LINKING PROPOSITIONS*

DAVIDA Y. TELLER†

Departments of Psychology and Physiology/Biophysics, Regional Primate Research Center and Child Development and Mental Retardation Center, University of Washington, Seattle, WA 98195, U.S.A.

Abstract—Linking propositions are statements that relate perceptual states to physiological states, and as such are one of the fundamental building blocks of visual science. A brief history of the concept of linking proposition is presented. Five general families of linking propositions—Identity, Similarity, Mutual Exclusivity, Simplicity and Analogy—are discussed, and examples of each are developed. Two specific linking propositions, involving the explanation of perceptual phenomena on the basis of the activity of single neurons, are explicated and their limitations are explored in detail. Finally, the question of the empirical testability and falsifiability of linking propositions is discussed.

Linking hypotheses   Linking propositions   Philosophy of science   Axioms   Neuron doctrine
Neural coding   Mach bands   Detectors   Opponent processes.

INTRODUCTION

Visual science is an eclectic discipline. Visual scientists are interested in what people see, as measured by psychophysical experiments and described by phenomenal reports. They are also interested in the physiology of the visual system, and in all elements of the substrate upon which human visual capacities are based. And finally, they are interested in the relationships between visual capacities and the neural substrates of vision; that is, in some sense, in explaining how we see on the basis of the properties of the machinery that makes seeing possible. This last is a very difficult problem, in part because of the inherent complexity of the first two topics, and in part because of the logical and philosophical difficulties involved in explanatory efforts which span between such different scientific disciplines (cf. Uttal, 1981).

The fact that physical/physiological statements and psychophysical/phenomenal statements are from such different sciences has often led to the presumption that some special kinds of logical links will be needed to bring them into a single domain. These special linking statements have been variously called *psychophysical axioms* (Müller, 1896), *psychophysical linking hypotheses* (Brindley, 1960, 1970) and *linking propositions* (Teller, 1981; Teller and Pugh, 1983). Statements that the shape of the dark-adapted spectral sensitivity curve is due largely to the absorption spectrum of rhodopsin, that the low-frequency falloff of the spatial contrast sensitivity function is caused by center-surround antagonism in receptive fields, that the mutual exclusivity of certain hues implies the existence of opponent neural coding, or that cortical simple cells are involved in form perception, may be cited as examples of linking propositions. Insofar as visual scientists are interested in explaining psychophysical data on the basis of the properties of the neural substrate, we must be said to be interested precisely in the formulation, testing, accepting and falsifying of linking propositions. Yet explicit, systematic discussions of linking propositions remain scarce in the vision literature.

The purpose of the present paper is five-fold. First, the concept of a linking proposition will be discussed and expanded briefly in the historical context and a general definition of the concept will be proposed. Second, a preliminary systematization of general linking propositions will be presented, including five families of propositions which are in common usage in visual science. Third, some specific linking propositions concerning the relation between single neurons and perception will be articulated, and used to develop a set of evaluative criteria for judging linking propositions involving single neurons. Fourth, linking propositions concerning the canonical forms that visual information must be assumed to take will be examined. And finally, a brief analysis will be attempted of the extent to which linking propositions are subject to empirical test.

LINKING PROPOSITIONS: EVOLUTION OF THE CONCEPT

The early psychophysicists, without the tools of modern electrophysiology, studied psychophysics in part as a way of discovering the properties of the brain. They developed the concept that phenomenal

states and physiological states must be, in some way,
lawfully related and worked out a list of relationships
that could be assumed by consensus to hold between
mental events and material events. These culminated
in the *psychophysical axioms* articulated by G. E.
Müller in 1896. Boring (1942, p. 89) quotes the first
three axioms as follows:

1. The ground of every state of consciousness is a material
process, a psychophysical process so-called, to whose occur-
rence the presence of the conscious state is joined.

2. To an equality, similarity, or difference in the consti-
tution of sensations... there corresponds an equality, simi-
larity, or difference in the constitution of the psychophysical
process, and conversely. Moreover, to a greater or lesser
similarity of sensations, there also corresponds respectively
a greater or lesser similarity of the psychophysical process,
and conversely.

3. If the changes through which a sensation passes have
the same direction, or if the differences which exist between
series of given sensations are of like direction, then the
changes through which the psychophysical process passes,
or the differences of the given psychophysical processes,
have like direction. Moreover, if a sensation is variable in
n directions, then the psychophysical process lying at the
basis of it must also be variable in n directions, and
conversely....

Two characteristics of these linking propositions
deserve special mention. First, they were assigned the
logical status of axioms, i.e. they had to be accepted
as fundamental, necessary and unprovable. And sec-
ond, there were several of them, including such
concepts as identity, similarity, difference, direc-
tionality and dimensionality, with the result that
many different sensory and perceptual characteristics
could be used to infer properties of the unobservable
"material process".

Over the ensuing years, technological progress has
allowed more direct knowledge of the optical, photo-
chemical, anatomical and physiological character-
istics of the visual system. Reliance on inferences
from psychophysical experiments as a direct source of
knowledge about physiological processes has dimin-
ished in consequence. In this new context, the ques-
tion of the relation between psychophysical and
physiological data was brought again to the attention
of visual scientists by Brindley (1960).

Brindley was writing as a visual physiologist, and
his primary interest was in the formulation and

testing of physiological hypotheses. He noted that in
visual science, in addition to purely physiological
data, one has available data from what he called
"sensory experiments; that is, experiments in which
an essential part of the results is a subject's report of
his own sensations" (p. 144). Brindley's question was:
What is the place and value of sensory reports in the
testing of physiological hypotheses? He formulated
the argument that phenomenal terms and phys-
iological terms are from different realms of discourse,
and that, if terms from the two different realms were
to be used together in a single sentence, explicit
bridging statements would be needed. He called such
statements *psychophysical linking hypotheses.*

Brindley was able to discover only one general
linking hypothesis which met his criteria for adequate
rigor. This linking hypothesis, which he suggested
may be a truism, was stated as follows (1960, p. 144):

... whenever two stimuli cause physically indistinguish-
able signals to be sent from the sense organs to the brain,
the sensations produced by those stimuli, as reported by the
subject in words, symbols or actions, must also be indistin-
guishable.

Other general linking propositions, use of which
would allow physiological inferences to be drawn
from experiments (p. 145) "in which the subject must
describe the quality or intensity of his sensations, or
abstract from two different sensations some aspect in
which they are alike ...", were deemed to be gener-
ally lacking in rigor, and were omitted from specific
discussion. Thus, through the adoption of extremely
strict criteria of admissibility, most of Müller's axi-
oms were discarded, and only a single general prop-
osition, concerning the indistinguishability of phys-
iological or perceptual states, survived*.

In addition to the single acceptable general linking
hypothesis, Brindley identified several very specific
ones that seemed to him well established. These dealt
mostly with the anatomical causes of various entoptic
phenomena; for example, he argued that the hypoth-
esis that the perceptual "retinal tree" is caused by the
shadow cast by the retinal blood vessels on the
receptors was extremely well established. He argued
that the truth of such special linking statements could
be established by "correlating very many features of
a sensory phenomenon with corresponding features
of an objectively determined one" (p. 149); in other
words, that special linking propositions were at least
in principle subject to empirical test.

Over the past 20 years, an enormous amount of
information has accumulated concerning the re-
sponse properties of single neurons at many levels of
the visual system. Visual physiologists have learned a
great deal about how and where information of
various kinds is lost, how it is carried at each level of
the visual system and how and where it is trans-
formed from one format to another. In this context
it is sometimes feasible, and always tempting, to try
to attribute psychophysical and perceptual phenom-
ena to the properties or activities of individual neu-

*Brindley's view subsequently became somewhat more
liberal. In the 1970 edition of his book, the chapter on
linking hypotheses is still much the same, but in the
chapter on color vision Brindley included for the first
time a discussion of Hering's contributions to color
vision theory. In the context of this discussion he
remarked (p. 208, parentheses added) that "(Hering's)
argument from the appearance of colours... though
non-rigorous is not necessarily bad; it is hard to judge
because it comes from a kind of thinking that is outside
the main tradition of natural science... (If) Hering's
argument is vindicated... it will help us decide whether
in future to pay attention to arguments from Class B
(i.e., subjective) sensory phenomena, or to disregard
them".

rons or sets of neurons at particular levels of the visual system, and explanatory efforts of this kind are now a major thrust of visual science.

A reformulation of the concept of a linking proposition, which encompasses earlier points of view and emphasizes use of the modern detailed knowledge of the activities of neurons at many levels of the visual system, has been put forward recently by Teller and Pugh (1983). For the sake of discussion, Teller and Pugh describe the process of vision as a sequence of deterministic maps, with information being mapped from the visual world to the retinal image, to the quantum catch states of all receptors, to the set of all receptor outputs, and so on through the visual system. In their formulation, the mapping from any earlier to any later stage may be thought of as a composite map.

Eventually the states of some large or small subset of these neurons must be mapped to subjective visual perception. Partly as a means of dealing with this final mapping, Teller and Pugh introduce the term *bridge locus*, as follows (p. 581, phrase in parentheses added):

Most visual scientists probably believe that there exists a set of neurons with visual system input, whose activities form the immediate substrate of visual perception. We single out this one particular neural stage, with a name: the *bridge locus*. The occurrence of a particular activity pattern in these bridge locus neurons is necessary for the occurrence of a particular perceptual state; neural activity elsewhere in the visual system is not necessary. The physical location of these neurons in the brain is of course unknown. However, we feel that most visual scientists would agree that they are certainly not in the retina (i.e. that they consist of a central rather than a peripheral subset of visual system neurons). For if one could set up conditions for properly stimulating them in the absence of the retina, the correlated perceptual state presumably would occur.

In this context, a linking proposition may be defined as *a claim that a particular mapping occurs, or a particular mapping principle applies, between perceptual and physiological states* (cf. Teller and Pugh, 1983, p. 581). Such a claim can be general, as were Müller's axioms and Brindley's general linking hypothesis, or it can be specific to a particular perceptual phenomenon or a particular set of optical, photochemical or neural elements. The location of the neural elements can be unspecified, and a claim made only that, since a particular perceptual state occurs, neurons having particular properties must exist somewhere within the visual system; or that visual information must somewhere be encoded in a particular form when a particular perceptual state occurs. Alternately, the linking proposition can involve specified neural elements at any location, peripheral or central, at or prior to the bridge locus. The more peripheral the neurons, the more complex is the implied composite map; and, most importantly, in all cases the implied composite map includes the mapping from the bridge locus neurons to perceptual states.

In addition to these reformulations, it is philosophically appropriate in the mid-1980s to adopt an analytical rather than a proscriptive attitude toward linking propositions (cf. Suppe, 1977, esp. pp. 650–730). Rather than prejudging the logical status of linking propositions, or condemning them for lack of rigor, one can simply state that there is a consensus among many visual scientists that physiological and perceptual facts together make up a viable scientific domain (Shapere, 1977; Suppe, 1977, pp. 686 ff.), and that linking propositions form some of the necessary elements of that domain. The question then becomes, what kinds of linking propositions do visual scientists use, what are their logical interrelationships and truth-values, and what roles do they play in the structure of modern visual science?

In the next three sections, three kinds of linking propositions will be examined: general linking propositions; linking propositions involving single cells; and linking propositions concerning the canonical forms that visual information must take before it can form the substrate of visual perception.

GENERAL LINKING PROPOSITIONS: A PRELIMINARY SYSTEMATIZATION

In the present section, a preliminary systematization of general linking propositions will be presented. Two main points will be made. The first is that general linking propositions come in logically organized groups or families. The second is that there are at least four such families, which can be called Identity, Similarity, Mutual Exclusivity and Simplicity, along with a less well organized proposition that can be called Analogy. As should be obvious, the point of the exercise is not to claim that the various propositions are true by logical necessity, or even that they are necessarily workable or even sensible; but rather to articulate them in order to make such evaluative judgments possible.

The logical structure of the first four families is shown in Fig. 1. By definition, a linking proposition is a statement that contains both physiological (ϕ) and psychological (ψ) terms. On the first line of Fig. 1 is portrayed a linking statement that ϕ implies ψ; i.e. that some characteristic of the state of the physiological substrate implies some characteristic of the perceptual phenomenon. On the second line is the Contrapositive (the negative of the converse) of the first line: $\bar{\psi}$ implies $\bar{\phi}$ ("not ψ" implies "not ϕ"). These two statements are logically equivalent because any statement implies the negative of its converse: If A implies B, $\bar{B}$ implies $\bar{A}$. On the third line is stated the Converse of the first line: ψ implies ϕ; i.e. that some characteristic of the psychophysical data implies some characteristic of the substrate. The truth of the Converse, of course, is not implied by the truth of the original statement. This point was explicitly made and appreciated by the early psychophysicists, as is shown by Müller's inclusion of the separate

1. Initial proposition $\phi \rightarrow \psi$
*2. Contrapositive $\bar{\psi} \rightarrow \bar{\phi}$
*3. Converse $\psi \rightarrow \phi$
4. Converse Contrapositive $\bar{\phi} \rightarrow \bar{\psi}$
   (CC)

Fig. 1. The logical structure of the first four families of linking propositions. Line 1 states a proposition that a particular characteristic of the physiological state (ϕ) implies a particular characteristic of the perceptual state (ψ). Line 2 is the Contrapositive of line 1 and is implied by line 1. Line 3 is the Converse of line 1 and states a proposition that a particular characteristic of the perceptual state implies a particular characteristic of the physiological state. Line 4 is the Contrapositive of the Converse and is implied by line 3.

phrase "and conversely" in the axioms quoted above. And on the fourth line is the Converse Contrapositive: $\bar{\phi}$ implies $\bar{\psi}$. The third and fourth lines, like the first and second, are logically equivalent to one another.

There is another interesting interrelationship among subsets of these propositions. Propositions 1 and 4 have to do with drawing conclusions about psychophysical phenomena from physiological data, while propositions 2 and 3, which are starred in Fig. 1, have to do with drawing physiological conclusions from psychophysical data. The latter direction of inference—from psychophysical data to physiological conclusions—has played the larger role in the history of visual science, largely because we have had access to psychophysical data for a much longer time than we have had access to any very detailed information about the physiological substrate. With the increasing availability of physiological data, the two directions of inference may well come to play more equal roles. Nonetheless, most of the examples that will be discussed below are examples of using propositions 2 and 3, the Contrapositive and the Converse, to argue from psychophysical data to physiological conclusions.

And finally, in deciding to do a psychophysical experiment, one commits oneself to the use of either the Contrapositive or the Converse, or both together. But one cannot choose which linking proposition one will use ahead of time, because that depends on the results of the experiment: on whether the outcome shows ψ or $\bar{\psi}$ to be true.

The first family of linking propositions, the Identity family, is shown in Fig. 2. The initial Identity proposition, line 1, is that identical ϕ imply identical ψ; i.e. that identical states of the nervous system must lead to identical sensations. The Contrapositive, line

2, is that if two sensations are discriminable the underlying physiological states must also differ.

A specific form of the Identity proposition, stated in the context of the physical stimuli used to produce the sensations in the psychophysical experiment, is familiar as the one Brindley felt was both acceptably general and acceptably certain, viz, "If two stimuli cause physically indistinguishable signals to be sent from the sense organs to the brain, the sensations ... will also be indistinguishable". The use of the Contrapositive was also made explicit by Brindley in his example of hypothesis testing. He argued that the non-identity of two sensations implies the non-identity of physiological states, and thus that the psychophysical finding of discrimination between two stimuli allows the rejection of any physiological hypothesis which predicts that the two stimuli cause physically indistinguishable signals to be sent from the sense organs to the brain.

The logical status of the Identity and Contrapositive Identity propositions is very strong. They are probably the most universally accepted of any of the general linking propositions to be discussed here. In Brindley's formulation—phrased in terms of the discriminability of stimuli—they are restatements of the more general scientific axiom that information that is lost at one level of an information processing system cannot be recreated by later levels. In the present formulation—phrased in terms of the non-identity of sensations—they amount to the statement that the mapping between brain states and phenomenal states is 1:1 or many:1, but not 1:many (Feigl, 1958). Most visual scientists would doubtless accept this statement also as axiomatic (but cf. Eccles, 1977).

At the same time, H. B. Barlow (personal communication) and others have argued convincingly that, in the original form, these propositions run the risk of being useless. That is, because of noise in both the stimulus and the neural system, even the same stimulus will generate a variable response from one trial to the next. For this reason, a proposition that requires the existence of two identical physiological states, generated by the same or different stimuli, even if logically unassailable, may refer to an empty set.

To alleviate this problem, one must speak of statistical rather than absolute identity or non-identity, of stimuli, sensations and neural states, over a series of trials. The reformulated identity proposition would be that statistically identical neural signals imply statistically indiscriminable sensations; and the reformulated Contrapositive Identity proposition would be that statistically discriminable sensations

1. Identity	Identical	ϕ	$\rightarrow$	Identical	ψ
*2. Contrapositive Identity	Non-Identical	ψ	$\rightarrow$	Non-Identical	ϕ
*3. Converse Identity	Identical	ψ	$\rightarrow$	Identical	ϕ
4. CC Identity	Non-Identical	ϕ	$\rightarrow$	Non-Identical	ψ

Fig. 2. The Identity family.

imply statistically discriminable physiological states. In this form the Identity proposition and its Contrapositive appear to be both logically strong and applicable to real psychophysical cases. The statistical discriminability of stimuli implies the statistical discriminability of the corresponding sensations, and hence the statistical discriminability of the corresponding physiological states.

The third line of Fig. 2 states the Converse Identity proposition (cf. Brindley, 1957; Boynton and Onley, 1962). It is that identical ψ imply identical ϕ; or, that statistically indiscriminable sensations imply statistically identical states of the visual system. The fourth line is the Contrapositive Converse Identity proposition: that statistically discriminable states of the nervous system imply statistically discriminable sensations.

The Converse and Converse Contrapositive identity propositions are clearly not analytically nor universally true, for several reasons. In the first place, there are many instances in which two different stimuli are believed to lead to different quantum catches, for example, and yet are psychophysically indiscriminable, precisely because the two neural signals are rendered indiscriminable at some subsequent stage of neural processing. Such cases of information loss within the visual system are among the most interesting and intensively researched phenomena in visual science. The elevations of threshold in the early stages of dark adaptation provide the classic example, and many other cases could be cited (cf. Ratliff and Sirovich, 1978).

This problem serves to illustrate the usefulness of the concept of a bridge locus and the bridge locus neurons, the central neurons that form the most immediate substrate of conscious perceptual events. In many cases, the bridge locus neurons will be the neurons whose states one will have the most direct means of inferring by using contrapositive or converse linking propositions in combination with psychophysical data. Inferences about the states of neurons at earlier levels of the serial processing system will necessarily be less direct, and the Converse Identity proposition needs to be reformulated to take the possibility of information loss in the earlier stages of visual processing into account.

*Walter Makous (personal communication) and others have argued that if two stimuli lead to behaviorally indistinguishable responses, the neural signals underlying these responses *must* be indistinguishable somewhere, even if only at the level of the motor neurons. Under this interpretation, the Converse Identity proposition, like the Identity proposition, is true and tautological. However, this is not the interpretation in common use. When visual scientists employ the Converse Identity proposition, I believe they specifically mean to assume that two signals are rendered identical within the *sensory* visual system. Where the sensory visual system leaves off and the central or motor systems begin is, of course, very much an unsolved problem in its own right.

To allow for the option of neural information loss, one may propose that the meaning of the Converse and Converse Contrapositive Identity linking propositions be stipulated as follows: if two different stimuli (or their resulting sensations) are statistically indiscriminable, the corresponding neural signals must be rendered statistically indiscriminable somewhere within the sensory system, at or prior to the bridge locus*.

Even so, the Converse and Converse Contrapositive Identity propositions are clearly not logical necessities. Few visual scientists would argue that indiscriminable sensations necessarily imply identical physiological states, even at the bridge locus, because to claim this is to claim that one is capable of sensing every different state of the bridge locus neurons; i.e. that the mapping at the bridge locus is 1:1. Visual scientists probably tend to believe that the mapping of physiological states to perceptual states is many:1, with there being more different states of the bridge locus neurons than there are sensations. This point is probably not addressable empirically; its logical consequences are explored further by Teller and Pugh (1983).

Despite this remaining logical difficulty, the Converse Identity proposition has often been used in the history of visual science and inference structures which have included it have often led to correct conclusions. It has provided us with some of our most striking successes, and is very much in use today. For example, the existence of metamers—stimuli of different wavelength composition which produce identical sensations—inspired the insight that color vision might be subserved by only a few kinds of receptors, each one of which loses wavelength information.

No instances in which the careful use of the reformulated Converse Identity proposition has led to false conclusions have yet come to the writer's attention. It is therefore proposed that Converse Identity be accepted as generally useful until at least one counterexample is found.

What kinds of psychophysical data are used in conjunction with Identity propositions? Clearly the relevant experiments are those which Brindley (1960, 1970) called Class A—experiments in which the subject is asked to judge the identity or non-identity of the sensations resulting from two different stimuli. Brindley put matching experiments and threshold experiments in this category, the latter presumably because in a threshold experiment the subject must judge whether or not the presence and the absence of the test stimulus lead to identical sensations (but cf. Boynton and Onley, 1962). In classical psychophysical terminology, Brindley's matching experiments can be identified with *discrimination* (since if two stimuli are indiscriminable they can be said to match) and Brindley's threshold category can be identified with *detection*.

When one tries, then, to draw physiological conclu-

sions from the data of detection or discrimination experiments, Identity propositions will always be involved. Which Identity proposition will be involved depends upon the outcome of the experiment. From failures to detect or discriminate, one can risk inferring the identity of physiological states. These inferences will employ the Converse Identity proposition. They would seem to be relatively safe if confined to the bridge locus, and more risky if used to infer the states of more peripheral sites. From correct detections or discriminations, one can infer the non-identity of physiological states. These inferences will employ the Contrapositive Identity proposition, and would seem to be remarkably safe.

The next family of linking propositions—Similarity—is shown in Fig. 3. Line 1 states the basic Similarity proposition, that similar states of the sensory system will lead to similar sensations. Phrased in terms of stimuli, and following Brindley's formulation of Identity, a more specific form would be, if two stimuli cause similar signals to be sent from the sense organs to the brain, the sensations produced by these stimuli will also be similar. Line 3 states the Converse, that similar sensations imply similar states of the nervous system.

Neither of these propositions is a logical necessity, nor universally true, and these propositions are subject to modified versions of all of the limitations of the Identity propositions discussed above. In particular, Similarity propositions would seem to be a better bet when used to relate central, rather than peripheral, physiological states to phenomenal states, because stimuli which appear similar can sometimes be very different physically, and hence often very different at the earliest stages of neural coding.

As is the case with the Identity family, it is Converse Similarity that has historically received the greatest usage in visual science. In fact, Converse Similarity is in common use, and is perhaps even indispensable to visual science. For example, we use the Converse Similarity proposition in discussions of the perceptual orderliness of the variation of hue with wavelength across the spectrum. It is in use when we assume that, since long wavelength and short wavelength light produce similar (reddish) sensations, these two wavelength bands must have similar neural codes; or when we assume that a blue sensation created by simultaneous contrast creates a neural state similar to that created by a 470 nm light. Even more strikingly, it is involved in discussions of the Benham top phenomenon, when we propose any explanation that assumes that the neural signals produced by the Benham top stimuli must be similar

to the signals produced by the chromatic stimuli they resemble. These assumptions are not logical necessities; yet, if one did not make these assumptions, it is not clear how one would go about searching for an explanation for the Benham top phenomenon, or recognize an explanation when one had found it.

There is also a second form of Similarity proposition, perhaps closely related to the n-dimensionality discussed by Müller and others. For example, in heterochromatic brightness matching—the matching of brightness across differences in hue—or studies of the Bezold–Brücke hue shift—matching of hue across differences in brightness—the subject judges two non-identical sensations to be matched on one perceptual dimension but not on another. The matching of hue across variations of size and eccentricity provides another example. Converse Similarity propositions are involved when we model these data by assuming that lights that match in hue will have important elements of neural coding in common, as will lights that match in brightness.

One also suspects that the Similarity family has a lot of cousins, concerned with the degree of similarity between two sensations. For example, there is probably a Continuity family, containing assertions to the effect that continuous (or discontinuous) changes in neural activity imply continuous (or discontinuous) changes in sensations and vice versa; a Montonicity family, with propositions to the effect that monotonic variation in neural activity implies a monotonic variation in sensation and vice versa; and perhaps also quantitative versions of these, for example, that a doubling of judged brightness implies the doubling of a value in a neural code and vice versa.

Experiments relevant to Similarity propositions will be a subset of the kind which Brindley called Class B; i.e. experiments in which the subject is asked to abstract from two distinguishable sensations ways in which they are alike. For the first kind of Similarity, judgments would concern subjective order. For the second kind, judgments would concern identity or similarity on one subjective dimension at a time. For the cousins of Similarity, the psychophysical data would be provided by other kinds of Class B experiments, such as the ordering or scaling of brightness or the matching or naming of hue.

The next two families of linking propositions—Mutual Exclusivity and Simplicity—both have their primary exemplars in opponent process theories of color vision. The origins of psychophysical opponent process theories rest heavily upon two subjective characteristics of colors: that some colors—red, yellow, green and blue—seem more *unique* or *simpler*

1. Similarity	Similar	ϕ	→	Similar	ψ
*2. Contrapositive Similarity	Dissimilar	ψ	→	Dissimilar	ϕ
*3. Converse Similarity	Similar	ψ	→	Similar	ϕ
4. CC Similarity	Dissimilar	ϕ	→	Dissimilar	ψ

Fig. 3. The Similarity family.

than others; and that these four colors divide themselves into two pairs, red–green and yellow–blue, the members of a pair being *mutually exclusive* in the sense that the two hues cannot be perceived in the same place at the same time. Opponent process theories typically assume that the mutual exclusivity is caused by color coding channels that take on either of two mutually exclusive states; and that the unique hues occur at wavelengths for which one of two such channels is at its null point.

The third family of linking propositions, then, can be called Mutual Exclusivity, and is shown in Fig. 4. The basic Mutual Exclusivity proposition is: Whenever two stimuli cause mutually exclusive neural signals, the sensations produced by these stimuli must also be mutually exclusive; or, mutually exclusive neural states will correspond to sensations which never occur together. The Contrapositive, which is easier to think about, is: if two sensations can occur together, their corresponding neural signals must also be compatible. The Converse Mutual Exclusivity proposition is: mutually exclusive sensations imply mutually exclusive states of the nervous system.

Psychophysical experiments relevant to the Mutual Exclusivity propositions would be those in which a subject is asked to judge the mutual compatibility or exclusivity of sensations. The judgment of mutual exclusivity of redness and greenness, and of yellowness and blueness, provides the major historical example. Reversible figures may be taken to provide a second example. Many subjects would judge that the two perceptual configurations of the Necker cube, for example, cannot happen at the same time. This observation has often been taken to imply that the Necker cube leads to the setting up of one of two alternative and mutually exclusive physiological states, or to a "bistable state" (cf. Marr, 1982, pp. 25–26). The author has thought for a while about whether seeing a cross excludes seeing the vertical and horizontal lines that make it up, but is unable to decide this issue on the basis of her own introspections. If the two sensations are mutually exclusive, then Mutual Exclusivity propositions might have some relevance to the question of whether cortical cells that respond to lines can form the physiological basis of the perception of complex patterns.

The logical status of the various Mutual Exclusivity propositions is difficult to assess. One of the problems is that, unlike the cases of Identity and Similarity, inferences from Mutual Exclusivity propositions do not seem to be safer when applied at the bridge locus than at more peripheral sites. That is, if the physiological states created by two stimuli are mutually exclusive at *any single* neural stage, the two sensations should never occur together, but as noted by Teller and Pugh (1983, p. 586, parentheses added): "... logic does not require that an opponent code be maintained all the way through the system to the bridge locus. At later stages the codes for ... redness and greenness need not be mutually exclusive: they simply (would) never occur together in normal visual experience, because of the mutual exclusivity of their causal precursors at a critical locus early in the visual system". A recent report of a set of highly unnatural conditions under which the perception of a reddish–greenish hue is reported by some subjects (Crane and Piantanida, 1983) lends further force to this argument.

Given these problems, the Mutual Exclusivity proposition is in need of a more circumscribed formulation, perhaps as follows: if two neural states are mutually exclusive at any single stage of neural processing, the corresponding sensations will ordinarily never occur together. However, if the coding is not mutually exclusive at the bridge locus, and if a way can be found to activate both neural codes together at the bridge locus, then the two sensations can be made to occur together. Similarly, the Contrapositive would be, if two sensations are ordinarily mutually compatible, their neural codes are compatible at all stages of neural coding.

The Converse Mutual Exclusivity proposition might be similarly elaborated, more or less as follows: if two sensations do not ordinarily occur togther, there exists a stage of neural processing, at or before the bridge locus, at which the neural codes for the two sensations are mutually exclusive. In this context, if stimulus conditions are found under which the two sensations do occur together, it would follow that the stage at which the two codes are mutually exclusive must be prior to the bridge locus, and that at the bridge locus the two codes are compatible. However, even in these modified forms, the truth values of mutual exclusivity propositions remain difficult to assess.

Another problem with the property of Mutual Exclusivity is that it appears to have strong and weak forms. In the strong form, as used in opponent process theory, mutually exclusive signals are those which depart in opposite directions from a null state. But in the weak form, any two different output states

1. Mutual Exclusivity	Mutually Exclusive ϕ	$\rightarrow$	Mutually Exclusive ψ
*2. Contrapositive Mutual Exclusivity	Mutually Compatible ψ	$\rightarrow$	Mutually Compatible ϕ
*3. Converse Mutual Exclusivity	Mutually Exclusive ψ	$\rightarrow$	Mutually Exclusive ϕ
4. CC Mutual Exclusivity	Mutually Compatible ϕ	$\rightarrow$	Mutually Compatible ψ

Fig. 4. The Mutual Exclusivity family.

for a neuron would seem to be mutually exclusive, in that they cannot both occur at the same time. Psychophysically, redness and greenness are strongly mutually exclusive, but yellowish–green and greenish–yellow would also seem to be mutually exclusive in some weaker sense. In sum, the concept of mutual exclusivity is a complex and slippery one, and it would doubtless profit from formal analysis at definitional and conceptual, as well as phenomenal and physiological, levels.

The fourth family of linking propositions, that of Simplicity (Fig. 5), also has its main historical illustration in opponent process theory. The basic hypothesis is: whenever a stimulus causes the occurrence of a particularly simple signal, the sensation produced by that stimulus will be particularly simple; or, simple states of the nervous system will produce simple sensations. The Converse is that sensations that are particularly simple imply the presence of particular elements of simplicity in the neural code.

Like Converse Mutual Exclusivity, Converse Simplicity plays a role in opponent process theory: the particularly simple sensation is a unique hue, and the particularly simple signal is a signal in which one of the two chromatic channels remains at its null point.

Simplicity propositions seem to be particularly intuitively appealing, but at the same time particularly weak, because the notion of what is simple may vary from context to context and from person to person. Visual scientists may not have the same implicit consensus on what is simple that we may have on identity, mutual exclusivity or similarity. For example, whether one thinks that the assumption that a neural network performs a crude Fourier analysis is a simplicity or not depends on how one feels about Fourier analysis. In addition to a possible lack of consensus at any one time in history, the notion of simplicity also seems to change from one decade to the next, and it will have to be redefined as the science goes along. The Simplicity propositions, however, will probably maintain an implicit and powerful role in our thinking processes, while our concrete notions of simplicity change around them.

The Simplicity family may well form the basis of much of the intuitive appeal that visual scientists of the present era have found in single cell models of visual phenomena. Line detectors have been appealing as candidates for the mapping of line stimuli into the perception of lines, because intuitively lines are among the simple elements from which a spatial pattern could be made up; and the firing of one particular cell out of a population of cells, for a line at each orientation at each position, has a ring of

simplicity about it. The Neuron Doctrine so beautifully explicated by Barlow (1972) illustrates the implicit involvement of Simplicity assumptions, as when he says that "the activity of each single cell is ... thought to be related quite simply to our subjective experience ..." (p. 371) or in his Fourth Dogma, that "... the active high-level neurons directly and simply cause the elements of our perception" (p. 381).

The last general linking proposition to be discussed here can be called Analogy. The Analogy proposition departs from the logical structure of the previous four familes. The proposition is shown in Fig. 6; it is: if ϕ "looks like" ψ, then ϕ explains ψ (see Teller, 1980, for an earlier discussion of the concept of "looking like"). This means, roughly, that if psychophysical and physiological data can be manipulated in such a way that they can be plotted on meaningfully similar axes, such that the two graphs have similar shapes, then that physiological phenomenon is a major causal factor in producing that psychophysical phenomenon. This kind of linking proposition appears in many forms, and is extremely commonly used, either explicitly or implicitly. The explanation of the shape of the scotopic spectral sensitivity curve by the absorption spectrum of rhodopsin (in combination with absorption in the optic media) is an example of a strong use of analogy, bolstered by many other elements of fact and logic. The explanation of Mach bands by the spatial pattern of activity in ganglion cells, discussed in detail below, is a good example of weak and implicit use of Analogy. The analogy between opponent cells and opponent colors has been explored recently by Hood and Finkelstein (1983). The explicit use of analogy in relating the states of central neurons to perceptual phenomena is best exemplified in the work of De Valois and his colleagues (e.g. De Valois, 1965).

The Analogy proposition is the first of the general propositions to contain both a psychophysical and a physiological referent on both sides of the arrow. This fact adds some conceptual complexity to the evaluation of Analogy propositions, because to use all the earlier families one had only to make judgments about the properties—identity, similarity, etc.—of physiological states with respect to each other and perceptual states with respect to each other. Here, one must judge the quality of the analogy between psychophysical and physiological data sets.

The degree of acceptability of Analogy propositions varies widely, and many potentially rigorous linking propositions probably enter the discipline by this route (cf. Suppe, 1977, p. 687). Ideally, the initial analogy will be followed up by exploring its impli-

1. Simplicity	Simple ϕ $\rightarrow$	Simple ψ
*2. Contrapositive Simplicity	Non-Simple ψ $\rightarrow$	Non-Simple ϕ
*3. Converse Simplicity	Simple ψ $\rightarrow$	Simple ϕ
4. CC Simplicity	Non-Simple ϕ $\rightarrow$	Non-Simple ψ

Fig. 5. The Simplicity family.

cations, and subjecting it to the usual kinds of tests that lead to the acceptance or rejection of other scientific propositions. On the other hand the fact that two curves look alike ought not to be taken as, in itself, an explanation of the psychophysical by the physiological phenomenon. It is the author's contention that the explication of criteria of acceptability and testability of Analogy propositions is an area in which additional analytical work is badly needed.

In summary, these five families—Identity, Similarity, Mutual Exclusivity, Simplicity and Analogy— are offered as the first step toward a modern systematization of general linking propositions. They are surely not an exhaustive list, and additional classes of general linking propositions will doubtless be added. A cataloging of the ways in which each proposition is used—as axiom, assumption, hypothesis, conclusion or speculation—needs to be undertaken, along with an examination of the mutual compatibility of the linking propositions that are in most common use. More examples of successful and unsuccessful uses of each proposition, both historical and modern, need to be developed. In particular, the scheme is in need of good counterexamples; i.e. instances in which one or another of the propositions has been found to be limited, misleading or useless in relating specific psychophysical to specific physiological phenomena. Some such specific counterexamples are explored in the next section.

SPECIFIC LINKING PROPOSITIONS INVOLVING SINGLE CELLS

Many of the linking propositions involved in visual theory in the last decade or two have had as their ϕ terms the activities of single neurons or regular populations of single neurons. To each visual scientist, some such linking propositions seem credible, while others seem foolish; but the criteria by which such evaluations are made have seldom been stated explicitly. Two cases of popular but faulty "textbook explanations" (Hood and Finkelstein, 1983) of perceptual phenomena will be analyzed here. The first case is the usual textbook explanation of Mach bands, and the second is the use of the concepts of *trigger features* or *detectors* as explanations of part of the neural processing that enters into form perception. The purpose of the exercise is to try, by analyzing these examples, to begin to develop a set of criteria for the evaluation of single unit linking propositions.

Mach bands are the illusory bright and dark bands seen at the intensity transition regions of a ramp pattern of intensity. The common explanation for Mach bands is that, if one could trace out the firing

rates of cells—say, a subclass of retinal ganglion cells—in the presence of a ramp pattern, one would find maxima and minima in firing rates in the regions where subjects report Mach bands. In fact, many people have tried this experiment in cat ganglion cells, and some of them have indeed found minima and maxima of activity in about the right places (cf. Robson, 1980). Under the right conditions, the physiological data "look like" the psychophysical data. The analogy is very appealing, but the question is, to what extent, or in what sense, do these results provide an explanation of why we see Mach bands?

Obviously there is an implicit linking proposition involved in this argument—that *the set of firing rates across a row of ganglion cells maps to the set of perceived brightnesses across the visual field*. At least three reservations about this proposition should be expressed: its intended range of applicability, its assumption of homogeneity, and its peripherality. They are illustrative of problems one can readily find implicit in many single cell models of visual phenomena.

First, what is the intended range of applicability of this linking proposition? Over how many of the phenomena of brightness perception would this correlation be expected to hold? If one were serious about this explanation, one would have to show a correspondence of details between the psychophysical and physiological phenomena (cf. Brindley, 1960, p. 149). For example, one would have to show that the maxima and minima are largest when the Mach bands are most prominent and absent when Mach bands are not reported, and that the correlation holds for a broader range of brightness phenomena, including the brightness of isolated spots of light, the perception of brightness across sinusoidal gratings, the growth of subjective contrast with stimulus contrast, and so on. In the absence of exploration of its range of applicability, such a linking assumption seems to be too *ad hoc* to have much explanatory power.

The second reservation concerns the implicit assumption, rarely tested empirically, that there exists a class of retinal ganglion cells with adequate homogeneity to support the linking proposition. Since it is impractical to record from each of several ganglion cells at once, physiological Mach band experiments are carried out by recording from a single cell, and placing the physical ramp pattern at a series of positions with respect to the cell being recorded. It is then assumed that many such cells exist, and that, if one could present the ramp pattern in a single location and record from each cell in the row in turn, one would see a pattern with maxima and minima at the ends of the ramp. But for this

ϕ "Looks Like" ψ → ϕ Explains ψ

Fig. 6. The Analogy proposition.

explanation to work, one must believe that the cells are homogeneous enough that the relatively small variations across the Mach band pattern will constitute a meaningful perturbation against the variation of firing characteristics from one cell to the next. This is to some degree an empirical question, but it is not often treated in the physiological literature.

The third reservation about this linking proposition is its peripherality: it states a causal or explanatory relationship between the activity of single cells at a peripheral level of the nervous system and a perceptual or behavioral phenomenon, without any accompanying proposal as to how this pattern maintains itself through the system. The proposition includes an implicit appeal to the "nothing mucks it up" proviso (Teller, 1980). But one could well argue that serious hypotheses proposing explanatory relationships between the activities of peripheral neural units and perceptual or psychophysical data should be accompanied by an explicit treatment of the problem of how the rest of the system manages not to muck up the relationship or, put more elegantly, of the constraints that the hypothesis puts on models of the composite map from the peripheral neural level to the bridge locus, and between the bridge locus and phenomenal states.

In the absence of any explicit treatment of these problems, the Mach band proposition would seem to amount to nothing more than a remote homunculus theory: the Mach band stimulus sets up the pattern of activity in the ganglion cells that "looks like" Mach bands, and the homunculus peers down at the ganglion cells through a magical Maxwellian telescope and sees Mach bands.

Many visual scientists acknowledge these problems, but argue that this explanation of Mach bands nonetheless provides an important heuristic for thinking about how Mach bands might come about. But one can argue the opposite: that it is all too common in visual science to find some superficial analogy between the activity of single cells and the results of perceptual or psychophysical experiments, and conclude that the former has something to do with, or even causes or explains, the latter. The crux of the argument concerns precisely the difference between a superficial analogy and an acceptable and well-founded explanation. Problems of range, homogeneity and peripherality would all seem to require further treatment before the value of the analogy can be assessed.

The second specific example to be discussed is the concept of trigger features or detectors, or more broadly, the role of individual cortical neurons in pattern perception. Many neurons in visual cortex seem to be "tuned" to respond with a rapid firing rate to light or dark bars at specific retinal orientations. Bars at non-optimal orientations lead to a lower firing rate, or no change from spontaneous activity levels. Since the discovery of these kinds of cortical cells (Hubel and Wiesel, 1959, 1968) it has been

commonplace in textbook visual theory to presume that these so-called "line detectors" play a fundamental (if not always clearly specified) role in the perception of oriented contours. Visual scientists have also been known to speculate that cells with ever more specific trigger features will be found at higher levels of the visual system, and that activity in these cells will underlie the perception of the specific patterns or objects that make them fire fastest.

The linking proposition involved in such arguments is that *a neuron's role in the neural code is to signal the presence of the stimulus which makes that neuron respond most vigorously*. And, as was the case with the linking proposition used in the Mach band explanation, several reservations again must be expressed. In addition to range and peripherality problems like those discussed above, this linking proposition would seem to be lacking in face validity with respect to reasonable principles of neural coding.

To discuss these problems, it is useful to start with a digression, to introduce the concept of an *Equivalence Class* (cf. Carnap, 1966, Chap. 5; Ratliff and Sirovich, 1978) as it applies to the activities of single neurons. An equivalence class is a set of objects or stimuli which may differ from each other in many ways, but which are rendered identical by a particular measurement operation or classification scheme. All objects that have the same weight are an equivalence class with respect to weight, and all stimuli that lead to the same output from a neuron are an equivalence class for that neuron. Each neuron, as a univariant information channel, collapses a large set of stimuli into a smaller set of neural outputs, and necessarily confounds all members of each equivalence class.

This concept is very familiar in the consideration of photoreceptors. Under steady-state conditions, an individual rod or cone is nothing more than a quantum counter. Many different patches of light of many different wavelengths and wavelength mixtures will lead to statistically equal quantum catches, and thus will yield statistically identical outputs from that receptor. These patches of light are an equivalence class for that cell. Similarly, if one doubles the intensity of each of those patches of light, each patch will each now lead to twice the average quantum catch, putting the receptor in an average state different from that of the first equivalence class; these new patches of light will be a second equivalence class.

Now, a receptor has a broad action spectrum. The probability of absorption of a quantum varies with wavelength, and given a set of stimuli of equal irradiances, there will be one stimulus that leads to the maximum output from the receptor. But there is no sense in which one would want to say that a receptor with maximum sensitivity at 570 nm is specialized to signal 570 nm light. It is not called a 570-nm detector, nor is its trigger feature said to be a light of 570 nm. The emphasis in color theory is rather on the fact that equivalence classes exist for

this cell; that the cell confounds wavelength and intensity information; and that, if wavelength information is to pass through a set of receptors, it must do so encoded in the relative activity levels of different receptor types.

Single simple cells in area 17 are often described as being triggered by, or detecting, the presence of bars of light at particular orientations. The reason that they are characterized in this way is that, of the stimulus set that has been tried, an oriented bar of light seems to make the cell fire fastest. But to say that the cell has a special role in encoding the stimulus which makes it fire fastest is to commit the same fallacy as to say that a cone with maximum sensitivity at 570 nm is a 570-nm detector. It would seem to make more sense to assume that each perceptual element is coded by a pattern of firing among many neurons, and that each different firing rate of each cortical cell is important to the neural code. To understand the part particular classes of cells are playing in the overall neural code, it will be necessary to use a large universe of stimuli, and to explore the equivalence classes for each different level of activity, for each different kind of cell (cf. Boynton, 1979; Daugman, 1980; Hood and Finkelstein, 1983).

In a recent detailed analysis of this kind, Peter Lennie and Anthony Movshon (personal communication) have undertaken experiments to establish whether or not cortical cells are univariant; i.e. whether the response to a stimulus having one spatial configuration is distinguishable from the response to a second stimulus having a different spatial configuration suitably adjusted in contrast. Neurons in area 17 of the cat were stimulated by moving gratings of a range of spatial frequencies and orientations that spanned the optimum for each cell. For each combination of orientation and spatial frequency, gratings were presented at a series of different contrasts, and the times of occurrence of impulses were recorded precisely. By any of several different analyses of the discharge, there existed sets of stimuli, differing in spatial frequency, orientation and contrast, that gave rise to identical responses. Hence, exact equivalence classes could be identified for cortical cells. Doubtless, had the range of stimulus parameters been expanded to include wavelength, field size, etc., larger equivalence classes would have been found. With such information available, one might begin to think more clearly about the roles of cortical cells, and the ways in which univariant cells with different equivalence classes work together to preserve information, just as the three receptor types do in color vision*.

To what extent can the existence of equivalence classes be reconciled with the concept of trigger features or detectors? To use the concept of a trigger feature appears to be to claim implicitly three things: that for any given cortical cell most of the stimuli in the universe are in the null class, i.e. they do not influence the firing of the cell; that the set of stimuli that makes any appreciable response in the cell is small and homogeneous; and that the later elements which receive inputs from the cell ignore variations in the firing rate of the cell and treat the cell as binary—either firing or not firing. These assumptions would seem to be important and empirically testable. But in the virtual absence of data it is hard to set aside the convictions that all of the possible firing rates of cortical cells play a role in the neural code; and that the use of broader universes of stimuli in physiological experiments would reveal the size and heterogeneity of the equivalence classes of neurons. Such findings would re-insert a justified complexity into our thinking about the functions of cortical neurons. If these arguments are valid, they seriously challenge the validity of the notion of detectors and the ultimate value of physiological experiments employing narrowly limited stimulus sets.

In regard to intended range, it is sufficient to point out that the linking proposition involved in the notion of a detector—that the presence of the stimulus that makes a cell fire fastest is coded by the activity of that cell—is not universally invoked. It is not applied to receptors, nor to ganglion cells; and, in fact, it contradicts the Mach band brightness assumption, which includes the idea that the graded firing rates of the neurons across the Mach band pattern encode the graded range of perceived brightnesses. It is also contradictory to one of the fundamental assumptions of opponent process theories of color vision; namely, that the null signal from one opponent channel is a fundamental element of the neural code for the unique hues represented by departures from the null state in the other opponent channel. In short, if the "fastest firing rate" linking proposition is to be saved, its intended range of applicability must be specified and justified on other than an *ad hoc* basis.

In regard to peripherality, it is interesting to note that our views on whether cortical cells are "central" or "peripheral" have changed drastically in the last 20 years. Twenty years ago, the properties of visual cortical cells were still shrouded in mystery. It was easy to think of these cells as being very "central", and perhaps even forming the immediate neural substrate of perceptual events. Today Area 17 is often regarded as being a relatively peripheral stage of neural processing (cf. Marr, 1982). In that case, if the properties of Area 17 cells are to be used to explain the elements of our perceptions, a "nothing mucks it up" theory is needed to bridge the gap between the Area 17 cell and the still more central sites which are now implicitly assumed to form the immediate neural

---

*These arguments would seem to be inconsistent with Barlow's (1972) fifth dogma; i.e. that: "The frequency of neural impulses codes subjective certainty; a high impulse frequency in a given neuron corresponds to a high degree of confidence that the cause of the percept is present in the external world". (p. 381).

substrates of conscious perceptions. This objection would not, of course, count against the use of the "fastest firing rate" linking proposition at the most central levels (i.e. the bridge locus), but the objections of range and face validity would still apply.

In sum, this section has been an attempt to use two common textbook explanations to explicate two common but mutually inconsistent linking propositions, and to begin to generate a set of criteria—range of applicability, homogeneity, peripherality, face validity and mutual consistency—by which one might attempt to evaluate a wider set of linking propositions. The suggestion is also made that the empirical analysis of equivalence classes for central neurons would be a worthwhile undertaking.

CANONICAL FORMS

We come finally to the question of canonical forms. The question is, does one wish to argue that there is a necessary form that the information coded in the visual neurons must eventually take, in order for particular perceptual states to occur? Claims of canonical form involve linking propositions, insofar as they claim that the existence of a particular brain state is a necessary condition for the occurrence of a particular perceptual state.

One area in which the question of canonical form manifests itself may be called the convergence dilemma, or the problem of pattern codes (cf. Barlow, 1972; Hood and Finkelstein, 1983). The question is, if two or more neurons are to act jointly to determine a perceptual state, must their outputs necessarily converge upon a successor neuron whose state uniquely determines the perceptual state? Must the physiological basis for each separate perceptual state be the activity of a different single neuron, or may it be the *pattern* of activity in a set of non-interacting neurons? In brief, if one chooses the first horn of the dilemma and allows the possibility of a pattern code, one must deal with the problem of how two or more cells can have a joint causal role without interacting. The best approach is probably through the concept of emergent properties (e.g. Hofstadter, 1980). But if one chooses instead the second horn, and argues that convergence is necessary, one must be ready to argue that there are enough eligible neurons to go around. The author knows of no comfortable place to hide from this dilemma.

An example from color vision theory may clarify the problem. As an opponent process theorist, one might argue that the joint states of three bridge locus neurons—a Redness/Greenness neuron, a Yellowness/Blueness neuron and a Blackness/Whiteness neuron—determine perceived hue, brightness and saturation in a region of the visual world; and further, that the outputs of three such neurons need not converge upon any other single neuron in order to act jointly to determine the perceptual state. Alternately, one might believe that there must exist a (later) stage of the visual system at which a different neuron would be active for the perception of each different combination of hue, brightness, and saturation, and that until this stage is found the neural correlates of perceived hue will not be properly understood.

A second example may be taken from the "second quantitative law of perception" formulated recently by Barlow (1983): "... the essential work done in perception is to construct a representation using symbols of high relative entropy from sensory messages in which the symbols have low relative entropy because of numerous forms of redundancy". It is not, however, clear that this proposition is a canonical claim, because it is also said to be subject to empirical verification or falsification.

A third example may be taken from the work of Marr (1982). Marr argues (p. 36) that the fundamental purpose of vision is that it "... tells about shape and space and spatial arrangement..." and therefore that (p. 326, parentheses added) "... it is clear in principle that the brain must construct (explicit) three-dimensional representations of objects and the space they occupy.... In order to recognize an object by its shape ... some kind of three-dimensional representation must be built from the image and matched in some way to a stored three-dimensional representation with which other knowledge is already associated". Although the physiological interpretation of the concept of explicit representation is purposely left unspecified (pp. 336, 343), this is clearly an abstract argument that there is a canonical form that visual information must take, and Marr's theory of pattern vision consists of ingeneous speculations about how the information could take that form. To a large degree the value of Marr's formulations to visual science will rest on the question of whether his argument for the explicit representation of form is true, whether it can be given a physiologically useful interpretation, and whether it can be rendered as a specific psychophysiological linking proposition.

As a final exercise, it is of interest to consider the following four questions, all of which involve assumptions about canonical forms. First, why do we see the world right side up when the retinal image is upside down? Second, why do we see the world as non-distorted when the map on the visual cortex is distorted? Third, why do we see a continuous visual scene, with no division at the midline, when left and right visual fields project to separate hemispheres? And fourth (Barlow, 1983), where in the brain are the cues of binocular parallex, motion parallex and texture gradients brought together to add the third dimension to our perceptual world? In modern visual science, the first two questions are usually considered to be pseudoquestions, not in need of answers. The third and fourth are usually considered to be real questions, and the postulated physiological communication pathways and convergence loci are taken as logical necessities. The point is that each of

the four questions involves a linking proposition. Our implicit acceptance or rejection of particular linking propositions forms the basis for very fundamental decisions about which questions require answers, and what properties are to be studied in anatomical and physiological experiments.

TESTABILITY OF LINKING PROPOSITIONS

We now return to the question of the empirical testability of linking propositions. A full treatment of this question remains to be undertaken, but a few brief comments can be made. It has been argued above that there are many kinds of linking propositions, and that different linking propositions take on different roles—as fundamental axioms, broad but non-universal generalizations, tentative assumptions, hypotheses or speculations—in the nomological network of visual science. If this is so, the testability or falsifiability of linking propositions will also vary widely, depending upon the role the specific proposition is playing in a given experimental or theoretical context.

At one extreme lie some general linking propositions dealing with information loss—Identity and Contrapositive Identity—which stand as axiomatic in visual science and its parent disciplines. These would not be subject to empirical falsification without major upheavals in a broad range of scientific disciplines. At the other extreme lie linking propositions to the effect that a specific mapping occurs between the states of particular, well-characterized visual neurons and particular perceptual states; i.e. propositions of the form: human beings perceive X whenever the set of neurons Y is in the set of states Z. Leaving aside the standard problems of anesthesia and cross-species generalizations, such a proposition seems eminently testable. To test it one would make the most inclusive possible collection of objects and stimulus configurations that produce the perception X, along with an equally varied collection that do not; and see whether or not presentation of each of them produces the predicted neural states. There does not appear to be any particular logical difficulty associated with such an approach, as long as the perceptual facts are available, the cell types can be identified unambiguously, and the linking proposition is stated with sufficient clarity and precision.

If such an experiment were carried out, and the proposed correspondence were shown to hold, one might well produce an empirically validated Analogy, with an established, broad range of applicability. The quality and explanatory value of the analogy would remain to be judged, and I have argued above that the criteria for judging the value of such analogies are greatly in need of explication and clarification.

CONCLUSION

Twenty years ago, Brindley pointed out that in using linking hypotheses, visual scientists often intro-duce unacknowledged, non-rigorous steps into their arguments. Brindley's remarks correctly sensitized us to the lack of rigor with which linking propositions have undoubtedly often been used, but led to few detailed, explicit discussions of linking propositions. It would seem useful to encourage such discussions, and to encourage visual scientists to make linking propositions explicit, so that linking propositions can be subjected to the requirements of consistency and the risks of falsification appropriate to the evaluation of all scientific propositions.

Perhaps we should include more often in our publications a paragraph or two making our linking propositions explicit. It will be important to define the ranges of intended applicability of individual linking propositions, the kinds of support they receive from prior experiments, their consistency with other broadly accepted linking propositions, the constraints they place on the composite map, the ancillary assumptions involved and the overall fit of the linking propositions into the current theoretical network of visual science. Within a few years, enough such descriptions should become available that it would be possible to refer to and build upon earlier explications, rather than starting from scratch each time. Similarly, perhaps reviewers could be encouraged to use the explicitness and potential values of linking propositions as one of the criteria of excellence for papers in which arguments containing linking propositions are explicitly or implicitly made.

REFERENCES

Barlow H. B. (1972) Single units and sensation: a neuron doctrine for perceptual psychology? *Perception* **1**, 371–394.

Barlow H. B. (1983) Perception: What quantitative laws govern the acquisition of knowledge from the senses. In *Functions of the Brain*, Wolfson College Lectures (Edited by Coen C.). Oxford University Press, Oxford. In press.

Boring E. G. (1942) *Sensation and Perception in the History of Experimental Psychology*. Appleton–Century–Crofts, New York.

Boynton R. M. (1979) *Human Color Vision*. Holt, Rinehart & Winston, New York.

Boynton R. M. and Onley J. (1962) A critique of the special status assigned by Brindley to "Psychophysical Linking Hypotheses" of Class A. *Vision Res.* **2**, 383–390.

Brindley G. (1957) Two theorems in colour vision. *Q. Jl exp. Psychol.* **9**, 101–104.

Brindley G. (1960) *Physiology of the Retina and Visual Pathways*, 1st Edition. Edward Arnold, London.

Brindley G. (1970) *Physiology of the Retina and Visual Pathways*, 2nd Edition. Williams & Wilkins, Baltimore.

Carnap R. (1966) *An Introduction to the Philosophy of Science* (Edited by Gardner M.). Basic Books, New York.

Crane H. and Piantanida T. (1983) On seeing reddish-green and yellowish-blue. *Science* **221**, 1078–1080.

Daugman J. G. (1980) Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Res.* **20**, 847–856.

De Valois R. L. (1965) Analysis and coding of color vision in the primate visual system. In *Cold Spring Harbor Symposia on Quantitative Biology, Vol. XXX. Sensory Receptors.* Cold Spring Harbor Laboratory of Quantitative Biology, Cold Spring Harbor, NY.

Eccles J. C. (1977) In *The Self and its Brain* (Edited by Popper K. R. and Eccles J. C.), Part II. Springer, Berlin.

Feigl H. (1958) The "mental" and the "physical". In *Minnesota Studies in the Philosophy of Science, Vol. II. Concepts, Theories, and the Mind–Body Problem.* University of Minnesota Press.

Hofstadter D. R. (1980) *Gödel, Escher, Bach: an Eternal Golden Braid.* Random House, New York.

Hood D. C. and Finkelstein M. A. (1983) A case for the revision of textbook models of color vision: the detection and appearance of small brief lights. In *Colour Vision: Physiology and Psychophysics* (Edited by Mollon J. D. and Sharpe L. T.). Academic Press, London.

Hubel D. H. and Wiesel T. N. (1959) Receptive fields of single neurons in the cat's striate cortex. *J. Physiol.* **148**, 574–591.

Hubel D. H. and Wiesel T. N. (1968) Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* **195**, 215–243.

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

Ratliff F. and Sirovich L. (1978) Equivalence classes of visual stimuli. *Vision Res.* **18**, 845–851.

Robson J. (1980) The physiological basis of spatial vision. In *Visual Coding and Adaptability* (Edited by Harris C. H.). Erlbaum Associates, Hillsdale, NJ.

Shapere D. (1977) Scientific theories and their domains. In *The Structure of Scientific Theories* (Edited by Suppe F.), 2nd Edition. University of Illinois Press, Urbana.

Suppe F. (1977) *The Structure of Scientific Theories,* 2nd Edition. University of Illinois Press, Urbana.

Teller D. Y. (1980) Locus questions in visual science. In *Visual Coding and Adaptability* (Edited by Harris C.) Erlbaum Associates, Hillsdale, NJ.

Teller D. Y. and Pugh E. N. Jr (1983) Linking propositions in color vision. In *Colour Vision: Physiology and Psychophysics* (Edited by Mollon J. D. and Sharpe L. T.). Academic Press, London.

Uttal W. R. (1981) *A Taxonomy of Visual Processes.* Erlbaum Associates, Hillsdale, NJ.