Seeing with the skin

BENJAMIN W. WHITE,[2] FRANK A. SAUNDERS, LAWRENCE SCADDEN, PAUL BACH-Y-RITA, AND CARTER C. COLLINS
PACIFIC MEDICAL CENTER

A system for converting an optical image into a tactile display has been evaluated to see what promise it has as a visual substitution system. After surprisingly little training, Ss are able to recognize common objects and to describe their arrangement in three-dimensional space. When given control of the sensing and imaging device, a television camera, Ss quickly achieve external subjective localization of the percepts. Limitations of the system thus far appear to be more a function of display resolution than limitations of the skin as a receptor surface. The acquisition of skill with the device has been remarkably similar for blind and sighted Ss.

Twelve years have passed since the publication of Geldard's (1957) provocative paper in which the potentialities of the skin as a communication channel were demonstrated. Since that time, Geldard and his colleagues have refined their technique of presenting highly discriminable vibrotactile stimuli to various parts of the body. In a recent report, Geldard (1968) describes the application of nine vibrators at widely separated body sites, each one triggered by one photocell in a linear array scanning typed and printed characters. It is hoped that this system will enable trained Ss to read at speeds comparable to those that Geldard reported earlier, using a set of five vibrators mounted on the chest and capable of delivering signals at three intensities and three durations. After a few hours of training with this sytem, Ss were able to receive material at a rate approximately

twice that of proficient Morse code operators.

Geldard's efforts have been oriented exclusively toward the use of the skin to receive a set of known and clearly discriminable characters, whose optimal coding has been determined by psychophysical techniques. The possibility of using the skin as a channel for *pictorial* material has not been systematically explored, though devices capable of presenting dynamic two-dimensional tactile images exist (Linville & Bliss, 1968, Strakiewicz & Kulizewski, 1965). The potential utility of such devices as visual substitution systems for the blind is immediately apparent. It is surprising that in this day of advanced technology, the blind are still moving about in the world using a cane, a guide dog, a sighted companion, or an outstretched hand.

In addition to its value as an aid for the blind, such a tactile image system provides an opportunity to explore a number of such perceptual phenomena as size constancy or space perception that heretofore have been considered uniquely visual. Also, the problems people encounter in gaining facility with this novel system should throw light on aspects of perceptual learning that have thus far been difficult to investigate.

For the past year, we have been working with a sensory substitution system that converts a visual image into a tactile one. It is the purpose of this report to present some of the initial findings from this work and to discuss their implications for perceptual theory.

METHODS

The theoretical neurophysiological basis (Bach-y-Rita, 1967) and the physical concept of the instrumentation (Collins, 1967) for the vision substitution system have been discussed previously. Preliminary results in training blind Ss to use the apparatus have been briefly reported (Bach-y-Rita et al, 1969a; Bach-y-Rita et al, 1969b; Scadden, 1969; Saunders, 1969).

As shown in Fig. 1, the "eye" of the system consists of a television camera. This camera, which is mounted on a tripod, is manipulated by the S, who can aim the zoom lens at any part of the room, in order to localize and identify objects or persons. Stimuli can also be presented on a back-lit screen by slide or motion picture projection. The video image is electronically transformed and sent to a 20 by 20 matrix of solenoid vibrators mounted in the back of a stationary dental chair. The 400 stimulators, spaced 12 mm apart with 1-mm-diam tips, cover an area approximately 10 in. square. Each solenoid is designed to vibrate at 60 Hz when its locus is within an illuminated region of the camera field. The on-off activity of the vibrators can be monitored visually on an oscilloscope as a two-dimensional pictorial display.

The Ss for this series of experiments were young adults, many of whom were from a nearby college. Twenty-five congenitally blind men and women have been tested in the apparatus, in addition to five adults who were blinded later in childhood. Over 50 sighted Ss have also been examined. Not all of these Ss have been put through all the procedures described below. Eight blind Ss have had over 40 h of training in the chair, while others have come in to serve as Ss for a single experiment.

RESULTS

Subjects are able to perceive certain simple displays with this tactile system, almost as soon as they have been introduced to it. If a white circular target, capable of activating approximately three-quarters of the tactors, is placed before the camera, Ss have no difficulty in centering the image on the tactor array by manipulating the camera. If a vertical white stripe is moved from side to side, the S is able to describe the motion
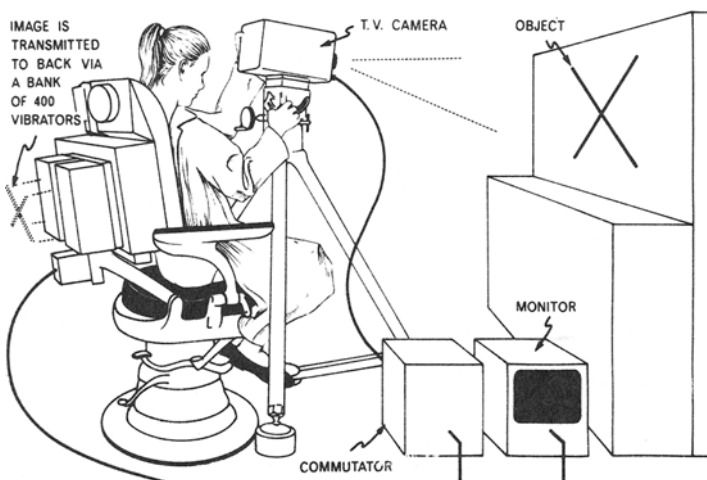

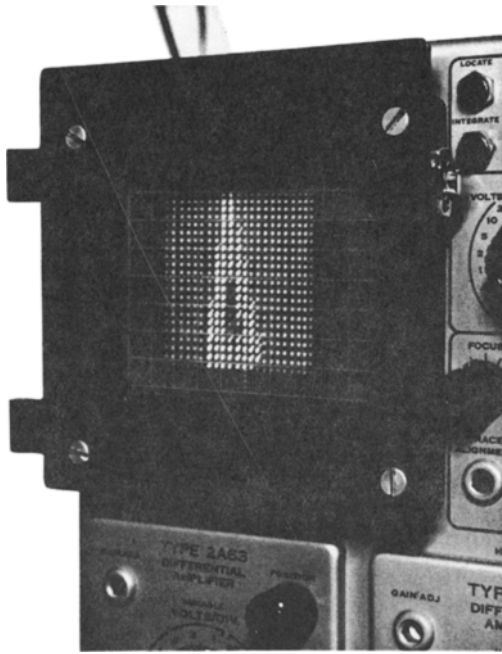Fig. 1. Schematic representation of the tactile television system.

not from the tactile input, but from the visual display seen on the oscilloscope monitor. Figure 2 shows the kind of visual display the sighted Ss were given.

The two groups showed remarkably similar performance. There was only one incorrect judgment out of a total of 480. The blind Ss had a mean latency to correct response of 1.2 sec as compared to 1.1 sec for the sighted group Though both groups were instructed in the control of the camera and encouraged to move it before making a judgment, neither group did so to any great extent.

Slant Judgments

The same groups of blind and sighted Ss were asked to make slant judgments based again upon slide-projected displays of a checkerboard that had been photographed at a 70-deg slant from frontal parallel. This slide was projected in one of four orientations by simply rotating the slide in the projector. On 40 such randomized presentations, the S was asked to judge whether the checkerboard tilted away to the top, bottom, left, or right. The results for this task are shown in Table 1. The blind Ss using the tactile display made significantly more errors and took significantly longer to make their judgments than the sighted Ss using the visual display. Unlike the vertical-horizontal bar orientation task, the blind Ss all used scanning of the camera to make the slant judgment. Since this scanning had to be done by turning two wheels on the tripod, this probably accounts for much of the latency difference. The accuracy of the experienced blind Ss is high on this task and the accuracy of the sighted Ss is not perfect. This would indicate that the limitations on the tactile system thus far probably are more a function of the resolution of the display than of the sensitivity of the back.

Observations of Experienced Blind Ss

Seven congenitally blind Ss have each had over 40 h of experience in the vision substitution system. These men and women have been studied intensively because they proved adept with the apparatus and were

accurately or to imitate it with an appropriate hand gesture. Similarly, a propeller motion of the stripe can be accurately followed or described. Ss can also discriminate between a stationary vertical or horizontal stripe. When given control of the wheels controlling camera position on the tripod, the horizontal-vertical discrimination becomes even easier since only motion orthogonal to stripe orientation produces a detectable change in the tactile display. If a diagonal stripe is presented, Ss are able to report whether the upper end is on the left or the right. They are also able to adjust the camera so that one end of the line is centered. Ss are able to report the orientation of a curved line (a 60-deg arc with a 4-deg radius of curvature) with reasonable accuracy.

On none of these tasks was the performance of the blind Ss significantly different from that of the sighted Ss who were, of course, blindfolded while being tested.

Form Discrimination

Initial experiments with form discrimination showed that performance rapidly improved when Ss were allowed to scan the figures by moving the camera and were given immediate correction after a wrong response. When asked to identify a circle, square, or triangle, Ss' performance remained near chance levels after 60 trials when no correction was given and no camera movement was allowed. With correction, accuracy reached 60% after 54 trials, with a

mean response latency of 6 sec. Significant improvement was found between each 18-trial block. When the Ss were allowed to pan the camera over the forms and given correction, they achieved 100% accuracy in the third block of 18 trials with a latency of 1 sec or less (Bach-y-Rita et al, 1969). The discrimination was even more rapidly established when the figures were initially presented in pairs for the S to inspect after telling him, "The square is on the right and the circle on the left." After as little as 10 min of such training, some Ss could achieve 100% accuracy in identifying these three simple forms. Particularly in early stages of training, Ss panned the camera so that the figure came into and passed out of the field, suggesting that the discrimination is largely based upon contour changes in the leading edge of the figure.

Acuity Judgments

Six blind Ss were asked to judge the vertical-horizontal orientation of slide-projected displays consisting of from 4 to 12 pairs of parallel black and white lines. The five displays were presented in random order for a total of 40 presentations, and the Ss were allowed to scan them with the camera before making a judgment. Five of these Ss were congenitally blind, and the sixth lost sight before the age of 4. They had had from 15 to 40 h of experience in the vision substitution system. The performance of the blind Ss using the tactile input was compared with the performance of six sighted Ss who made the same judgments,

Table 1
Comparison of Tactile
and Visual Slant Judgments

	Accuracy	Latency
Blind Os Tactile Display	83% (n = 240)	8.4 sec
Sighted Os Visual Display	97% (n = 240)	2.8 sec

able to provide valuable insights about its capacities and limitations.

These Ss were given extended practice in the identification of a collection of some 25 "things"—a coffee cup, a telephone, a stuffed animal. They were encouraged to scan these objects by manipulating the camera and to try to describe them. Initial scanning was prolonged, a new object often requiring up to 15 min of exploration before correct identification. This latency dropped steadily with repeated presentations until the Ss were often able to identify an object within 10 sec on its fifth presentation. A "learning to learn" phenomenon was evident, since new objects took less and less time to identify as the vocabulary of objects increased.

The experienced Ss, after they had gained facility in identifying a number of objects, were asked to describe arbitrary arrangements of objects placed upon a table top. The table was placed so that the camera looked down at it from an angle of approximately 20 deg off horizontal. The Ss were able both to identify the objects on the table top and to describe their arrangement even though their placement on occasion was such that objects at the rear were partially occluded by those in front.

Obviously judgments of this sort must make use of the sort of information that in classical texts is called "cues for distance." Particularly useful in making judgments about the position of the object in depth was its vertical position on the display. The higher up it was, the farther back on the table top. This is precisely the information that Roberts (1963) found so useful in his program enabling a computer to construct a three-dimensional representation of an object using only the two-dimensional representation of it as input.

The congenitally blind person has never directly experienced the relationship between the visual angle subtended by an object and its distance from the observer. One of the experienced blind Ss was a psychologist who had explained this relationship to introductory psychology classes for several years. One day, while in the chair, he experienced the change in size of a tactile image as an object was brought closer and closer to the camera, and the sudden perceptual realization of this size-distance relationship came as a genuine "aha" experience. This evidence of size constancy has been found in other Ss as well. Perhaps of greater interest have been the occasions when Ss have given startled ducks of the head when the tactile image was suddenly magnified by a sudden turn of the zoom lever on the camera.

One discrimination that some of the Ss found most difficult to achieve was that of internal detail. For such Ss, the detection of

facial features in photographs or folds in fabric followed by many hours the ability to recognize objects on the basis of outside contour alone. There appear, however, to be profound individual differences in this respect, and one S, on his first session in the chair, was able not only to detect an internal hole in an object, but to describe its shape accurately.

Further evidence of the existence of powerful three-dimensional organization of tactile information presented in this unusual manner is seen in the response some Ss have made spontaneously to a kinetic depth display. A modified version of the Metzger apparatus was used, consisting of a turntable on which two vertical white rods were mounted. This was rotated slowly before the camera and the Ss were asked to describe what they "saw." Some sighted Ss, upon first tactile presentation of this moving display, have spontaneously described it as moving in depth. Several blind Ss were given experience with a yoked pair of turntables. On one of these, an object was placed within view of the camera, while the S turned the other freely, experiencing the transformations that the object underwent with the rotation. After an hour's experience with this equipment, they could report accurately the eccentric placement of two and three objects on the turntable, and could also experience rotation in depth with the Metzger display.

DISCUSSION

This sytem was constructed in order to find out whether or not people could make sufficient use of it to warrant its serious consideration as a visual substitution system for the blind. The most striking feature of the initial results with the system is that Ss, blind and sighted, are able, after only relatively short training periods, to identify familiar objects and to describe their arrangement in depth. Such results are the basis for cautious optimism about the ultimate utility of a visual substitution system. Evidently even a crude 400-tactor system is capable of providing sufficient information to permit construction of what is usually called the visual world. The limits thus far encountered are attributable to the poverty of the display rather than to limitations in the capacity of the tactile perceptual system.

It will be important in subsequent models of the system to employ a lighter sensor than the present cumbersome television camera, a sensor that can be head-mounted so the S can scan his environment in a more congenial fashion. It will also be important to explore new factor configurations. Questions of optimal spacing of tactors and the resolution required in order to perform various tasks will require extensive

investigation. Recent experiments in detection of small inner detail indicate that turning off half the tactors in the array makes surprisingly little difference. A second version of the system, presently under construction, will embody both a light head-mounted camera and a tactor array built in a wheel chair, thus giving the user a modest degree of mobility and making it possible to evaluate the system under a much wider range of conditions than has been possible thus far.

It is clear from these first tests with the visual substitution system that three-dimensional organization of the information in the dynamic tactile array is easily achieved. The demonstrations of the kinetic depth effect on the skin, and the instances of startle response to tactile looming are clear examples of such organization in this modality.

In addition to the three-dimensional interpretation of motion in tactile displays, the results thus far point to the great importance of self-generated motions on the part of the observer. When asked to identify static forms with camera fixed, Ss have a very difficult time; but when they are free to turn the camera to explore the figures, the discrimination is quickly established. With fixed camera, Ss report experiences in terms of feelings on their backs, but when they move the camera over the displays, they give reports in terms of externally localized objects in front of them. The camera motion here is analogous to eye movements in vision and this finding raises the interesting possibility that external localization of percepts may depend critically upon such movements.

It is at least a plausible hypothesis that a translation of the input that is precisely correlated with self-generated movement of the sensor is the necessary and sufficient condition for the experienced phenomena to be attributed to a stable outside world. The converse of this hypothesis should also hold—that a lack of correspondence between a translation of the input array and self-generated sensor movements should result in experiences that are attributed either to nonrigid conditions in the external world, or to phenomena that have their perceived origin within the observer. It would be of some interest to examine protocols from stabilized retinal-image experiments. Such a stabilization eliminates all correlation between eye movement and translations of the optic array. The finding that stabilized images fade has usually been interpreted to mean that fine eye movements are necessary in order to generate stable visual forms. It is possible that the stabilized retinal-image results may be due to the fact that the percepts arising under such conditions of stimulation are
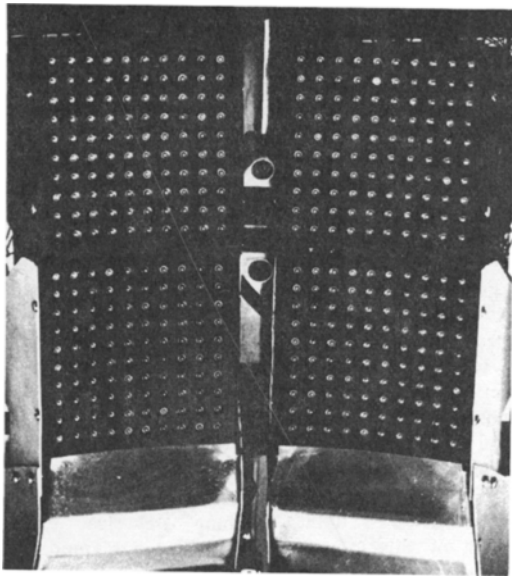
Fig. 3. The 20 x 20 tactor array.

localized within the observer rather than in the outside world. As von Békésy (1967) points out, we are adept at tuning out such internal information as the taste of our own saliva or the sound of the blood in our ears. Perhaps we are equally adept at tuning out visual phenomena that are uncorrelated with eye movements. Certainly afterimages are apt to disappear abruptly and then reappear several times before they fade for good, and the motes in our eyes are seen only under special circumstances.

Whether or not this hypothesis about the basis for external localization of percepts is valid, the fact remains that Ss with this visual substitution system were seriously handicapped when they could not move the camera. At first we erroneously assumed that little or no significant form discrimination was possible with fixed camera, but later experiments disproved this when experienced blind Ss were able to make quick and accurate judgments of grill orientation without moving the camera. Static judgments of orientation of the checkerboard patterns were rare, however. The experienced blind Ss found it relatively easy to determine the plane of the tilt, a judgment similar to the bar-orientation task, but had considerable difficulty telling which edge of the display was near or far. For this, they all employed many scanning movements of the camera. This suggests that there may be some aspects of tactile information pickup that depend heavily upon changes in the tactile array, and in this respect operate differently from vision.

The tactor matrix in this system was arranged in four quadrants, as shown in Fig. 3. With such an arrangement, the space between the center columns of tactors was at least twice as wide as the spacing of the other columns. Interestingly enough, this wide silent area was not evident to any of the Ss. If a narrow vertical stripe was moved horizontally from one side of the field to the other, its path was perceived perfectly smooth with no jump in the middle, even though the tactile image had travelled twice as far in the same time when it crossed the center of the tactor matrix. The fissure in the tactor array is not comparable to the blind spot in the eyes since in the latter case this is a gap in the sensor so that information is not received in this region, while in the tactor-array case there is no gap in the camera field, just an irregularity in the spacing of the tactors. Both, however, seem to be invisible. It would be of some interest to see how far apart the two halves of the tactor matrix could be placed on the body before a S would report a gap.

Thus far there have been no salient differences between the blind and the sighted Ss in their acquisition of skill on this visual substitution system. As was mentioned previously, the relationship between visual angle subtended by an object and distance from the S is one that the blind have not experienced directly and thus have to be introduced to as a "cue" for distance. Some blind Ss have never been taught the shapes of the letters of the alphabet, so in order to teach the discrimination of such forms they have to be given an opportunity to handle cutout letters. But on the whole, the similarities between the performance of the two groups are far more striking than the differences.

The visual substitution system offers a unique vehicle for the study of perceptual learning. If one accepts a Hebbian position on the manner in which visual form discrimination is established, one would be inclined to expect that proficiency with this tactile system would be tediously acquired, especially for congenitally blind Ss who have had no previous opportunity to create the cell assemblies and phase sequences that Hebb postulates as the basis for visual form discrimination. Another more recent line of evidence on the neurophysiological basis of certain visual discriminations in lower animals suggests that they may be dependent upon highly specialized and surprisingly peripheral feature detectors in the visual system (Hubel & Wiesel, 1962; Lettvin et al, 1959). If such detectors are also to be found in the human eye, it might be predicted that detection of such features from tactile inputs would be difficult, if not impossible, since such detectors in all probability would not be found in the skin.

The Gibsonian (1966) view of perceptual learning would probably predict that, to the extent a tactile array contained the same temporal and spatial adjacencies to be found in the optic array, learning to respond to "the higher order invariances" in tactile stimulation should present no overwhelming difficulty. Also, Gibson would expect performance with the system to improve markedly when the observer was free to probe the environment to pick up information. Gibson has repeatedly stressed the importance of such exploratory activity in perception.

The results to date with the visual substitution system support the conclusion that facility is quite rapidly attained and that some Ss are able to make highly sophisticated discriminations with it almost as soon as they are put in the chair. One completely naive S, for example, was able to report accurately the shape of a hollow tetrahedron, complete with shape of the space that formed an internal detail. The learning seen thus far is certainly not of the prolonged sort postulated by Hebb in initial visual form discrimination. Nor have there been marked differences between the congenitally blind and sighted Ss, which might also be expected by empiricists like Hebb or Taylor (1962). Also, the results do not lend support to a theory of visual form perception that postulates highly specialized feature detectors at the retinal level, though the drastically impoverished display in the present tactual system makes any conclusions in this regard only tentative.

In the past, many efforts at providing information to the blind have been based upon hopelessly old-fashioned ideas about the way the perceptual system works. In psychological prehistory there used to be a distinction between sensation and perception. The former had to do with stimulation of end organs that sent their

messages to the brain where they were synthesized and correlated through long experience until a percept emerged. Many efforts at creating sensory aids are still hung up on this antique notion and set out to provide a set of maximally discriminable sensations. With this approach, one almost immediately encounters the problem of overload—a sharp limitation in the rate at which the person can cope with the incoming information. It is the difference between landing an aircraft on the basis of a number of dials and pointers that provide readings on such things as airspeed, pitch, yaw, and roll, and landing a plane with a contact analog display. It is the difference between Skinner and John Holt. It is the difference between Titchener and Koffka. Visual perception thrives when it is flooded with information, when there is a whole page of prose before the eye, or a whole image of the environment; it falters when the input is diminished, when it is forced to read one word at a time, or when it must look at the world through a mailing tube. It would be rash to predict that the skin will be able to see all the things the eye can behold, but we would never have been able to say that it was possible to determine the identity and layout in three dimensions of a group of familiar objects if this system had been designed to deliver 400 maximally discriminable sensations to the skin. The perceptual systems of living organisms are the most remarkable information-reduction machines known. They are not seriously embarrassed in situations where an enormous proportion of the input must be filtered out or ignored, but they are invariably handicapped when the input is drastically curtailed or artificially encoded. Some of the controversy about the necessity of preprocessing sensory information stems from disappointment in the rates at which human beings can cope with discrete sensory events. It is possible that such evidence of overload reflects more an inappropriate display than a limitation of the perceiver. Certainly the limitations of this system are as yet more attributable to the poverty of the display than to taxing the information-handling capacities of the epidermis.

REFERENCES

BACH-y-RITA, P. Sensory plasticity: Applications to a vision substitution system. Acta Neurologica Scandinavica, 1967, 43, 417-426.
BACH-y-RITA, P., & COLLINS, C. C. Sensory plasticity and tactile image projection (abstract). Investigations in Ophthalmology, 1967, 6, 669.
BACH-y-RITA, P., COLLINS, C. C., WHITE, B. W., SAUNDERS, F., SCADDEN, L., & BLOMBERG, R. A tactile vision substitution system. American Journal of Optometry, 1969, 46, 109-111.
BACH-y-RITA, P., COLLINS, C. C., SAUNDERS, F., WHITE, B. W., & SCADDEN, L. Vision substitution by tactile image projection. Nature, 1969, 221, 963-964.
BÉKÉSY, G. von. Sensory inhibition. Princeton, N.J.: Princeton University Press, 1967.
COLLINS, C. C. Tactile image projection (abstract). National Symposium on Information Display, 1967, 8, 290.
GELDARD, F. A. Adventures in tactile literacy. American Psychologist, 1957, 12, 115-124.
GELDARD, F. A. Body English. Psychology Today, 1968, 2, 42-47.
GIBSON, J. J. The senses considered as perceptual systems. Boston: Houghton Mifflin Co., 1966.
HEBB, D. O. Organization of behavior. New York: Wiley, 1949.
HUBEL, D. H., & WIESEL, T. N. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. Journal of Physiology, 1962, 160, 106-154.
LETTVIN, J. Y., MATURANA, H. R., McCULLOCK, W. S., & PITTS, W. H. What the frog's eye tells the frog's brain. Proceedings of the Institute of Radio Engineers, 1959, 47, 1940-1951.
LINVILLE, J. G., & BLISS, J. C. A direct translation reading aid for the blind. Proceedings of the Institute of Electrical & Electronics Engineers, 1966, 54, 40-51.
ROBERTS, L. G. Machine perception of three-dimensional solids. Technical Report No. 315, 1963, Lincoln Laboratory, M.I.T., Cambridge, Massachusetts.
SAUNDERS, F. Tactile television: Characteristics of visuo-spatial information. Proceedings of the A.F.B. Tactile Display Conference, Palo Alto, Calif., April 1969.
SCADDEN, L. The reception of visual images through the skin. New Scientist, 1969, 41, 677-678.
STARKIEWICZ, W., & KULIZEWSKI, T. Progress report on the electroftalm mobility aid. Proceedings of the Rotterdam Mobility Research Conference. New York: American Foundation for the Blind.
TAYLOR, J. G. The behavioral basis of perception. New Haven, Conn.: Yale University Press, 1962.

NOTES