Visual mediation and the haptic recognition of two-dimensional pictures of common objects

SUSAN J. LEDERMAN

Queen's University, Kingston, Ontario, Canada

ROBERTA L. KLATZKY

University of California, Santa Barbara, California

CYNTHIA CHATAWAY

Harvard University, Cambridge, Massachusetts

and

CRAIG D. SUMMERS

Queen's University, Kingston, Ontario, Canada

A set of three experiments was performed to investigate the role of visual imaging in the haptic recognition of raised-line depictions of common objects. Blindfolded, sighted (Experiment 1) observers performed the task very poorly, while several findings converged to indicate that a visual translation process was adopted. These included: (1) strong correlations between imageability ratings (obtained in Experiment 1 and, independently, in Experiment 2) and both recognition speed and accuracy, (2) superior performance with, and greater ease of imaging, two-dimensional as opposed to three-dimensional depictions, despite equivalence in rated line complexity, and (3) a significant correlation between the general ability of the observer to image and obtained imageability ratings of the stimulus depictions. That congenitally blind observers performed the same task even more poorly, while their performance did not differ for two- versus three-dimensional depictions (Experiment 3), provides further evidence that visual translation was used by the sighted. Such limited performance is contrasted with the considerable skill with which real common objects are processed and recognized haptically. The reasons for the general difference in the haptic performance of two- versus three-dimensional tasks are considered. Implications for the presentation of spatial information in the form of tangible graphics displays for the blind are also discussed.

The haptic perceptual system uses both cutaneous and kinesthetic inputs to derive information about the properties of objects and about their spatial layout (Gibson, 1966; Loomis & Lederman, 1986). Past research has consistently represented haptics as an ineffective system for learning about the concrete world. It has been compared to vision in particularly unfavorable ways. The evidence derives largely from three sets of perceptual tasks: the perception of extent, direction, and position of simple point and line stimuli, the recognition of raised-line drawings of familiar objects, and the recognition of planar non-sense shapes. We will consider each in turn.

This research was supported by an undergraduate summer student award (to C. Chataway) and an operating grant, both from the Natural Sciences and Engineering Research Council of Canada, and by contracts from the Ontario Information Technology Research Centre and the US Office of Naval Research. The contributions of S. J. Lederman and R. L. Klatzky are equal, and are not reflected in the order of authorship. Address reprint requests to either S. Lederman, Queen's University, Department of Psychology, Kingston, Ontario K7L 3N6, Canada, or R. Klatzky, University of California, Department of Psychology, Santa Barbara, CA 93106.

First, the haptic system appears to produce substantial error when perceiving position, distance, and direction. Errors are both variable (i.e., there is no directional bias) and constant (i.e., they are signed in a particular direction). There are likely to be multiple sources of such errors under haptic exploration. The low spatial resolution of the system, as well as memory and integration demands of exploration over time, should contribute to uncertainty about spatial layout, thereby producing variable error. Constant errors, we have argued elsewhere (Lederman, Klatzky, Collins, & Wardell, 1987), may reflect the haptic explorer's use of "heuristics" in an attempt to compensate for perceptual deficits, with systematic distortions as a result. Heuristics would include the referring of positions or orientations to anchor points of an explicit or implicit nature. Evidence for such processes was found by Lederman and Taylor (1969), who required subjects to reproduce the position of a raised dot on a rectangular card or to reproduce the orientation of a raised-line radius in a semicircle. Errors indicated that position responses were drawn toward the corners of the card, and that orientation responses were drawn to implicit static axes. Similar

orientation errors were observed by Lederman, Klatzky, and Barber (1985), with the pull toward the vertical axis as great as 50%. A different type of spatial distortion was demonstrated by Wong (1977), who showed that distances extending radially from the body were haptically overestimated with respect to those explored tangentially. He proposed that the distortion was due to the use of movement inertia to assess extent, the inertia being greater in radial limb movements. More recently, Klatzky and Lederman (Lederman et al., 1985; Lederman et al., 1987) have documented another "length distortion" error, in which haptic explorers in both manipulatory and ambulatory space overestimate the distance of the straight-line path between two points as a function of the actual "detour" pathway explored. For the longest straight-line paths, an eightfold increase in the detour pathway resulted in a 100% increase in overestimation error. The magnitude of the haptic spatial errors described here indicates that the haptic system is subject to considerable limitations.

Haptic recognition of raised-line drawings would seem to be even poorer. This has been informally reported by educators of the blind (e.g., Schiff & Foulke, 1982); we too have consistently observed the same phenomenon with blind and blindfolded visitors to our laboratories. More formally, Magee and Kennedy (1980) examined the ability of sighted, blindfolded observers at recognizing raised-line drawings of 12 visually identifiable objects (a butterfly, boat, crown, fish, goblet, claw hammer, hand, body, rabbit, umbrella, swan, and flower). In the active (free) exploration condition that is most relevant to the current study, only 17% of the drawings were recognized; performance in an earlier study was even lower, although details were not provided. Poor recognition performance is no doubt influenced by factors such as a lack of familiarity with the task or the tactual unfamiliarity of the stimulus objects. However, we suggest that constraints inherent in the haptic system itself limit performance with two-dimensional displays.

Finally, a considerable amount of research has been devoted to the haptic recognition of planar shapes (for a summary and references, see Jones, 1981). This work has dealt primarily with issues pertaining to the intersensory integration of visual and haptic inputs. However, the use of transfer and matching paradigms also provides comparative information pertaining to intrasensory performance. The data indicate that haptic recognition of form by children and adults is consistently poor and almost always inferior to vision.

In all three of these sets of experimental tasks, we note that observers were required to base their judgments on the structural properties of two-dimensional objects. We call them "two-dimensional," inasmuch as the lack of variation in the third dimension restricts relevant structural information to only two. Typically, the stimuli are raised points or lines, outline depictions of common objects, and planar solid nonsense objects. The haptic system must extract information that is spatially and tem-

porally distributed, in order to judge shape, size (including extent, area, and volume), position, and layout.

Throughout the work just described, there seems to exist the implicit (if not explicit) assumption that the haptic system processes information as an inferior form of vision would: It translates the low-resolution, temporally sequential, haptic inputs into a visual image, then reperceives the image through the use of visual processors (Figure 1, top). We (Klatzky & Lederman, 1987; Lederman & Klatzky, in press) call this an "image-mediation" model of haptics, and we have argued that it is incomplete, since it focuses primarily on tasks that highlight the system's weaknesses, without recognizing or accounting for its strengths. Not only do these tasks impose tremendous constraints on performance with respect to spatial resolution, memory, and integration, they also minimize the role of haptically relevant cues, such as weight, hardness, and thermal characteristics.

In marked contrast to the performance of two-dimensional spatial tasks, we recently offered an existence proof that the haptic system can perform remarkably well. We (Klatzky, Lederman, & Metzger, 1985) confirmed that subjects recognized 100 common objects with almost perfect accuracy, typically within only 1 or 2 sec! Such a task clearly highlights the strengths of the haptic system. In this initial study, our subjects noted a variety of properties that led them to correct recognition: texture, hardness, thermal characteristics, shape, size, weight, and function.

This work has led us to propose an alternate view of haptics, known as the *direct haptic apprehension* model (Figure 1, bottom). Here, haptics is viewed as a separate perceptual system, with its own physiological apparatus and its own style of processing. Only during later stages might there be some sharing or common processing on the part of both vision and touch. An important feature of our interpretation is the distinction between substance and structure. Substance properties include those relating to surface texture (e.g., roughness), hardness (e.g., compliance, elasticity, brittleness), thermal characteristics (e.g., conductivity, capacity), and weight (as determined by density). Structural properties include those pertaining to the two- and three-dimensional shape and size of objects, and to weight (as determined by volume). The anatomical, sensory, and motor capacities of the human hand particularly lend themselves to an emphasis on substance and three-dimensionality. We have argued that it is the multidimensionality of the haptic cues, and the efficiency with which it is possible to extract local substance information and gross three-dimensional structure, that results in such excellent performance. These, we argue, represent the particular strengths of the haptic system.

In most two-dimensional tasks, there is consistently little of the potentially rich information normally available to the exploring hand. There is typically no variation in substance, leaving only limited planar contour information that must be sequentially apprehended, recalled, and integrated. Given the severely restricted cues available, we

Figure 1. "Image-mediation" and "direct apprehension" models of haptic processing. From "The Intelligent Hand" (p. 129) by R. L. Klatzky and S. J. Lederman, in *The Psychology of Learning and Motivation* (Vol. 21), 1987, New York: Academic Press. Copyright 1987 by Academic Press. Reprinted by permission.

hypothesize that the haptic system may be forced to operate as suggested by the image-mediation model, because imaginal processing is one of the few (although not necessarily successful) options available. Accordingly, the present study sought to provide evidence for an image-mediation process and to assess its efficacy under optimal conditions.

To maximize recognition, we used only highly familiar objects that are commonly held in the hand, and we depicted them in highly prototypical form. (Note that this was not the case in previous work by Magee & Kennedy, 1980.) We were able to compare performance with respect to two- and three-dimensional objects directly, because we used raised outlines of many of the same objects that were recognized in real, three-dimensional form in the experiment of Klatzky et al. (1985).

Our second purpose was to identify sources of error in the task and to identify them with components of an image-mediation process. We therefore assessed the effects of variables that should influence different components. For example, shape complexity and imageability might be expected to influence image formation; the presence or absence of three-dimensional cues in the picture should influence image parsing and interpretation; and object-name frequency should affect lexical access subsequent to conceptual recognition.

In Experiment 1, blindfolded, sighted subjects attempted to recognize a set of raised-line drawings of common objects as quickly and as accurately as possible. In

Experiment 2, we report independent ratings of the imageability of each drawing, which were used for purposes of assessing the role of visual imaging in this two-dimensional haptic task. In Experiment 3, we compared the recognition performance of congenitally blind observers with that of the blindfolded, sighted observers from Experiment 1, to assess the role of visual experience in haptic picture recognition.

EXPERIMENT 1 Haptic Recognition of Raised-Outline Drawings of Common Objects by Sighted, Blindfolded Observers

In Experiment 1, we required sighted, blindfolded subjects to haptically recognize raised-line drawings of highly familiar objects. Their performance was very poor. The data suggest that the subjects adopted an image-mediation model of haptic recognition, inasmuch as visual factors were strongly associated with both speed and accuracy of response in this two-dimensional task.

Method

Subjects. Twelve males and 12 females, ranging in age from 22 to 32 years of age, participated in the experiment. All were English-speaking university students with no reported tactual deficits.

Stimulus materials. A set of 22 raised-line pictures depicting highly familiar objects was prepared manually, using a braille stylus on plastic sheets (27.9 × 29.6 cm). This technique was selected because of the relative ease with which it could reproduce sharp,

textured (dotted) contours from a reverse print image, which had been transferred to the back of the plastic sheet. The pictures were drawn to a standard size: The length along the object's largest axis was within 12.7–15.2 cm. The criteria for inclusion were: (1) the pictures depicted object concepts at the "basic" level (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976); (2) the depicted objects had themselves been used in the Klatzky et al. study; (3) the objects represented could be examined by hand (e.g., a carrot was acceptable, but a house was not); and finally, (4) with respect to Snodgrass and Vanderwart (1980), print drawings of the objects were available, and the objects had high intersubject name agreement (i.e., with the exception of the tie—69%—and the baseball bat—56%—all items had greater than 80% name agreement), greater than average concept familiarity, and ratings greater than 2.78 with respect to image agreement (i.e., agreement between the subject's image of an object and the associated picture; 1 = *low*, 5 = *high*). The 22 selected Snodgrass and Vanderwart drawings were further divided into two- versus three-dimensional representations, according to whether there were perspective cues or internal lines depicting edges. (We simplified the original Snodgrass and Vanderwart drawings by eliminating lines that depicted visual cues for shading and texture.) The two-dimensional drawings included a sweater, pencil, sock, lightbulb, comb, envelope, tennis racquet, screw, carrot, key, and baseball bat. The three-dimensional drawings included a bowl, candle, screwdriver, cup, hammer, plug, tie, lock, book, whistle, and ashtray. The relevant drawings in the Snodgrass and Vanderwart study that had an ambiguous top-bottom orientation were oriented with the functional end down; long, narrow objects were oriented about equally to the left and right. Unfortunately, the orientation of the book was reversed in presentation by mistake, and the book was therefore not used in Experiment 2; it was represented as opening from the left, rather than from the right. A number of practice items were also prepared. A pretest with 6 additional subjects confirmed that all the line drawings were visually recognized with 100% name agreement.

Procedure. The participants began the session by completing the Visual Vividness Imagery Questionnaire (VVIQ; see Marks, 1973) with their eyes closed. The questionnaire is a self-report of the vividness of a generated image.

The subjects were blindfolded and asked to explore freely a set of raised-line drawings of common objects with one or both hands. They were to identify the drawings as quickly and as accurately as possible. If they did not know an answer, they were told to indicate this. The experimenter explained that the objects were not represented as they were normally encountered in everyday life, but were produced to a standard size, which was described. They were further told that in real life, the objects depicted would be no larger than a pair of pants and no smaller than a ring. They were given a maximum of 2 min for each picture. A warning that 20 sec remained was given if an answer was not forthcoming by that time. Reaction times to naming were recorded with a stopwatch. When an answer was not given, the subject was asked to report his or her impressions. At the end of the trial, each picture was further rated in terms how difficult it was to form a mental image of the object as it was felt (1 = *not at all difficult*; 5 = *very difficult*).

The subjects were given a number of timed practice trials with additional line drawings; this was the only time during which feedback was provided. Both the practice and the test stimuli were presented in a different random order for each subject.

Results

The mean percent of correctly identified pictures was 33.5% ($SD = 24.5$). Mean reaction time for the set of stimuli was 91.2 sec ($SD = 22.6$). Reaction time, accuracy, and rated difficulty in imaging are reported by dimensionality (two vs. three) of representation, and by stimulus, in Table 1.

Table 1
Experiment 1: Performance by Item

	Mean Reaction Time	Mean Percent Accuracy	Mean Imaging Difficulty
Two-Dimensional Representation			
sweater	106.5	16.7	3.2
pencil	57.4	50.0	2.3
sock	95.2	45.8	3.1
lightbulb	58.8	83.3	2.4
comb	47.7	75.0	2.1
envelope	93.7	37.5	2.7
tennis racquet	75.0	33.3	2.5
screw	94.9	29.2	3.2
carrot	113.5	12.5	3.4
key	96.1	37.5	3.4
baseball bat	74.1	33.3	2.3
Three-Dimensional Representation			
bowl	49.9	75.0	2.0
candle	93.6	45.8	3.2
screwdriver	104.8	12.5	3.3
cup	82.1	54.2	2.7
hammer	84.5	50.0	2.8
plug	111.9	12.5	4.0
tie	96.8	16.7	3.0
lock	116.8	8.3	4.0
book	116.7	4.2	4.0
whistle	119.9	4.2	4.3
ashtray	117.0	0.0	4.2
Total mean	91.2	33.5	

Note—Reaction time is given in seconds. Imaging difficulty was rated on a scale from 1 to 5, where 1 = *very easy to image* and 5 = *very difficult to image*.

With stimulus as the unit of observation, simple intercorrelations among reaction time, accuracy (number correct), and imaging difficulty were determined. For purposes of later theorizing, we also added four other variables. Of these, three were alternate measures of familiarity, since familiarity has been shown to influence visual object recognition (e.g., by Bartram, 1973). They included the Kučera-Francis (1967) word-count frequency, familiarity ratings of the concepts depicted in the pictures (1 = *low*; 5 = *high*), and percent name agreement among subjects asked to name the corresponding pictures of the objects. The numeric values for the three familiarity variables for items in the current study were obtained from Snodgrass and Vanderwart (1980). The fourth variable added to the intercorrelation matrix was an independent measure of line complexity (amount of detail and complexity of lines; 1 = *very simple*, and 5 = *very complex*). The values, taken from the Snodgrass and Vanderwart study, represented rated visual complexity.

The correlations among subject-generated measures were all highly significant—between reaction time and accuracy, reaction time and imaging difficulty, and accuracy and imaging difficulty [$r(20) = -.91, .94$, and $-.84$, all $ps < .001$], respectively. Of the three familiarity measures, rated concept familiarity was statistically correlated with the three factors above ($ps = .04-.06$), indicating that objects that were least frequently encoun-

tered were recognized more poorly. The Kučera-Francis measure also correlated with imaging difficulty ($r = .42$, $p < .03$). The half intercorrelation matrix is shown in Table 2.

Differences in performance on two- versus three-dimensional representations were evaluated, using a t test for independent means on both dependent variables. Mean reaction time for two-dimensional items was significantly lower than for three-dimensional items [$t(20) = 1.79$, $p < .05$, one-tailed]. Furthermore, mean number of items correct tended to be higher for two- than for three-dimensional items [$t(20) = -1.53$, $p = .07$, one-tailed]. Thus, two-dimensional pictures were processed faster and with greater accuracy than were three-dimensional pictures. Two-dimensional items were also easier to image than three-dimensional items [$t(20) = 2.06$, $p < .05$, one-tailed], despite the fact that there was no statistical difference in the visual complexity ratings of two- versus three-dimensional pictures [$t(20) = .76$, $p > .10$, one-tailed].

The significant correlation between subjects' VVIQ scores and their corresponding reaction times [$r(22) = .44$, $p < .05$] indicated that high ability to image (low scores) was associated with low reaction time; likewise, high image ability tended to be associated with high percent accuracy. The correlation approached statistical significance [$r(22) = -.38$, $p < .10$].

Discussion

Sighted, blindfolded subjects haptically recognized common objects portrayed as raised two-dimensional outlines very poorly. Their success rate was only about 33%, and it took them an average of 90 sec to respond. Given the paucity of sensory inputs under such circumstances, we proposed that subjects would adopt an image-mediation

model, whereby the haptic inputs are translated into a visual image, which is subsequently processed by the visual system. The current results provide some initial evidence for this hypothesis.

First, imageability (as measured by the rated imaging difficulty) was correlated with both performance measures—reaction time and accuracy. That is, the less difficult it was to form an image of the drawing, the less time it took subjects to respond, and the more accurate they were. Second, VVIQ and item imageability were significantly correlated, indicating that if a subject was a good imager, he or she tended to rate the various drawings as easier to image. Third, the two-dimensional pictures were recognized better than the three-dimensional pictures were, as would be the case if it was necessary to interpret three-dimensionality from a two-dimensional visual image. In support of this interpretation is the additional finding that two-dimensional pictures were easier to image. This difference in performance occurred despite the fact that the two sets of pictures were judged to be equal in simple two-dimensional line complexity.

What do the strong correlations between imaging difficulty and the two measures of performance signify? A causal interpretation would suggest that imageability (as measured by imaging difficulty) strongly determined recognition performance (as measured by accuracy and reaction time). An alternate interpretation is that subjects may have believed that the speed and accuracy of their responses actually reflected the difficulty they had in trying to imagine the drawing. Presumably, they would not have, had they felt there was no relation between their performance and their ability to image. Thus, either way, these correlations support the hypothesis that visual imagery played an important role in this task.

Table 2
Experiment 1: Intercorrelation Matrix

RT	Accuracy	Imaging Difficulty	Kučera-Francis Familiarity	Conc. Fam.	Name Agreement	Visual Complexity	Vividness
RT	.91 $p = .00$	.94 .00	.13 .30	-.35 .06	.18 .21	-.00 .50	.74 .00
Accuracy	—	-.84 .00	-.14 .29	.37 .04	.17 .47	-.11 .32	-.65 .00
Imaging Difficulty		—	.16 .26	-.35 .05	.31 .08	.07 .37	.87 .00
Kučera-Francis Familiarity*			—	.49 .02	.12 .31	-.20 .20	-.20 .43
Conc. Fam.*				—	.03 .45	-.28 .11	-.58 .00
Name Agreement*					—	.13 .26	.28 .22
Visual Complexity*						—	.20 .38
Vividness (Experiment 2)							—

* Taken from Snodgrass and Vanderwart (1980).

Next, we wished to perform a multiple regression on the performance data to determine the contribution of the variables mentioned above. All factors involving familiarity and line complexity had been obtained from an independent source (Snodgrass & Vanderwart, 1980). However, given the possible dependence of obtained imagery ratings on performance, we decided to obtain an independent assessment of the imageability of each item.

EXPERIMENT 2

An Independent Assessment of Image Vividness of Raised-Line Drawings of Common Objects

Accordingly, in Experiment 2, we asked a new group of subjects to rate the vividness of the visual image they could form of each of the line drawings used in Experiment 1, with the exception of the book item.¹ Each was rotated 180° to discourage recognition; we intended vividness to reflect imageability uninfluenced by recognition or naming.

Method

Subjects. Eight females and 2 male undergraduates, ranging in age from 19 to 46 years, were paid for their participation. All were normally sighted.

Stimuli and Procedure. Depictions of the 21 objects from Experiment 1 (and Experiment 3) were rated for image vividness by each subject.² The actual order of stimulus presentation was completely random for each subject, with the single exception that two inverted stimuli that were "recognized" during pilot work, a lightbulb and sock, were presented alternately in the last two positions. This was done to prevent subjects from trying to recognize any of the other stimulus items.

The blindfolded subjects were told they would be presented with a number of raised-line patterns. They were told that the patterns did not depict real objects, and, consequently, that they should not try to name them. To prevent recognition, the displays were also rotated 180°. On each trial, subjects explored the pattern freely with one or both hands and tried to form a visual image of the pattern in their minds. Five practice stimuli that ranged in difficulty of imaging were initially presented. For these drawings and subsequent test items, the subjects were required to rate the vividness of their visual images on a 1 (very vivid) to 5 (not at all vivid) scale. As in the VVIQ test, images were to be considered vivid if the subjects had very clear views of them in their mind's eyes; they should believe they could almost see them. Whole numbers were used to make the ratings. The subjects were told that speed was important, but that it was also necessary to form each image as well as possible. They had up to 1 min to explore each pattern. Any spontaneous object identifications were noted. At the end of the experiment, the subjects completed the VVIQ questionnaire.

Results

The mean reaction time for subjects to report the vividness of their visual images was 51.8 sec ($SD = 15.5$). The Pearson product-moment correlation between reaction time scores in this experiment and the mean vividness rating was highly significant [$r(48) = .51, p < .0005$]. Not surprisingly, this indicates that as the vividness of an image decreased, it took longer to make the judgment. More importantly, correlations between vividness ratings and both the mean reaction time and the ac-

curacy data obtained in Experiment 1 were again highly significant [$r(20) = .74$ and $-.65, ps < .001$]. Since the vividness scores were also highly correlated with the image difficulty scores obtained at the end of each trial in Experiment 1, these correlations suggest that a very sizeable portion of the variance in the latter scores actually reflected the contribution of factors pertaining to imageability [$r(20) = .87, p < .0001$]. Table 2 presents the full set of correlations between vividness and all variables previously examined, again with stimulus items as the unit of observation.

Since subjects judged visual vividness almost twice as fast as they haptically recognized the objects depicted (Experiment 1), the data suggest that we were probably effective in minimizing attempts to recognize and name the stimuli. In keeping with this interpretation, only the lightbulb was recognized spontaneously, and this by a single subject.

We tentatively propose a four-stage process model of the haptic recognition of two-dimensional representations of common objects. At this early phase in our research, its main function is to provide a reasonable justification for choosing predictor variables in linear multiple regression models of performance data. During Stage 1, subjects engage in haptic exploration of the raised-line pattern. Factors affecting performance in this initial stage would most likely relate to tactual and kinesthetic acuity. None of the variables in Experiment 1 are likely to affect Stage 1. Stage 2 represents the visual translation phase, during which subjects form a visual image from the haptic inputs. We include image parsing/interpretation at this stage as well, since they may be theoretically equivalent and impossible to separate empirically. Potential predictor variables would include visual line complexity and our alternate measures of imageability (vividness and imaging difficulty). We reserve the process of concept identification for Stage 3, which may be reflected by rated concept familiarity. Stage 4 represents the final process of name retrieval, as reflected perhaps by name agreement and Kučera-Francis word frequency.

With this theoretical guide, we tested two linear regression models for the reaction time and accuracy data obtained in Experiment 1. In the first one, we used predictor variables that were obtained independently of that experiment. As previously mentioned, there was no predictor variable representing Stage 1. For Stage 2, we included visual picture complexity from the Snodgrass and Vanderwart (1980) study (since it seemed likely to affect both image formation and parsing/interpretation) and vividness (to reflect image formation). For each of Stages 3 and 4, we chose the factor with the largest simple correlations with both reaction time and accuracy. Accordingly, the following linear model for reaction time and accuracy was tested: reaction time (or accuracy) = constant + visual line complexity + imageability (vividness) + concept familiarity + Kučera-Francis word frequency. The R^2 values were .56 and .51 for reaction time and accuracy, respectively. For both dependent measures, only vivid-

ness proved to account significantly for the variability ($t = 2.97$ and -2.52 , $p < .01$ and $.03$, respectively).

In the second regression model, the vividness variable was replaced by the imaging difficulty variable generated by the subjects in Experiment 1. No doubt this variable is contaminated by others. (We note, for example, that it was significantly correlated with concept familiarity.) Nevertheless, imaging difficulty was strongly correlated with vividness, supporting its validity, and with each dependent measure. Thus, we reevaluated the four-predictor model above, substituting rated imaging difficulty for vividness. The resulting R^2 rose to .89 and .84 for models of reaction time and accuracy, respectively. Similar to the previous results, imaging difficulty was the only factor that contributed significantly to the variance ($T_s = 8.57$ and -6.98 , $p_s = .0001$).

The fact that imageability (as measured by vividness or image difficulty) was the only significant factor in both models provides further evidence that a visual translation stage is important in the haptic recognition of raised two-dimensional pictures of common objects. In the next experiment, we test the visual image-mediation model in a different manner, by having congenitally blind subjects perform the task that the sighted, blindfolded subjects performed in Experiment 1. For those blind from birth, an imagery process like that of sighted individuals is unlikely to be available (although spatial components of imagery may be available to the blind; see, e.g., Kerr, 1983). The convention of portraying three-dimensional objects through projection to a picture plane is likely to be unnatural for those who learn about such objects primarily through three-dimensional haptic exploration, with free manipulation of the object's orientation in space. To the extent that familiarity with pictorial depictions and the ability to visually imagine such depictions are necessary for recognition in this task, we would therefore expect congenitally blind individuals to perform even more poorly than blindfolded, sighted subjects do.

EXPERIMENT 3 Haptic Recognition of Raised-Line Drawings of Common Objects by Congenitally Blind Observers

Method

In Experiment 3, a total of 7 congenitally blind adults (3 males and 4 females; no subject had more than light perception) from the Canadian National Institute for the Blind in Toronto participated. All subjects had at least high school education. The subjects attempted to identify the 22 stimulus items used in Experiment 1. The experimental procedure was also identical to that used in Experiment 1, with two exceptions. The participants were not asked to complete the VVIQ, since the questions were clearly inappropriate. For the same reason, they did not provide imaging difficulty ratings. Because the blind subjects had no idea about the identity of most of the line drawings, the experimenter stopped the timer when subjects "gave up," rather than continuing until the end of the 2-min period. Hence, the reaction times are not directly comparable to those from Experiment 1.

Results

To permit comparison with the results of the sighted group, 7 normally sighted subjects from the previous experiment were randomly chosen. The data from these subjects were used in a t test of independent means for comparison with the blind. The mean accuracy (standard deviation) for the sighted group was 6.00 (2.45), compared with 2.29 (2.43) for the congenitally blind. This difference is statistically significant [$t(12) = 2.85$, $p < .01$]. Thus, the sighted were more accurate than the congenitally blind (means of 27.3% and 10.4%, respectively) in haptically recognizing common objects drawn in raised outline form. Performance by the blind was profoundly poor in an additional sense. When correct answers were provided at the end of the experiment to cue the previously unrecognized items, the blind were unable to make use of this feedback to subsequently recognize the object portrayed in the drawing in a second attempt.

Unlike the sighted (Experiment 1), the blind subjects showed little difference in the accuracy with which they recognized two- as opposed to three-dimensional representations. Of the 11 two-dimensional drawings, only 6 were recognized correctly by 1 or more of the 7 subjects; in fact, there were only 8 correct responses out of a total of 77. Of the 11 three-dimensional drawings, 4 were recognized correctly by 1 or more subjects, and again, 8 out of 77 responses were correct. The trend is in the expected direction.

Intercorrelations among scores for reaction times and accuracy were not as high as for the sighted subjects, but they were in the same direction. Thus, the Pearson product-moment correlations between reaction time and accuracy (0 = correct, 1 = incorrect) were $-.91$ and $-.42$ for the sighted and blind, respectively ($p_s < .001$ and $.05$). This is not surprising, since the blind usually stopped early, believing that continued exploration would be of no further assistance. Furthermore, the correlations were based, respectively, on a considerably different sample sizes (24 and 7 for Experiments 1 and 3, respectively).

The current results further support the use of a visual translation model by the sighted in this two-dimensional haptic recognition task. Without visual image mediation, as indicated by the congenitally blind performance, the effectiveness of the haptic system was even more severely curtailed. The blind subjects were forced to depend on the very limited nonvisual spatial cues available in these displays, the value of which was further constrained by sensory integration and/or memory demands.

GENERAL DISCUSSION

The results of this study highlight the remarkable contrast between two- and three-dimensional haptic object recognition. The recognition of real common objects approaches 100%, and is most frequently performed in just 1–2 sec (Klatzky et al., 1985). Yet, despite the fact that the subjects were highly familiar with the objects depicted

in the current study, the two-dimensional recognition rate was only 33% for blindfolded, sighted observers (Experiment 1), and 10% for the congenitally blind (Experiment 3). The sighted subjects also took (on the average) as long as 90 sec to perform the task. (We were unable to consider corresponding reaction times for the blind subjects, for they usually chose to discontinue the trial early, when they felt that recognition was impossible.) Performance by the sighted subjects was generally higher than that obtained by Magee and Kennedy (1980). This may be due to the fact that all objects depicted in the current study are not only highly conceptually familiar, but also capable of being manipulated in the hand. However, performance is still very poor, especially when compared to the effectiveness with which real common objects are haptically identified.

Can such poor performance be attributed simply to inexperience with tangible graphics? After all, the sighted usually read such displays by eye, not by hand. Doubtless, substantial learning in this task can be demonstrated. More importantly, however, we believe such poor performance reflects fundamental constraints that many two-dimensional tasks such as this one impose on the haptic system. Data just recently collected (Loomis, Klatzky, & Lederman, 1989) support our interpretation. We compared visual and haptic recognition of the raised-line pictures used in the current study. In a haptic condition, subjects freely explored the raised-line displays with a single finger. However, in a corresponding visual condition, subjects explored an equivalent computer-generated pattern by moving a pressure-sensitive pen over a graphics tablet; the image was stored internally and was exposed in real time in a constant location on a visual monitor as the pen contacted the corresponding coordinates on the tablet. The visual output was blurred as much as was required to equate the acuity of the two perceptual systems—Loomis (1981) has argued that passive touch functions much like a low-resolution vision system. In an attempt to further match exploration conditions, the size of the visual “window” also approximated that of the exploring finger. The results of this experiment indicated that visual recognition was not only identical to haptic recognition but equally poor. The performance of both groups resembled that of the blindfolded subjects in Experiment 1. These data clearly argue against the hypothesis that inexperience with tangible graphics displays underlies poor two-dimensional haptic recognition, since the visual system too performed more poorly when highly familiar objects were depicted in sequence through a reduced viewing aperture.

What, then, are the factors that typically limit haptic performance of this two-dimensional task? And why, in contrast, is the haptic system so effective at recognizing natural common objects? Let us first consider the three-dimensional task. Once again, we emphasize the multidimensional nature of real objects. A pencil varies from an ashtray in many different ways. The pencil is long and thin. It may have a compliant part at one end, which often lies adjacent to a harder and cooler metal section. The

other end may be pointed if the pencil has been used, or flat if it has never been used before. Finally, it is relatively light and warm, except on the metal part. In comparison, an ashtray is heavier, colder, harder, larger, and bulkier in shape. It also has distinguishing slots to hold cigarettes, and so forth.

In previous work, we (Klatzky & Lederman, 1987; Lederman & Klatzky, 1987) have shown the importance of different classes of purposive hand movements for the extraction of various object dimensions, as was noted earlier by Lotz (1856/1885), Katz (1925), and Gibson (1966). We have described nine different categories of object knowledge, pertaining to an object's substance, structure, function, and motion of a part (or parts). Each category is typically associated with one or two movement classes, which we call “exploratory procedures.” For example, textural properties are usually extracted by relative motion between skin and surface (*lateral motion* procedure). Additional experimental results indicated that the procedure most closely associated with a designated category of object knowledge was the most efficient means, in terms of accuracy and speed, of acquiring such information. Procedures also varied in terms of their relative specialization. That is, some procedures were highly specialized to extract information about one or two dimensions, whereas others provided low-level information about almost all of the different attribute categories simultaneously. Our most recent work (Lederman & Klatzky, 1989) documents how people effectively use knowledge of highly diagnostic object properties to guide the selection and sequence of exploratory procedures during the haptic classification of real objects.

Other research in our laboratory (Klatzky, Lederman, & Reed, 1987) has emphasized the particular importance to the haptic system of substance dimensions and their corresponding procedures. We have demonstrated that when subjects freely sort multidimensional objects into similar piles haptically without vision, they choose attributes pertaining to substance (hardness and texture) over those relating to planar structure (shape and size) as the primary basis for their sorts. And just recently, we reported that subjects integrate texture and hardness cues more than they do either one of them with shape, when the information from the different dimensions is made perceptually redundant for purposes of object classification (Klatzky, Lederman, & Reed, 1989). In both studies, the choice of exploratory procedures clearly reflected these important characteristics of haptic processing. In interpreting these results, we noted that local substance (e.g., texture and hardness) properties of homogeneous objects may be extracted quickly and accurately, using their associated procedures. In contrast, more global structural information about planar shape and size may be extracted quickly and imprecisely with an *enclosure* procedure (molding fingers to the object's contours), or very slowly and again relatively inaccurately by a *contour following* procedure (edge following).

In summary, subjects explore multidimensional objects purposively. They direct their hand movements in a highly

efficient manner for extracting information relevant to the designated task. Further, haptics appears to be particularly well suited to extracting information pertaining to substance (see also Gibson, 1966; Katz, 1925). Elsewhere (Klatzky & Lederman, 1987; Lederman & Klatzky, *in press*), we have argued that this research supports the use of a direct apprehension model of haptics in the recognition of common objects. In this model, no visual translation stage is postulated. Rather, it is assumed that the haptic system uses its cutaneous, kinesthetic, and thermal apparatus to process and represent information independently of vision, at least during the early stages.

Now let us consider the corresponding two-dimensional, object recognition task. Gibson (1966) has argued that in creating pictures, the artist structures light in the same way that it is structured in the environment, and that this structured light carries information. Since only variables in the optic array are sensed, by fiat the haptic system must be denied access to the medium of pictures. Certainly, the results of the previous and current studies are not very encouraging. However, our goal was to specifically consider the demands that such a task imposes on the haptic system, and further, how it is performed, in order to learn more about the particular limitations of this perceptual system.

In the current task, only structural information about planar shape and size was available, and it could only be extracted by a contour following procedure. However, the latter is decidedly slower than any of the other procedures. Consequently, it is subject to difficulties inherent in the integration of sequentially presented spatial information, and in addition, to considerable memory constraints. In keeping with much previous thinking about the limitations on the haptic system (e.g., Berla', 1982), we argue that the extent to which the haptic system is forced to extract information sequentially is one critical factor that impairs the performance of many other two-dimensional spatial tasks as well (perception of distance, direction, and position, and the recognition of planar nonsense shapes that are made of a homogeneous material). (We note that the detrimental effects of sequential presentation are not limited to haptic processing either; recall, for example, the poor performance of the visual group in the Loomis et al., 1989, study.) Normally, however, the entire visual array is available in a single view.

We had anticipated that the blindfolded sighted subjects would adopt a visual imaging process to haptically recognize raised-line drawings of common objects. We argue that what subjects are given is only identifiable by extracting local contour segments and if possible integrating them, and by comparing the resulting representation to what is known about the projections from three-dimensional objects. It is likely that this can be done most effectively by using visual imagery as a medium of representation. A number of results provide converging support for this prediction, as follows:

First, both measures of imageability—imaging difficulty (Experiment 1) and the independently obtained image-

vividness rating (Experiment 2)—were strongly correlated with both measures of performance—reaction time and accuracy. They were strongly intercorrelated as well. When either was used in a multiple regression model of the performance data of Experiment 1, it was the single significant predictor of haptic recognition performance.

Second, the congenitally blind performed this spatial task even more poorly than the normally sighted did, which is in keeping with previous research (McLinden, 1988). Although there may be many reasons for this result, one that is consistent with our assumptions is that unlike the sighted, the blind had no access to imagery as a medium for representing and interpreting the stimuli.³ The limited number of object properties portrayed in the displays constrained the nature of purposive haptic exploration and essentially eliminated other strategies for task performance. Under the circumstances, the poorer performance by the blind is readily understood.

Third, the raised lines in our displays were sometimes used to depict an object's three-dimensional structure. How are visual pictorial cues such as linear perspective, occlusion, and texture gradients interpreted haptically by the normally sighted? By visually imagining the haptic inputs, they may have attempted to derive information about the third dimension. We would expect three-dimensional drawings to be recognized more slowly and less accurately than two-dimensional drawings, since additional processing is required. This prediction was confirmed. Furthermore, the former displays were judged to be more difficult to image, despite the fact that the line displays were perceived to be equally complex. In contrast, there was no difference in two- versus three-dimensional performance by the congenitally blind, as we further predicted, since they cannot use visual imagery to interpret the third dimension. In keeping with this interpretation, Wake, Shimuzu, and Wake (1980) have noted that relative to sighted, blindfolded subjects, the congenitally blind tended to interpret pictorial cues in the picture plane, assuming that interior lines represented texture rather than protruding or receding edges.

Fourth, subjects who had a high ability to image (as measured by the VVIQ test) tended to take less time to respond and to be more accurate. Although the correlations were not exceptionally high, the trends were in the expected direction, and they should be considered in conjunction with the other converging evidence.

To the extent that haptically accessed information in raised pictures can be visually imaged (Kennedy, 1982, has argued that some tangible pictures are comprehensible to the haptic system; see below), this mode of processing is potentially available to the sighted. In the current task, nevertheless, the power of visual imagery was necessarily limited by the sequential nature in which the haptic information was accessed. This same limitation is also likely imposed on many other two-dimensional tasks that present limited spatial cues and little else to the haptic system. Thus, whereas visual imaging presumably improved haptic picture recognition by the sighted, relative to the

congenitally blind, performance was still very limited in comparison with the haptic recognition of three-dimensional common objects. We have argued that in the latter circumstance, the haptic system can capitalize on its strengths by using direct haptic apprehension, by means of which many different object attributes can be extracted quickly and accurately, without the need for visual imagery.

The current results also offer some insights into the design of graphics displays (e.g., maps, graphs, and pictures) for the blind. Designers should be aware that much of the information that makes haptics effective (see previous discussion) is either misrepresented or missing from current displays. For example, extending Gibson's (1966) argument above, contemporary displays often contain cues that are inappropriate for the haptic system. The lines in the displays are intended to represent edges, protruding and receding corners, and surface layout in two or three dimensions. However, the pressure distributions created by touching these lines do not differ, in contrast to the distinctly different pressure patterns produced by touching these features in the natural world. To the extent that these properties can be extracted from a tangible picture, the process is necessarily indirect, consciously intellectualized, and very slow.

One relatively minor recommendation for designing raised pictures would be to incorporate diagnostic substance differences (e.g., texture, thermal properties, hardness) whenever possible. In this respect, the technique of attaching different materials to the surface of a display may offer a wider range of values than does working solely in a single plastic medium, as is common today. Often, however, this is not feasible. One must then carefully weigh the advantages and disadvantages of providing only spatial information to the haptic system against the decision to present such information in an alternate, nongraphic format (e.g., written or aural text). When only crude spatial information is being presented, a graphic format can prove useful. For example, two-dimensional, raised-line graphs may be effectively employed to provide simple information about line functions (e.g., monotonically increasing) and to permit the extraction of point-coordinate information from grids (Barth, 1984; Lederman & Campbell, 1982). In addition, point symbols included on simple maps can be designed more effectively, so that their features inherently suggest the properties of the object (e.g., stairs) or concept (e.g., direction) that the symbol is intended to represent (e.g., see Lambert & Lederman, 1989; Schiff, Kaufer, & Mosak, 1966).

But when the blind user must fully configure the spatially and temporally extended haptic inputs, the displays may be of considerably less value. We are therefore somewhat less optimistic than Kennedy (1982) about the ability of congenitally blind individuals to effectively recognize raised-line drawings across a wide range of viewing conditions. Arguing against Gibson's (1966) complete rejection of haptic pictures, he noted several techniques that

the blind are capable of using haptically, without training, to understand depiction—that is, the low spatial frequency information in the outline shape, certain cues regarding perspective (such as occlusion), and cognitive metaphor. But even though the blind can understand the medium of depiction in principle, and though they are able to use it to produce their own drawings, they may still encounter considerable difficulty in effectively and reliably interpreting the idiosyncratic representations of others.

REFERENCES

- BARTH, J. (1984). Incised graphs: Enhancing the readability of tangible graphs for the blind. *Human Factors*, **26**, 61-70.
- BARTRAM, D. J. (1973). The effects of familiarity and practice on naming pictures of objects. *Memory & Cognition*, **1**, 101-105.
- BERLA, E. (1982). Haptic perception of tangible graphics displays. In W. Schiff & E. Foulke (Eds.), *Tactual perception: A sourcebook* (pp. 364-387). New York: Cambridge University Press.
- GIBSON, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.
- JONES, B. (1981). The developmental significance of cross-modal matching. In R. Walk & H. Pick, Jr. (Eds.), *Intersensory perception and sensory integration* (pp. 109-136). New York: Plenum.
- KATZ, D. (1925). Der Aufbau der Taswelt. *Zeitschrift für Psychologie*, **11**, 1-270.
- KENNEDY, J. (1982). Haptic pictures. In W. Schiff & E. Foulke (Eds.), *Tactual perception: A sourcebook* (pp. 305-333). New York: Cambridge University Press.
- KERR, N. (1983). The role of vision in "visual imagery" experiments: Evidence from the congenitally blind. *Journal of Experimental Psychology: General*, **112**, 265-277.
- KLATZKY, R. L., & LEDERMAN, S. J. (1987). The intelligent hand. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 21, pp. 121-151). New York: Academic Press.
- KLATZKY, R. L., LEDERMAN, S. J., & METZGER, V. A. (1985). Identifying objects by touch: An "expert system." *Perception & Psychophysics*, **37**, 299-302.
- KLATZKY, R. L., LEDERMAN, S. J., & REED, C. (1987). There's more to touch than meets the eye. *Journal of Experimental Psychology: General*, **116**, 356-369.
- KLATZKY, R. L., LEDERMAN, S. J., & REED, C. (1989). Haptic integration of object properties: Texture, hardness, and planar contour. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 45-57.
- KUČERA, H., & FRANCIS, N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- LAMBERT, L., & LEDERMAN, S. J. (1989). An evaluation of the legibility and meaningfulness of potential map symbols. *Journal of Visual Impairment & Blindness*, **83**, 397-403.
- LEDERMAN, S. J., & CAMPBELL, J. (1982). Tangible graphics for the blind. *Human Factors*, **13**, 18-21.
- LEDERMAN, S. J., & KLATZKY, R. L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology*, **19**, 342-368.
- LEDERMAN, S. J., & KLATZKY, R. L. (1989). *Haptic classification of common objects: Knowledge-driven exploration*. Manuscript submitted for publication.
- LEDERMAN, S. J., & KLATZKY, R. L. (in press). Haptic exploration and object representation. In M. Goodale (Ed.), *Vision and action: The control of grasping*. Norwood, NJ: Ablex.
- LEDERMAN, S. J., KLATZKY, R. L., & BARBER, P. (1985). Spatial and movement-based heuristics for encoding pattern information through touch. *Journal of Experimental Psychology: General*, **114**, 33-49.
- LEDERMAN, S. J., KLATZKY, R. L., COLLINS, A., & WARDELL, J. (1987). Exploring environments by hand or foot: Time-based heuristics for encoding distance in movement space. *Journal of Experimental Psychology: General*, **113**, 606-614.

- LEDERMAN, S. J., & TAYLOR, M. M. (1969). Perception of interpolated position and orientation by vision and active touch. *Perception & Psychophysics*, **6**, 153-159.
- LOOMIS, J. M. (1981). On the tangibility of letters and braille. *Perception & Psychophysics*, **29**, 37-46.
- LOOMIS, J. M., KLATZKY, R. L., & LEDERMAN, S. J. (1989). *Equivalence of tactual and visual picture recognition with limited field of view*. Manuscript submitted for publication.
- LOOMIS, J. M., & LEDERMAN, S. J. (1986). Tactual perception. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of human perception and performance* (chap. 31, pp. 1-41). New York: Wiley.
- LOTZ, H. (1885). *Mikrokosmos*, (Vol. 1; E. Hamilton & E. E. C. Jones, Trans.). Edinburgh: T. T. Clark. (Original work published 1856)
- MAGEE, L., & KENNEDY, J. (1980). Exploring pictures tactually. *Nature*, **283**, 287-288.
- MARKS, D. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, **64**, 17-24.
- McLINDEN, D. (1988). Spatial task performance: A meta-analysis. *Journal of Visual Impairment & Performance*, **82**, 231-236.
- ROSCH, E., MERVIS, C., GRAY, W., JOHNSON, D., & BOYES-BRAEM, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, **8**, 382-439.
- SCHIFF, W., & FOULKE, E. (1982). *Tactual perception: A sourcebook*. New York: Cambridge University Press.
- SCHIFF, W., KAUFER, L., & MOSAK, S. (1966). Informative tactile stimuli in the perception of direction. *Perceptual & Motor Skills*, **23**, 1315-1335.
- SNODGRASS, J. G., & VANDERWART, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, **6**, 174-215.
- WAKE, T., SHIMUZU, Y., & WAKE, H. (1980). Perception of tactile three-dimensional information and visual aids for blind persons. *Japanese Journal of Ergonomics*, **16**, 27-36.
- WONG, T. (1977). Dynamic properties of radial and tangential movements as determinants of the haptic horizontal-vertical illusion with an L-figure. *Journal of Experimental Psychology: Human Perception & Performance*, **3**, 151-164.

NOTES

1. Experiment 2 was actually performed subsequent to Experiment 3 and to an additional experiment not reported here, to obtain an independent assessment of item imageability for all of the pictures used by sighted and blind subjects. We have chosen to present the results out of chronological order, so as to maintain the continuity of our theoretical argument.

2. Vividness ratings were obtained for 10 of the drawings exactly as they were presented in Experiments 1 and 3. Depictions of the remaining 11 objects (excluding the image of the book) in those experiments were presented, along with 9 additional objects, in both prototypical and distorted (either enlarged or reduced) sizes. No subject judged more than one size per pair. These 20 pairs of object depictions were used in an additional unreported experiment. Here, we report only the results pertaining to objects used in the current study. Where the vividness of an object depiction was evaluated in two sizes, the mean of the two numeric ratings is reported.

3. We thank Mort Heller for noting that in this study, item familiarity was based on visual, as opposed to haptic, assessments (by either sighted or blind subjects). It might be argued, therefore, that the inferior performance by the congenitally blind group occurred as a result of their relative inexperience with some of the objects in our stimulus pool. However, the empirical data do not support this interpretation. The hammer was identified most frequently (by 5 of 7 blind subjects), although it is likely to be used considerably less often than a cup, key, plug, or envelope, none of which was ever correctly identified. In addition, the bowl was only identified by 1 of 7 blind subjects.

(Manuscript received February 14, 1989;
revision accepted for publication July 19, 1989.)