earlier, strategy length was shown to influence basic levelness in Hoffmann and Ziessler (1983, Hierarchy I) and Gosselin and Schyns (1998, Overall).

Table 4
Variations of Strategy Length: Numerical Predictions and Observations

Source/model	Level		
	H_{-2}	H_{-1}	H
Hoffmann & Ziessler (1983, Hier. I)			
Observation[a]	~700	~500	~500
SLIP[b]	2.036	1.6	1.6
Utility	0.25	0.375	0.375
Compression[c]	865	2,277	2,759
Context	0.524	0.523	0.516
Possession	1	1	3
Gosselin & Schyns (1998, Overall)			
Observation[a]	1,184	1,012	819
Compression[c]	0	73	149
SLIP[b]	2.164	1.875	1.5
Possession	1	1	3
Utility	0.188	0.25	0.25
Context	0.516	0.516	0.516

Note. The cells under Level are ordered from left to right according to increasing generality. For each taxonomy, models are ranked from best to worst and ties are ordered alphabetically. H = highest; negative subscript numbers following H = the distance away from the H level of generality; Hier. = hierarchy; SLIP = strategy length and internal practicability.
[a] Data in milliseconds. [b] Data in attempts. [c] Data in bits.

However, these two experiments did not dissociate strategy length from level of abstraction. Experiment 1 was designed to dissociate level of categorization and speed of access. In the HIGH_FAST taxonomy, shorter strategies should lead to a faster access at the higher level. In LOW_FAST, the opposite applies: shorter strategies should access the low-level categories faster. In both conditions, the longer strategies arose from overlap between the attributes of categories (volumetric primitives; Biederman, 1987). SLIP predicts that shorter strategies are completed faster, irrespective of categorization levels (i.e., a faster access to the high level of HIGH_FAST, and to the low level of LOW_FAST).

Method

Participants. Twenty University of Glasgow students with normal or corrected vision were paid to take part in the experiment.

Stimuli. Stimuli were computer-synthesized chains of four geometric primitives similar to those of Tarr, Bülthoff, Zabinski, and Blanz (1997). We designed the stimuli with a 3D modeling software on a Macintosh computer. Five geometric elements defined the categories of the HIGH_FAST taxonomy. One different geon defined each one of three high-level categories. Each one of six possible low-level categories was further specified by one of the two remaining geons. The top taxonomy of Figure 8 illustrates this.

In HIGH_FAST, strategy length equals one for the higher level categories. This means that only one feature must be tested to access these categories. Strategy length equals two at the lower levels, because these categorizations require two feature tests. The overlap of features across lower level categories produced the longer conjunctive strategies.

To create the experimental stimuli, we substituted the letters in Figure 8 with their corresponding geometric elements. To these two geons, we added two supplementary geons that served as fillers. Fillers were identical across objects and so could not be used to distinguish them. We created two exemplars per low-level category by changing the location of the diagnostic geons in the chain (see Figure 9 for examples).

Nine geons defined the LOW_FAST taxonomy. A unique combination of two geons, sampled from a set of three, defined each one of three top-level categories (see Figure 8, bottom taxonomy). High-level strategies had Length 2 because a combination of two geons defined categories at this level. A unique diagnostic geon further specified the categories at the low level. However, low-level categories had Length 1 strategies because a single-feature test on a diagnostic geon determined membership. Figure 8 shows the LOW_FAST taxonomy. One filler was added to generate six four-geon chains. From these, we created two exemplars per category.

Procedure. The procedure was similar to that of Murphy (1991). In a learning phase, participants were evenly split between the learning of the HIGH_FAST and LOW_FAST taxonomies. We instructed participants to learn the names and the defining geon(s) of nine categories (see the specific names and corresponding geon combinations in Figure 8). Participants saw their taxonomy on a sheet of paper; this learning phase was not constrained in time.

We tested participants' knowledge of the taxonomy by asking them to list the features associated with each category name. Criterion was reached when participants could list twice in a row, without any mistake, the attributes defining each category. Corrective feedback was provided.

When participants knew the taxonomy, a category verification task measured categorization time at each level. Each trial began with the presentation of a category name. Participants pressed one keyboard key to see the list of all learned definitions on the screen (each definition corresponded to a set of geons). Participants had to identify the definition corresponding to the previously shown category name. With this procedure, we wanted to keep a close control of participants' knowledge of their taxonomy[15] and to minimize the contribution of access time to a strategy in memory to the overall verification time. After a 200-ms delay, an object appeared on the screen. Participants had to decide as quickly as possible whether the named category and object matched by pressing the *yes* or *no* keys on the computer keyboard. We recorded response latencies. Note that low-level categories are more numerous than high-level categories. We normalized the number of positive and negative trials with the constraint of equating the number of trials per level.

Results and Discussion

We performed the analysis of the logarithm of RTs on all correct positive trials (error rate = 6.5%). Table 8 reports the mean RTs at the low and high levels for the two taxonomies tested (see Observations in Table 8).

A two-way (Group × Strategy Length) analysis of variance (ANOVA) of the logarithm of RTs with repeated measures on one variable (strategy length) revealed a main effect of strategy length, $F(1, 18) = 91.64$, $p < .0005$, (mean Length 1 strategies = 2.946 feature tests; mean Length 2 strategies = 3.084 feature tests), meaning that participants systematically verified Length 1 strategies faster than Length 2 strategies, irrespective of the considered level (low vs. high). All participants verified the categories associated with Length 1 strategies faster. Neither the interaction between group and strategy length, $F(1, 18) = 0.03$, *ns*, nor the main group effect, $F(1, 18) = 0.03$, *ns*, was significant. The error

[15] Although this identification was not a feature of Murphy's (1991) experiment, we believe that the author had the similar intention of ensuring that participants knew what the categorization strategy was. From Murphy (1991): "In the learning phase, subjects were given a cover sheet for each category that gave the category name and explained why all the patterns were in the same category. That is, it mentioned the features distinctive of that category. For example the description for the category LAR, a middle-level category, was: 'There are LARs, because their edge is serrated, they have squares inside, and the squares are solid' " (p. 429).

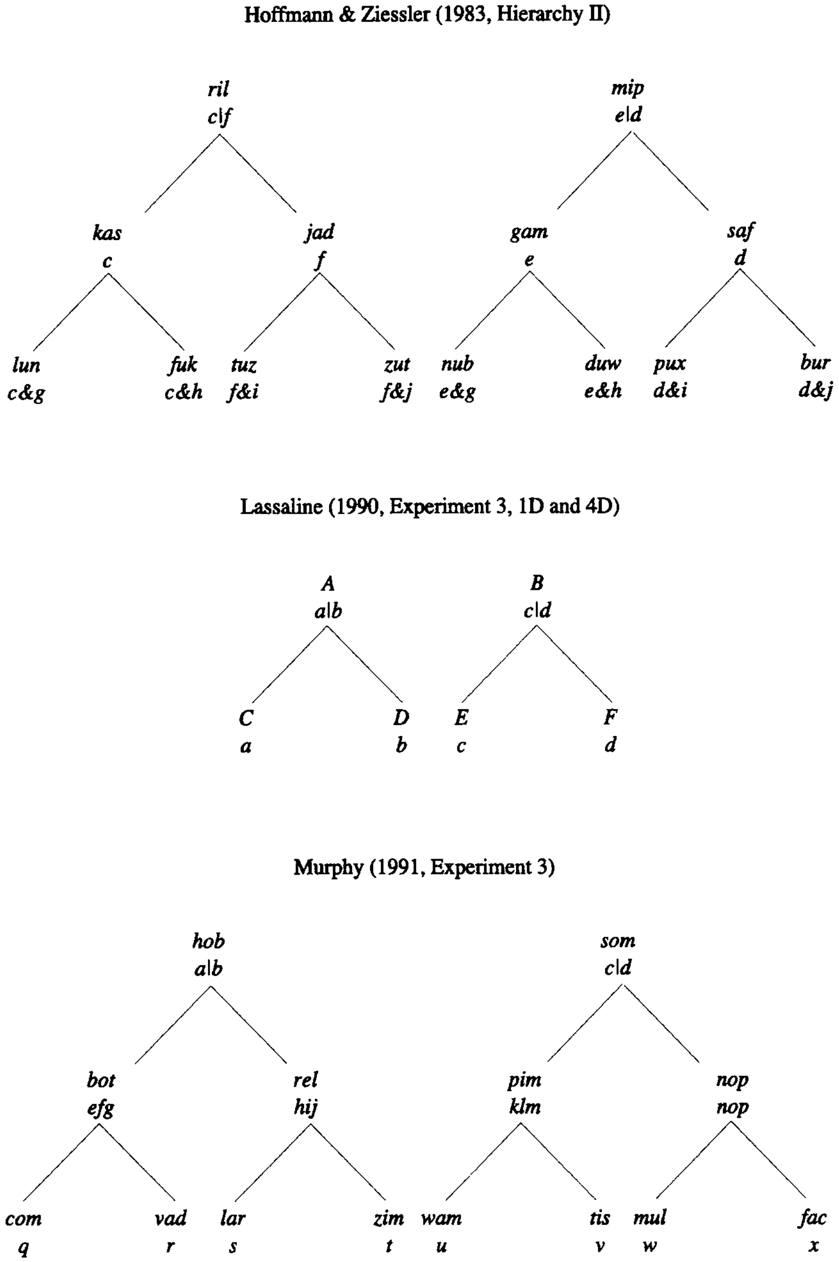


Figure 6. Taxonomies of all experiments with simple disjunctions. From top to bottom: Hoffmann and Ziessler (1983, Hierarchy II); Lassaline (1990, Experiments 3, 1D, and 4D); and Murphy (1991, Experiment 3). Below the category names are the optimal computational strategies of the strategy length and internal practicability model. The features of each exemplar are given underneath each taxonomy. Even though Lassaline's one-dimension and four-dimension taxonomies are identical in terms of strategy length and internal practicability, they differ in one important respect: In the one-dimension taxonomy, all attributes came from the same dimension, whereas in the four-dimension taxonomy, none of the attributes came from the same dimension. An index for these abstract features is given in Appendix B.

rate was low overall and did not correlate with RT ($r = -.17$, *ns*), ruling out a speed–accuracy trade-off.

SLIP predicts that Length 1 strategies should be completed faster than Length 2 strategies, irrespective of categorization level (see SLIP in Table 8 for numerical predictions). The data reported here confirms that strategy length, rather than categorization level, determines the basic levelness of a category.

Experiment 2

Practicability refers to the ease with which features identify a category at any level of a taxonomy. A category has high practicability whenever many of its defining features are uniquely diagnostic of this category. It has low practicability when a single feature defines the category. If this variable influences the basic

Table 5
Simple Disjunctions: Numerical Predictions and Observations

	Level		
Source/model	H_{-2}	H_{-1}	H
Hoffmann & Ziessler (1983, Hier. II)			
Observation[a]	~700	~500	~700
Compression[b]	865	2.277	189
SLIP[c]	2.667	2	4
Utility	0.25	0.375	0.125
Context	0.509	0.539	0.516
Possession	1	2	2
Corter, Gluck, & Bower (1988)			
Observation	3.045	2,567	3.115
Compression[b]	865	2.277	189
SLIP[c]	2.667	2	4
Utility	0.25	0.375	0.125
Context	0.509	0.539	0.516
Possession	1	2	2
Lassaline (1990, Exp. 3, one-dim.)			
Observation		1st	2nd
SLIP[c]		2	4
Utility		.259	.167
Compression[b]		119	135
Context		0.511	0.521
Possession		1	6
Lassaline (1990, Exp. 3, four-dim.)			
Observation		2nd	1st
Context		0.511	0.522
SLIP[c]		2	2
Compression[b]		527	389
Possession		9.25	0
Utility		.316	.048
Murphy (1982, Exp. 3)			
Observation[a]	776	688	779
SLIP[c]	2	1.143	4
Possession	1	3	2
Utility	0.531	0.719	0.563
Context	0.536	0.554	0.516
Compression[c]	865	2.277	3.569

Note. The cells under Level are ordered from left to right according to increasing generality. The number of levels varies across taxonomies; some taxonomies do not have an H_{-2} level. For each taxonomy, models are ranked from best to worst and ties are ordered alphabetically. H = highest; negative subscript numbers following H = the distance away from the H level of generality; Hier. = hierarchy; SLIP = strategy length and internal practicability; Exp. = experiment; dim. = dimension.
[a] Data in milliseconds. [b] Data in bits. [c] Data in attempts.

levelness of a category, then it should apply equally to all levels of a taxonomy.

In Experiment 2, all strategies had Length 1, but the high and low levels differed in practicability. In the HIGH_FAST condition, high-level strategies had greater practicability than low-level strategies. The opposite applied to the LOW_FAST condition, with low-level strategies having higher practicability. SLIP predicts that categories with higher practicability will be verified faster, irrespective of their level in the taxonomy.

Method

Participants. Twenty students from University of Glasgow with normal or corrected vision were paid to take part in the experiment.

Stimuli. Stimuli were similar to those of Experiment 1: four-geon chains synthesized with a 3D modeling software on a Macintosh computer. In the HIGH_FAST condition, 10 diagnostic geons were used. Three different geons defined each one of two high-level categories; one different geon defined each low-level category (see Figure 10, top taxonomy). We generated two exemplars per category by changing the location (either rightmost or leftmost of the chain) of the three geons defining the high-level categories (see Figure 10, top taxonomy).

The LOW_FAST taxonomy comprised 14 diagnostic geons. One diagnostic geon defined each one of two high-level categories, and 3 different geons defined each one of four low-level categories (see Figure 10, bottom taxonomy). As before, we created two category exemplars by changing the location (either far right or far left of the object) of the triplets defining the low-level categories (see Figure 10, bottom taxonomy). Practicability is greater for high-level categories in the HIGH_FAST condition and for the low-level categories in the LOW_FAST condition. These levels have more unique features associated with them.

Procedure. The procedure followed in all respects that of Experiment 1: Participants were randomly assigned to the HIGH_FAST and LOW_FAST conditions. They were taught their respective taxonomy before entering a verification task in which we measured speed of access to the two levels of categorization. Each one of the 280 trials consisted in the initial presentation of a category name followed by an object. Participants had to decide as quickly as possible as whether the two matched. We then recorded response latencies.

Results and Discussion

We performed an ANOVA on the logarithm of RTs of positive trials (error rate = 5.4%). Table 8 shows the mean RTs at the low and high levels for HIGH_FAST and LOW_FAST.

A two-way (Group × Practicability) ANOVA on the logarithm of RTs with repeated measures on one variable (practicability) revealed a main effect of practicability, $F(1, 18) = 24.88$, $p < .0005$ (mean verification time = 2.838 for high practicability strategies; 2.894 for low practicability strategies). Of the 20 participants, only 1 did not respond faster to categories with greater practicability; a sign test showed that this was significant ($p < .00003$). Neither the Group × Practicability interaction, $F(1, 18) = 4.15$, *ns*, nor the main group effect, $F(1, 18) = 1.50$, *ns*, was significant. The error rate was low overall and did not correlate with RTs ($r = .05$, *ns*), ruling out a speed–accuracy trade-off.

In summary, SLIP predicted that strategies with greater practicability should yield faster categorization decisions, irrespective of categorization level (see SLIP in Table 8 for numerical predictions). The results of Experiment 2 confirmed the prediction.

Experiment 3

Experiments 1 and 2 revealed that changing either strategy length or internal practicability of any level of a taxonomy changes its basic levelness. In Experiment 3, we explored how these two variables interact to determine performance. There are many possible interactions to investigate and we did not investigate them all. Instead, we examined three main scenarios that changed the fastest level by modifying either strategy length or internal practicability.

EQUAL was our neutral, baseline scenario. Strategies at the high and low levels had an equal length of one and the same constant practicability. SLIP predicts that categorization speeds should be equal across levels. In the SL_DOWN scenario, we augmented the length of the strategies that access the high-level categories to produce faster categorizations at the lower level. In the IP_UP scenario, we kept the difference of strategy length just

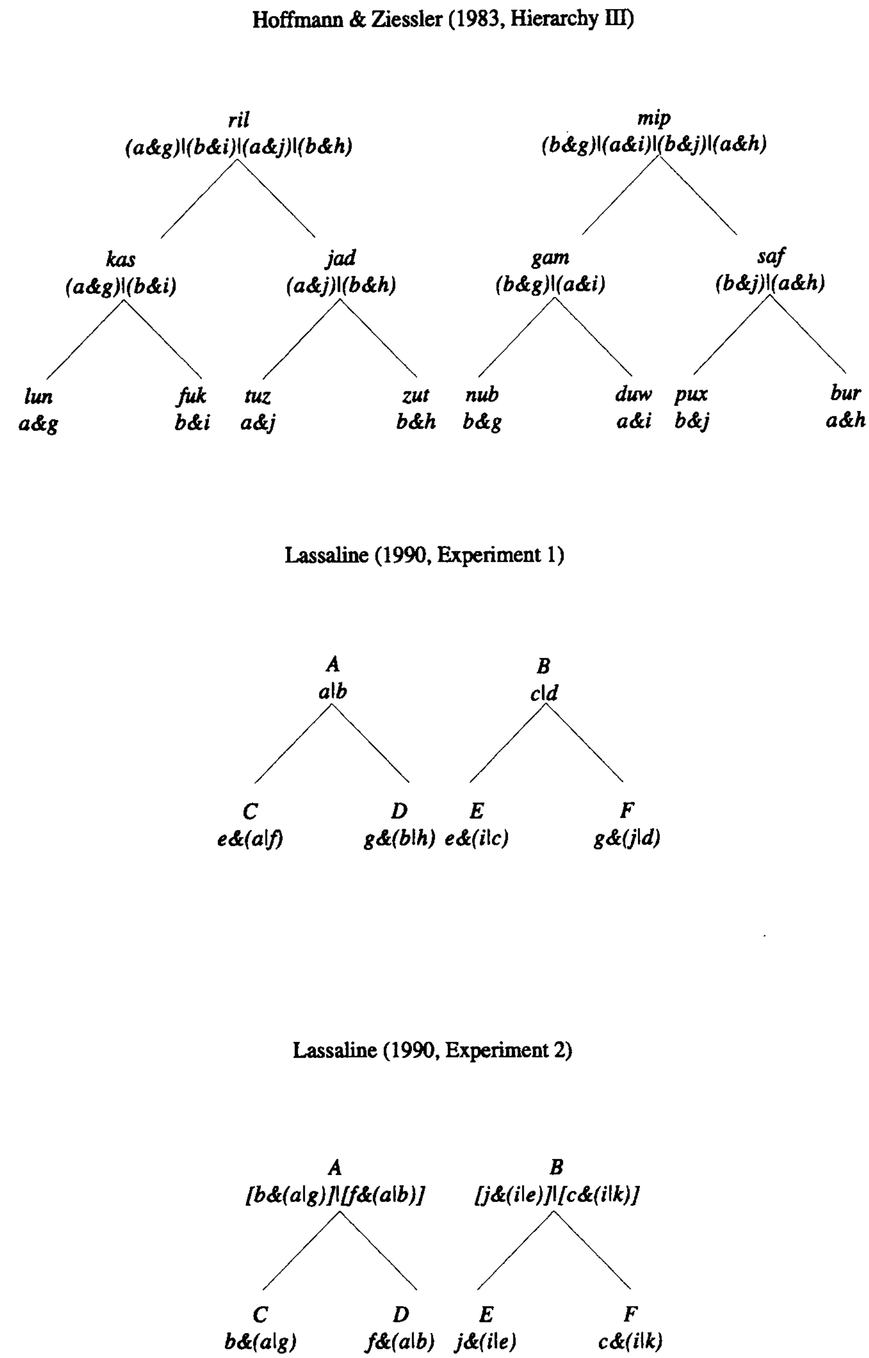


Figure 7. Taxonomies of all experiments with mixtures of disjunctions and conjunctions. From top to bottom: Hoffmann & Ziessler (1983, Hierarchy III); Lassaline (1990, Experiment 1); Lassaline (1990, Experiment 2). Below the category names are the optimal computational strategies of the strategy length and internal practicability model. Whenever possible, an index for these abstract features is given in Appendix B.

discussed, but decreased the practicability of the low level to speed up access to the high level. In summary, starting from an equal access to two levels of a taxonomy, a change of strategy length in the SL_DOWN condition produces faster categorizations at the low level. From this taxonomy, decreasing the internal practicability of the low level in the IP_UP condition should produce faster categorization at the high level.

Method

Participants. Thirty students from University of Glasgow with normal or corrected vision were paid to take part in the experiment.

Stimuli. Stimuli were similar to those of Experiments 1 and 2: geon chains designed with a 3D-object modeling software. Nine diagnostic geons entered the composition of categories in the EQUAL, SL_DOWN, and IP_UP conditions. In the equal condition, one geon defined each one

Table 6
Mixtures of Disjunctions and Conjunctions: Numerical Predictions and Observations

Source/model	Level		
	H_{-2}	H_{-1}	H
Hoffman & Ziessler (1983, Hier. III)			
Observation[a]	~700	~1,050	~1,475
Compression[b]	865	2	−22
SLIP[c]	2.036	4.504	9.329
Utility	0.25	0.125	0
Possession	3	0	3
Context	0.501	0.516	0.546
Lassaline (1990, Exp. 1)			
Observation		2nd	1st
Compression[b]		1,359	1,479
Possession		2.5	3
SLIP[c]		1.708	1.164
Utility		.127	.238
Context		0.526	0.523
Lassaline (1990, Exp. 2)			
Observation		1st	2nd
Compression[b]		380	106
SLIP[c]		1.708	3.882
Utility		.209	.119
Possession		2	2
Context		0.508	0.521

Note. The cells under Level are ordered from left to right according to increasing generality. The number of levels varies across taxonomies; some taxonomies do not have an H_{-2} level. For each taxonomy, models are ranked from best to worst and ties are ordered alphabetically. H = highest; negative subscripts following H = the distance away from the H level of generality; Hier. = hierarchy; Exp. = experiment; SLIP = strategy length and internal practicability.
[a] Data in milliseconds. [b] Data in bits. [c] Data in attempts.

of the nine categories of the taxonomy (see the top taxonomy of Figure 11). To this defining geon, we added four fillers to form six six-geon chains. We placed the geons defining the high-level categories at the far left of the chains, and those defining the low-level categories at the far right (see the top taxonomy of Figure 11).

In the SL_DOWN condition, a unique combination of two geons taken from a set of three geons defined each top-level category. The addition of one different geon further specified each lower level category. We produced six six-geon chains by adding three fillers. We placed the geon pairs defining the high-level categories at the far left of the chains, and those defining the low-level categories at the far right (see the bottom taxonomy of Figure 11). These chains also served to construct the exemplars of the IP_UP condition. Here, we generated four exemplars per category by changing only the location in the chain of the single geon defining the low-level categories (one of the four rightmost positions in the six-geon chains; see Appendix D for relevant formal changes of SLIP).

Procedure. The procedure was identical to that of Experiments 1 and 2. Participants were randomly assigned to one of three conditions (EQUAL, SL_DOWN, and IP_UP). Following a learning of their taxon-

Table 7
Comparison of Models for the Predicted Percentage of Nominal Data of a Database of 22 Basic-Level Experiments (Basic) and Experiments 1–3 (New)

Model	Strategy length		Internal practicability		Mixture	*M*		
	Basic	New	Basic	New	Basic	New	Basic	*M*
SLIP	100	100	82	100	85	100	85	88
Utility	83	50	51	80	80	65	63	64
Possession	50	50	77	70	40	60	63	62
Compression	83	50	23	50	60	50	40	42
Context	33	40	38	20	35	30	37	35

Note. SLIP = strategy length and internal practicability.

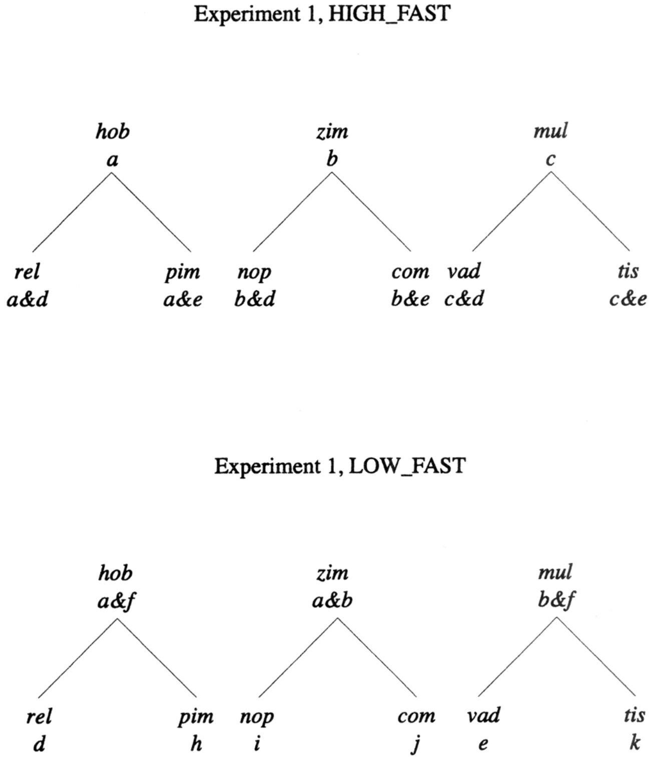


Figure 8. Taxonomies of Experiments 1, HIGH_FAST and LOW_FAST. Strategy length is the only varying variable. Below the category names are the optimal computational strategies of the strategy length and internal practicability model. An index for these abstract features is given in Appendix B.

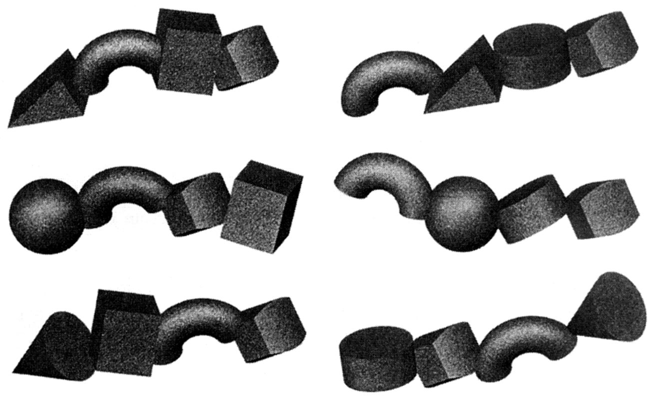


Figure 9. A sample of computer-synthesized objects used in Experiment 1, HIGH_FAST (one exemplar per low-level category).

Table 8
Observed Mean Response Times for Positive Trials in Experiments 1–3 and Numerical Predictions

Model	Level: Lowest	Level: Highest
Experiment 1		
HIGH_FAST		
Observation[a]	1,373	976
Compression[c]	0	30
Context	0.361	0.531
Possession	2	3
SLIP[b]	3.692	2.667
Utility	.195	.222
LOW_FAST		
Observation[a]	1,049	1,329
SLIP[b]	2.667	3.692
Compression[c]	0	30
Context	0.376	0.531
Possession	1	3
Utility	.25	.333
Experiment 2		
HIGH_FAST		
Observation[a]	854	708
Compression[c]	0	5
Possession	1	3
SLIP[b]	2.667	1.323
Utility	.375	.500
Context	0.607	0.531
LOW_FAST		
Observation[a]	890	970
Possession	3	1
SLIP[b]	1.323	2.667
Utility	.624	.500
Compression[c]	0	5
Context	0.581	0.592
Experiment 3		
EQUAL		
Observation[a]	733	763
Compression[c]	0	30
Possession	1	5
SLIP[b]	1.714	1.714
Utility	.176	.260
Context	0.376	0.531
SL_DOWN		
Observation[a]	1,036	1,192
SLIP[b]	1.714	2.219
Compression[c]	0	30
Context	0.376	0.531
Possession	1	5
Utility	.250	.333
IP_UP		
Observation[a]	1,013	909
Compression[c]	0	30
Context	0.376	0.531
Possession[b]	1	5
SLIP	4.8	2.219
Utility	.250	.333

Note. The cells under Level are ordered from left to right according to increasing generality. For each taxonomy, models are ranked from best to worst and ties are ordered alphabetically. SLIP = strategy length (SL) and internal practicability (IP).
[a] Data in milliseconds. [b] Data in attempts. [c] Data in bits.

omy, participants performed 240 verification trials. Each trial consisted in the presentation of a category name followed by an object. Participants had to decide whether these matched and we measured response latencies.

Results and Discussion

We performed an ANOVA on the logarithm of the RTs of correct positive trials (error rate = 2.3%). Table 8 shows the mean RTs. A two-way (Group × Level) ANOVA with repeated measures on the level variable revealed a significant interaction between group and level, $F(2, 27) = 5.67$, $p < .0095$, and two significant main effects of Group (SL_DOWN) × Level, $F(1, 27) = 4.25$, $p < .0495$ (mean high-level strategies = 3.001; mean low-level strategies = 2.953; only 1 participant responded faster for the high-level categories, $p < .011$) and Group (IP_UP) × Level, $F(1, 27) = 6.95$, $p < .0145$ (mean high-level strategies = 2.893; mean low-level strategies = 2.955; but 2 participants responded faster for the low-level categories, *ns*). The last main effect was not significant: Group (EQUAL) × Level, $F(1, 27) = 0.16$, *ns*. The error rate was low overall and was positively correlated with RT ($r = .31$, $p < .05$), ruling out a speed–accuracy trade-off.

SLIP predicted all the results observed in Experiment 3 (see SLIP in Table 8 for numerical predictions). Participants categorized equally fast at both levels in the EQUAL condition. Increasing the strategy length of the higher level in the SL_DOWN condition induced faster categorizations of the lower level. Diminishing practicability at the lower level then made the high level faster. The two computational factors of SLIP predicted speed of categorization in these taxonomies.

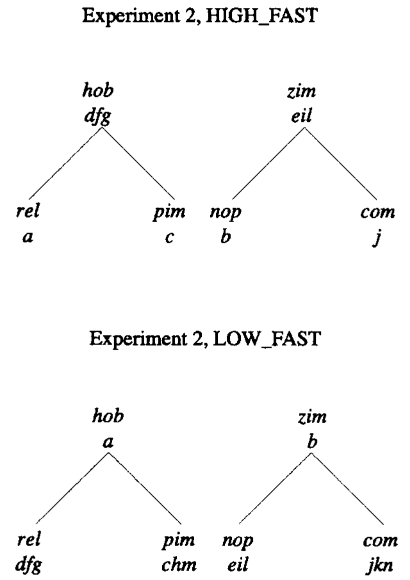


Figure 10. Taxonomies of Experiments 2, HIGH_FAST and LOW_FAST. Internal practicability (redundancy) is the only varying factor. Below the category names are the optimal computational strategies of the strategy length and internal practicability model. An index for these abstract features is given in Appendix B.

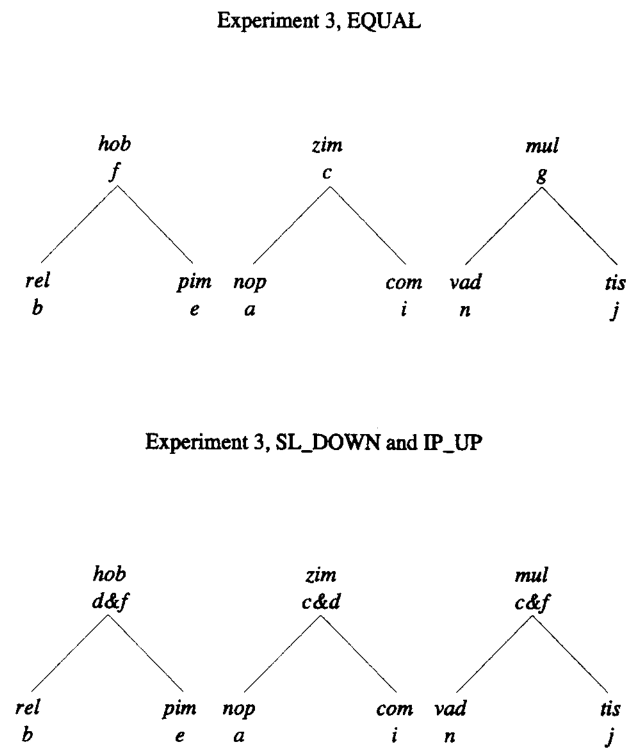


Figure 11. Taxonomies of Experiment 3, EQUAL, SL_DOWN and IP_UP. SL_DOWN and IP_UP differ in one important respect: The attributes of SL_DOWN always occupied the same positions, whereas this was not the case for those of IP_UP. Specifically, the high-level attributes of IP_UP always occupied the same positions in the input object, whereas the low-level attributes could occupy one of four possible positions. To model this, we integrated position in our formalism (see Appendix D). Below the category names, we provide the optimal strategies of the strategy length (SL) and internal practicability (IP) model. An index for these abstract features is given in Appendix D.

Comparison of Models of Basic Levelness for Experiments 1–3

SLIP predicts all of the qualitative results of these three experiments if it is expanded to take into account feature configurations (see Appendix D for developments). Category utility follows with 65% of the data explained, followed by category feature possession with 60%, compression with 50%, and the context model with 30%. The decomposition of these global scores into strategy length and redundancy scores is as follows: For the conditions in which only strategy length is tested (Experiment 1: HIGH_FAST and LOW_FAST; Experiment 3: EQUAL and SL_DOWN), category feature possession, category utility, and compression each predicts 50% of the data; the context model predicts 40% (see Table 7). This confirms the argument made earlier that all models have so far neglected strategy length as a specific factor of basic-level performance. This is a serious problem because attributes do overlap between categories in the real world, so strategy length is an important factor of recognition outside the laboratory.

For the conditions in which only practicability is tested (Experiment 2: HIGH_FAST and LOW_FAST; Experiment 3: EQUAL and IP_UP), category utility *measure* predicts 80% of the data, category feature possession, 70%, compression measure, 50%, and the context model, 20%. (We included all of the Experiment 3 conditions in both strategy length and internal practicability.)

General Discussion

In this article we presented SLIP, a measure of basic-level performance that implements two computational constraints on the organization of information in taxonomies: strategy length, the number of feature tests necessary to place the input in one category, and internal practicability, the ease with which these tests can be performed. We reviewed 22 published experiments and examined how the two constraints varied in each one of them. We tested the empirical validity of the two constraints in three experiments. In Experiment 1 we isolated the effect of strategy length on basic levelness, in Experiment 2 the effect of internal practicability, and in Experiment 3 the interactions of the two constraints. We compared the predictive power of SLIP with that of four established models of basic-level performance: the context model (Estes, 1994; Medin & Schaffer, 1978), category feature possession (Jones, 1983), category utility (Corter & Gluck, 1992), and compression (Pothos & Chater, 1998). When combining the predictions for 22 published experiments (see Tables 1 to 6) with those for our three new experiments (see Table 8), it appears that SLIP predicts 88% of the data, category utility 64%, category feature possession 62%, the compression measure 42%, and the context model 35% (see Table 7).

To the extent that any model of categorization implements computational constraints, even if these are not well specified, the conclusion is that SLIP's are closest to those underlying the speed of access to the hierarchy of categories. It is worth noting that these two constraints are strictly of a categorical nature: In SLIP, speed of access is a direct function of the representation of object information in memory. We purposefully normalized the presentation of taxonomies here to clarify how a change in their structures modifies the relative speed of access to their levels of categorization. To our knowledge, this is the first time that such computational analysis has been performed.

We should be careful to point out that we do not believe that the two constraints of SLIP completely predict the speed of access to categories in the real world. As discussed in Schyns (1998), the phenomenology of recognition, including speed of access, arises from interactions between the information demands of a categorization task and the availability of object information in distal stimuli. SLIP formalizes the information demands of categorization tasks with different strategies, which are composed of a number of sets of redundant tests on object attributes. We are aware that some of these attributes will be more difficult to extract than others from the input. It will be interesting to study further how speed of access arises from an interaction between the information demands of strategies and the relative availability of this information in the input. In fact, SLIP enables a study of the respective contributions of these two sources of influence.

Generalization to Other Correlates of Basic Levelness

We designed SLIP to model category verification. This is by far the most common method to assess the basic levelness of a category. However, we pointed out earlier that a critical aspect of basic levelness is that it optimizes a number of indexes of performance. It is important to show that SLIP is not limited to model category verification. For example, in a naming task—in response to "What is this object?"—SLIP can apply most of its strategies in parallel, and output the name associated with the first completed strategy. The formal models for verification and naming are therefore strictly identical. Thus, SLIP predicts the same qualitative order of speed of access in naming and verification. However, parallel testing of multiple strategies increases the likelihood that many more f-agents will compete for the same information channels in naming than in verification (when only one strategy is resolved at any time; see Appendix C for a discussion of competition in the context of disjunctions). With this competition, SLIP predicts that it will take more time on average to name than to verify (Rosch et al., 1976).

Another correlate of the basic levelness of a category is that people tend to list many more features, especially shape features, at this level than at others (Rosch et al., 1976; Tversky & Hemenway, 1984). Remember that most features of one basic-level category do not overlap with those of contrasting categories (e.g., Tversky & Hemenway, 1984, Tanaka & Taylor, 1991). Following the principles of SLIP, the addition of such diagnostic features in a category increases its internal practicability, and its basic levelness. SLIP therefore predicts a proliferation of listed features at the basic level. The fact that these features are mostly shape rather than color and texture could reflect the organization of our perceived world.

A similar reasoning applies to the discovery of Tanaka and Taylor (1991) and Johnson and Mervis (1997) that expertise induces faster verification times, and number of listed features, for subordinate categories. It also applies to the observation of Jolicoeur et al. (1984) and Murphy and Brownell (1985) that atypical subordinates (e.g., *penguin, electric knife*) behave more like basic-level categories than the other subordinate categories (e.g., *robin, Swiss knife*). Murphy and Brownell have shown that these atypical subordinates are more informative (i.e., they have more listed features) and distinctive (i.e., they share fewer of these listed features with contrasting categories) than other subordinates. In other words, atypical subordinate categories have more internal practicability, are more redundant, than other subordinate categories. In summary, the computational principles of SLIP can account for the most important correlates of basic levelness: faster verification, naming, and number of listed features.

Turning to development, it has been suggested that children have a *comprehension* bias (innate or learned) for the basic level. Because adults show the bias in *production,* this would enable the children and adults to resolve the level of categorization ambiguity and understand each other (Markman, 1989; Markman & Hutchinson, 1984). We have argued here that the production bias for basic names in adults arises from the organization of their taxonomic knowledge and the resulting strategies that access the categories. The development literature is unclear about the origin of the bias for children to comprehend at the basic level. SLIP suggests that infants acquire concept taxonomies (e.g., Eimas & Quinn, 1994) and access them following the general principles of SLIP. Adults would produce basic names because they are first accessed in their "mental race," and children would connect these names with basic concepts because the latter are also first accessed in their mental race. This does not imply that the taxonomic organizations of adults and children are identical, only that the same categories are first accessed. In other words, adults and children can differ markedly in the number of categories and levels of categorization they have in memory, but still access the same basic-level categories.

Concluding Remarks

The two computational principles of SLIP appear to explain a large proportion of the variance of basic-level experiments. The model itself has more predictive power than its competitors and it allows explicit predictions of the sequencing of information intake in categorization and recognition. For example, if some animal categories are more equal than others (e.g., in verification tasks, *dog* is superior to *mammal* as well as to *Doberman*), we would say that it is because they have shorter strategies or strategies with greater internal practicability.

References

Adelson, B. (1985). Comparing natural and abstract categories: A case study from computer science. *Cognitive Science, 9,* 417–430.

Anderson, J. R. (1990). *The adaptative character of thought.* Hillsdale, NJ: Erlbaum.

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

Anglin, J. M. (1977). *Word, object and conceptual development.* New York: Norton.

Berlin, B. (1992). *Ethnobiological classification: Principles of categorization of plants and animals in traditional societies.* Princeton, NJ: Princeton University Press.

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

Brown, R. (1958). How shall a thing be called? *Psychological Review, 65,* 14–21.

Cantor, N., & Mischel, W. (1979). Prototypes in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 3–52). New York: Academic Press.

Cantor, N., Smith, E. E., French, D. S., & Mezzich, J. (1980). Psychiatric diagnosis as prototype categorization. *Journal of Abnormal Psychology, 89,* 181–193.

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

Corter, J. E., Gluck, M. A., & Bower, G. H. (1988). Basic levels in hierarchically structured categories. In V. L. Patel & G. J. Groen (Eds.), *Proceedings of the 10th Annual Conference of the Cognitive Science Society* (pp. 118–124). Hillsdale, NJ: Erlbaum.

Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually-based basic-level categories in young infants. *Child Development, 65,* 903–917.

Estes, W. K. (1994). *Classification and cognition.* New York: Oxford University Press.

Fisher, D. (1987). Knowledge acquisition via incremental conceptual clustering. *Machine Learning, 2,* 139–172.

Fisher, D. (1988). A computational account of basic level and typicality effects. In *Proceedings of the Seventh National Conference on Artificial Intelligence* (pp. 233–238). Saint Paul, MN: Morgan Kaufmann.

Gosselin, F., & Schyns, P. G. (1998). Contingency of parts in object concepts. In S. J. Derry & M. A. Gernsbacher (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (p. 1222). Hillsdale, NJ: Erlbaum.

Hoffmann, J., & Ziessler, C. (1983). Objectidentifikation in kunstlichen

begriffshierarchien [Object identification in artificial concept hierarchies]. *Zeitschrift fur Psychologie, 194,* 135–167.
Horton, M. S., & Markman, E. M. (1980). Developmental differences in the acquisition of basic and superordinate categories. *Child Development, 51,* 708–719.
Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General, 126,* 248–277.
Jolicoeur, P., Gluck, M., & Kosslyn, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology, 19,* 31–53.
Jones, G. V. (1983). Identifying basic categories. *Psychological Bulletin, 94,* 423–428.
Kersten, D. (1990). Statistical limits to image understanding. In C. Blackermore (Ed.), *Vision: Coding and efficiency* (pp. 32–44). Cambridge, England: Cambridge University Press.
Lamberts, K. (1994). Flexible tuning of similarity in exemplar-based categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1003–1021.
Lassaline, M. E. (1990). *The basic level in hierarchical classification.* Unpublished master's thesis, University of Illinois.
Lassaline, M. E., Wisniewski, E. J., & Medin, D. L. (1992). Basic levels in artificial and natural categories: Are all basic levels created equal? In B. Burns (Ed.), *Percepts, concepts and categories: The representation and processing of information* (pp. 326–378). New York: North-Holland.
Malt, B. C. (1995). Category coherence in cross-cultural perspective. *Cognitive Psychology, 29,* 85–148.
Markman, E. M. (1989). *Categorization and naming in children: Problems of induction.* Cambridge, MA: MIT Press.
Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meanings: Taxonomic vs. thematic relations. *Cognitive Psychology, 16,* 1–27.
Marr, D. (1982). *Vision.* San Francisco: Freeman.
Medin, D. L. (1983). Structural principles of categorization. In B. Shepp & T. Tighe (Eds.), *Interaction: Perception, development and cognition* (pp. 203–230). Hillsdale, NJ: Erlbaum.
Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85,* 207–238.
Mervis, C. B., & Crisafi, M. A. (1982). Order of acquisition of subordinate-, basic-, and superordinate-level categories. *Child Development, 53,* 258–266.
Morris, M. W., & Murphy, G. L. (1990). Converging operations on a basic level in event taxonomies. *Memory & Cognition, 18,* 407–418.
Murphy, G. L. (1982). Cue validity and levels of categorization. *Psychological Bulletin, 91,* 174–177.
Murphy, G. L. (1991). Parts in objects concepts: Experiments with artificial categories. *Memory & Cognition, 19,* 423–438.
Murphy, G. L., & Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 70–84.
Murphy, G. L., & Lassaline, M. E. (1997). Hierarchical structure in concepts and the basic level of categorization. In K. Lamberts & D. R. Shanks (Eds.), *Knowledge, concepts and categories: Studies in cognition* (pp. 93–131). Cambridge, MA: MIT Press.
Murphy, G. L., & Smith, E. E. (1982). Basic level superiority in picture categorization. *Journal of Verbal Learning and Verbal Behavior, 21,* 1–20.
Newport, E. L., & Bellugi, U. (1978). Linguistic expression of category in a visual-gestural language: A flower is a flower is a flower. In E. Rosch & B. B. Lloyd (Eds.), *Semantic factors in cognition* (pp. 137–168). Hillsdale, NJ: Erlbaum.
Nosofsky, R. M. (1986). Attention, similarity and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115,* 39–57.
Pothos, E. M., & Chater, N. (1998). Rational categories. In S. J. Derry & M. A. Gernsbacher (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 848–853). Hillsdale, NJ: Erlbaum.
Rifkin, A. (1985). Evidence for a basic level in event taxonomies. *Memory & Cognition, 13,* 538–556.
Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Semantic factors in cognition* (pp. 137–168). Hillsdale, NJ: Erlbaum.
Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.
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. (1998). Diagnostic recognition: Task constraints, objects information, and their interactions. *Cognition, 67,* 147–179.
Shaver, P., Schwarz, J., Kirson, D., & O'Connor, D. (1987). Emotion knowledge: Further explorations of prototype approach. *Journal of Personality and Social Psychology, 52,* 1061–1086.
Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.
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.
Tarr, M. J., Bülthoff, H. H., Zabinski, M., & Blanz, V. (1997). To what extent do unique parts influence recognition across changes in viewpoint? *Psychological Science, 8,* 282–289.
Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352.
Tversky, B., & Hemenway, K. (1983). Categories of environmental scenes. *Cognitive Psychology, 15,* 121–149.
Tversky, B., & Hemenway, K. (1984). Objects, parts and categories. *Journal of Experimental Psychology: General, 113,* 169–191.
Wisniewski, E. J., & Murphy, G. L. (1989). Superordinate and basic category names in discourse: A textual analysis. *Discourse Processing, 12,* 245–261.

Appendix A

A Serial Version of SLIP

A strategy to categorize an object at a given level of abstraction is defined here as a series of feature tests. Typically, some of these feature tests are unique to this category, and some overlap with the defining features of other categories. An optimal strategy is the shortest series of tests on the features defining the category. We posit that the strategy length and internal practicability (SLIP) categorizers always use optimal strategies. We call redundant features, or set of redundant features, the collection of features that, individually, provide exactly the same information as to the category membership of objects. In other words, testing one, two, or more redundant features does not provide more information.

Formally, we say that a strategy is a series of sets of redundant feature tests. It has succeeded whenever all sets of redundant features tests have been completed in a specific order. Furthermore, a set of redundant features is completed as soon as a test on the presence of one of its redundant features has been performed.

This usually happens after a succession of misses. The probability of having $t - 1$ successive misses is given by $(1 - \psi_j)^{t-1}$, where ψ_j—when redundancy of sets of features and the number of possible configurations that these can take in objects are taken into account—is equal to $C_j(1 - S) + C_jSR_j$; that is, the practicability of set of redundant features j or the probability that it will be completed after a single attempt. S is the probability of a random slip (it was arbitrarily set to .5 throughout the simulations), and C_j is the probability that the target features will be in the expected configuration (one divided by number of configurations). Thus, the first term of ψ_j is the probability that the SLIP categorizer will guess the feature configuration correctly and that it will not slip. R_j is the probability that a random slip will result in a diagnostic test, (cardinality of j)/(number of features in objects). The second term of ψ_j is the probability that the categorizer will slip, but that it will guess the correct configuration and will perform a diagnostic feature test.

The probability of a hit is one minus the probability of a miss. Thus, the probability that the set of redundant features j will be completed after t trials is

$$(1 - \psi_j)^{t-1}\psi_j,$$

and the probability that a strategy of length n will have succeeded after t trials in a certain configuration of hits and misses is

$$\prod_{j=1}^{n} (1 - \psi_j)^{\phi}\psi_j,$$

where ϕ is a function of j, it will remain unspecified, which gives the number of misses for the jth set of redundant features for that particular configuration. Usually, more than one such configuration exist. In fact, the number of possible configurations is easy to compute. The last hit necessarily happens at the tth trial; the $n - 1$ other hits, however, can happen anywhere in the $t - 1$ trials left, in order. Therefore, the number of possible configurations is the number of combinations of $t - 1$ items taken $n - 1$ by $n - 1$ that is,

$$\lambda = \binom{t-1}{n-1} = \frac{(t-1)!}{(t-n)!(n-1)!}.$$

We can now give the global shape of the probability that a strategy of lengths n will succeed after t trials:

$$\sum_{i=1}^{\lambda} \prod_{j=1}^{n} (1 - \psi_j)^{\omega}\psi_j,$$

where ω is a function of i, and j specifies the number of misses for the jth set of redundant features for the ith configuration of hits and misses. We call this the response time function (RTF). We still have to specify ω. We establish a connection between this function and multinomial expansions. The multinome $(a_1 + a_2 + \ldots + a_n)^{t-n}$ expands into λ different terms, and the sum of the n exponents of each term is equal to $t - n$. It follows that ω gives the jth exponent of the ith term in this multinomial expansion.

As a global measure of basic levelness, we use *t_mean*, the mean number of tests required to complete a strategy. When internal practicability is constant within a strategy (this is true for all experiments reported in this article), the RTF is a Pascal density function and, thus, t_mean is equal to n/ψ.

Appendix B

The Figures Attribute Index

Murphy and Smith's (1982; see Figures 2 and 4) and Murphy's (1991, Experiments 4–5; see Figures 2 and 3) artificial tools: a = pounder, b = cutter, c = hammer head, d = hammer shaft, e = hammer handle, f = brick head, g = brick shaft, h = brick handle, i = knife head, j = knife shaft, k = knife handle, l = pizza cutter head, m = pizza cutter shaft, n = pizza cutter handle, o = wide hammer head, p = narrow hammer head, q = one-part brick handle, r = two-part brick handle, s = straight knife edge, t = serrated knife edge, u = short pizza cutter shaft, v = long pizza cutter shaft, w = large, x = small, y = stirrer, z = scrapper, a' = wedge head, b' = wedge shaft, c' = wedge handle, d' = carrot head, e' = carrot shaft, f' = carrot handle, g' = scoop head, h' = scoop shaft, i' = scoop handle, j' = rake head, k' = rake shaft, l' = rake handle, m' − b'' = small variations, c'' = red, d'' = dot texture, e'' = yellow, f'' = circle texture, g'' = blue, h'' = solid color, i'' = green, j'' = stripe texture, k'' = broken edges, and l'' = continuous edges. Lassaline (1990) used the same artificial tools with two added dimensions for her own experiment.

Mervis and Crasifi's (1982; see Figure 2) artificial objects: a = straight lines, b = curved lines, c = sharp corners, d = smooth corners, e = overall square shape, f = interior vertical line, g = interior losange, h = lined texture, i = overall triangle shape, j = interior oblique line, k = interior cross, l = shaded texture, m = overall concave cell shape, n = elongated nucleus, o = interior fat floating body, p = slim excroissance, q = overall convex cell shape, r = Y-shaped nucleus, s = interior slim floating body, t = fat excroissance, u − b' = small configural changes, and c' − z' = minor variations.

Hoffmann and Ziessler's (1983; see Figure 5) PacMan's ghosts: a = rectangular shell, b = curved shell, c = interior square, d = interior triangle, e = interior star, f = interior circle, g = triangular saw teeth bottom edge, h = broken vertical lines bottom edge, i = rectangular saw teeth bottom edge, j = circular saw teeth bottom edge, k = shaded texture, and l = lined texture.

Gosselin and Schyns's (1998; see Figure 5) three-dimensional artificial objects in the color, texture, geon (CTG) condition: a = red, b = green, c = smooth, d = rough, e = cone, and f = pyramid; in the geon, color,

(*Appendixes continue*)

texture (GCT) condition, a = cone, b = pyramid, c = red, d = green, e = smooth, and f = rough; and, in the texture, geon, color (TGC) condition, a = smooth, b = rough, c = cone, d = pyramid, e = red, and f = green.

Murphy's (1991, Experiment 3; see Figure 6) artificial stamps: a = blue, b = yellow, c = red, d = green, e = texture squares, f = solid (texture squares), g = wavy edges, h = texture lines, i = wavy (texture lines), j = straight edges, k = texture stripes, l = thick (texture stripes), m = no edge, n = texture circles, o = empty (texture circles), p = jagged edge, q = large, r = small, s = broken edge, t = continuous edge, u = vertical stripes, v = horizontal stripes, w = large circles, and x = small circles.

Our geon chains (Experiments 1–3; see Figures 8, 10, and 11): a = wedge, b = sphere, c = cone, d = cube, e = cylinder, f = macaroni, g = fat cube, h = slim cylinder, i = trumpet, j = magnet, k = fat cylinder, l = slim cube, m = hook, and n = pyramid.

Appendix C

SLIP and Disjunctions

The strategy length and internal practicability (SLIP) model can handle disjunctive categories with one additional assumption: Information channels can only be probed by a single feature agent (f-agent) at a time (this does not change any of our previous ordinal predictions). As with any other strategy, each attribute in a disjunctive strategy is assigned an f-agent for testing. There are two types of disjunctive strategies: In the first type, the attributes of a disjunction share the same information channel; whereas in the second type, the attributes scan different information channels. This difference is important because f-agents will compete for the same information channel in the first, but not in the second case. With this assumption of uniqueness of f-agent per channel at any given time, competing f-agents in a disjunction will increase the time of its completion compared with one of its features tested in isolation.

To illustrate, consider the two taxonomies of Lassaline's (1990) Experiment 3. Remember that two-attribute either-or disjunctions defined the high-level categories and one different feature defined each low-level category, in the two taxonomies. In the four-dimension condition, low-level features were one value of four different stimulus dimensions (e.g., *hammer head, pizza cutter handle, dotted texture,* and *square internal shape*). In the one-dimension condition, low-level features were four different values of the same dimension (e.g., *hammer, brick, knife,* and *pizza cutter heads*).

That is, attributes in the disjunctions of the one-dimension condition share their information channels, whereas attributes in the four-dimension do not. Thus one-dimension disjunctions should be verified slower than their individual attribute (because of competition for the unique information channel), whereas four-dimension disjunctions should be as fast to complete as their individual attributes (because each scans a different information channel and either one or the other will succeed).

Formally, we need to develop SLIP to account for shared-channel disjunctions—the other type of disjunctions is formally equivalent to the set of redundant features. We start with the assumption that the n f-agents of one disjunction can access the single information channel with equal probability. Then, I, the probability that one f-agent cannot access the channel is $1 - 1/n$. In Lassaline's (1990) one-dimension disjunctions, this probability is equal to one half. With S the probability of a slip, two independent events can lead to a failure of testing one attribute: (a) The f-agent does not slip but does not access its information channel. This event occurs with a probability of $(1 - S)I$; (b) the f-agent slips on a wrong information channel (with probability $S - SQ$), or its information channel, but its access is blocked (with probability SQI). The probability of a nonsuccessful slip is therefore $(S - SQ) + SQI$, where Q is one divided by the total number of attributes in the input object (i.e., one fourth in the case of Lassaline). To conclude, φ, the probability that a f-agent in a shared-channel disjunction successfully tests its attribute is $1 - [(1 - S)I + (S - SQ) + SQI]$.

Generalizing to $\psi_j = 1 - (1 - \varphi)^{m_j}$, the equation becomes $\psi_j = 1 - (I + S + SQI - SI - SQ)^{m_j}$. For the high- (the shared-channel disjunctions) and low-level categories of Lassaline's (1990) one-dimension condition, with $S = .5$, we obtain ψ_j equals 0.292 and 0.417, respectively.

Appendix D

SLIP and Configurations

To model the data of Experiment 3, we need to develop the strategy length and internal practicability (SLIP) to take feature configurations into account. Remember that each f-agent looks for its feature in one information channel. So far, information channel was loosely defined, but in Experiment 3, it can be assigned a specific meaning: a position in an oriented string of features. SLIP needs to be expanded to account for the interactions between information channels and features. This is the goal of this appendix.

The SL_DOWN and IP_UP conditions in Experiment 3 have the same taxonomy (see bottom taxonomy of Figure 11). The only difference between them is the number of positions that geons can occupy in the six-geon strings. In the SL_DOWN condition, geons always occupy the same position; in the IP_UP condition, however, the two geons of the conjunction that defines the high-level categories always occupy the same position, whereas the unique geons that define the low-level categories can each occupy four positions.

C_j is the probability that the f-agent guesses the correct position j, where its geon should appear. It is one divided by the number of possible positions of geons (i.e., $C_j = 1$ at the high level of IP_UP, and one fourth at the low level). If S is the probability of a slip, two independent events can lead to a successful test of one geon: (a) a right guess of the position j, with a probability of $C_j(1 - S)$—in the IP_UP example, this probability becomes $(1 - S)$ at the high level and $.25*(1 - S)$ at the low level; (b) a slip, by change alone, on the right position j, with probability SQ, where Q is one divided by the total number of attributes in the input object (i.e., one sixth in the earlier examples). In sum φ, the probability that a f-agent successfully tests its geon is $C_j(1 - S) + SQ$. Generalizing to $\psi_j = 1 - (1 - \varphi)^{m_j}$, the equation becomes $\psi_j = 1 - \{1 - [C_j(1 - S) + SQ]\}^{m_j}$. With $S = .5$, we have ψ_j equals 0.417 and 0.042, respectively, for the high- and low-level categories of the IP_UP conditions.

Received December 30, 1999
Revision received March 8, 2001
Accepted March 22, 2001 ■