VISUAL PERCEPTION OF THE THIRD DIMENSION.

BY PROFESSOR CHAS. H. JUDD.
New York University.

Since the time of Berkeley the most important particular problem in the general discussion of space-perception has been the explanation of the visual perception of depth. The objective conditions on which such perception of depth depends may be isolated more easily than the conditions for any other form of space-perception. As a result, the analysis of the subjective phenomena is here very much easier than in other cases. The conclusions reached in regard to this particular form of perception are, however, applicable to the more general problem; for the visual perception of depth must be regarded as similar in kind to all other forms of space-perception. When the particular case is taken up for discussion, then, as it is to be in the present paper, the aim must be to reach general conclusions which shall bring the whole problem nearer to a final solution. And it will be possible in the course of such a discussion to test in this particular sphere some of the general theories that have been advanced.

First of all, then, let us raise the question, Does the presence of a third dimension in ordinary binocular and monocular fields of vision justify the doctrine that the sensation factors, which are the elements of the completed percepts, have an original attribute of extension, analogous to quality and intensity but differing from these? This doctrine has been advanced by a number of recent writers, and it must be granted that it seems a simple and direct hypothesis. It is true that we never observe a sensation except as part of a percept, and every visual percept has its projection into the third dimension. It follows from this that we can never observe visual sensations apart from some spacial relation. But while granting this, it is not necessary to conclude further that these spacial attributes are attributes of the sensa-

tion factors themselves rather than of the percepts in which they appear. Empirical evidence of an indirect character can be found which will aid in supporting this position.

If we take a thread of several meters in length and let it extend from the bridge of the nose to a distant point which is somewhat above or below the plane of the horizontal position of the optical axes, we shall observe double images, whatever point on the thread we fixate. If we fixate any point other than the extreme end of the thread we shall see a figure *X*. This *X* is due to the fact that the double images fuse at the two points of clear vision and do not fuse at other points. Now fixate some point along the thread such that the nearer and further halves of the *X* shall appear to be equal. The distance will vary according to the angle of elevation or depression of the thread, but it will always be found that the nearer half of the *X* is in reality very much shorter than the further half. I find, for example, that with a thread 330 cm. long and sloping at an angle of about 30° upward, the point of fixation is only 45 cm. from the bridge of the nose. Figure 1 represents the thread (*AB*) and the image (*ab*) in one eye. If now the point of fixation coincided with the real middle point (*G*) of the thread, it would follow that one part of the image (*ag*), while representing no more depth than the other part of the image (*gb*), would have greater extension on the retina. If the thread were fully recognized as a three-dimensional object, this relation of the parts of the image would also be recognized. As a matter of fact, under ordinary circumstances, the thread is regarded as projecting somewhat into depth. If, however, the associated factors that come from a familiar field filled with objects are removed, this projection into depth becomes less and less clear. In no case is the true relation between the parts of the image fully recognized. This failure to recognize the meanings of the two parts of the retinal image goes to show that the third dimension is, under the circumstances described, very incompletely recognized. This can

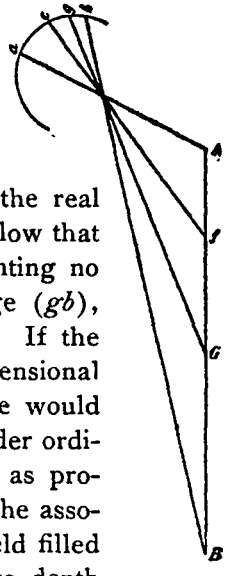

Fig 1.

be still further emphasized by noticing the real principle of division. If the retinal image, instead of the thread, is divided into two equal halves by the center of clearest vision, the point of fixation on the thread (f) will lie very much nearer A than it does to B. This is what actually takes place in the experiment described. If AB is taken as 330 cm. in length, Af will be 45 cm. But the principle that one-half the retinal image equals the other half holds only for two-dimensional interpretation. We have here then a series of facts which indicate that three-dimensional attributes are secondary and derived forms of perception rather than original attributes of sensation. The experiment is not direct or complete, but it shows that, if some of the factors of an ordinary percept are removed, so that we approach more nearly to the original sensation elements, we find the spacial attributes much less complete. The earlier interpretation is clearly two-dimensional, not three-dimensional.

Further evidence which goes to show that all the objects in the monocular field lie in a single plane, that is, are not distinguished as having different degrees of depth, may be found in an earlier article devoted to that special discussion.[1]

In none of the cases brought forward has it been possible to show that sensation factors do not have some three-dimensional attributes, but it has been shown that the complete perception of the third dimension is the product of complex rather than simple conditions. The simpler the conditions, the less complete the three-dimensional attributes. It must be insisted that this incompleteness of the third dimension is not merely a matter of measurement and subdivision of a perceived dimension. But while in some cases it expresses itself in the form of a judgment which does not correspond to objective measurement, it is always an incomplete perception. The measurement is not uncertain or vague to the mind of the observer, it is incorrect because it is essentially two-dimensional rather than three-dimensional in its principle.

This conclusion will be very much strengthened if it is possible to give an account of the conditions that explain the development of this third dimension as the product of complex relations, and it is to this task that we now turn our attention.

[1] An Optical Illusion, PSYCH. REVIEW, Vol. V., p. 286 seq.

The theory that space is a product of a complex of sensations is as old as Berkeley. Even before his time there were attempts of a rather crude type to explain space-perceptions as complexes. In all such attempts there has been a marked tendency to lay great stress on sensations of movement. Berkeley himself wrote: "This disposition or turn of the eyes is attended with a sensation which seems to me to be that which in this case brings the idea of greater or less distance into the mind."[1] The question of the relation of movements to visual perception of depth has been made the subject of elaborate experimentation. These experiments have dealt mainly with movements of accommodation. The earliest experiments were made by Wundt, and recently Hillebrand and Arrer have taken up the discussion and renewed many of the earlier experiments, making additions also in the way of methods and results.

The methods of these investigators may be subjected to careful examination. Wundt tried most of his experiments by allowing an observer to see with one eye a field of vision with a plain background and no objects other than threads which were suspended vertically in front of this background and were long enough so that they always reached from the top to the bottom of the visual field. The distance of the thread from the observer was varied, or else two threads were shown and the observer was called on to judge the relative distances. Hillebrand varied the method by using, instead of threads, mathematical lines which he was able to produce by taking the boundary between two surfaces as his object. In these latter experiments, as well as in Wundt's, the observer used only one eye, the other being covered or closed so that its retinal image was eliminated. In neither of these experiments, however, was it possible to eliminate movements of convergence. In spite of this fact, however, Wundt assumes that the convergence is not an important factor. Arrer tends more towards Hillebrand's position. Hillebrand is very explicit. He holds that the movements of convergence are in these cases the same as in normal vision, and that the results of the experiments apply, especially if they are negative results, to convergence as

[1] Essay toward a New Theory of Vision, § 16

well as to accommodation. The whole method is seriously complicated by the presence, in some degree at least, of movements of convergence in the closed eye. In *Science*, of February 25, 1898, I reported some experiments which showed that the movements of convergence which actually do take place are very complex. Until it can be definitely shown what are the results in sensation of such complex movements, it is evident that these movements cannot be ignored on the one hand or regarded as similar to those of ordinary experiences on the other. That the conditions of these experiments are exceedingly complex appears furthermore in the failure of the various investigators to agree on interpretations, even when their empirical data are the same.

The conditions for experimentation are very much better when we take up convergence rather than accommodation, and instead of trying to eliminate all other factors try rather to keep them constant. It can be shown in this way that movements of convergence are the sources of sensation which give rise to perceptions of depth.[1] But it is important to notice in this connection that, while it is shown that sensations of movement are factors in the perception of depth, it is not by any means shown that they are the only factors or even that they are the essential factors.

We turn to a criticism of the general type of theory which makes sensations of movement the essential factors of all space-perception. And, first of all, it may be argued that space is not identical with sensations of movement. This may be put in a still more general form: Space is not identical with any sensation quality. If we try to represent to ourselves absolutely empty space, we find that the nearest approximation to this which we are able to reach is a space which is only relatively devoid of content. It will always have some vague visual or muscle sensations, qualities in addition to its formal spacial attributes. Stumpf[2] has argued this point in detail, and we need only to accept his position as agreeing with the results of intro-

[1] 'Some Facts of Binocular Vision,' PSYCHOLOGICAL REVIEW, Vol. IV., p. 374.

[2] Ueber d. Psych. Urspr. d. Raumvorst., p. 114.

spection. Space, then, always requires some content of which it is merely the form or attribute. But it is an essential characteristic of sensation qualities that they can be isolated from each other and represented as distinct. In the necessary relation of space to some quality, we have, accordingly, evidence that space is not itself a quality but rather in some sense a secondary factor. This general argument applies to sensations of movement. Sensations of movement are characterized by certain qualities of strain, but these qualities are not space, they are contained in space.

This leads to a second point. Why should sensations of movement be any more directly the sources of space-perception than any other sensation qualities? This question cannot be answered without finding in sensations of movement some special characteristic that will differentiate them from other sensations, and this special spacial characteristic, as we may call the undiscovered factor, must be something other than quality, for, as we have just shown, the quality of muscle-sensation cannot be identified with space. A reference of space to sensations of movement is not then a final explanation. This fact has been overlooked too often. Writers have been satisfied if they could explain every case of space-perception by showing that the sensation complexes which entered into them contain some factors of movement. Movement in some form has been regarded as a necessity because it has been held that this alone could explain the spacial character of the whole complex. This has led to the assumption of movement-sensations in cases where it is difficult if not impossible to show their presence. Take, for example, cases of binocular vision in which the two eyes are not moved at all. We are capable of very fine discriminations of depth under such circumstances. But it is answered, there have been movements in the past experience, and the result is that the retinal sensations at any given moment are supplemented by a large number of revived associations. This hypothesis is necessary if the importance of muscle-sensations is to be maintained in the cases where there are certainly no immediate sensations of this type present. The difficulty with this hypothesis is that introspection reveals no such revived muscle-sensation. Take

stereoscopic vision, for example. Here the third dimension of a body must, on the movement-sensation hypothesis, be due to movements of convergence. In the original and crude form of the hypothesis, it was assumed that these movements actually took place. It is merely a more refined form of the same hypothesis to say that they are associated sensations from convergence called up by the retinal images. The preception of the third dimension is certainly vivid enough, and one would think that the sensation factors which are necessary to the formation of these percepts must also be clearly perceived in consciousness. But this is not the case if movement-sensations are in reality such elements. For not only are movement-sensations not revived as conscious factors, but the keenest introspective analysis cannot succeed in bringing them out. Some observers may not agree with the statement here made, holding that they are conscious of such ideas of change in convergence. The difference in opinion is frequently to be explained by the fact that movements in the objects are mistaken for movements in the eyes. A plain figure that may be seen as a cone or as a hollow funnel may be used as a convenient figure to illustrate what is meant. As the object is seen passing from one position to the other, there is a very striking impression of movement, but it is the movement of the object, not of convergence.

Another strong motive which has always supported theories based upon sensations of movements is to be found in the fact that there is no other factor common to all sensation complexes which yields ideas of space. This common factor has been either consciously or unconsciously demanded in all explanations of space, because there must be some means of accounting for the fact that spacial attributes in percepts are always the same attributes whether they are connected with visual, tactual or other sensations. Take a concrete illustration of the difficulty that has arisen here. The blind man's space is primarily a space of tactual and muscle sensation. Normal space is predominantly a visual space. The most divergent opinions have been advanced on the question, Does the blind man's space correspond to that of the normal man? Or, take a case from ordinary normal experience. Does the space we perceive in the

dark differ in any way from the space we see when with perfectly stationary eyes we look at the same scene? The only possible answer to these questions seems to me to be that there is but one space whatever the sensations through which it is known. I am well aware that the contrary opinion is widely accepted. Berkeley's dictum on the subject is, "The extension, figures and motions perceived by sight are specifically distinct from the ideas of touch, called by the same names; nor is there any such thing as one idea or kind of idea, common to both senses."[1] Why, then, I ask, if these extensions are different, 'specifically distinct,' do we fail to recognize the difference between them? It is often answered that the visual and tactual spaces have become so thoroughly associated that they loose their identity. But it is not the function of association to identify associated contents. A and B are associated because they are similar or contiguous, but not because they are identical. The difference between association and space-perception may be well illustrated by looking at a rough surface. We associate certain tactual quality with these visual qualities, but we never fail to recognize the two groups of qualities as distinct from each other. The spacial attributes, on the other hand, like the object itself as distinct from its qualities, are the same whether perceived by touch or by vision. We may recognize the tactual content as distinct from the visual content, but both contents are referred alike to the same object and to the same points in space, and their extensions are recognized as identical. It was this fact that was brought out by the distinction that Locke made between primary and secondary qualities. Primary qualities must be known in the same way whatever the sense through which they are known. It is this fact which finds expression in the discussions that identify space and substance. The one factor common to all percepts representing objects is the spacial character of these percepts. The qualitative attributes vary indefinitely. This essential unity of space is one of the most important facts in the psychology of space-perception. Any theory which is to be at all adequate must be prepared to explain this fact. Explanation is, however, diffi-

[1] Essay, § 127.

cult, because it is not easy to find in the various sensation complexes which give rise to spacial percepts any common factor. Some writers have found this common factor in muscle-sensation. But muscle-sensations, as already argued, will not serve to explain space even if it could be shown that they are always present, and there is much evidence which points to the conclusion that they are not present in all such percepts.

One other method of accounting for this unitary character of space has been the adoption of nativistic theories of various types. Take, for example, Lotze's discussion in the *Medicinische Psychologie*.[1] He writes: "Space is an original and *a priori* function belonging to the nature of the mind itself. It is not produced by external impressions; these serve merely to determine the particular application of the subjective function."

Such a solution of the problem is evidently no explanation of space. It merely refers space to the subject as an unanalyzable function and devotes itself to the discussion of the particular applications of this subjective activity. But here the old problem presents itself in just about the same form. Why should this subjective function apply to one content and not to another? There must be some attribute common to all of the sensation complexes to which this subjective function applies, and this common attribute must be present in various degrees and particular modifications, for not all sensation complexes are equally spacial in character.

This common attribute, which seems necessary, then, from any point of view, has been sought in all of the various theories in some concrete sensation quality. In place of any particular quality or qualities of sensation, there is good empirical ground for holding that the only common factor in all sensation complexes leading to spacial percepts is *a particular kind of relation*. The qualities may vary indefinitely. They may be derived from one sense or from several senses. They depend, however, for their spacial attributes not upon their character as quality, but upon the way in which they are related in the whole complex. *The common factor is not a concrete factor, but it is a relation.*

[1] P. 335.

Let us test this hypothesis in connection with the problem in hand, namely, the visual perception of depth. We may assume the existence in retinal sensation of two-dimensional data, and raise merely the question, How can two-dimensional data be related so as to give rise to a percept with an entirely new attribute, namely, the attribute of solidity? Some of the experiments to which we must refer are familiar, but we shall review them briefly for the sake of systematic treatment. It is a well-known fact that two plain figures which are exactly alike may be fused in a stereoscope and that the result will be the perception of a single flat object. Absolutely nothing is added to the original monocular data. Complete identity of the two retinal images will result in a single two-dimensional percept. We may try to imagine solidity, but we can never perceive it directly. We may call up all the associations possible, and they will serve to make the imagined solidity vivid, but they will never result in a direct perception of depth. Suppose now, instead of these two figures that are exactly similar, we fuse in the stereoscope the familiar figures which will produce a truncated pyramid. The two figures differ in their horizontal attributes. This difference in horizontal attributes must not, however, be too great. If the inner squares are brought too near to the boundaries on the outer square, that is, if the difference between the two images is made too great, there will be no fusion at all. It is only slight differences that can be overcome. Complete identity of the two-dimensional attributes of the two groups of sensations is not demanded, but, on the other hand, the differences must not be too striking either in their degree or kind. If the difference is in the diagonal or perpendicular direction, even though it be relatively very small, it will, on account of its unfamiliarity, prevent the fusion of the two figures. Two groups of sensations with a similarity approximating identity, but with a slight difference in certain familiar directions, seems to be a full description of the conditions. This may be made somewhat clearer by a detailed examination of the figures. The broad side of the right figure and the narrow side of the left figure are the elements which combine in the completed percept. If these two sides were exactly alike they would fuse without

difficulty into a two-dimensional image. If the difference is too great they will not fuse at all. If the difference is slight they will fuse in such a way as to represent a side extending into the third dimension. *The third dimension is not a new content; it is a form in which the contradiction between the two-dimensional attributes of the two groups of sensations is eliminated.*

This description is continually hampered by the necessity of using terms which apply to conceptual rather than perceptual thought. Of course, there is no actual scrutinizing of the sensation data with a view to finding out the likenesses and unlikenesses, but there is an immediate synthesis of these factors. In order that they may be united into a single percept there must be some recognition of the differences as well as the similarities between the various factors. If two groups of sensation data belonging in the same system and disagreeing with each other are to be united, it must be possible to find some formal attribute, some relation between these factors, which shall permit identity and difference to be reconciled.

The experiment may be made very striking by covering up one of the sides of one of these figures after they have been fused in the stereoscope. The result is, first, that the corresponding side of the pyramid loses its third dimension. This is due to the fact that it is now a monocular object. Secondly, the width of the side thus seen as flat will be the exact width of the figure. It is impossible to get a symmetrical figure out of the pyramid. If space is the product of association this ought not to be so. The horizontal extension in the monocular figure means depth just as much when it is not united with the other monocular figure as it would in binocular vision, and if space is merely a matter of association, the associated movements should operate in this case to bring out the meaning as well as when the fusion with the other image takes place. But the side of the pyramid cannot by any possible turn of imagination be thrown out into the third dimension in such a way as to give a symmetrical image. The moment the covering is withdrawn so that the side fuses with the corresponding side of the other figure, all is different. The figure is symmetrical. The two sides of unequal width are no longer merely imagined as depth,

with horizontal extension based on the real width of the figure only, but they are fused into the third dimension, and this third dimension is an additional form of arrangement in which the contradiction between the two-dimensional attributes is eliminated and in which the width of the figure is interpreted in its full meaning of width and depth both.

The fact that the difference between the two images must not be too great can be illustrated by a variety of experiments. If we take in the stereoscope the two figures which produce a truncated pyramid and after fusion erase part of one or more of the lines in either of the figures, we find that the perception of solidity is immediately impaired even though it is impossible for the observer to determine which retinal image is incomplete. Or take the simplest possible figure, one made of dots placed at the corners of the larger and smaller squares, without any connecting lines in the figure anywhere. Such figures will fuse if the dots are similar in quality. Differentiate two or more of the dots by making one red and the other blue, and that part of the figure will no longer be solid. Or in such dotted figures draw one of the diagonal lines in one of the figures only. Again the figure will lose its depth in that particular part. And so on indefinitely, the groups of sensation factors must closely resemble each other or they will not fuse.

It may be objected that we are in these cases dealing with spacial attributes from the start, and are, therefore, begging the question by assuming two-dimensional data. But, as was pointed out in the earlier part of the paper, we can never hope to trace out completely all the processes by which space is developed in all its concrete phases. A single dimension added to data which did not originally possess it is just as much a development of space as would be the derivation of the first two dimensions. Furthermore, after reaching these conclusions in regard to the third dimension, we may take up the more difficult problem of the first and second dimensions, with better hope of final success.

This hypothesis covers the facts of movement-sensations as well as the facts of binocular parallax just discussed. Take the movements of convergence. The sensations of movements from

the two eyes differ slightly in quality and intensity. These differences enter into the total complexes of sensation, which are in the main alike, and give rise to spacial attributes. Their influence is not due to any qualitative attributes, but rather to the way in which the two groups of sensations coming from the two eyes are related.

In regard to accommodation the case is somewhat more complex, for here there are no simultaneous differences in sensations. The differences here are all differences between the memory image and the present sensation. Localization through mere changes in accommodation is possible, but this localization is very incomplete. There is only a relative localization, that is, a localization due to a recognition of the difference between a given state of accommodation and one recalled out of past experience. When a person is forced to make use of these factors to a greater extent than is ordinarily the case, as, for example, the one-eyed man, it may become a valuable means of discriminating depth.

Visual depth is thus conditioned by a certain relation of contradiction between the two-dimensional data of the retinal images or between the muscle-sensations which are connected with these images. The agreement between the elements in which the contradiction appears must be great, but it must not be complete, for the contradiction is an essential element. Depth, then, may be properly defined as a form rather than as a content of perception. It is not in any sense an original attribute of sensation. It is not especially related to any particular sensation quality. It is a complex form of perception in which any sensation quality which is presented in the proper relation may be arranged. It is furthermore an empirical rather than *a priori* form, since it depends upon the objective relation between sensations for its origin and character. It is, of course, like all psychical experience, subjective, but its source is just as much objective as is that of the sensation qualities which are arranged in space.