In J.J. Rieser, D.H. Ashmead, F.F. Ebner, A.L. Corn (Ed.), Blindness and Brain Plasticity in Navigation and Object Perception (2008), New York: Lawerence Erlbaum Assoc.

10

Object Recognition by Touch

Roberta L. Klatzky and Susan J. Lederman

This chapter is organized around a series of questions. At the most general level, we pose the following question: With respect to fundamental research, what do we know about how people recognize objects by touch? Another general question is of broad concern throughout this volume: Does what we know have implications for informing and educating people who have low vision or are blind? With respect to the first question, psychological research has led to a substantial body of knowledge about haptic object recognition. With respect to the second question, we acknowledge considerable challenges in converting this work into a program for education and communication. However, the work offers the promise that if educational tools are designed to take advantage of perceptual capabilities, they will be highly useful to people who are blind or visually impaired.

When we talk informally about using touch we mean using what is technically called *the haptic system*. Haptics is the perceptual system that incorporates sensory information from the skin (cutaneous sensing) and from muscles, tendons, and joints (kinesthetic sensing). Haptic perception is typically active and under the perceiver's own control. As a result, it incorporates information about motor intentions and the outflow of movement commands. For general reviews, the reader is directed to chapters by Loomis and Lederman (1986) and by Klatzky and Lederman (2003).

Haptic object recognition involves a stream of processing that begins with exposure to a real, tangible object and ends with the formation of an internal representation of its properties. The object may or may not be familiar, and recognition may or may not include naming. As an outcome of this process, we know about haptically accessible properties of the object: what it is shaped like, how warm or cool it feels, its roughness, and so forth. Our representation may be inaccurate or incomplete, but it is the culmination of the perceptual processing stream. Here are the questions we address about haptic object recognition, as we pose them throughout the chapter.

1. *Does visual object recognition provide a model for haptic object recognition?*

Visual and haptic object recognition share basic mechanisms, such as a progression from sensory primitives to abstract representations and use of prior knowledge and context where possible. But these general similarities notwithstanding, the two channels turn out to be quite different, and the emergent model of haptic object identification is fundamentally different from its visual counterpart.

2. *How does haptic object recognition depend on the way people explore freely?*

Patterns of manual exploration play a critical role in what people know about objects, how they recognize them, and how they think about them once they are apprehended and named. We characterize manual exploration as involving haptic "exploratory procedures" with particular patterns, each with their own costs and benefits.

The costs and benefits of exploration lead to further questions, including:

Are some properties of objects more salient, that is, stronger in our conscious experience, than others?

Would we use touch for object recognition if vision is available, and if so, why?

Does exploration of an object for purposes of recognition proceed in a stereotyped way, and if so, what determines the progression?

Can we capitalize on the fact that multiple properties of objects predict their identity?

3. *What happens when we constrain manual exploration in space and/or time?*

We will consider how exploration and object identification change when the finger is covered with a sheath or when people explore with a rigid probe, both of which limit the spatial information available. In both cases, the direct relationship between the pressure layout on the fingertip and objects' spatial properties is eliminated, and the sensing is described as "indirect." A second type of constraint is temporal in nature. We will consider what can be learned about an object with only a very brief exposure.

4. *Based on the fundamental research on haptic object recognition, what are the implications for education and communication in people who are blind or have low vision?*

The research we describe makes the general point that touch and vision serve complementary roles in recognizing objects. A direct implication is that although there is some utility to adapting visual displays for touch, the nature of the modality favors displaying the 3-D and material properties of objects. We will consider the challenges and opportunities of this approach. One direct implication is that 3-D displays are preferred, when possible, over raised-line drawings for conveying spatial information to the people who are blind or have low vision.

DOES VISUAL OBJECT RECOGNITION PROVIDE A MODEL FOR HAPTIC OBJECT RECOGNITION?

The first of our questions asks whether models of visual object recognition can be directly adapted to account for object recognition by touch. This is important when considering how teaching materials designed for sighted individuals should be adapted for those who are blind or have low vision. But before asking how visually based models should be adapted for haptic object recognition, we must first ask how well people can recognize objects by touch alone. Some years ago, we addressed this preliminary point in a simple study in which blindfolded sighted participants were handed a series of 100 objects—a comb, a mitten, or a tea bag—and asked to name each in turn (Klatzky, Lederman, & Metzger 1985). They could explore in any way they liked. At the time this study was performed, it was not obvious a priori what the outcome would be because touch had been thought of primarily as a poor substitute for vision. Somewhat surprisingly, then, performance was nearly 100% correct, and most responses occurred in 2 s or less. From an information-theoretic perspective, the results pointed to a virtually boundless potential for transmitting information about real, common objects over a haptic channel, given free exploration.

Figure 10–1 presents a general overview model of visual object recognition. It divides processes into three broad classes: Sensory-level processes extract primitive properties, like edges and regions of uniform color or texture. Intermediate-level processes may group these elementary features together, enlarge them, or break them apart. The familiar Gestalt psychological processes such as "grouping by similarity" are good examples. A higher-level process compares the ongoing object description to prior knowledge, ultimately producing feelings of familiarity (or not) and naming.

At first glance, one might feel that this model could be directly adapted for touch. In terms of the broad division of processes, the correspondence is fairly direct. In the details, however, the two modalities reveal themselves to be quite different. One of the principal differences is the emphasis on edge descriptions in vision. Simple line drawings can be identified readily even after brief visual exposure; these stimuli lack color, texture, or movement but still give ample information for the process to be carried to completion.

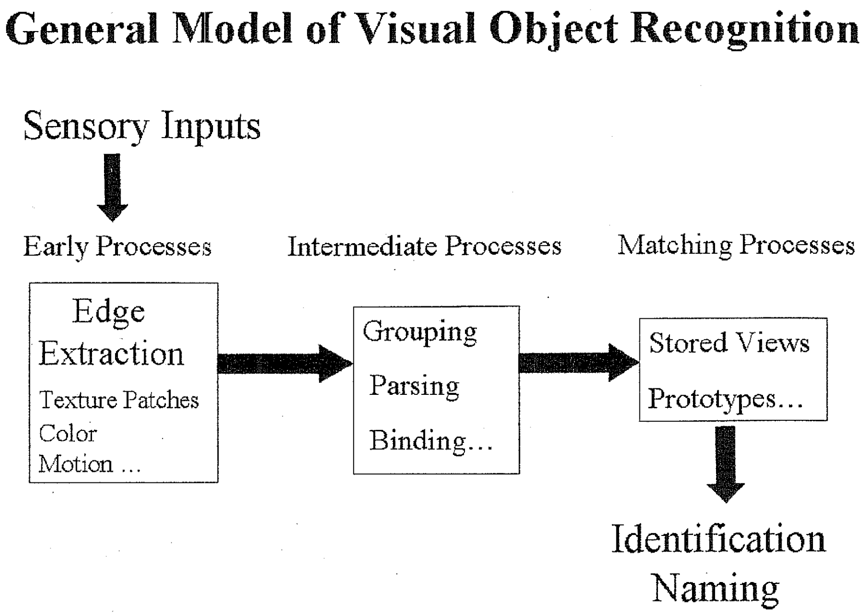
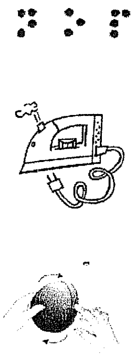


FIGURE 10–1. A general model of visual object recognition.

Edge information, as we will see, is much harder to come by in haptic object processing and may play a subsidiary role to nonedge properties like roughness or softness when it comes to recognition. Consider the examples shown in Figure 10–2. Extracting the geometry of Braille symbols relies on a tight compromise among multiple constraints: The symbols must be big enough to overcome the blurring of shape by the fingerpad, they must be small enough to fit within the limited size of a single fingerpad, and they must be complex enough to convey the information (number of bits) required for all the alphanumeric characters (Loomis, 1990). People who are blind do read Braille, of course, but it is difficult to attain fluency. As for the second example, recognizing raised-line drawings without vision is difficult for sighted people as well as those who are blind (Lederman, Klatzky, Chataway, & Summers, 1990; Magee & Kennedy, 1980). A major problem is with intermediate-level processes (those that further process the initial pattern of light and dark elements passed forward from the retina of the eye), like parsing the lines into regions and separating the figure from the background. Another problem is that the displays must be explored over time and space, imposing the need for perceptual and/or cognitive integration. Fully 3-D forms, as in the third example, present another problem: controlling the movements that bring new surfaces into reach of the exploring fingers while controlling the exploration itself. In addition, 3-D displays demand spatiotemporal integration, much as do 2-D displays.

What's wrong with this model for touch?

- edge information is primary in vision
- edge information from touch is severely limited



Braille reading
→ limited by finger size, fingertip filtering

Recognizing raised drawings
→ limited by spatial incoherence, temporal integration

Encoding 3D contours
→ limited by motor constraints, kinesthetic precision, temporal integration

FIGURE 10–2. Models of visual object recognition are not appropriate for haptic object recognition.

In the sections to come, we return to the issue of modeling haptic object recognition in comparison to visual recognition. Before doing so, however, we must present considerable data relating to the ways in which people explore objects they are trying to learn about and to identify.

HOW DOES HAPTIC OBJECT RECOGNITION DEPEND ON THE WAY PEOPLE EXPLORE FREELY?

When people want to learn about an object or name it using only the sense of touch, they must explore manually. That is, they manipulate it while moving their hand(s) and finger(s) actively over the object, probing and rubbing it. The regularity of their actions is striking and, as we will see, has implications for the capacities and limitations people have for perceiving objects by touch. We have developed a catalogue of touching actions (Lederman & Klatzky, 1987) that not only describes the patterns of movements that people use to explore objects but also shows how each type of movement can be linked to one or more object properties. To develop this catalogue, we used a match-to-sample task: Participants were first given a sample object and then three comparison objects. They were asked to pick the best match on a targeted object property, like surface roughness or hardness. Their exploratory hand movements were videotaped and proved to be reliably classified into what we called *exploratory procedures* (EPs). An EP is a way of feeling an object that has some invariant characteristic. For example, when people want to find out about an object's roughness, they perform an EP we call *lateral motion*. The invariant aspect of that EP is that the skin moves tangentially across the local surface of the object. In other respects, the EP can vary. People may use one finger or more, they may move quickly or slowly, or they may rub in circles or make a short sweep. The main point is that there is always tangential motion relative to the object's surface. If people are asked about some other property, say, the object's volume, they will not typically execute lateral motion. More likely is that they will enclose the object in one or more hands, which has the invariant characteristic that it maximizes the skin surface contacting the hand.

Figure 10–3 shows, in schematic form, the set of exploratory procedures that we developed. Listed with each EP is the property that elicits it. The figure shows what property is associated with what EP, but it does not explain why the association exists. That question merits a longer discussion. In another source (Klatzky & Lederman, 1999), we have addressed the causes for EP–property relationships to the extent that we have some understanding of them. For example, lateral motion is known to enhance the response of specialized sensory structures under the skin that underlie the sensation of roughness (Johnson & Lamb, 1981). It is this increase in the sensory signal from the relevant receptor population that presumably leads people to use tangential movements when they are asked to make roughness comparisons.

In a further experiment in the same series, we sought to determine what we call *costs* and *benefits* of using different EPs for learning about objects. The benefits have to do with information gathering. On this side of the equation, we ask not only how well an EP delivers information about the associated property (e.g., how well lateral motion provides information about roughness) but also how much information it delivers about other properties as well. On the other side of the equation, the costs of an EP include how long it takes to execute and whether, when it is being performed, other manual movements are locked out, precluding the performance of more than one EP at the same time.

To determine costs and benefits, we asked people to explore in a particular way while they were comparing objects on a particular property. For example, they might be asked to use lateral motion to explore when they were to compare for shape. Or they might be asked to enclose the object when they were to compare for roughness. As you might imagine, not all combinations worked. Rubbing an object on just part of its surface is not very conducive to learning about its overall shape. What that means is that any one EP usually does not provide exhaustive information about all the properties of an object.

The results of this constrained-exploration experiment are shown in Figure 10–4. The rows of the figure are EPs. The first column lists object proper-

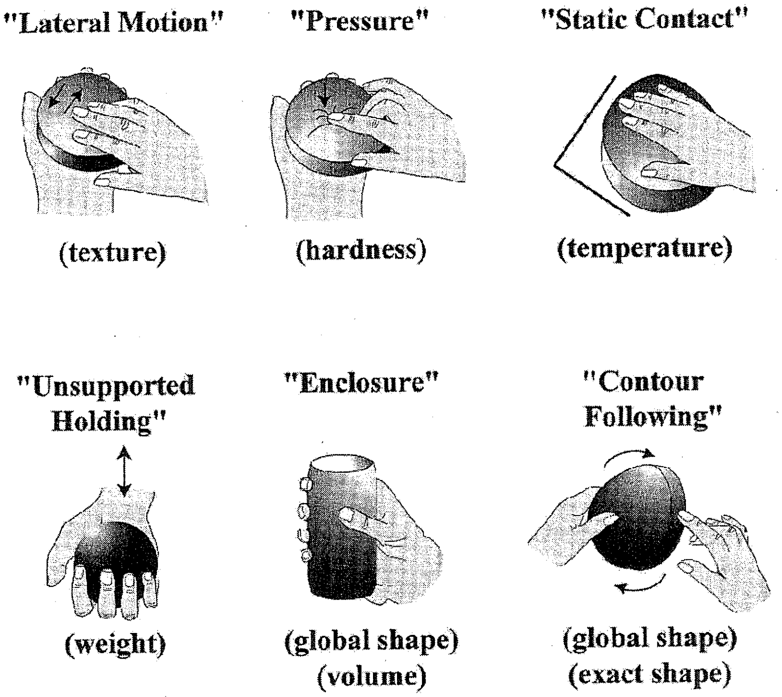
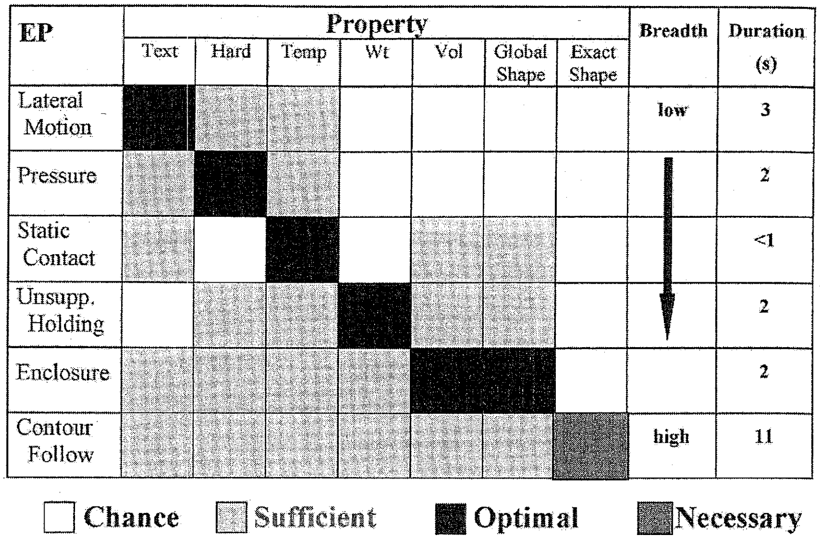


FIGURE 10–3. Depictions of six "exploratory procedures" together with the property(ies) for which each is optimal. From S. J. Lederman and R. L. Klatzky (1987)."Hand movements: A window into haptic object recognition," *Cognitive Psychology, 22*(3), 342–68. Copyright 1987 by Elsevier. Reprinted with permission.

ties. The legend indicates how well people performed when a particular EP was required and a particular property was tested. Performance was classified as at chance (people did no better than guessing), above chance but not maximal for the given property (i.e., it was "sufficient"), maximal for the given property (i.e., the EP was "optimal"), or—in one case only—whether the EP was "necessary," meaning that it was the only EP to give above-chance information about that property.

Several things can be learned from this figure. By looking down a column, one can see how many EPs were sufficient or better for a given property. The columns are arranged so that the properties toward the left can be extracted by more EPs than the properties on the right. Looking at the column headings, it emerges that the material properties of objects (roughness, hardness, and temperature) are more accessible than the geometric properties (weight, volume, and shape), where "accessible" means that they can be extracted by multiple EPs.

By looking across a row, one can see how many properties a given EP was sufficient to extract. Some EPs extract many properties or have considerable



EP	Property							Breadth	Duration (s)
	Text	Hard	Temp	Wt	Vol	Global Shape	Exact Shape		
Lateral Motion	■	░	░	☐	☐	☐	☐	low	3
Pressure	░	■	░	☐	☐	☐	☐	↓	2
Static Contact	░	☐	■	☐	░	░	☐		<1
Unsupp. Holding	☐	░	░	■	░	░	☐		2
Enclosure	░	░	░	░	■	■	☐		2
Contour Follow	░	░	░	░	░	░	▓	high	11

☐ **Chance** ░ **Sufficient** ■ **Optimal** ▓ **Necessary**

FIGURE 10–4. Exploratory procedure (EP) costs and benefits in terms of relative EP precision (chance, sufficient, optimal, or necessary), breadth of sufficiency, and average duration.

"breadth." Others are more specialized. The rows are arranged so that broader EPs occur closer to the bottom, as shown by the descending arrow. Finally, the last column shows the average duration of the EP in the original match-to-sample task, where subjects explored freely. From this one can see that the broadest EP, contour following, was typically much slower than the others.

Contour following is the EP that is spontaneously used to extract exact shape and is necessary to achieve precise shape matching. However, the slowness of contour following is part of the message we prefaced above: There is a high cost to extracting shape information by touch. It takes a lot of time. We should also note that shape matching performance is generally much more accurate with vision than with touch (e.g., Walk & Pick, 1981).

ARE SOME PROPERTIES OF OBJECTS MORE SALIENT (STRONGER) IN OUR CONSCIOUS EXPERIENCE THAN OTHERS?

We turn next to a set of questions raised by our cost/benefit analysis of EPs. The first of these questions has to do with the salience, or impact on conscious experience, of one object property as compared to another. The cost–benefit relationships shown in Figure 10–4 indicate that material properties are more broadly available than geometric ones. That is, there are more patterns of exploration that reveal material than geometry. Moreover, precise shape, which is a geometric rather than material property, is encoded relatively inaccurately and very slowly. Does this mean that material is more salient when we feel an object freely, without vision?

To address this question, we and our colleagues Reed and Summers used a sorting task (Klatzky, Lederman, & Reed, 1987; Lederman, Summers, & Klatzky, 1996). We constructed objects that varied along several properties and asked people to sort them into bins according to their similarity. The objects were constructed by simultaneously varying geometric properties (e.g., size, shape), and material properties (e.g., roughness, hardness). There was no objectively correct response, and the dimensions were all equally discriminable. We used the pattern of sorting to determine which properties were most cognitively salient or important to the observer. For example, if size was salient, people should put all the large objects in one bin and all the small ones in another. If so, they would be mixing the objects within a bin with respect to the other properties: shape, roughness, and hardness. On this basis, the sorting pattern could be used to determine a "salience score" for each property, representing the extent to which the objects varying in that property were segregated into different bins.

There is an added manipulation in these experiments; namely, the instructions that participants were given with respect to what constituted similarity. Some subjects were told to think about similarity in terms of how the objects felt. Others were given no particular definition of similarity, but they could only feel the objects. Another group had no definition of similarity but could see the objects as well as feel them. A final group could only feel the objects, but they were told to imagine how the objects would look.

Both studies confirmed our expectations about relative dimensional salience based on costs and benefits. Considering the material properties, the salience score was higher when people felt the objects without vision or visual bias than when they saw them or imagined how they would look. When it came to the geometric properties, the salience score was higher when people saw the objects or imagined seeing them than when they merely touched them.

In short, costs and benefits affect how we think about objects. The cost of extracting geometric properties by haptic exploration means that material properties of objects become more salient when they are touched than when they are seen or visually imagined. Geometric properties become more salient when vision is present—or even imagined—than when the object is only touched.

WOULD WE USE TOUCH FOR OBJECT RECOGNITION IF VISION IS AVAILABLE, AND WHY?

Another implication of EP costs and benefits is that touching carries a cost in terms of exploration time. Why should people use touch, then, if vision is available? The answer is, of course, that it provides information that is needed but is otherwise unavailable or unreliable. Klatzky, Lederman, and Matula (1993) hypothesized that people would touch an object to encode its properties if vision did not provide reliable and accurate information, and if the information was not already available in factual memory. To test this hypothesis, we showed people pairs of real objects and asked them to say which was greater with respect to a particular property: roughness, hardness, apparent warmth, weight, size, or shape. The comparison was either easy or difficult, as in the table below.

Compared property	*Difficult pair*	*Easy pair*
Roughness	toast vs. sponge	pineapple vs. plum
Size	golf ball vs. marshmallow	penny vs. CD diskette

The results were striking. When confronted with difficult comparisons about an object's material (roughness, hardness, warmth) and also about its weight (which depends in part on material), participants touched one or both objects more than 60% of the time. When the question was about size or shape, or when it was an easy question about material properties or weight, participants rarely if ever touched the objects. As we expected, the relatively greater cost of haptic exploration than visual examination means that touch will be used only when it is particularly needed for difficult questions about properties that are not readily encoded with vision.

DOES EXPLORATION OF AN OBJECT FOR PURPOSES OF RECOGNITION PROCEED IN A STEREOTYPED WAY, AND IF SO, WHAT DETERMINES THE PROGRESSION?

Another implication of costs and benefits is that some methods of exploration quickly provide a lot of information, at a coarse level, while others more slowly provide precise information. EPs that are applicable broadly (the bottom rows in Figure 10–4) convey information about many properties, but not optimally. Among the broad EPs, contour following has a high cost in duration, leaving unsupported holding (lifting) and enclosure (grasping) as EPs that have high breadth and take relatively little time. Does this mean that those EPs will be executed early in exploration? Alternatively, if people want to find out about a particular property, such as hardness, will they execute the relevant EP (pressure) first? Coexecution is another factor in the mix: Perhaps people execute EPs early when they can be performed in tandem, which again would favor unsupported holding and enclosure (grasping and lifting).

To address this issue, Lederman and Klatzky (1990) gave participants a targeted identification task for which one property of an object was particularly diagnostic. For example, we asked them whether a noodle was a cooked noodle, for which purposes its hardness would be particularly diagnostic. (In order to determine the relative diagnosticity of properties in such questions, we performed an initial questionnaire study.)

We found that when participants were presented with a task of identifying an object for which one property was particularly relevant, they did not initially go after just that property. Rather, participants almost uniformly began the task by grasping and lifting the object—maximizing the properties that could be extracted coarsely. Only when this broadly applicable, low-cost, co-executable pair of EPs were performed did they go after other EPs that were more directly relevant to the question being asked. Thus exploration followed a two-stage sequence: first grasp and lift, then perform specialized exploration for targeted properties. When the targeted properties were associated with enclosure and unsupported holding—that is, they were revealed by grasping and lifting alone—exploration stopped at the first stage.

CAN WE CAPITALIZE ON THE FACT THAT MULTIPLE PROPERTIES OF OBJECTS PREDICT THEIR IDENTITY?

Suppose you are feeling a pear. What properties lead you to know it is a pear? It has a well-known, distinctive shape, it is smooth, and it is hand-sized. Presumably all these features act together to lead you to recognize it as a pear. Shape alone might be sufficient, and in that sense the other features are redundant, but redundant features help. It is well known from visual categorization research (e.g., Garner, 1974) that redundancy speeds object categorization, a phenomenon called redundancy gain.

With touch, redundancy gain is potentially limited by the costs of exploration. Our work has considered two kinds of costs: (a) When EPs cannot be executed together, using one EP to encode an object feature may preclude using another EP. This could block information from target features associated with the second EP. Not being able to perform the second EP means the information will not be available, and however redundant, it cannot speed categorization. (b) When EPs are executed on different regions of an object, the same locking-out can occur, not because they are motorically incompatible but because they are regionally so. In particular, exploring on a sharp edge to encode exact shape may preclude encoding texture, which requires sampling a broader region than is provided by the edge.

We demonstrated both types of blocking effects in a series of experiments. We asked subjects to categorize a set of objects by touch. In some cases, distinct categories were defined by just a single feature, such as size (all As are large, all Bs are medium, all Cs are small). In other cases, either of two features defined a category (e.g., As are large and rough), and in other cases three features provided a redundant definition (As are large, rough, and circular). Klatzky, Lederman and Reed (1989) conducted such a manipulation in which the stimuli were wafer-like shapes with textured surfaces. The shape information could be found only on the edges, which provided relatively little information about texture and hardness. Accordingly, we found that redundancy gains occurred when there were two redundant features, but when texture, shape, and hardness were all redundant, there was no further improvement relative to the two-feature case. We attribute this effect to both types of incompatibility described above. That is, exploring for texture and hardness precluded exploring for shape, due to movement incompatibility and concentration on different regions of the object. The failure of three-feature redundancy to improve over two-feature redundancy is not a general rule, then, but depends on how compatible the actions are that elicit the features in question.

To test our ideas further that incompatibilities prevent redundancy gain, we moved to a new set of 3-D stimuli varying in curvature. Shape information was now available all over the surface. This allowed shape and roughness to be encoded by a single motion, and redundancy gains between the two features were then seen (Lederman, Klatzky, & Reed, 1993).

We (Lederman et al., 1993; Reed, Lederman, & Klatzky, 1990) also introduced experimental manipulations that we called *redundancy withdrawal* and *orthogonality insertion*. In the redundancy-withdrawal paradigm, we asked subjects to categorize the same set of stimuli over a long series of trials. We instructed them that a category was defined by only a single property, like roughness, but in fact (without mention) there was a second property that covaried with roughness (e.g., all rough stimuli were also round; all smooth stimuli were oval). After subjects had performed the task for a while, we covertly changed the stimulus set, so that the second property was no longer a reliable cue to the category (now rough and smooth stimuli might both be round, and only roughness defined category membership). We asked whether categorization time would increase once previously redundant features were held constant. If so, this would indicate that participants had been relying on the redundant secondary dimension without having been instructed to do so. The increase in response time would be the cost of withdrawing redundancy. This manipulation confirmed the incompatibilities we had observed. Some combinations of redundant features produced a cost when redundancy was withdrawn. These were features that were extracted by compatible EPs in the sense they could be performed together on the same region. Other feature combinations showed little or no effect, indicating that participants had been ignoring the redundant secondary dimension. This occurred when the features were extracted by EPs that were motorically and/or regionally incompatible.

Orthogonality insertion is the converse of redundancy withdrawal. Now subjects began the experiment by categorizing stimuli according to a single dimension, with other properties held constant (e.g., by categorizing by roughness with constant shape). At some point in the series of trials, a previously constant attribute began to vary in a manner that was irrelevant to the targeted classification (e.g., shape began to vary while categorization continued on the basis of roughness). If participants spontaneously encoded the newly varying attribute, the time to do the classification should increase. Indeed, just such increases were found when the second attribute was extracted by the same EP that was performed to encode the first attribute.

In short, covariations among object features can speed classification when they are both diagnostic of the objects' categories. Irrelevant patterns of variation among objects from a common category can impair categorization. But whether these effects are seen when objects are categorized by touch depends on the patterns of exploration. The expected interactions among object features—positive or negative—occur only when those features are simultaneously encoded by the same EP or two compatible EPs. These constraints of commonality and compatibility are reflected in our cost–benefit analysis. Clearly, how you choose to explore manually constrains the available information and the perceptual consequences for haptic object recognition.

WHAT HAPPENS WHEN WE CONSTRAIN MANUAL EXPLORATION IN SPACE AND/OR TIME?

Up to this point, we have emphasized free exploration of objects in the service of learning about and identifying them. We turn now to what happens when exploration is constrained, either by limiting the time allowed for exploration or by limiting access to the spatial attributes of objects. Why study such constraints? We do so because the manipulations help us understand more about how object properties are encoded. Limiting the time informs us about the sequence in which object properties emerge and what can be achieved on the basis of early information. Limiting access to objects in a spatial sense tells us about the importance of shape and material information to object identification.

Cutaneous Spatial Constraints

We have approached spatial constraints in two ways, either by blocking the array of stimulation at the fingertip or by locking the joints. To understand the first of these approaches, one must be aware that some sensory receptors underneath the skin of the fingertip respond increasingly as more pressure is applied (within limits of saturation). These receptors are densely packed underneath the skin. When the responses of an entire population of the receptors beneath a skin area are considered at a single point in time, the result is an instantaneous map of the surface pressing against the skin (LaMotte & Srinivasan, 1993). For example, a Braille symbol would excite the receptors that lie below the skin contacting the raised dots more than the receptors that lie below the skin contacting the flat base. Mapping out the excitation would show higher responses in the skin areas where the dots were pressing. This spatial map of pressure is known to be important for perceiving not only geometric features of surfaces, as in Braille, but also for their nongeometric textural features as well (Hsiao, Johnson, & Twombly, 1993).

We asked what would happen to perceptual processing if the spatial pattern of forces on the fingertip were no longer available and people had access only to the summed forces as they excited receptors in skin, muscles, tendons, and joints. This is what happens, for example, when an object is explored with a tool such as a pencil. Our interest in this question stemmed in part from newly developed haptic interfaces, which similarly eliminate the pressure array on the fingertip as people explore with thimble-shaped coverings on the finger or pencil-like probes.

We eliminated the array map by covering the fingertip with a rigid sheath (Lederman & Klatzky, 1999) that capped the fingerpad from the tip of the finger to just above the most distal joint. The participants then performed in a battery of simple sensory tests and more complex haptic tests. The results showed that when the spatial deformation pattern on the fingertip is eliminated, people retain their ability to sense vibration, and they remain sensitive to gross pressure differences, but they become unable to discriminate patterns pressed against the skin. Figure 10–5 shows the decrements in performance that we found. These results, then, show that haptic perception must include the spatially distributed array of cutaneous information if fine spatial details of objects within the scale of the fingertip are to be apprehended.

Kinesthetic Spatial Constraints

In another series of studies, we considered larger-scale spatial variations that would ordinarily be encoded with contour following or the enclosure EP (Lederman & Klatzky, 2004). The subjects' task was to identify a set of common objects that were selected so that all were rigid, fairly smooth, and could not be lifted. The result of this selection was that shape information was critical to identification. We then constrained access to spatial properties by locking the subject's finger joints with splints, thus preventing execution of the enclosure EP, and by reducing the number of fingers that could be used in exploration. In the principal control condition, participants had intact kinesthetic perception from muscles, tendons, and joints, but because they wore a thick glove, they experienced sharply reduced cutaneous information relative to normal. The glove not only eliminated the precise spatial array, as in the previous study, but also damped vibration from contact. Comparisons across conditions, then, allowed us to ask how reductions in kinesthesis affected object recognition while keeping the contribution of cutaneous sensing to a minimum.

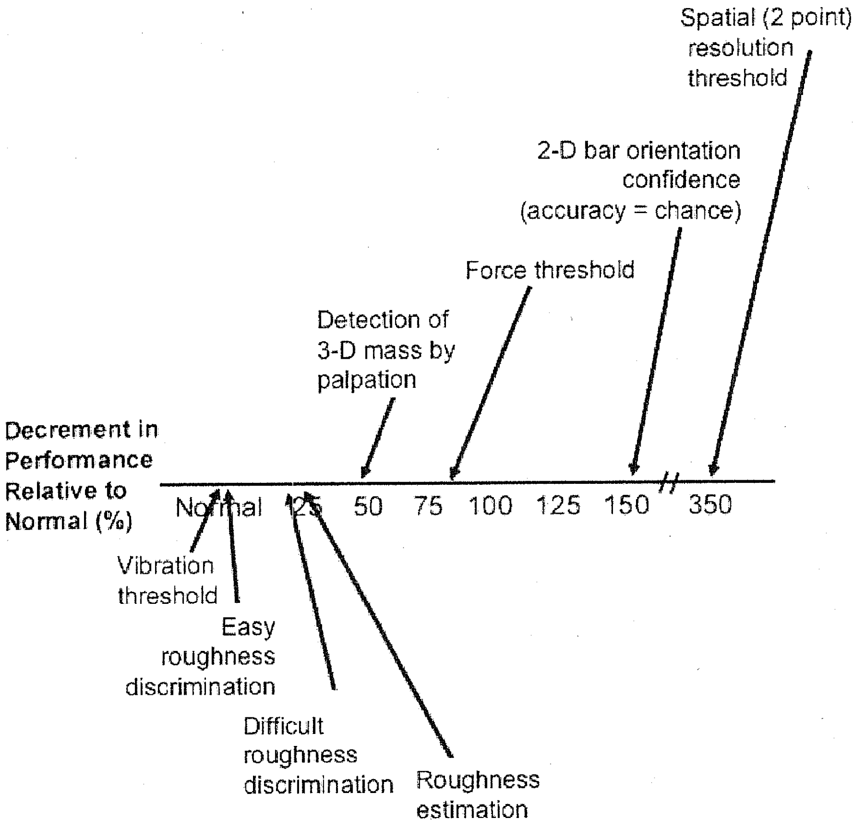


FIGURE 10–5. Consequences of removing spatially distributed cutaenous cues. Percentage deficit in performance relative to normal baseline for multiple sensory and perceptual tasks (based on data from Lederman & Klatzky, 1999).

The results were straightforward. Losing kinesthesis impairs recognition, whether the constraint comes from reducing the number of fingers contacting the object or from splinting it so that enclosure is impossible. And the more constraints, the more the impairment. Thus, for example, exploring with a single splinted finger (maximum reduction in number of fingers as well as eliminating enclosure) led to worse performance than exploring with five splinted fingers or one unsplinted finger. However, even with such constrained kinesthetic information and substantially limited cutaneous information, subjects were still able to perform with about 80% accuracy or better.

Additional conditions in this study introduced a rigid interface between the hand and the object. They required subjects to wear a rigid sheath capping their finger or to explore the objects with a pencil-shaped probe. Performance now degenerated to about 40%. Further research is needed to understand the reasons the rigid interface reduced performance. One likely contribution is that subjects found it far more difficult to explore the objects under these conditions. In contrast to the rigid probe, a thick but compliant glove no doubt provided limited cutaneous information that helped them to maintain contact with the object while moving around it.

Temporal Constraints: Feature Availability Over Time

The costs and benefits of EPs suggest that features become available at different points in time depending on the duration of exploration and the processes that encode the explored features. In order to determine the relative availability of haptically encoded features, Lederman and Klatzky (1997) adapted a procedure called *visual search* (after Treisman & Gormican, 1988). Our adaptation is called *haptic search.* In the search paradigm, the subject must say whether a target stimulus is present in a display that contains nontarget (i.e., distractor) stimuli. A sample task is to find a horizontal line among vertical distractors. The target (horizontal) may or may not be in the display, and the number of distractors is varied.

The subject's response time for the search task is plotted as a function of the total number of items in the display including the target—if present—and the distractors. If the response time does not vary with the number of items in the display—that is, the response takes the same amount of time, whether there are few distractors or many—the target is said to "pop out." Vertical lines, for example, pop out from among horizontal ones in visual displays. A vertical line is detected at the same speed whether it is alone or among several distractors. The logic of this paradigm is that when a particular feature pops out, that feature is assumed to be extracted by specialized detectors at some point in the perceptual channel. When a feature does not pop out, it must be processed with some cost by nonspecialized processors.

The cost of processing a single item can be estimated from the response-time data as follows. The relation of response time to the number of items in the display is typically linear. That is, each additional item adds some amount of time to the total. The increment in time from a single item is estimated by the slope of the response-time function. The slope, then, indicates how much processing time must be devoted to a particular target in a particular set of distractors. It indicates the relative availability of that target among those distractors. When targets and distractors are maximally different from one another, the slope of the response-time function can be taken as a general estimate of the cost of processing the targeted feature. (The intercept of the function is also relevant but will not be discussed here; see Lederman & Klatzky, 1997.)

In order to adapt visual search to the haptic modality, we used a motorized display (Moore, Broekhoven, Lederman, & Ulug, 1991) that could transport from one to three stimuli up to simultaneously contact the fingers of each of the two hands. The fingers were positioned in finger rests with the fingerpads exposed. On any trial, from one to six fingers were contacted by a stimulus plate. The plate could have a texture, an edge, or a 3-D variation such as a ramp. It could be made of metal or wood. These variations were used to define targets and distractors for a series of trials.

The subject's task was to say whether any of the fingers were in contact with a designated target, such as a rough surface, made of raised elements, among smooth distractors. When present, the target could be the only stimulus, or it could be presented with up to five distractor stimuli (in our example, a rough target could occur with up to five smooth surfaces). On target-absent trials, from one to six fingers were contacted by distractors. This task was performed with 13 target/distractor pairings, all chosen to be highly discriminable. The target/distractor pairs fell into four sets: (a) material discriminations, such as rough surface versus smooth surface or warm (wood) versus cool (aluminum) surface; (b) discrimination of a flat surface from a surface with an abrupt spatial discontinuity, such as a raised bar; (c) discrimination between 2- and 3-D spatial layouts, such as discriminating whether a raised dot was on the left versus the right of an indentation; and (d) discrimination between two continuous 3-D contours, such as a curved surface versus a flat surface.

For each target/distractor combination, we constructed two response-time functions, one for trials in which there was a target present and one in which it was absent. The slopes of these functions indicated first whether the target popped out, and if not, the cost of processing that type of target on each additional finger. Figure 10–6 shows the two most extreme response-time functions and the associated stimuli.

The slopes of the response-time functions varied considerably, from essentially zero (i.e., pop-out) to a half second. The distribution of slope values across the various types of stimuli indicated a progression in the availability of haptic object properties, from material properties to geometric ones. The slopes of the response-time functions for material properties were small (none was greater than 36 ms and several were close to zero). Somewhat higher slope values were obtained for geometric properties that would produce a sharp discontinuity in the surface (e.g., detecting a bar in the middle of an otherwise flat surface among distractors that were entirely flat). These properties could be identified quickly by the differential intensity of pressure across the surface. Finally a third group of slopes occurred when the spatial arrangement of the surface differentiated between targets and distractors (e.g., discriminating between a dot on the right versus left of a central indentation). The slopes were highest in this third group.

These results, then, provide clear convergence with our original cost–benefit analysis. They indicate a distinction between the accessibility of material and geometric properties within the haptic system in terms of their availability over time. There is little or no cost to extracting material information from one additional finger. There is a lot of cost to determining the spatial layout of

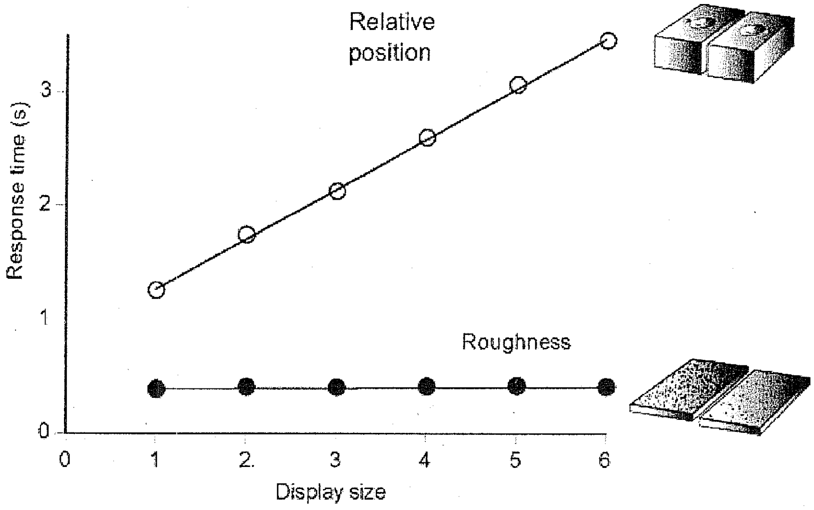


FIGURE 10–6. Haptic search functions for material (i.e., rough-smooth) and geometric (relative position) tasks. Response time is plotted in seconds as a function of haptic display size. Linear functions are fit to the data. From S. J. Lederman and R. L. Klatzky (1997). "Relative availability of surface and object properties during early haptic processing," *Journal of Experimental Psychology: Human Perception and Performance, 23*(6), 1–28. Copyright 1997 by American Psychological Association. Reprinted with permission.

surface edges in contact with a single finger. When we feel an object, its material emerges before the spatial arrangement of its surface features, even when those features lie within the span of a fingertip.

Temporal Constraints: A Haptic Glance

Another way in which we examined the time-course of haptic feature perception was to sharply limit the amount of time that people had to contact an object and to ask them to identify the object. We call this brief exposure a *haptic glance* (Klatzky & Lederman, 1995). To implement the haptic glance, we asked subjects to move their hand downward along a rod until they contacted the object with the middle fingers. After 100 ms, a tone sounded, indicating that they should withdraw the hand. Total exposure time was approximately 200 ms. The task was then to name the object.

The objects were selected so that either shape or texture was particularly diagnostic, that is, was a strong cue to their identity. For example, a sponge could be identified by texture, or a paper clip by shape. If there was a particularly informative region, such as the lip on a measuring cup, it was placed below the hand so that it would be contacted. An additional variable in the experiment was whether the object was large or small. A large object extended beyond the outstretched fingers, although the most diagnostic area was situated for contact; a small object had an informative feature that lay entirely within the finger span.

The results are shown in Figure 10–7. The first point to note is that people could identify a substantial percentage of the objects despite such time- and space-limited exposure. The second point is that performance depended on both the size of the object's surface and its diagnostic property. Where texture was diagnostic, identification was greater when the object was larger and hence provided more surface area for the fingers to sample. Where shape was diagnostic, performance was better when the object was small, in which case the contours fell entirely within the finger span.

These results suggest that familiar objects with distinctive, diagnostic contours, along with objects for which material is highly diagnostic, can sometimes be processed sufficiently for identification even given brief exposure. A paper clip touched with the fingers, for example, provides a unique, frequently identifiable shape cue. More often, however, identification requires extended exploration, giving rise to the advantages for material properties that we have described.

EMERGENT MODEL OF OBJECT IDENTIFICATION

The research reviewed above allows us to return to the issue with which we began, namely, how people identify objects by touch in comparison to vision.

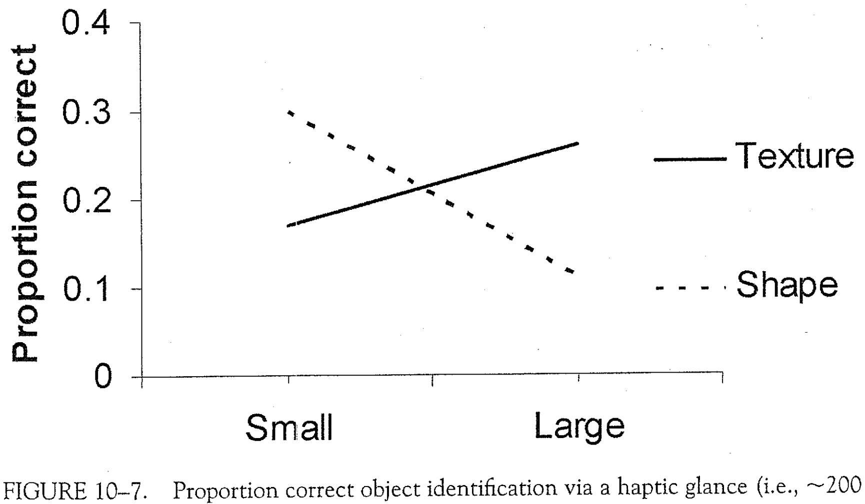


FIGURE 10–7. Proportion correct object identification via a haptic glance (i.e., ~200 ms) as a function of size of contact area and most diagnostic attribute. From R. L. Klatzky and S. J. Lederman (1995). "Identifying objects from a haptic glance," *Perception and Psychophysics*, 57(8), 1111–23. Copyright 1995 by Psychonomic Society. Reprinted with permission.

Fundamental properties of the haptic model emerge as follows (see Klatzky & Lederman, 2000, 2003, for more extended discussion of the model):

1. Edge information is primary in visual object recognition, and a 3-D object representation can readily be obtained from a 2-D projection. In contrast, material information is, generally speaking, more accessible than geometric information in touch, and fully 3-D information is more accessible than a raised 2-D projection.
2. Touch capitalizes on redundant information about material and geometric properties in order to achieve object categorization. Common object categories are often defined by shape features, as are, for example, tables and chairs (Rosch, 1978). However, material is often diagnostic, and its greater accessibility to touch means that it can play an important role in haptic recognition.
3. Haptic object recognition unfolds over time as objects are explored. Exploration is used as needed; sometimes a simple glance may be sufficient to do the job. The exploratory sequence is governed by general cost–benefit principles and expectations about the object, which render certain features particularly diagnostic and therefore make certain EPs advantageous.
4. Object recognition requires comparison between perceptual input and memory. Models of visual recognition have suggested representations in the form of stored prototypes (Biederman, 1987) or stored views (Tarr & Bülthoff, 1995). The haptic model must, however, consider comparison to previous haptic experiences, or stored "touches."

Based on the Fundamental Research on Haptic Object Recognition, What Are the Implications for Education and Communication in People Who Are Blind or Have Low Vision?

A take-home message from our report is that touch, without vision, richly conveys a sense of real, tangible objects. Moreover, the most accessible and salient features pertain to object material and 3-D geometry. These features are conveyed by texture, heat flow, compliance, weight, and pressure differences that accompany edges or continuous 3-D contours. Communicating such information should be an important goal for purposes of education, aesthetics, or entertainment.

We envision educational programs in which simulation of tangible objects is used for courses in human anatomy, astronomy, mechanics, or molecular biology, for example. Museum artifacts might be presented over the web in arts classes. People who are blind could have access to tangibly guided tools for computer-aided design and modeling applications. Free and flexible exploration of these simulations is enabled in our ideal educational technology.

Unfortunately, this goal is far easier to express than to accomplish. There are many technical challenges. In the past decade or so, a number of devices have been developed for delivering forces to the skin that simulate contact in virtual or remote environments. The world of such force-feedback interfaces is evolving, but the predominant design at present simulates contact with an object at a single point. Textural information is limited; thermal information is nonexistent; objects are not entirely rigid. Maintaining extended contact with the object is difficult when it is felt at a single point. The interactions with these interfaces, in short, resemble those we have described in which people explore objects with a rigid probe or a sheath over the finger. Vibration and gross pressure variations can be felt, but the spatial array is absent, as are non-force-related properties.

Haptic interfaces are needed that stimulate the finger with an array of forces, that provide thermal changes, and that simulate the transition of pressure as the finger crosses a surface boundary. The simulation model must incorporate complex end-effectors and not just contact with a rigid point. To portray complex objects over remote connections, the devices will have to meet increased demands for bandwidth, storage and transmission. Libraries of objects' tangible properties must be amassed.

While these goals are enormously challenging, they are also very exciting because researchers have been developing promising new haptic technologies that may prove valuable. We look forward to educational developments that unite the needs of persons who have low vision or are blind with the marvelous perceptual capabilities of the human haptic system and the dedication of the haptic research community.

REFERENCES

Biederman, I. (1987). Recognition by components: A theory of human image understanding. *Psychological Review, 94*, 115–145.

Garner, W. P. (1974). *The processing of information and structure.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Hsiao, S. S., Johnson, K. O., & Twombly, I. A. (1993). Roughness coding in the somatosensory system. *Acta Psychologica, 84*, 53–67.

Johnson, K. O., & Lamb, G. D. (1981). Neural mechanisms of spatial tactile discrimination: Neural patterns evoked by Braille-like dot patterns in the monkey. *Journal of Physiology, 310*, 117–144.

Klatzky, R. L., & Lederman, S. J. (1995). Identifying objects from a haptic glance. *Perception and Psychophysics, 57*(8), 1111–1123.

Klatzky, R. L., & Lederman, S. J. (1999). The haptic glance: A route to rapid object identification and manipulation. In D. Gopher & A. Koriat (Eds.), *Attention and performance. XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 165–196). Mahwah, NJ: Lawrence Erlbaum Associates.

Klatzky, R. L., & Lederman, S. J. (2000). L'identification haptique des objets significatifs [The haptic identification of everyday life objects]. In Y. Hatwell, A. Streri, & E. Gentaz (Eds.), *Toucher pour connacetre: Psychologie cognitive de la perception tactile manuelle* [Touching for knowing: Cognitive psychology of tactile manual perception] (pp. 109–28). Paris: Presses Universitaires de France. English translation 2003, Philadelphia: Benjamins.

Klatzky, R. L., & Lederman, S. J. (2003). Touch. In A. F. Healy & R. W. Proctor (Eds.), Experimental psychology (pp. 147–76). Vol. 4 in I. B. Weiner (Editor-in-Chief). *Handbook of psychology*. New York: John Wiley & Sons.

Klatzky, R. L., Lederman, S. J., & Matula, D. (1993). Haptic exploration in the presence of vision. *Journal of Experimental Psychology: Human Perception & Performance, 19*(4), 726–743.

Klatzky, R. L., Lederman, S. J., & Metzger, V. (1985). Identifying objects by touch: An "expert system." *Perception & Psychophysics, 37*(4), 299–302.

Klatzky, R., Lederman, S. J., & Reed, C. (1987). There's more to touch than meets the eye: The salience of object attributes for haptics with and without vision. *Journal of Experimental Psychology: General, 116*(4), 356–369.

Klatzky, R., Lederman, S. J., & Reed, C. (1989). Haptic integration of object properties: Texture, hardness, and planar contour. *Journal of Experimental Psychology: Human Perception & Performance, 15*(1), 45–57.

LaMotte, R. H., & Srinivasan, M. A. (1993). Responses of cutaneous mechanoreceptors to the shape of objects applied to the primate fingerpad. *Acta Psychologica, 84*, 41–51.

Lederman, S. J., & Klatzky, R. L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology, 19*(3), 342–368.

Lederman, S. J., & Klatzky, R. L. (1990). Haptic classification of common objects: Knowledge-driven exploration. *Cognitive Psychology, 22*, 421–459.

Lederman, S. J., & Klatzky, R. L. (1997). Relative availability of surface and object properties during early haptic processing. *Journal of Experimental Psychology: Human Perception and Performance, 23*(6), 1–28.

Lederman, S. J., & Klatzky, R. L. (1999). Sensing and displaying spatially distributed fingertip forces in haptic interfaces for teleoperator and virtual environment systems. *Presence: Teleoperators and Virtual Environments, 8*(1), 86–103.

Lederman, S. J., & Klatzky, R. L. (2004). Haptic identification of common objects: Effects of constraining the manual exploration process. *Perception & Psychophysics, 66*(4), 618–628.

Lederman, S., Klatzky, R., Chataway, C., & Summers, C. (1990). Visual mediation and the haptic recognition of twodimensional pictures of common objects. *Perception & Psychophysics, 47*(1), 54–64.

Lederman, S. J., Klatzky, R. L., & Reed, C. L. (1993). Constraints on haptic integration of spatially shared object dimensions. *Perception, 22*, 723–743.

Lederman, S., Summers, C., & Klatzky, R. (1996). Cognitive salience of haptic object properties: Role of modality-encoding bias. *Perception, 25*(8), 983–998.

Loomis, J. M. (1990). A model of character recognition and legibility. *Journal of Experimental Psychology: Human Perception & Performance, 16*, 106–120.

Loomis, J. M., & Lederman, S. J. (1986). Tactual perception. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance* (pp. 31-1–31-41). New York: Wiley.

Magee, L., & Kennedy, J. (1980). Exploring pictures tactually. *Nature, 283*, 287–288.

Moore, T., Broekhoven, M., Lederman, S., & Ulug, S. (1991). Q'HAND: A fully automated apparatus for studying haptic processing of spatially distributed inputs. *Behavior Research Methods, Instruments and Computers, 23*(1), 27–35.

Reed, C., Lederman, S. J:, & Klatzky, R. L. (1990). Haptic integration of planar size with hardness, texture, and planar contour. *Canadian Journal of Psychology, 44*, 522–545.

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), Cognition and categorization (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum Associates.

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*(6), 1494–1505.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95, 15–48.

Walk, R. D., & Pick, H. L., Jr. (1981). *Intersensory integration and sensory integration*. New York: Plenum.