TR-H-204

Diagnostic Recognition: Task Constraints, Object Information, and their Interactions.

Philippe SCHYNS

1996.11.13

ATR人間情報通信研究所

〒619-02　京都府相楽群精華町光台2-2　TEL: 0774-95-1011

ATR Human Information Processing Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02, Japan

Telephone: +81-774-95-1011

Fax　　　: +81-774-95-1008

Diagnostic Recognition: Task Constraints, Object Information, And Their Interactions

Philippe G. Schyns
University of Glasgow

Running Head: Diagnostic Recognition

Please send all correspondance to:
Philippe G. Schyns
Dept. of Psychology
56, Hillhead St.
University of Glasgow
Glasgow G12 8QB
United Kingdom
philippe@psy.gla.ac.uk

ABSTRACT

Object recognition and categorization are both concerned with the understanding of how input information matches with memorized object information. It is therefore surprising that these two fields have evolved independently, without much cross-fertilization between their research. It is the main objective of this paper to lay out the basis of a dialogue between object recognition and categorization research, with the hope of raising issues the could cross-fertilize both domains. To this end, this paper develops *diagnostic recognition,* a framework which formulates recognition performance as an interaction of task constraints and object information. We argue and present examples suggesting that diagnostic recognition could be applied to the understanding of everyday object recognition.

Object recognition and categorization research are both concerned with the question "what is this object?" To recognize an object as a car is not very different from placing the object in the *car* category. In both cases, the problem is to understand how input information matches with memorized information. For example, if *handle* and *container* or equivalent information represent cups in memory, a *cup* recognition/categorization is only possible if this information is perceptually available.

Given such profound similarity, it is surprising that object recognition and categorization have evolved separately, without much cross-fertilization between their research. The reason for this could be a difference of focus: Typical categorization studies have sought to understand the rules governing the formation of categories (e.g., the representation of *cup* is the combination of the features *handle* and *container)*, while recognition researchers have mostly looked into the perceptual aspects of the recognition process (e.g., the perceptual representations of *handle* and *container* that enable the *cup* recognition). However, recent debates on the possible interactions between categorization and perception have suggested that the principles governing the formation of categories should be more tightly coupled with the perceptual aspects of recognition (Goldstone, 1994; Schyns & Murphy, 1991, 1994; Schyns, Goldstone & Thibaut, in press; Thibaut, 1991). It is proposed that such interactions will promote the emergence of new, more powerful theories of visual cognition.

It is the main objective of this paper to lay out the basis of a dialogue between object recognition and categorization studies, to raise issues that could cross-fertilize both fields. To this end, the first section develops *diagnostic recognition*, a framework which integrates two main factors: The task constraints of categorization studies, and the perceptual information of recognition theories. It is important to stress from the outset that diagnostic recognition is *not* a new theory of object recognition. Instead, it is a broad framework which proposes one possible answer to the question: "How could we frame issues common to object recognition and categorization theories?" The first section illustrates with two examples how recognition/categorization performance can result from interactions between task constraints and perceptual information. The remaining sections discuss implications of diagnostic recognition for the study of "everyday object recognition."

The diagnostic recognition framework: Interactions of task constraints and object information. Even everyday observation reveals that a single object fits into many possible categories. For example, an object may be recognized as a *Porsche*, a *car*, or a *vehicle*. On other occasions, it may be called a *toy*, an *expensive gift*, a *public nuisance*, or a *public danger* which sometimes leads to *scrap metal*. Categorization is highly flexible and people tend to place an object into one category or another depending on their goals and actions.

It is always worth stressing that different classifications of an identical object tend to change the information requirements of recognition tasks. For example, when assigning a visual event to the *Porsche, collie, sparrow, Mary* or *New York* category comparatively more specific information might be necessary than when categorizing it as a *car, dog, bird, human face* or *city*. I will not consider all possible object classifications here, but instead focus on the information constraints associated with the hierarchical organization of categories; the idea that an object belongs to a sequence of progressively more inclusive categories such as *Porsche, car, vehicle*. Within this hierarchy, I will concentrate on the initial, or so-called "perceptual" classifications (e.g., *Porsche* or *car*), instead of the abstract classifications (e.g., *vehicle*) which are arguably more detached from the perceptual input. Henceforth, *task constraints* will denote the visual information needed to place the input into the hierarchy of perceptual categories. Although task constraints have traditionally been the province of categorization research, they are an irreducible factor of *any* recognition task, and the first factor considered in diagnostic recognition.

The second factor is the nature of the perceptual information available to form hierarchically organized categories. Objects form categories because they "look alike"--i.e., they share cues such as a similar silhouette or global shape, distinctive sets of parts similarly organized (e.g., nose, mouth, eyes, ears, hair and their structural relationships), or characteristic surface properties (e.g., smooth vs. discontinuous, symmetric vs. asymmetric, and textural, color and illumination cues). Generally speaking, there are perceptual limitations to the extraction of image cues. For example, two objects could share a similar global silhouette, but have very different internal features (think, e.g., of Navon's letters, 1977). The Global-Precedence effect predicts that the availability of the silhouette should precede in time the availability of the internal features (e.g., Navon, 1977; though see also Grice, Graham & Boroughs, 1983). Such perceptual limitation constrains the spectrum of object similarities and the categories that can be spontaneously formed. The perceptual availability of object information has traditionally been the province of perceptually-oriented object recognition researchers. However, perceptual cues are an irreducible factor of any object categorization, and the second factor of diagnostic recognition.

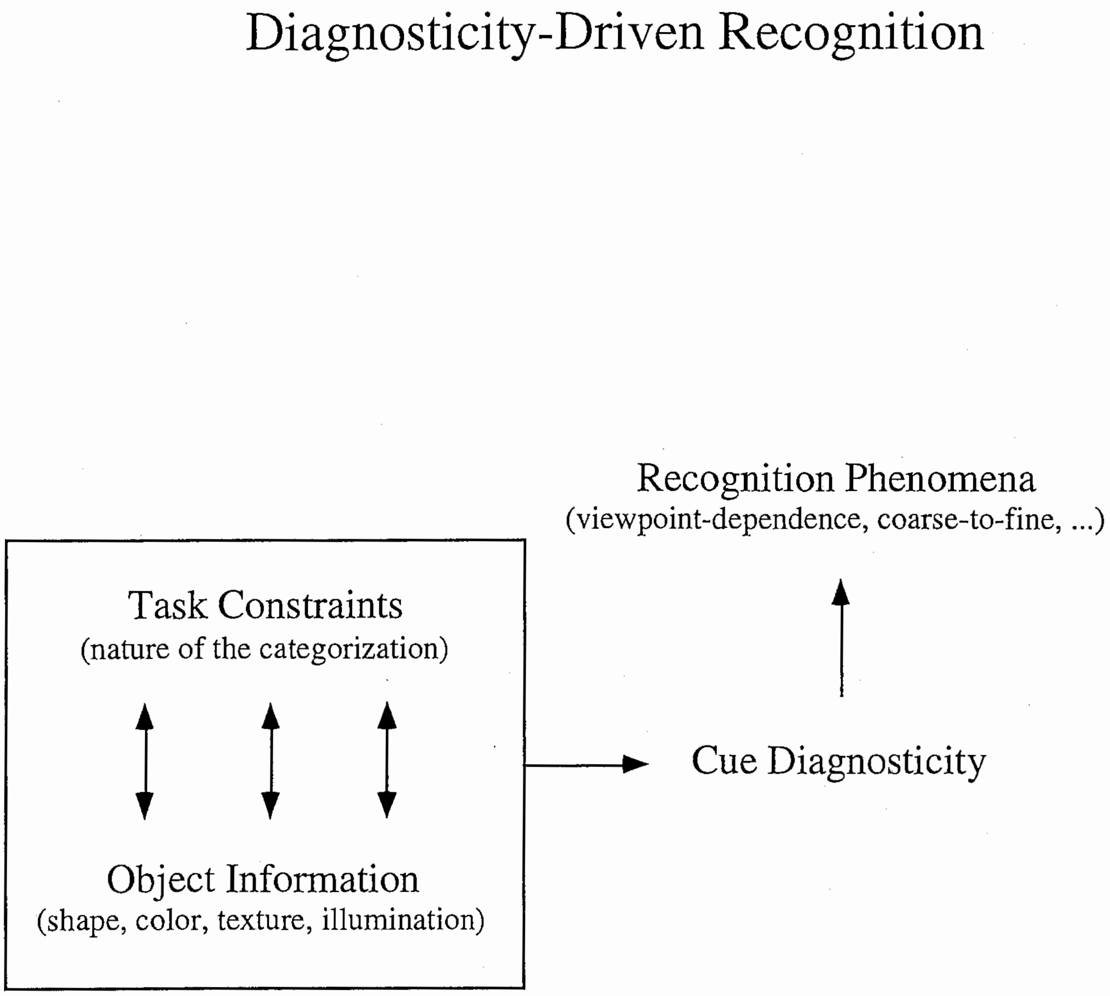

Figure 1. This figure illustrates the main components of diagnostic recognition: task constraints and object information, whose interactions give rise to cue diagnosticity. Diagnostic recognition claims that cue diagnosticity could explain the usage of image information subtending many recognition phenomena.

Diagnostic recognition seeks to frame recognition *performance* as an interaction of task constraints and perceptual object information, the two factors just discussed (see Figure 1). This is how it would work: When the information required to assign an object to a category matches with input information, a subset of object cues become particularly useful (i.e., diagnostic) for the task at hand. Diagnosticity is the first component of recognition performance. However, perceptual limitations on the extraction of this diagnostic information could also affect recognition performance. Thus, the interaction of task-dependent diagnosticity and perceptual limitations could explain the usage of information in the input image. The main claim of diagnostic recognition is that differences in usage of information could account for a wide range of recognition phenomena (e.g., viewpoint-dependence vs. independence, basic-level recognition and scale-dependent recognition).

We should be careful in pointing out that diagnosticity is already a component of most object recognition and categorization theories. However, we believe that its impact has not always been fully appreciated. For example, the diagnosticity of features for different classifications has been thoroughly studied in categorization research (e.g., Anderson, 1991; Elio & Anderson, 1981; Estes, 1986; Gluck & Bower, 1988; Kruschke, 1992; Nosofsky, 1984, 1986; and many others), but these theories often adopt a stance of: "You tell me what the object cues are, and I will tell you how they are integrated to perform the object categorization" (Schyns et al., in press). Consequently, they place little constraint on what may count as an object cue, and they tend not to incorporate perceptual limitations in their explanations of performance. However, the second illustration of diagnostic recognition presented below will show that perceptual factors will need to be integrated in complete explanations of even very simple categorizations.

Opposedly, object recognition researchers are well aware of the importance of perceptual object cues. However, they sometimes overlook the influence of task constraints. When object recognition researchers use performance to infer the format of an object representation, there is always a danger of "over-representing" the information demands of a particular task in the theory, and "under-representing" the object cues that could be diagnostic of other tasks. Example 1 will illustrate the importance of studying multiple task constraints when inferring representation formats from recognition performance. In sum, we believe that complete theories of object recognition and categorization could benefit from a thorough integration of the two aspects of diagnostic recognition.

Example 1: Diagnostic recognition and viewpoint-dependence. One of the most challenging problems of object recognition is to explain the relative invariance of recognition to changes of an object's orientation. This is not to say that object recognition is fully viewpoint-invariant; there are now many independent evidence suggesting that a large number of objects are better recognized when shown from particular viewpoints (e.g., Bülthoff & Edelman, 1992; Edelman & Bülthoff, 1992; Palmer, Rosch & Chase, 1981; Rock & Di Vita, 1987; Perrett, Oram, Harries, Bevan, Benson & Thomas, 1991; Tarr & Pinker, 1989; Vetter, Poggio & Bülthoff, 1994). Subjects label such object views as "better" and are faster to categorize the objects shown in these views. Typically, "viewpoint-dependent recognition" refers to a monotonic increase in recognition performances (reaction times and/or error rates) with increasing misorientation from the preferred views. Evidence of such viewpoint-dependent recognition has been reported for familiar (e.g., Palmer et al., 1981), unfamiliar (e.g., Perrett & Harries, 1988; Rock & Di Vita, 1987), realistic and artificial objects (e.g., Bülthoff & Edelman, 1992; Tarr & Pinker, 1989), and conditions of viewpoint-dependence have become a central issue in object recognition.

However, the debate is still open about the interpretation of the phenomenon. One interpretation uses viewpoint-dependence to tease apart formats of object representation (see Biederman & Gerhardstein, 1995; Tarr & Bülthoff, 1995). For example, the "view-based approach" claims that objects are stored in memory as collections of discrete views and that dependence to a view reveals that it is effectively stored in memory. In contrast, structural, or model-based accounts argue that recognition is viewpoint-independent, at least over a limited range of viewpoints, mostly because recognition uses viewpoint-independent perceptual cues (non-accidental properties of object edges, Lowe, 1987) which form geons, the primitives of memorized object representations (Biederman, 1987).

Alternatively, diagnostic recognition is agnostic about the specific object representation format: It emphasizes the available image information (shape, part, color, or other derived cues) that is diagnostic for a particular object categorization. Viewpoint-dependence to one (or a subset of) view(s) would then be determined by the availability of diagnostic information in these views, rather then by how that information is represented in memory (Hill, Schyns & Akamatsu, in press). To illustrate, many views would convey sufficient information to categorize Mary's face as "face." However, many fewer views would authorize to classify the same face as "Mary" (because Mary's features are visible only from a restricted subset of views). These two categorizations of an identical face might change the information requirement of the task, the face cues that are diagnostic, and the subset of views that are preferred for recognition. Within the range of views in which diagnostic cues are visible, there could be geometrical and perceptual limitations on their extraction from the image. For example, although the nose of a face is visible in all views between the two profiles, its length (if it was important to identify Mary) might be easier to measure from a 3/4, or profile view than from the frontal view. These limitations, together with task constraints, would in the example predict that *length-of-nose* (in fact, its perceptual implementation) would only affect viewpoint-dependence when this cue was diagnostic of the categorization--i.e., in the "Mary," not the "face" classification.

Hill, Schyns and Akamatsu (in press) tested for the possibility of such relative patterns of viewpoint-dependent face recognition. Their subjects learned one view of a face and were tested on their generalization to other views of the same face. One set of experiments used shaded models of 3D laser-scanned faces, to isolate the influence of shape-from-shading cues (see Figure 2, the top left picture).

A second set of experiments added color and texture to the shaded models, to analyze the role of these supplementary cues (see Figure 2, the top right picture). Together, these two experiments framed viewpoint-dependence as an interaction between fixed task demands and variable stimulus information.

As a class, faces share geometrical object information to which perception could be attuned. One such property is their approximate bilateral symmetry (Vetter, Poggio & Bülthoff, 1994) which allows occluded cues to be inferred from a single learned view. Consequently, the learned view and its symmetric might be identified with equal accuracy, and possibly better than any other unseen views. Such effect of symmetric object information should be particularly salient with shaded face models, for which no other cue than shape is available from the image.

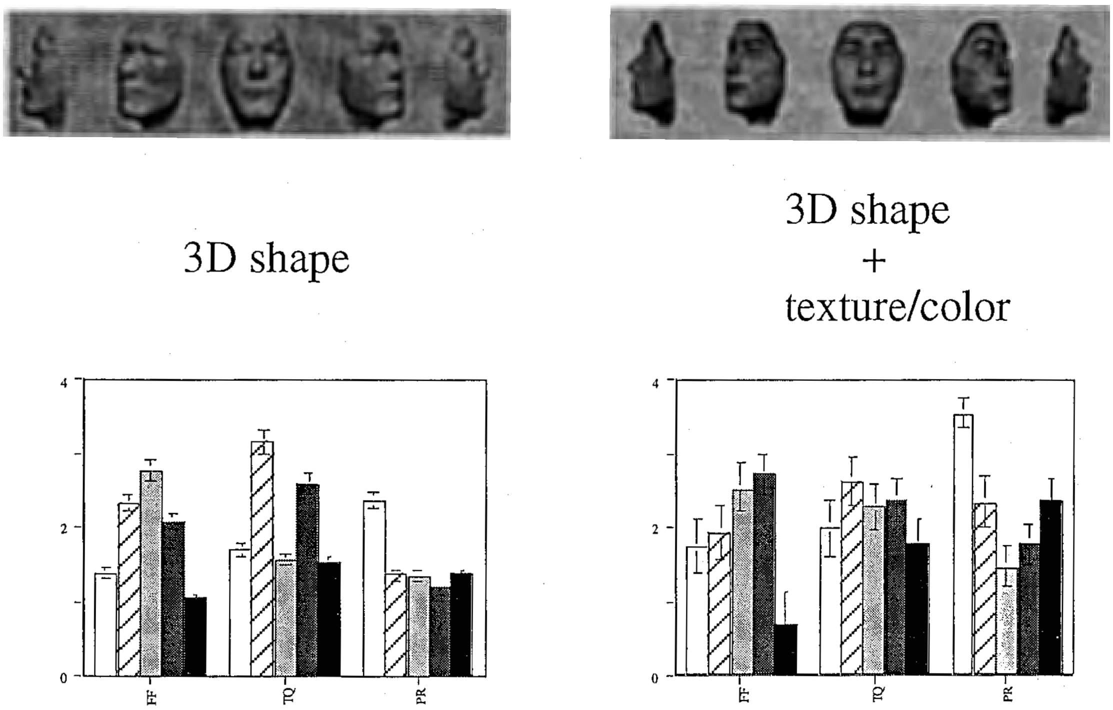

Figure 2. This figure (adapted from Hill, Schyns & Akamatsu, in press) illustrates the main results of an experiment on viewpoint dependence in face identification. The top left picture shows the face views of the shape-from-shading condition. The bottom left histogram presents patterns of viewpoint-dependent performance for the identification of shaded stimuli. The top right

picture shows the face views of the shape plus texture conditions, and the bottom right histogram presents the viewpoint-dependent performance for the identification of these stimuli.

Results of Hill et al. (in press) are summarized on Figure 2. The bottom left histogram illustrates that subjects who learned a three-quarter view recognized almost as efficiently the symmetric three-quarter, while performance decreased monotonically with rotation in depth for subjects who learned the profile, or the full-face view (see also Schyns & Bülthoff, 1994), confirming the role of symmetric face information in explanations of viewpoint-dependent identification.

The addition of color and textural cues in the same identification task affected performance (see Figure 2, the bottom right histogram). It was found that learning a three-quarter view now elicited good generalization to all views (of those tested). Also, a symmetric peak appeared to the other profile when learning a colored profile. Color and textural cues offered supplementary object information which reduced the overall viewpoint-dependence.

Evidence of such preferred views could prompt an object recognition researcher to hypothesize that these views (or their information content) actually represent faces in memory. However, there would be a difficulty with this strategy if each change categorization changed the overall pattern of performance. Object representations should be quite independent of the considered task. That is, they should ideally support many (not just one) categorizations of an object. Pilot studies were run that changed the categorization task, the other factor of diagnostic recognition. Using face sets identical to those of the identification task, subjects were now instructed to solve a generalization of gender across viewpoints. Performance was at chance with shaded faces, but it was near ceiling with textured faces, with no marked dependence on viewpoint in either case. This contrasted with the viewpoint-dependent performance observed for identity judgments.

In summary, this example suggested that viewpoint-dependent vs. independent recognition performance might be fruitfully framed as an interaction between the multiple categorizations of an object and its perceptual information. If a categorization requires selective input information, and if its extraction depends on viewpoint, recognition performance might reflect the requirement of "getting a good view" of the diagnostic cues. We come back in the General Discussion to the implications of changing task constraints for the study of object representations.

Example 2: Object information and categorization. As explained earlier, categorization research as a whole may have underestimated the importance of perceptual information in their explanations of performance. In typical experiments, (see, e.g., Bruner, Goodnow & Austin, 1956; Bourne, 1982; Shepard, Hovland & Jenkins, 1961), subjects were shown simple colored geometric shapes and were instructed to learn the rules for their categorization. These rules were logical combinations of the features that were clearly demarcated in the stimuli. For example, subjects could learn that the rule "red and circle" categorized the objects, but there was little doubt about the featural analysis of the stimuli.

Even though perception is not a *tabula rasa*, there are occasions when a relevant perceptual analysis of the incoming stimulus is not readily available. For example, novices reading chest X-rays (e.g., Christensen, Murry, Holland, Reynolds, Landay & Moore, 1981), sexing chicken (Biederman & Shiffrar, 1987), and categorizing dermatosis (Norman, Brooks, Coblenz & Babcock, 1992) do not seem to grasp the relevant object cues that structure these categories. Even when told what the diagnostic cues are, novices are not always able to see them in the image. In such cases, a significant part of learning the categories could involve learning the object information that structure them.

Here is how it would work: When a fragment of a stimulus distinguishes between members and nonmembers of categories, the fragment could become a new cue of perceptual analysis (Schyns & Murphy, 1994). That is, object cues would not necessarily be always fixed and perceptually given, as is often assumed in categorization theories, but rather flexible and perceptually adjustable to the experience and the categorization history of the individual. Consequently, two individuals with different categorization histories (think, e.g., of experts vs. novices of chest X-rays or dermatosis) might differently perceive and analyze an identical stimulus for its categorization. This could be a potential difficulty for existing object categorization theories which typically assume that category learning consists of selecting and weighting fixed perceptual evidence.

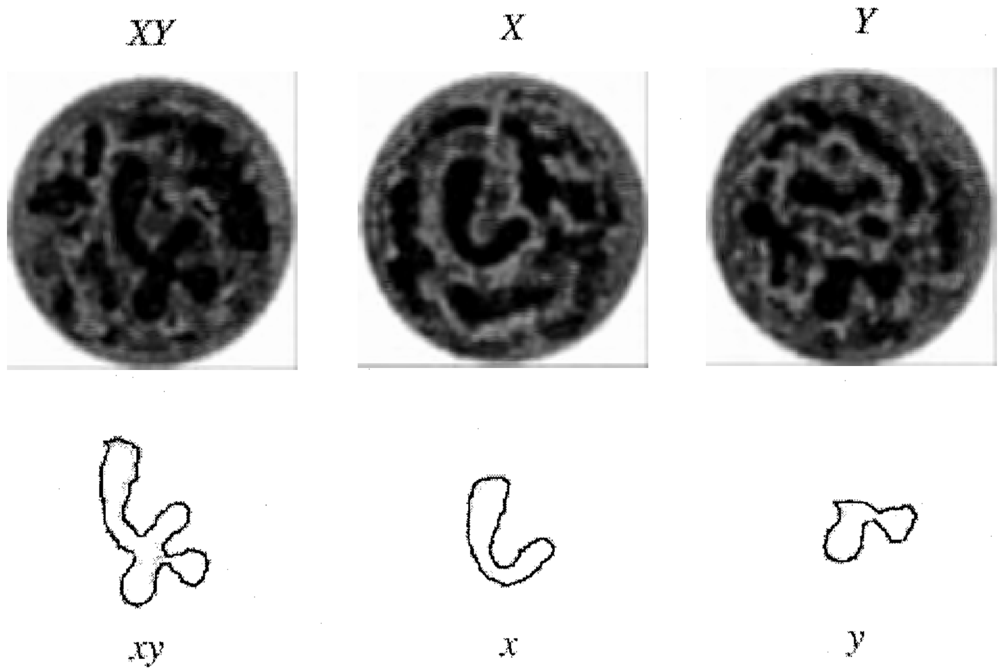

Figure 3. This figure illustrates the design of Schyns and Rodet's (in press) feature creation experiment. From left to right, the top pictures are Martian cell exemplars from the *XY*, *X*, and *Y* categories, respectively. From left to

right, the bottom pictures are the features *xy*, *x*, and *y*, defining the categories. Note that the feature *xy* is a composition of feature *x* and feature *y*. Subjects in the *XY->X->Y* (vs. *X->Y->XY*) group learned the category in this order.

Schyns and Rodet (in press) provided an "existence proof" that identical stimuli could be orthogonally perceived as a result of different categorization histories. Their experiments used three categories of unknown stimuli called "Martian cells" (see Figure 3). Categories were defined by specific blobs common to all members to which irrelevant blobs were added (to simulate various cell bodies). Figure 3 shows, from left to right, an exemplar of the *XY* category, and exemplar of *X*, and a *Y* exemplar. Figure 3 also shows the cues *x*, *y* and *xy* defining each category. Note that *xy* is the conjunction of *x* and *y*. One subject group was asked to learn *X* before *Y* before *XY* (*X->Y->XY*); the other group learned the same categories, but in a different order (*XY->X->Y*).

Results showed that while all subjects reliably classified *X*, *Y* and *XY* test stimuli (revealing that both groups saw, attended to and used the *x* cue and the *y* cue), their categorizations of *X-Y* test cells were orthogonal. *X-Y* tests were used to tease apart a conjunctive (*x&y*) vs. a configural (*xy*) perceptual analysis of *XY*: the components *x* and *y* were presented independently--i.e., non adjacently. Only one subject group (*X>Y->XY*) categorized *X-Y* cells as *XY*; the other group categorized these cells as either *X* or *Y*. Together, results of this experiment suggested that different object cues were learned to perceptually analyze and categorize identical visual stimuli.

As explained earlier, object categorization theories tend to place little constraints on what may count as an object cue. However, Example 2 demonstrated that one cannot just assume the features on which classification operates. Even very simple categorization problems such as distinguishing between three categories *X*, *Y* and *XY* might not have an obvious featural solution. If the individual's history of categorization affects the object cues that perceptually analyze the input, complete categorization theories will not only have to explain the ways in which object features are combined to form concepts, they will also have to explain the development of the perceptual object information which subtend these categories.

In summary of these two examples, it appears that there is much to be gained by considering in greater depth the possible interactions of task constraints and perceptual object information in object recognition/categorization theories. The face recognition example insisted on the impact of task constraints on viewpoint-dependent performance, and the Martian cell example stressed the role of object information on explanations of simple categorization rules.

"Everyday object recognition" Although the categorization of faces and Martian cells are clear instances of recognition, they do not appear to be typical of a more generic form of recognition which occurs when people categorize a car, a chair, a dog or similarly common objects. Instead, face (and Martian cell) recognition is often pictured as a more specialized (or expert) form of recognition. This questions the applicability of diagnostic recognition to the explanation of "everyday object recognition." But what is everyday object recognition? More precisely, what information demands does it impose on common categorization tasks, and which object information does it use in the input? These issues are essentially intertwined, but we will first explore the task demands of everyday object recognition before discussing its object information.

The task demands of "everyday object recognition." As explained earlier, classic categorization research has shown that the interactions between the human perceiver and the objects of his/her world specify several hierarchical levels of categorization. Following Rosch, Mervis, Gray, Johnson, and Boyes-Braem's (1976) seminal research, three of these levels are often isolated: the superordinate (*animal*, *vehicle, furniture*), the basic (*dog, car, chair*), and the subordinate (*collie, Porsche, Chippendale chair*). Although these categorizations are all important, Rosch et al. showed that one of them had a privileged status. When subjects were asked to spontaneously name pictures of common objects, Rosch demonstrated that they preferentially used basic-level names (see also Jolicoeur, Gluck & Kosslyn, 1984). Similarly, when asked to verify that a picture belonged to a particular category, subjects' decisions were faster for the basic-level (Rosch et al., 1976). Together, these findings suggested that the initial contact between the object percept and its semantic information occurs at the *basic level*, also know in object recognition research as *primal access* (Biederman, 1987), or *entry point* (Jolicoeur et al., 1984). Current thinking attributes supplementary processing on visual input to achieve subordinate categorizations. Consequently, a significant portion of object recognition research has focused on basic-level categorizations, or been criticized for not addressing the basic-level.

If we agree that basic-level categorizations reflect everyday object recognition, then it is particularly important to understand the structure of object categories at this level, to understand their information demands on recognition tasks. The basic level is often pictured as the most inclusive level at which objects "look alike" in terms of their shape. One determinant (but not the only one) of shape is part structure: Objects with common parts tend to have a common shape. Tversky and Hemenway (1984) found a dramatic increase in the number of parts listed from the superordinate to the basic level; non part cues increased from the basic to the subordinate level, but little increase was found for parts. Thus, Tversky and Hemenway suggested that "the natural breaks among basic-level categories are between clusters of parts" (1984, p. 186). This claim is the basis of Biederman's (1987) very influential Recognition By Components (RBC) theory, which represents basic-level categories with part-based descriptions--specifically, with different Geon

Structural Descriptions, Biederman & Gerhardstein (1993). Hence, the widely held assumption in object recognition research that shape, as is represented by parts, constitutes the information demands of everyday, basic-level categorizations.

Difficulties with part information for all basic-level tasks. Two issues should be distinguished in explanations of the basic-level phenomenon: (1) The information demands of everyday recognition: Are parts really necessary and sufficient to distinguish between all memorized basic-level categories? (2) The available object information: Is perception so organized that it actively seeks parts in the input image? On the one hand, if parts were indeed the information demands of everyday recognition, it would make good evolutionary sense that perception would have evolved to become primarily attuned to the parts of objects. On the other hand, if parts were not strictly necessary to distinguish between basic-level categories, a perceptual primacy for this information would be harder to justify in visual development.

We should be careful in stressing that the issue here does not concern the relevance of object shape in general for the basic level, but the relevance of one particular shape representations--namely, part structure. There is evidence that a "part-centric" conception of basic level could be misguided. For example, Murphy (1991) has demonstrated that other cues than parts could determine the basic-level. Research in expert categories (Tanaka & Taylor, 1991) has also suggested that the basic level is neither absolute nor unimodally specified, but that it could instead fluctuate with category expertise. Finally, available evidence that parts are the point of contact for basic-level are either based on feature listings, or on reaction times. Unless part representations were theoretically related to the basic-level, evidence that suggests parts should be experimentally confronted to alternative shape descriptions--e.g., silhouette, 2D edge configurations, or representations of the image at a coarser descriptive level. The following sections review each of these arguments in turn.

1. *A basic level without parts.* Since Rosch's seminal research, an important issue has always been the extent to which the basic-level is determined by the organization of categories, or by the perceptual constraints of visual cognition. Murphy (1991; see also Murphy & Brownell, 1985) suggested that the basic level was a consequence of the *informativeness* and *distinctiveness* of a category representation in memory. Representations are informative when they are linked with a lot of concrete object features; a representation is distinctive to the extent that it differs from contrast representations. In general, the more specific a representation the more informative it is, but the less distinctive it is from other representations (Murphy, 1991). Thus, 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 distinction, superordinate categories score low on informativeness, but high on distinctiveness (e.g., vehicle and furniture have different shapes, parts, colors, textures, and so forth). On this account, the basic level would simultaneously be the most informative and distinctive; it is a compromise between the accuracy of categorization at a maximally general level and the predictive power of a maximally specific level (Murphy & Lassaline, in press).

It is worth stressing that the constraint of optimizing category informativeness and distinctiveness bears on the memory organization of object categories. In principle, any object cue (or subset of cues) could be used, as long as it achieves the optimum. This "opportunistic" conception of visual cognition predicts that part descriptions would only specify the basic level when they optimized the informativeness and distinctiveness (i.e., the diagnosticity) of the considered category organization.

In a series of experiments with artificial stimuli, Murphy (1991) questioned the necessity of parts for basic-level structure. His reasoning was that addition of nonpart information (here, mainly color and texture) to a part-based basic structure should not speed up the basic level advantage, if parts were the sole determinant of performance. Alternatively, if the basic level depended on the informativeness and distinctiveness of categories, the addition of other cues to the basic-level should make these categories more informative and distinctive, therefore faster to identify. Results showed that adding color and texture to a part-based basic-level enhanced the basic-level advantage. Furthermore, subordinate categorizations times increased with the addition of new cues (because subordinate categories were now less distinctive, sharing similar color and texture across exemplars), and superordinate categorization times decreased (because they were more distinctive). A separate experiment showed that massing nonpart information (color, size and texture) at the superordinate level eliminated the advantage of a basic-level defined by parts: The diagnosticity of nonpart cues at one level suppressed the diagnosticity of part information at the basic level. These results lead Murphy (1991, p. 436) to conclude that "parts are neither necessary, nor sufficient to establish a basic-level structure ... categorization into basic categories uses all kinds of information, not just part-based information." In sum, this suggests that the information demands of basic-level, everyday recognition could be whichever object information (not just parts) that happens to maximizes category informativeness and distinctiveness in memory.

Although these conclusions could have profound implications for object recognition, Murphy's (1991) results could also have a limited impact. First, we must stress that the experiments demonstrate that a basic-level *effect* can be obtained with other object cues than parts. Although this contrasts with the standard assumption that parts are necessary, the effect was obtained with artificial stimuli whose part structure might not tap the same perceptual processes as real-world object. Secondly, it must be pointed out that shape could very well optimize the informativeness and distinctiveness of real-world object categories, and so these could still be primarily organized around their shape at the basic level. However, it is important to stress that the optimum can in principle involve any object information that is available to the individual categorizer. Consequently, any experimental set up testing for the basic-level should contrast multiple object cues (different, not just one, types of shape representations such as an object's silhouette, its description at a coarse scale, its luminance contours, its shading, plus textural and chromatic cues) to determine the actual information demands of everyday recognition. The following discussion suggests that these demands could change with the individual's perceptual expertise with an object category.

2. *A relative basic-level.* If parts were the only "search keys" of perception into basic-level object memory, and if possible parts were always sampled from a fixed set as in Biederman (1987), the entry-level to recognition might appear as being fixed once and for all in the hierarchy of categorization: The basic-level would systematically be this level at which categories differ in terms of their Geon Structural Description (Biederman & Gerhardstein, 1993), irrespectively of the considered category (e.g., cars, rocks, valleys, kitchen appliances, and so forth).

However, research in conceptual expertise has questioned the absolute character of the basic level. Tanaka and Taylor (1991) showed that conceptual expertise enhances the speed of access of subordinate categories, which become at least as accessible as basic-level categories. In a category verification task, their subjects (*dog* and *bird* experts) first heard a category label (superordinate, basic and subordinate) and then were asked to indicate whether a picture was an exemplar of the category. For expert categories (*bird* or *dog*), the subordinate categorization was as fast as the basic-level categorization. Furthermore, the authors also discovered that experts' subordinate categories were associated with more cues than were their novice categories. As explained earlier, these supplementary features could increase the informativity and distinctiveness of the expert subordinate category, and correspondingly change the entry point of recognition.

Although this "basic-to-subordinate shift" was only reported for *dog* and *bird* experts, it nonetheless demonstrates how different expertise levels with identical object sets affects their categorization. In a similar vein, although we probably all are experts in faces, the "other race effect" in which people perceive faces of their own race with greater facility than those of another race (Brigham, 1986) reveals that this expertise is limited to the subcategory of faces of our own race. If experts and novices used different sets of cues to describe identical objects (and Example 2 demonstrated that this was possible), the informativeness and distinctiveness of their categorization levels would differ, as might the entry point to recognition. As Tanaka and Taylor (1991) suggested, the basic-level could be in the eye of the beholder.

Together, the conclusions of Tanaka and Taylor (1991) and the "existence proof" of new feature learning of Schyns and Rodet (in press) question the origin of relevant object cues for recognition. There is evidence that new categorizations can induce the extraction of new object information, but evidence is strictly limited to artificial and unfamiliar "Martian" categories. Could the principles of feature learning be extended to more familiar object categories (e.g., dogs, chairs, cars and so forth)? For example, could the requirements of distinguishing between many man-made object categories and the perceptual principles of edge description interact to progressively constrain the development of a more complicated vocabulary of regular shape primitives--e.g., Biederman's (1987) geons? Future research could reveal that the learning of new object cues only applies to the learning of very specialized (or subordinate) categories such as X-rays, dermatosis, birds (or dogs) for *bird* (or *dog*) experts and so forth. Alternatively, it may turn out that feature learning mechanisms pervade the very early stages of conceptual development (when the first categories and their structuring features must be learned), but that mature categorizers (who tend to know the relevant perceptual analysis of most objects) only synthesize new features when they become experts in specific categories (Schyns & Rodet, in press). It is now an empirical issue of developmental psychology and computational modeling to seek out the perceptual and task factors that could promote the development of everyday object representations.

Lessons from task constraints. Everyday object recognition probably occurs at the basic-level. In attempting to characterize the information demands of a basic-level task, it appears that there might not be a single criterion that is necessary and sufficient to determine the entry point to recognition. In contrast to the idea that the entry point is rigidly determined by a demands of part information, there is suggestive evidence (1) that the basic level is the optimal level of informativity and distinctiveness of a category, (2) that parts are neither necessary, nor sufficient to determine this level, but that (3) other cues (e.g., color, texture, size, and so forth) could determine entry level, and (4) that the individual's perceptual experience with a category could change the defining cues of, and the entry level to, this category. Thus, the information demands of a basic task could be relative and dependent on the individual's experience with this particular object category.

Task constraints suggests that different categorization levels of an identical object (minimally the basic and subordinate) should be considered before the basic-level can be assessed. If the entry level is the most inclusive level at which objects "look alike," the entry point could change. To illustrate, the identification of a face is often thought to be a clear-cut subordinate categorization (because faces have a similar global shape). However, different views of an individual look more alike than the same views of different persons. Thus, our basic categorization of faces of our own race could be at the level of the individual (the level at which face views look more alike) instead of at their assumed "basic" level, the "face" categorization (the level at which faces views look more different). Note that this effect might be inverted with faces of another race: Faces with which we are generally less familiar could look more similar from the same viewpoint. Of course, all the implications of a flexible basic-level are pure speculations, but they nonetheless highlight the potential impact of interactions between information demands and perception.

The perceptual object information of everyday recognition. Everyday recognition does not occur in thin air; it uses the perceptual information available from the visual array. We know that object information is outwardly bounded by the retinal output (a million-dimensional space), but neuroanatomical and computational data suggest that this space is gradually projected (recoded) onto a much smaller dimensional space of object cues. Unfortunately, too little is known about this dimensionality reduction and the perceptual constraints it imposes on classification processes. Instead of attempting the impossible task of listing all likely information sources, the next section examines the reduced set of object cues that determine performance in leading object recognition theories. We then relate these cues to the information demands of the categorization tasks they subserve, and discuss how task constraints and object information interact to determine performance in current recognition theories.

Current thinking in object recognition predicates that perception delivers shape cues to match against spatio-visual object representations in memory. To illustrate, in Biederman's (1987) RBC theory the assumption that parts are the information demands of basic-level recognition justifies the projection of the 2D retinal input onto a space of Geon Structural Descriptions (GSD, Biederman & Gerhardstein, 1993). This reduced object description is obtained in two stages. The first stage computes the edges of the input image. The second stage seeks 2D, non-accidental, viewpoint-invariant properties of edge descriptions (e.g., collinearity, curvilinearity, symmetry, parallelism, cotermination, see Lowe, 1987; Kanade, 1981). Collection of viewpoint-invariant features individuate the geons composing the input object.

Together, these mechanisms suggests viewpoint-invariant recognition performance at the basic-level. More precisely, RBC predicts viewpoint-invariance performance whenever (1) the input object is decomposable into geons, (2) different GSDs represent different objects in memory, and (3) the same GSD is recoverable from different viewpoints (Biederman & Gerhardstein, 1993). It is important to note that explanations of performance involving geons will, by construction of RBC, necessarily overlap with explanations involving the viewpoint-invariant features that compose geons. To date, there is no conclusive evidence that real-world object recognition uses geons as a reduced object description space.

In opposition to RBC, the view-based approach predicts viewpoint-dependent performance to stored object views. In attempting to define more precisely what constitutes a "view," Tarr and Kriegman (1996) recently began to explore how available viewpoint-dependent shape information could determine performance. As an observer changes its vantage point, drastic changes often occur in the qualitative appearance of an object. For a given geometrical object class (e.g., smooth objects), qualitative changes may be described with a vocabulary of viewpoint-dependent features (local and multi-local edge configurations, see Tarr & Kriegman, 1996). These features partition the viewpoint space of an object into stable regions (in fact, into the views of an *aspect graph* of this object, Koenderink & Van Doorn, 1979). Hence, viewpoint-dependent recognition performance could partially result from enhanced perceptual sensitivity to viewpoint-dependent shape information. Tarr and Kriegman's (1996) psychophysical experiments demonstrated that this was the case, at least for some of the visual cues predicted by their theory.

Viewpoint-dependence/independence revisited. From the above discussion, it appears that although RBC and the view-based approach use shape information for recognition, they radically differ on the exact nature of the perceptual object cues that represent objects in memory. RBC suggests that perception detects 2D, viewpoint-independent cues that serve to reconstruct 3D geons, but Tarr and Kriegman showed that humans were sensitive to 2D, viewpoint-dependent edge configurations. Consequently, both theories should in principle[1] predict different recognition performance.

It is important to emphasize that the presence in the image of viewpoint-dependent vs. independent cues is entirely determined by the geometry of the input object and its projection on the retina (the image formation process). Consequently, both types of cues could be available to object recognition mechanisms, independently of whether recognition uses viewpoint-dependent (views) or independent (geons) object representations. Thus, whether performance will be viewpoint-dependent or independent shifts to the recognition conditions in which one type of image cue, or the other, is being used. Diagnostic recognition suggests that image cues are used when they are diagnostic of a particular object categorization. What are, then, the object categorizations that would preferentially require viewpoint-dependent vs. independent cues? More specifically, what kind of information demands would underlie the usage of each cue class?

Object recognition suggests a generic answer: Basic-level categorizations would generally require viewpoint-independent image cues, and subordinate categorizations would demand viewpoint-dependent cues (e.g., Biederman, 1987;

[1] It should also be noted that RBC effectively predicts viewpoint-dependent performance in many practical situations of recognition. Rotation in depth of many real-world objects is such that different views will often convey different information (part included, think, e.g., of a human body rotating in depth). Consequently, Biederman and Gerhardstein's (1993) condition number three for viewpoint-independent performance (that the same part structural description is recovered

Farah, 1992; Jolicoeur, 1990; Tarr & Pinker, 1990). However, the previous analysis of task constraints at the basic level would question the generality of this claim, because basic information demands might flexibly tune to different types of object cues, some of which resulting from perceptual expertise with an object class. Although speculative, this suggests that viewpoint-in/dependent performance might not so much depend on a generic categorization level (basic or subordinate) as it would depend on the particular cues that structure the entry point of the individual categorizer, and whether the extraction of these cues is local and/or viewpoint dependent, or global and/or viewpoint invariant. If basic-to-subordinate shifts occur when people gain expertise with real-world objects, then it becomes difficult to associate either viewpoint-dependence, or independence with a generic categorization level. Viewpoint-in/dependent should depend on the object class, its geometric, chromatic and textural information, and the cues that optimize the informativeness and distinctiveness of this category for the individual categorizer.

To summarize, our speculation, then, is that differences in the information demands of the categorizer might explain much of viewpoint-dependent/independent performance in recognition studies, because available object information is a function of the theory-independent image information process (see also Liu, 1995; Tarr & Kriegman, 1996, for a similar view). Further studies of perceptual learning in object recognition might illuminate these issues.

GENERAL DISCUSSION

It was the main goal of this paper to establish a dialogue between object recognition and object categorization theories, with the hope of raising issues that could cross-fertilize their research. To this end, I developed diagnostic recognition, a framework in which object recognition and categorization phenomena are expressed as interactions between the information demands of specific

---

from different viewpoints) will not always be met, and both RBC and view-based categorization tasks and the rich perceptual information available from the input object. Diagnostic recognition insists on the task-dependent diagnosticity of perceptual object cues to understand object recognition and categorization phenomena. Two examples illustrated the opposite benefits that object recognition and categorization theories could obtain from considering the two factors of diagnostic recognition. The face recognition example showed how object recognition could benefit from more extensive studies of multiple task constraints, and the categorization of Martian cells example suggested that perceptual object information might need further considerations in very simple categorization problems.

The second part of the paper extrapolated the approach of diagnostic recognition to the account of "everyday recognition" performance. Everyday recognition was equated with the basic-level of the categorization hierarchy. Examination of information demands at the basic-level from a categorization perspective suggested that it was the optimal level of informativity and distinctiveness of a category, that parts were neither necessary nor sufficient to structure this level, but that many other cues could elicit a basic level phenomenon. It was also suggested that the individual's perceptual experience with an object category could change the defining perceptual cues of this category, as well as its entry level.

Turning to the object information used in leading object recognition theories, it was first observed that perception is often assumed to initially deliver shape cues to match against memory representations. These shape cues are either viewpoint-dependent or viewpoint-invariant, but this is a function of the image formation process. The usage of one or another type of cues that account for recognition performance was then related to the information demands of different

---

theories will predict viewpoint-dependence.

categorization tasks, suggesting that diagnosticity did not only apply to expert categories.

There is a similarity between the framework presented here and the ideal observer approach to recognition (Bennett, Hoffman & Prakash, 1993; Liu, Knill & Kersten, 1995). Both diagnostic recognition and the ideal observer insist on available object information and the perceptual constraints on its extraction that could influence performance (see also Liu, in press). However, actual developments of the ideal observer do not include task constraints which "acts" on different object information to assign them different diagnosticities. For example, Liu (in press, p. 5) states that "... the prediction from a viewpoint-independent representation should be that the performance only depend on the information content of the input image." However, there is suggestive evidence that recognition performance is also dependent on task constraints. Extensions of ideal observers might need to include the notion of flexible task constraints.

Implications of diagnostic for studies of object representations. Diagnostic recognition is a framework in which the information goals of object categorization tasks are considered before their perceptual representations. Although this is good, generally recommended hygiene of theory construction (e.g., Marr, 1982), it presents nonetheless serious limitations for the study of object representations.

Ideally, an object representation should offer sufficient information to solve many, not just one categorizations--with the additional possibility of incrementally adding new perceptual information to the representation, when this information represents a new categorization (Schyns & Murphy, 1994). The recognition task and its associated information demands should then tap into this or that facet of the object representation, and elicit this or that aspect of recognition performance.

An argument could be made that diagnostic recognition is ill-suited to the goal of studying object representations. The reason is simply that thinking from task constraints to their perceptual representations could over-represent the considered information demands in the proposed representation. For example, if it was discovered that the information demands of an object categorization were X, then it would be an easy step to assume that the representation of this object was effectively that X. But then, how would we know whether X represents the object, or the task? For this reason, a first methodological recommendation would be to multiply task constraints before inferring representation formats. At the limit, representations and task demands would be more independent.

It is interesting to note that leading recognition theories are two poles on the spectrum of independence of representations from task constraints. RBC 's representations (geons) directly mirror the assumed information demands (parts) of the basic-level. However, even if parts were functionally required for the basic-level, it would still remain an empirical issue that they are explicitly represented in memory. The view-based approach stands on the opposite side of the spectrum of independence from task constraints. Unless more clearly specified, an object view could potentially represent all the information that can ever be demanded from an object seen from this view--parts included. In other words, views are too powerful a representation, and it is an important research goal to attempt to reduce the high-dimensionality of a view to a low-dimensional subset of image features (e.g., Tarr & Kriegman, in press).

It nonetheless remains that diagnostic recognition might be better suited to explain recognition performance than representation formats. Performance involves both object *and* input representations: The input image must first be encoded with object cues for matching against memorized representations. Because there is no general theory specifying the information content of the 2D projections of 3D objects, behavioral performance might not be sufficiently powerful to isolate issues of object formats from issues of input information (see also Liu, in press). Performance that might be attributed to a particular object format might also be attributed to the interaction between task demands and the usage of specific image cues. For these reasons, diagnostic recognition suggests that object and input representations should be unified and constitute the set of image cues that are available for different object categorization tasks.

Diagnostic recognition and scenes. We conclude this discussion with a specific example of interactions between perceptual information and task constraints taken from scene recognition. Although a scene is not an object, but many objects, the issues the example will raise can be directly transposed to object recognition and categorization studies. We start with a discussion of image information for scene categories, and then show how its usage could change with the diagnosticity of cues.

Computational vision and psychophysics studies have emphasized the importance of simultaneously processing images at multiple spatial resolutions, or scales (Blackemore & Campbell, 1969; Breitmeyer & Ganz, 1976; Burt & Adelson, 1983; Campbell & Robson, 1968; Canny, 1986; de Valois & de Valois, 1990; Ginsburg, 1986; Mallet, 1989; Marr, 1982; among many others). Starting with the observation that recognition algorithms could hardly operate on the raw pixel values of digitized images, vision researchers investigated multi-scale representations to organize and simplify the description of events. *Coarse-to-fine processing* summarizes the idea that it may be computationally more efficient to first derive a coarse and global (albeit imprecise) description of the image before extracting more detailed (but considerably noisier) information (Marr, 1982; Watt, 1987). Evidence of coarse-to-fine processing in humans was reported for face (e.g., Breitmeyer, 1984; Fiorentini, Maffei & Sandini, 1983; Sergent, 1982, 1986), object (e.g., Ginsburg, 1980) and scene recognition (e.g., Parker, Lishman & Hughes, 1992; Schyns & Oliva, 1994), and for simpler patterns (see de Valois & de Valois, 1990, for a review).

Figure 4. This figure (adapted from Schyns & Oliva, 1994) shows examples of the hybrid stimuli. Hybrids were constructed by combining the Low Spatial Frequency (LSF) components of the amplitude and phase spectra of one scene (e.g. a highway) with the High Spatial Frequency (HSF) components of another scene. The top picture mixes the LSF of a highway and the HSF of a city. The bottom picture mixes the LSF of a city with the HSF of a highway.

To illustrate perceptual spatial scales, consider the two pictures presented on Figure 4 (from Schyns and Oliva, 1994). Combinations of fine-grain edges at the fine scale should reveal that the top picture is a city scene and the bottom picture a highway scene. However, if you squint or blink while looking at Figure 3, the top picture should become a highway and the bottom picture a city (if the demonstration does not work, step back from Figure 4).

Even though it is now well established that the visual system operates at multiple scales, their selection for recognition is still a matter of on-going research. One possibility is that perception extract the coarse before the fine (a constraint on available image information) and therefore coerces a mandatory coarse-to-fine recognition scheme. Alternatively, the information demands of a categorization could bias recognition to operate at the task-diagnostic scale. For example, while coarse scale information might be sufficient to categorize a picture of New York as "city," a "New York" categorization of the same picture might require comparatively finer scale cues (see Figure 4).

Figure 5. This figure (from Oliva & Schyns, 1996) illustrates the stimuli used in the sensitization phase of the reported experiment. The top picture is a LF/Noise hybrid composed of the LF of a city added to structured Noise in HF. The bottom picture is a HF/Noise resulting from the addition of the HF of the same city and LF structured noise.

In a two phase experiment, Oliva and Schyns (1995) tested that such interaction between mandatory scales perception and categorization demands could promote orthogonal classifications of identical hybrid stimuli. In a sensitization phase, two subject groups (the Low Spatial Frequency, LSF, and the High Spatial Frequency, HSF, groups) were instructed to categorize hybrids which were meaningful at only one scale, the other scale being structured noise (see Figure 5). It was expected that these stimuli would sensitize visual processes to operate at the diagnostic scale (either LSF or HSF). The testing phase followed immediately, without any transition. Without subjects being aware, the two scale components of the test hybrids were both meaningful (as in Figure 4). Results revealed that the groups categorized test stimuli at the scale diagnostic of the task (either LSF or HSF, depending on the group). Thus, identical pictures elicited mutually exclusive categorizations, without subjects being even aware of the other meaningful scene.

These results illustrate the point of this manuscript that recognition performance (here, coarse-to-fine recognition) might be better explained as interactions of information demands (locating diagnostic cues) and available image information (here, perceptual spatial scales). The task-dependent, orthogonal classifications of identical hybrids raise important issues for low-level perception research. There is evidence that the percept of the hybrids was scale-specific (coarse *or* fine), although independent evidence suggests that all scales (coarse *and* fine) were effectively registered. Would a nonlinear, task-driven tuning of low-level scale perception scotomized the subjects' percept to the diagnostic scale? Although this would mean that task constraints can have a strong impact on low-level perception (see also Example 2), hybrids stimuli could be used to address empirically such issues.

Hybrids multiplex scene information in scale space. Thus, they offer dense perceptual information to classification processes; information that could be differentially repackaged if task constraints necessitated such flexible encodings. Scale information also raises issues for object recognition researchers. For example, the results suggest that diagnostic scene cues are already available at coarse scales. What are these cues? How are they structured and represented? What categorizations do they subtend? How do categorizations integrate information across scales? Although expert categorizations would probably seek detailed, fine scale information, there might cases of expertise that require discriminations of the overall structures of objects and scenes. Which ones of these global and local cues are viewpoint-dependent/independent? In sum, the study of perceptual spaces, their component features, the conditions of their availability, and the information they bring for different categorizations seem to be the way forward in object recognition studies.

REFERENCES

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

Bennett, B. M., Hoffman, D. D., Prakash, C. (1993). Recognition polynomials. *Journal of Optical Society of America, A*, **10**, 759-764.

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

Biederman, I. & Gerhardstein, P. C. (1995). Viewpoint-dependent mechanisms in visual object recognition: a critical analysis. *Journal of Experimental Psychology: Human Perception and Performance*, **21**.

Biederman, I. & Gerhardstein, P.C. (1993). Recognizing depth rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 121-133.

Biederman, I., & Shiffrar, M. M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. Journal of Experimental Psychology: Learning , *Memory and Cognition*, **13**, 640-645.

Blackemore, C., & Campbell, F. W. (1969). On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology (London)*, **203**, 237-260.

Bourne, L. E., Jr. (1982). Typicality effects in logically defined categories. *Memory & Cognition*, **10**, 3-9.

Breitmeyer, B.G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression and information processing. *Psychological Review*, 83, 1-35.

Bruce, V., Valentine, T., & Baddeley (1987). The basis of the 3/4 view advantage in face recognition. *Applied Cognitive Psychology*, **1**, 109-120.

Bruner, J. S., Goodnow, J., & Austin, G. A. (1956). *A study of thinking*. New York: Wiley.

Bülthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view theory of object recognition. *Proceedings of the National Academy of Science USA*, **89**, 60-64.

Burt, P., & Adelson, E. H. (1983). The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, **31**, 532-540.

Campbell, F. W., & Robson, J. G. (1968). Application of the Fourier analysis to the visibility of gratings. *Journal of Physiology London*, **88**, 551-556.

Canny, J. F. (1986). A computational approach to edge detection. *IEEE Pattern Analysis and Machine Intelligence*, **8**, 100-105.

Christensen, E. E., Murry, R. C., Holland, K., Reynolds, J., Landay, M. J., & Moore, J. G. (1981). The effect of search time on perception. *Radiology*, **138**, 361-365.

De Valois, R. L., & De Valois, K. K. (1990). *Spatial Vision*. Oxford University Press: New York.

Edelman, S., & Bülthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research*, **32**, 2385-2400.

Elio, R. & Anderson, J. R. (1981). The effect of category generalizations and instance similarity on schema abstraction. Journal of Experimental Psychology: *Human Learning & Memory*, **7**, 397-417.

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

Farah, M. J. (1992). Is an object an object an object? Cognitive and neuropsychological investigations of domain-specificity in visual object recognition. *Current Directions in Psychological Science*, **1**, 164-169.

Ginsburg, A. P. (1986). Spatial filtering and visual form perception. In K. R. Boff, L. Kaufman and J. P. Thomas (Eds.). *Handbook of Perception and Human Performance, II: Cognitive Processes and Performance.* NY: Wiley.

Gluck, M. A., & Bower, G. H. (1988). Evaluating an adaptive network model of human learning. *Journal of Memory and Language,* **27**, 166-195.

Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General,* **123**, 178-200.

Grice, G. R., Graham, L. & Boroughs, J. M. (1983). Forest before trees? It depends where you look. *Perception and Psychophysics,* **33**, 121-128.

Hill, H., Schyns, P. G., & Akamatsu, S. (in press). Information and viewpoint dependence in face recognition. *Cognition.*

Jolicoeur, P. (1990). Identification of disoriented objects: A dual-systems theory. *Mind and Language,* **5**, 387-410.

Jolicoeur, P., Gluck, M., & Kosslyn, S. M. (1984). Pictures and names: Making the connexion. *Cognitive Psychology,* **19**, 31-53.

Kanade, T. (1981). Recovery of the three-dimensional shape of an object from a single view. *Artificial Intelligence,* **17**, 409-460.

Koenderink, J. J. & Van Doorn, A. J. (1979). The internal representation of solid shapes with respect to vision. *Biological Cybernetics,* **32**, 211-216.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review,* **99**, 22-44.

Liu, Z. (in press). Viewpoint-dependency in object representation and recognition. *Spatial Vision: Special issue on perceptual learning and adaptation.*

Liu, Z., Kersten, D., & Knill, D. C. (1995). Structural organization improves object discrimination. *Vision Research,* XX.

Lowe, D. G. (1987). The viewpoint consistency constraint. *International Journal of Computer Vision,* **1**, 57-72.

Mallet, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Pattern Analysis and Machine Intelligence*, **11**, 674-693.

Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.

Murphy, G. L. (1991). Parts in object concepts: Experiments with artificial categories. *Memory & Cognition*, **18**, 407-418.

Murphy, G. L., & Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constrations on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory & Cognition*,**11**, 70-84. .

Murphy, G. L., & Lassaline, M. E. (in press). Hierarchical structure in concepts and the basic level of categorization. To appear in K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories.* London: UCL Press.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, **9**, 353-383.

Norman, G. R., Brooks, L. R., Coblentz, C. L., & Babcock, C. J. (1992). The correlation of feature identification and category judgments in diagnostic radiology. *Memory & Cogintion*, **20**, 344-355.

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

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

Oliva, A. & Schyns, P. G. (1995). Mandatory scale perception promotes flexible scene categorization. *Proceedings of the XVII Meeting of the Cognitive Science Society*, 159-163, Lawrence Erlbaum: Hilldsale, NJ.

Oliva, A. & Schyns, P. G. (1996). Scale perception interacts with scene categorization. Submitted.

Palmer, S., Rosch, E., & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *Attention and Performances IX*. Hillsdale, NJ: Lawrence Erlbaum.

Perrett, D. I., Oram, M. W., Harries, M. H., Bevan, R., Benson, P. J., & Thomas, S. (1991). Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *Experimental Brain Research*, **86**, 159-173.

Rock, I., & Di Vita, J. (1987). A case of viewer-centered object representation. *Cognitive Psychology*, **19**, 280-293.

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

Schyns, P. G., & Bülthoff, H. H. (1994). Viewpoint dependence and face recognition. *Proceedings of the XVI Meeting of the Cognitive Science Society*, 789-793.

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

Schyns, P. G., & Murphy, G. L. (1991). The ontogeny of units in object categories. *Proceeding of the XIII Meeting of the Cognitive Science Society*, 197-202, Lawrence Erlbaum: Hilldsale, NJ.

Schyns, P. G., & Murphy, G. L. (1994). The ontogeny of part representation in object concepts. In Medin (Ed.), *The Psychology of Learning and Motivation*, **31**,301-349.

Schyns, P.G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time and spatial scale dependent scene recognition. *Psychological Science*, **5**, 195-200.

Schyns, P. G. & Rodet, L. (in press). Categorization creates functional features. *Journal of Experimental Psychology: Learning, Memory & Cognition.*

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, **75**.

Tanaka, J., & Taylor, M. E. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, **15**, 121-149.

Tarr, M. J., & Bülthoff, H. H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? *Journal of Experimental Psychology: Human Perception and Performance*, **21**.

Tarr, M. J., & Kriegman, D. J. (in press). Toward understanding human object recognition: Aspect graphs and view-based representations. *Psychological Review*.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology*, **21**, 233-282.

Thibaut, J. P. (1991). *Récurrence et variations des attributs dans la formation de concepts*. Unpublished doctoral thesis. Department of Psychology. University of Liège, Belgium.

Tversky, B., & Hemenway, K. (1984). Objects, parts and categories. *Journal of Experimental Psychology: General*, **113**, 169-193.

Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition*, **32**, 293-254..

Vetter, T., Poggio, T., & Bülthoff, H. H. (1994). The importance of symmetry and virtual views in three-dimensional object recognition. *Current Biology*, **4**, 18-23.

Watt, R. (1987). Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *Journal of Optical Society of America*, A 4, 2006-2021.