An Information Theoretic Approach to Veridical Hallucination

By

Brian Trimmer

B.S. Philosophy (1995)
B.S. Computer Science (1995)

North Carolina State University

Submitted to the Department of Linguistics and Philosophy
in Partial Fulfillment of the Requirements for the Degree of
[Master of Science]
Master of Arts in Philosophy
At the
Massachusetts Institute of Technology
[February 2004]
September, 2003

Signature of author:______________________________
Department of Linguistics and Philosophy
September 8th, 2003

Certified by:______________________________
Alex Byrne
Associate Professor of Philosophy
Thesis Supervisor

Accepted by:______________________________
Vann McGee
Professor of Philosophy
Chair, Committee on Graduate Students

An Information Theoretic Approach to Veridical Hallucination

By

Brian Trimmer

Submitted to the Department of Linguistics and Philosophy
On September 8th, 2003 in Partial Fulfillment of the
Requirements for the Degree of Master of Arts
in Philosophy

ABSTRACT

David Lewis, in "Veridical Hallucination and Prosthetic Vision", outlines his views on seeing. He discusses, by way of several examples, unusual visual conditions and gives explanations of why one does or does not see in those conditions. However, it is not always clear exactly how Lewis' views apply to unusual cases. He also admits that he has made mistakes in applying his criteria to examples, in the Postscript to the original article. However, I think Lewis' ideas are worthwhile and would like to expound upon them. In what follows, I hope to provide clearer criteria that are compatible with Lewis' views, and show how such criteria do or do not apply to unusual circumstances.

The criteria I will use in place of Lewis derive from a branch of signal theory, called Information Theory. Information Theory is a formal calculus for quantifying and computing the information content of a source or a signal carrying information about a source. It is an attempt to formalize an intuitive notion of information that we all work with. The goal will be to look for discrepancies between the information theoretic criteria and Lewis' conclusions, so cases where there is substantial agreement between Lewis and the information theoretic criteria will be only briefly glossed. Clarification of both views can be obtained by seeing how and why they differ and which view is plausibly correct about the case.

Thesis Supervisor: Alex Byrne
Title: Associate Professor of Philosophy

Table of Contents

I. Introduction

David Lewis, in "Veridical Hallucination and Prosthetic Vision", outlines his views on seeing. He discusses, by way of several examples, unusual visual conditions and gives explanations of why one does or does not see in those conditions. However, it is not always clear exactly how Lewis' views apply to unusual cases. He also admits that he has made mistakes in applying his criteria to examples, in the Postscript to the original article. However, I think Lewis' ideas are worthwhile and would like to expound upon them. In what follows, I hope to provide clearer criteria that are compatible with Lewis' views, and show how such criteria do or do not apply to unusual circumstances.

The criteria I will use in place of Lewis derive from a branch of signal theory, called Information Theory. Information Theory is a formal calculus for quantifying and computing the information content of a source or a signal carrying information about a source. It is an attempt to formalize an intuitive notion of information that we all work with. Even Lewis mentions information in outlining his view. For, he describes a visual experience as "matching the scene before the eyes" when the experience has correct *information content* about the scene before the eyes. An experience matches the scene before the eyes only insofar as the information content of the experience is (mostly) correct with regards to the scene before the eyes[1]. Lewis relies heavily on this notion of the scene before the eyes producing matching visual experiences, so an exploration of what exactly this means is worthwhile.

I propose to approach the topic in the following way. Section II will briefly sketch Lewis' views on seeing. In Section III, I will supply a brief technical introduction to Information Theory and outline information theoretic criteria that can be applied to the same kind of problem that Lewis is considering. Section IV will run through the cases presented by

---

1 Lewis, David: Philosophical Papers, v. 2, p 274 All future page references are from the same text unless otherwise noted.

Lewis, in the order he presents them. For these cases, I will discuss how Lewis responds to the case, and how the information theoretic criteria might decide the case. The goal here will be to look for discrepancies between the information theoretic criteria and Lewis' conclusions, so cases where there is substantial agreement between Lewis and the information theoretic criteria will be only briefly glossed. Clarification of both views can be obtained by seeing how and why they differ and which view is plausibly correct about the case. Finally, in Section V, I will turn to a case that Lewis recognizes, but does not really address. I will examine the paradox of seeing that it is too dark to see, and explain how this seemingly paradoxical statement can be disambiguated to resolve the paradox.

II. The Counterfactual Dependence Theory of Seeing

Lewis starts off by making quite clear what notion he intends to analyze. The word 'seeing' has several important and subtle variations in meaning, which can be confused quite easily. The notion of 'seeing' Lewis means to be discussing is the grammatically intransitive sense of the verb 'to see'. We are not here concerned with whether such-and-such a thing *in* the scene before they eyes is seen or not. Lewis is only considering whether or not one sees, and not what in particular one sees while they are seeing[2]. Also, the notion of 'seeing' Lewis is interested in precludes the possibility of seeing what is not a scene before the eyes. One may see, in some sense of the word, that which is not there. Hallucinatory daggers and other such false visual experiences are sometimes described at seeing. However, this is not the sense of 'seeing' Lewis is discussing. For Lewis' purposes, if one "sees that which is not there", then one is not seeing. In other words, if one is genuinely seeing, then the content of their visual experience is correct.[3]

Lewis builds his view progressively, considering some criteria and then applying them to examples to see if they hold up or not. I shall not discuss the examples in detail here, but will go through the steps that Lewis does in reaching his view. A first, intuitive criteria would be that one sees just in case the scene before the eyes causes a matching visual experience.[4] Lewis remarks that this is good enough for most normal cases, but may break down when considering some more unusual scenarios. The example of a veridical hallucination comes up as a counter-example to this case. One may be hallucinating, the hallucination might well be caused by something in the scene before the eyes (i.e. the Wizard or Brain before the Eyes cases), and the visual experience matches the scene before the eyes when interpreted in the normal way. If something in the scene before the eyes causes one to hallucinate, and the

2 However, when we turn to Section V, seeing such-and-such a thing or seeing that so-and-so is the case will be important, since both these senses of "seeing" entail the sense we address here.
3 Or, at least mostly correct. Some amount of error in visual experience must be tolerable even in genuine cases of seeing.

hallucinatory experience matches the scene, then this criteria is satisfied. The scene before the eyes causes a matching visual experience, namely the hallucinatory experience.

Intuitively, the person is *not* seeing in such cases. In these examples, the individual is hallucinating. When someone is hallucinating, that person does not see in the sense Lewis has in mind. Unfortunately, there is nothing about the concept of a hallucination that precludes the possibility of the hallucinatory experiences matching the actual scene before the eyes[5]. A hallucinatory visual experience might very well match the scene before the eyes, even though it is hallucinatory. Distinguishing genuine seeing from veridical hallucination caused by the scene before the eyes drives most of the discussion that follows.

A first attempt at ruling out veridical hallucination is to require that the visual experience be caused "in the normal way". In the examples of veridical hallucination, the visual experience is not caused by light reflected in to the eye of the individual, but by some other means. Indeed, calling the experience 'hallucinatory' suggests that it is not caused by normal visual processing, but by some other means (drugs, blows to the head, psychiatric or neurological disorders, etc...) This, as Lewis points out, rules out too much. For one thing, it rules out the possibility of artificial eyes or other means of vision replacement. It also means that creatures that do not see in the same way that we do (the "normal" way), in fact do not see at all. Again, this is intuitively wrong. One would hope that medical science will some day be able to restore sight to the blind by means of prosthetic eyes or other means that do not involve the same kind of processes that go on in normal sighted humans. One would also hope that extraterrestrial creatures are in principle capable of sight, even if the mechanisms by which their visual experiences arise are different from those in human beings.

4 p273

5 If you find the notion of an experience "matching the scene before the eyes" to be problematic, the problem of veridical hallucination can be re-cast in terms that do not rely on matching the scene before the eyes. If the hallucinatory experience matches the experience you would have had, were your experiences formed in the way they normally are, the same problem arises. There should be no difficulty in the notion of ones actual experience matching a counterfactual experience.

Another attempt to rule out veridical hallucination is based on the fact that any matches between the scene before the eyes and the hallucinatory experience would be rare. One would not expect hallucinatory experiences to match the scene before the eyes very often. Few, if any, of the hallucinatory experiences would match the scene before the eyes. Again, this will not do. In principle, one could have a very long running series of veridical hallucinations. Even though longer series of veridical hallucinations are considerably more improbable than shorter series, the probability never reaches zero. So it *could* happen that someone had a very long run (maybe even a life-long run) of veridical hallucinations. It still seems that one is not seeing if one's experiences are hallucinatory[6]. Also, as Lewis points out in the Deathbed Cure example, one may be granted sight for a very brief period of time. The fact that the period of sight is very short, and so produces only a few matching visual experiences, does not seem to detract from it being a genuine case of seeing. There is no reason to think that creatures with a life span of a few seconds can not see, for no other reason than the fact they live such a short time and so would have very few visual experiences.

Lewis then moves to his proposal for ruling out veridical hallucinations. This is the counterfactual dependence criteria of seeing. Not only must the scene before the eyes cause a matching visual experience, it must also be the case that were the scene before the eyes to be suitably different, the visual experience would be different in similar ways [7]. It is not at all clear that in cases of veridical hallucination, a counterfactual situation in which the scene before the eyes was different would result in a different visual experience. Indeed, changes to aspects of the scene which have no causal role in producing the experience should have no impact whatsoever on the visual experience produced. This is most or all of the scene in cases

---

6 Intuitions on this may vary. It may be incorrect to describe someone who has a life-long run of veridical visual 'hallucinations' as not seeing. They may simply be seeing in some non-standard way. (i.e. The Minority case).

7 Lewis does not address the question of what constitutes a "suitable difference" is, claiming that this is an empirical matter. Lewis claims the only thing his analysis provides is the assertion that there must be many "suitably different" scenes available to discriminate. Once we have the theoretical machinery from Section III, I

of veridical hallucination. The match in cases of veridical hallucination is accidental. The scene might have been quite different without affecting the experie nce in any way. This also allows us to properly explain the cases ruled out by the other two attempts. Artificial eyes or momentary restorations of vision count as cases of seeing, since in those cases a counterfactual scene would have resulted in different visual experiences.

However, Lewis is not quite done yet. Not just any old counterfactual dependence between the visual experience and the scene before the eyes will do. For one thing, the dependence must not be a backtracking dependence from the experience to the scene. In other words, the scene must be the cause of the experience, rather than the experience the cause of the scene. This is needed to handle The Screen case that Lewis discusses. We shall look at the example in question later on. Briefly, if some situation were to arise in which an individual's visual experiences were causally responsible for matching changes in the scene before the eyes, it seems unusual to describe this as a case of seeing. Seeing is generally passive with respect to the scene before the eyes; visual experiences resulting from seeing what is before you are caused by the scene, rather than causing that scene to come into existence.

Furthermore, Lewis suggests that there must be a "large class of alternative possible scenes (that might have been) before the subject's eyes".[8] This large class must be divided into many mutually exclusive and jointly exhaustive subclasses, such that any two scenes in a sub-class produces identical visual experiences which closely match the scenes, and scenes from different subclasses produce different visual experiences.[9] Ideally, there would be precisely one scene per subclass, but this is too much to ask for. The subclasses, however, should be considerably smaller than the overall class. Thus, the requirement that there be

think we can do away with terms like 'many' or 'suitably different' alternatives and other terms which are not well defined.

8 p283

9 p283

many subclasses. In order to ensure the subclasses are small, and yet cover the large space of possible scenes, there have to be many of them. The questions that immediately arise are (a) how large must the class of alternate scenes before the eyes be, (b) how many subclasses must there be, and (c) how small do the subclasses have to be?[10] Again, Lewis defers on questions regarding these criteria. He simply stipulates that the criteria are vague, and that there are bound to be borderline cases.

This seems to be Lewis' final suggestion for criteria on whether or not one sees. One sees just in case the scene before the eyes is the cause of a matching visual experience, where a visual experience matches a scene jus t in case the information content of the experience with respect to the scene is largely correct, and the visual experience is appropriately counterfactually dependent (no backtracking) on the scene before the eyes, such that were the scene to be in a different subclass of the large class of alternate scenes, the experience one has would be different.

---

10 p283

III. Information Theory and Seeing

A. Preliminaries

I now turn to a discussion of Information Theory. Information Theory is one branch of a larger engineering discipline known as signal theory. It was compiled into current form by a Bell Telephone company engineer [11], for use in evaluating equipment needs for telephone systems. It is a theory that is used to describe various aspects of information flow, including how much information is generated by an event (also referred to as a *source*), how much information is transmitted to another location about an event, and how much information a particular transmission system is capable of sending. This theory has extensive practical use in the engineering of transmission systems, which need not concern us here. Rather, for our purposes, we will focus on what information is and how to quantify how much information there is.

Information theory typically assumes the following kind of set-up: There is some type of event about which we want to know things. There are two or more distinct possible outcomes for this event[12]. Upon completion of the event, one possibility is realized. Of course, after an event realizes one of two or more possible outcomes, the unrealized outcomes are no longer possible. After outcome A is in fact the outcome that occurs, it is no longer possible that outcome B or C will in fact be the outcome that occurs. The removal of these other outcomes from the set of possible outcomes is what we shall call 'information'.
A situation in which one or more formerly possible outcomes are removed from the set of possible outcomes is a situation in which information is generated[13]. Such an event is referred

---

11 c.f. Shannons The Mathematical Theory of Communication

12 The possible outcomes I will discuss with respect to seeing are in general the same as the alternate scenes Lewis refers to. I will not consider whether or not there are possible outcomes which are, physically or logically, impossible included in the set. For now, assume all possible outcomes or alternate scenes are physically possible.

13 Technically, this is not quite right. Information is generated when the probability distribution among the possible outcomes shifts further from a flat distribution(each possible outcome is equiprobable), and thus closer to a boundary distribution (single outcome with a probability of 1, all others probability of 0). Thus an event in

to as a *source* of information. The realized outcome is the source state.

Just as information can be generated, it can be transmitted from one location to another. A transmission system, for information theoretic purposes, is the following: Suppose we have some apparatus - a device with some manner of detector at one end and an indicator at the other end. A detector is something which is sensitive to the outcomes of some type of event. A detector takes on different states, depending on the state of the event it is a detector of. A particular detector state occurs just in case the outcome of the detected event is one particular outcome among the possible outcomes of the event. The responses of the detector travel (by wires, radio waves, or what have you) to the indicator, causing the indicator to realize different states based on the responses of the detector. The indicator realizes some particular state if and only if the detector responds in a particular way, and the detector in turn responds in that way if and only if the particular type of event the detector responds to realizes a particular outcome. Such an apparatus is referred to as a *transmission system*. The state of the indicator (called the 'output state') indicates the state of another event (the source). The output state may be very far removed in space-time from the source about which it carries information. This is usually the case in the more useful transmission systems we have, such as telephones, television, and photographs of long ago events. It need not always be so, such as with microscopes or medical imaging devices. Human visual perception falls in to this latter category, as we shall see later on.

One thing to notice is that transmission systems also fit the definition of a source; they are events that realize one of two or more possible outcomes. The crucial difference between a source and a transmission system is that an ideal transmission system generates zero additional information. Any information content present in the states of the indicator is information about the source, not about the apparatus or the indicator itself. In the ideal case,

---

progress may generate some information, even before the final outcome is realized. In any event, this is accurate enough for the purposes of this paper.

given the source is in a particular state, the output state of the transmission system could not have been anything other than what it actually is. If there are no other possible states the transmission system could be in, the transmission system generates no information. There must be two or more possible outcomes for an event realizing one outcome to generate information. This is not the case in an ideal transmission system. [14]

At least two questions arise: (a) How much information does a source generate, and (b) how much of that information is available in the output state of some transmission system? There are other questions we might have as well. How much information can a system transmit[15]? Is it even possible for the system to transmit all the information generated by some source? How good is our transmitter? Does the transmission system really generate zero information? If not, how much of the information in the output state is generated by the transmission system, and how much by the source? All of these questions, in one way or another, rely on comparing "how much" information content there is. Some system is needed to allow us to compare information quantities. Information Theory is an apparatus for providing a quantitative answer to these questions. We will see shortly that the quantity of information is proportional to the number of possible outcomes of an event. Since this can be an extremely large number for some kinds of events, Information Theory represents quantities of information as a logarithmic value of the number of possible outcomes. For base 2 logarithms, the unit of information quantity is called a "bit".

I now turn to formulas for calculating how many bits of information are generated by an event, or transmitted by a transmission system. Readers who do not want to read through the formal calculus may skip ahead to Part D of this section, without severely impacting their

---

14 Unfortunately for us, actual transmission systems typically do generate some amount of excess information. When the output state of a transmission system carries both information about the source and information generated by the transmission system, we have a so-called noisy signal. 'Noise' is the term used to refer to any information content of an output state that is not generated by the source. 'Signal' is used to refer to the information which *is* generated by the source. The Signal to Noise ratio, or S/N ratio, indicates how good a transmission system is. Systems with a higher S/N ratio are better transmission systems. An ideal transmission

comprehension of the rest of the material. I include the calculus for completeness sake, since I will be discussing quantity of information at points later in the paper. However, no calculations or attempts to quantify the quantity of information in actual scenes before the eyes is attempted.

B. Discrete Case formulas

Consider the following as a starting point. The quantity of information generated by an event where one of a discrete number of possible and equally likely outcomes occur is N bits, where N is equal to the base 2 logarithm of the number of possible outcomes *before* the actual outcome of the event occurs[16]. This is written in mathematical notation as:

[1] $N=\log2(X)$

where X is the number of possible outcomes. So, flipping a fair coin generates one bit of information, since there were two possible outcomes of this event that might be actual before the coin landed. N=log2(2)=1 bit. Three fair coin flips in a row, taken as a single event where each sequence of heads and tails is considered different[17], generates 3 bits of information (8 possibilities: N=log2(8)=3 bits). Rolling a fair 6-sided die generates 2.585 bits of information (6 possibilities: N=log2(6)=2.585 bits). Drawing one card from a fairly shuffled deck of 52 distinct cards generates 5.7 bits of information (52 possibilities: N=log2(52)=5.7 bits), and so on.

Of course, this formula for information quantity has very restricted uses. It is only

---

system has a S/N ratio of infinity.

15 The amount of information a system can transmit is called the 'bandwidth' of the transmission system.

16 Unless mentioned, all of these formulas apply to calculating the amount of information transmitted by some transmission system as well. For a transmission system, replace "generated" with "transmitted", "event" with "system" and "outcomes" with "output states" to describe a transmission system, rather than an event about which information is to be gathered.

17 In other words, the sequence 'Heads, Tails, Heads' is different than the sequence 'Heads, Heads, Tails'. If particular sequences are *not* considered different, such as if any sequence containing 2 heads and 1 tail were counted as the same, then this composite event would generate less than 3 bits of information. We shall see how to calculate how much information is generated by that case shortly.

applicable when the following three conditions apply:

[a] there are a finite number of possible outcomes

[b] each of those outcomes is equally likely to occur

[c] one particular outcome is clearly indicated as the result

We require a more general formula for the computation of the information generated by a particular kind of event or transmitted by a system. I shall address how to eliminate each of these three conditions, but not in the above order. Let us first address the question of non-eqiprobable outcomes (condition [b]). Notice that Formula 1 [N=log2(X)] can be transformed by law of logarithms into this equivalent formula:

[1a] N= -1 * log2(1/X)

where X is the number of possible outcomes before the event completed. Also notice that, in the case where each outcome is equally likely, (1/X) is equal to the probability of a particular outcome occurring. Using probabilities, however, presents a problem. If the outcomes are *not* equally probable, then what probability value ought to be used in the formula? If we use the probability for outcome A, does this sayanything about the information generated when the outcome of the event is B? In the case where the outcomes have different probabilities, the information generated by an event varies, depending on the probability of the resulting outcome. So, the formula for the information generated by a particular outcome of an event is

[1b] Nx = -1 * log2 (P(X))

when P(x) is a function from outcomes of an event to the probability of the specified outcome. In the discrete case, this function is called a probability distribution function, henceforth abbreviated as 'p.d.f', for the event. X is a particular outcome, which is usually a number representing the position of the outcome in a list.

Now we know how to calculate the information generated by a particular outcome of an event. But what we want to know is how to calculate the average amount of information generated by an event, regardless of how it turns out. To do this, we weight the information generated for each possible outcome by multiplying the quantity of information produced for each possible outcome by the probability of that outcome Then we add all the weighted information values up. In formal notation, when there are a discrete number of possible outcomes [a] and one outcome is clearly the result [c]:

$$[2]\ N = \sum_{X=1}^{M} P(X)*Nx$$

where P(x) is a probability distribution function, and Nx is Formula [1b]. Substituting [1b] for Nx yields:

$$[2a]\ N = \sum_{X=1}^{M} P(X)* -1 * \log2(P(X))$$

Sigma is the standard mathematical symbol for summation. The notation '$\sum_{X=1}^{M} F(X)$ is read as 'The summation from X equals 1 to M of F(X).' It is shorthand notation for the sum F(1)+F(2)+F(3)+....+F(M) [i.e. the value of the function F(X) at every integer from 1 to M inclusive]. The numbers 1,2,3,...,M indicate the positions of possible outcomes for the event on an enumerated list. The same formula can be used for a countable infinity of possible outcomes (M=∞), although the calculus of infinite summations will need to be deployed in evaluating such sums.

Let us now turn to removing condition [c]. Since we are talking about information loss, this restrictions only applicable when discussing the amount of information transmitted by some system.[18] Suppose that there is an event which has X possible outcomes. We have a transmission system that sends information about the outcome of the event to another location. However, the output states of our information-carrying system eliminate only some possibilities for the outcome. One case where this happens is when the number of output

---

[18] I assume that when an event ends, the result is one particular outcome, and all others are no longer possible. If this is not true for some events, then what follows may apply to such sources. It may also apply to events with

states is less than the number of source states. For example, suppose we have some system that turns a light bulb on or off depending on the outcome of a dice roll. A dice roll has six possible outcomes, while the light bulb system only has two possible output states (on or off). Such a system cannot unambiguously indicate which of the six possible source states was realized. There must be at least six distinct output states to do this[19].

Suppose that, in the above system, the light turns on if the outcome is 1, 2 or 3. It remains off if the outcome is 4, 5 or 6. Of course, after the state of the light bulb is set (the output state), it will not be the case that one of the possibilities now has a probability of 1. However, in order to remain probability values, it must still be the case that the sum of the p.d.f. over all outcomes is 1; it will simply be that some of the 'outcomes' (i.e. the ruled out ones) now have a probability of 0, and so no longer help to satisfy the requirement that the p.d.f. sum up to 1. The probability of some or all the remaining outcomes must increase to ensure the p.d.f. satisfies this requirement. This means that, after our output state occurs, we have a *new* probability distribution among the remaining possibilities (the ones with a non-zero probability), which must still add up to 1. With this new p.d.f., we can calculate how much information we would gain, were we to have an additional output state that ruled out all but one of the remaining possibilities, in exactly the same way we did before. Since the output states in our above example do not carry this extra information, this value represents how much information about the source is *missing* from the output state of the transmission system. In other words, when there are a discrete number of possible outcomes [a],

$$[3]\ N = \left[\sum_{X=1}^{M} -1 * P(X) * \log2(P(X))\right] - \left[\sum_{Y=1}^{L} -1 * Q(Y) * \log2(Q(Y))\right]$$

Where 1,2,..,L represent positions on a numbered list of the remaining possibilities after

intermediate stages, where some but not all of the final outcomes are eliminated at some intermediate step.

19 Again, this is not quite right. It is true only when there is no redundancy in the source states. We need not consider this problem for the purposes at hand

output state occurs, and Q(Y) is the p.d.f. over those remaining possibilities.

C. Continuous Case formulas

Now, we turn to the remaining condition [a]; infinite possible outcomes. I mentioned, in discussing how to eliminate condition [b], that we could handle a countable infinity of possible outcomes using the calculus of infinite sums. All that remains to completely eliminate condition [a] is to deal with non-countable infinite sets of possible outcomes. Things are more complex when there is a non-countable infinity of possible outcomes, however. Unlike the discrete case, where the probability of a particular outcome was the value of the p.d.f. at that outcome, the p.d.f. and the probability measure function (the function that, given a range of outcomes, yields the probability of the outcome lying in the supplied range) yield different values when applied to single outcomes in the continuous case. By definition, the probability of one particular outcome, or any finite or countably infinite set of possible outcomes, occurring from a non-countable infinity of possible outcomes is zero. This is easy enough to see. In the discrete case, probability theory requires the p.d.f. to sum up to 1 (i.e. $\sum_{X=1}^{M} P(X) = 1$), and the probability of the actual outcome being one member of some set of outcomes (even if the set includes only one outcome) is the sum of the p.d.f. values at those outcomes[20]. A similar restriction applies to continuous probability distributions. The difference is that, for a continuum of possible outcomes, the definite integral of the p.d.f. over the set of all possible outcomes must equal 1, and the probability of any subset of outcomes is the definite integral of the p.d.f. over those outcomes. Since the definite integral of a function at points is always zero, the probability value assigned to any countable subset of a non-countable set of outcomes must also be zero. We can only meaningfully calculate the probability of the outcome lying within a range. However, the

20 Throughout this paper, I am assuming that the probabilities are independent. Thus, the probability of a source realizing either outcome X, Y or Z [P(X or Y or Z)] is equal to P(X)+P(Y)+P(Z).

p.d.f. value of a particular outcome is *not* zero at every point. Indeed, it cannot be zero at *every* point, since it must integrate to 1 over the range of outcomes. Thus, the value of a continuous p.d.f. at a particular outcome is not generally the probability value of that outcome.

So, by parity of construction (replacing summations in Formula 3 with integrals), we get the following as the corresponding formula for the quantity of information content in a continuous case:

$$[4]\ N= \int_{A}^{B} -1 * P(X)*\log2(P(X))\, dX \quad - \quad \int_{C}^{D} -1 * Q(Y)*\log2(Q(Y))\, dY$$

A and B replace 1 and M respectively, in the first half of formula [3]. C and D replace 1 and L respectively, in the second half of [3]. There is one very important difference between this and the discrete case, however. Unlike the discrete case, where the second summation drops out (has value zero) when there is one particular outcome which is clearly the final outcome, we cannot drop the second integral in the formula and say that the average information generated by a source, or transmitted by a system that indicates one definite outcome of an event with a non-denumerable infinite set of outcomes is

$$[4a]\ N= \int_{A}^{B} -1 * P(X)*\log2(P(X))\, dX$$

The problem with [4a] is that we can not assign to one particular outcome of an event, with a non-countable infinity of possible outcomes, a probability of one without violating the restriction on p.d.f.'s that they must equal one when integrated over the whole domain of possible outcomes. To demonstrate this, I offer a reductio argument. Suppose there were a p.d.f. for the case of one particular outcome of a non-countable infinity of possibilities being actual. If one particular outcome of the non-countable infinity is certainly the outcome (probability 1) and all other possible outcomes are certainly not the outcome (probability 0), then the definite integral of the resulting p.d.f. over any range of outcomes containing that particular outcome must have value 1, and the definite integral of that p.d.f. over any range of outcomes *not* containing that particular outcome must have value 0. In particular, when integrated over the range containing only that one particular outcome with a probability of 1, the definite integral must equal 1. However, any definite integral over a single point has a value of zero (c.f. previous paragraph). This contradicts the assertion that the definite integral over any range containing the particular definite outcome has a value of one. A similar reductio can be constructed for reducing the possible outcomes from a non-countable infinity to any finite number, or even a countable infinity, of remaining possible outcomes. Thus, we cannot simply drop the second integral, without getting a meaningless result. Removing the second integral assumes that the p.d.f. is a function which has the above mentioned contradictory properties. We can only calculate the information quantity of a state that reduces the final outcome to a member of some non-finite subset of the original possible outcomes.[21]

I will now show that, from the general formulations in [3] and [4], the more specific formulations follow as special cases. Take Formula [1], applicable to cases where there is a finite number M of possible outcomes, each outcomes is equally likely and all but one possible outcome has probability zero after the event occurs. We can derive our initial formula for this veryspecialized case from the general discrete formula:

$$N = \left[\sum_{X=1}^{M} -1 * P(X) * \log2(P(X))\right] - \left[\sum_{Y=1}^{L} -1 * Q(Y) * \log2(Q(Y))\right] \qquad \{\text{Formula [3]}\}$$

21 Intuitively, to learn that one particular outcome from a non-countable infinity of outcomes is actual would require a non-countable infinite amount of information. We almost certainly never have that much information available about anything, so the problem with the mathematics should never cause difficulties in actual cases. Whether or not this presents problems in principle for information theory is not a topic I will address here.

$$= \left[\sum_{X=1}^{M} -1*1/M*\log2(1/M)\right] - \left[\sum_{Y=1}^{1} -1*1*\log2(1)\right] \qquad \{\text{substitute constant p.d.f's}\}$$

$$= M*[-1*1/M*\log2(1/M)] - [-1*1*0] \qquad \{\text{replace summations with multipliers}\}$$

$$= -1 * (M/M) * \log2(1/M) - 0 \qquad \{\text{simplification}\}$$

$$= -1 * \log2(1/M) \qquad \{\text{simplification}\}$$

$$= \log2(M) \qquad \{\text{law of logarithms}\}$$

D. Information Theory application

I now turn to the application of Information Theory to the issue of seeing. How can Information Theory be put to use in providing criteria for whether or not an individual sees? Let us start with the following details about a typical case of seeing. There is some actual scene before the eyes. There is also some large number of counterfactual scenes that might have been before the eyes. Since there are several possible scenes that might have been before the eyes, only one of which is realized, an actual scene before the eyes generates information. An actual calculation of the information content of the scene before the eyes would be difficult to carry out, since one would have to know the p.d.f. for the possible scenes before the eyes. However, it should be clear that there is a p.d.f. for possible scenes before the eyes, and that not all possible scenes need be equally likely[22]. So, an actual scene before the eyes, since it rules out alternative possible scenes, is a source of information. There is also a system (i.e. the human visual system) by which information about the scene is carried by way of my visual experiences. The properly functioning visual system is a transmission system, with detectors that respond to the scene before the eyes, and indicators in the form of visual

22 For example, scenes which include a live Bengal tiger sitting on my office desk are less likely than scenes which include a pile of papers sitting on my office desk. I will not explore the details here, but there is some correlation between the probability of a possible scene being actual and the remoteness of counterfactual worlds from the actual world. Possibilities realized in more remote counterfactual worlds (such as the world in which there are live Bengal tigers roaming about Cambridge, MA and wandering in to offices) are less probable than those realized in worlds nearer to the actual world (where there are no live Bengal tigers roaming about

experiences

This transmission of information from the scene to the content of the visual experience seems to be the crucial element for a criteria of seeing. One sees just in case their visual experiences are the output states of an information-transmitting system whose source is the scene before the eyes. Again, an ideal visual system is one which generates no new information; all the information present in some visual experience should be information transmitted from the scene before the eyes. This is almost certainly not the case for actual human visual systems. But, the question of how good an actual visual system is need not concern us here. I will assume an ideal visual system (which adds no new information) in the discussion to follow.

We can answer the question of whether or not an individual sees by answering the question of whether or not the individual's visual experiences are output states of a transmission system, which has the scene before the individual's eyes as the source. In addition, our answers to the above question can be finer grained than merely a 'yes' or 'no' answer. We can, in principle if not in practice, describe *how well someone sees.* You and I may both be seeing, and yet you see much better than I do. Both our visual experiences are output states, carrying information about the scene before our respective eyes, but your visual experiences carry *more* information than mine do. It is unclear to me that Lewis appreciates this about the grammatically intransitive sense of 'seeing' he is concerned with, and I believe this leads him to make some mistakes in applying his criteria to the cases he brings up.

Some additional notes before I move on to the examples that Lewis presents. We should be clear that the objective is to provide a definition of seeing, such that we can decide whether or not a particular individual sees. This definition will almost certainly differ from conventional use of the grammatically intransitive sense of seeing. In general, for it to be

---

Cambridge, MA)

conversationally appropriate to say that someone sees, the individuals visual experiences must do more than just carry some amount of information about the scene before their eyes. The experiences must carry 'enough' information, in some poorly defined sense of 'enough'.

Consider the example of an individual who is so near sighted that he cannot discern anything about the scene before his eyes further away than the tip of his nose[23]. For objects less than a nose length away from his eyes however, he can see well enough to read and forge signatures on documents. The correct general usage of the word 'seeing' would probably be to say this person cannot see. Yet, if he can duplicate signatures well enough to be an accomplished forger, how can he do this without seeing? If pressed, I think most people would agree that he does see, but sees extremely poorly. So poorly that in typical conversation, we ought to just say he cannot see. The sense of seeing we are interested in is the former one - the one that people will typically acknowledge when pressed with examples like these.

Another advantage over Lewis is the lack of dependence on an undefined 'large' number of alternate scenes or 'many' subclasses. While the requirement of a large number of alternate scenes may work for most cases, it raises problems in the case of impoverished visual environments. If there are not a 'large' number of alternate scenes, it seems Lewis would have to describe the individual as not seeing. It seems odd that a sighted person could be rendered blind, simply by reducing the number of alternate scenes that might be before their eyes[24]. Information Theory only requires there to be two scenes in the class of alternates, one of which is the actual scene before the eyes. In such situations, the person may still be seeing, although they see very poorly. Their poor sight is merely a result of the fact that there is little to see. But, having very little to see should not be enough to render one sightless. Only the case where there is quite literally nothing to see, can one be rendered

23 The example I have in mind is a character from the 1963 Steve McQueen movie The Great Escape.
24 The exception, of course, is when the number of possible scenes is reduced to one.

sightless by changing the environment. It may not be possible for there to be a situation in which there is nothing to see (no alternate scenes to the actual one), so perhaps not much turns on this. In any event, I find the number 'two' to be a much easier number to comprehend than the number 'large'.

IV. Lewis versus Information Theory: The Examples

A. Preliminaries

The title of this section is misleading. I do not mean to suggest that Lewis and Information theory are in opposition; indeed, I find there is a great deal of overlap between the two. However, I think a few points are clearer in Lewis' way of describing whether one sees or not, while many other points are clearer with an information theoretic criteria of seeing. My intention here is to compare the two criteria, clarifying points and pointing out possibly misleading cases. Lewis himself, in the Postscript to "Veridical Hallucination…", suggests he has misapplied his own criteria to some of the cases in the original paper. I wish to run through the cases presented by Lewis in that paper now, considering what Lewis says about them and comparing it to what an information theory account might say. Some disagreements will arise, and I will look at these as they come up in the discussion.

The first two cases do not require much examination. They are clearly cases in which the subject does not see. Neither of these examples is a situation in which the scene currently before the eyes is a cause of the visual experience the subject has, and so are not cases of seeing by Lewis' criteria. In the first case (The Brain), the visual experiences is not of something in the scene before the eyes. The second example (Memory), although the visual experiences are experiences of scenes before the eyes, they are not caused by the *current* scene before the eyes.

This is an important distinction to make. Visual experiences produced in the Brain or Memory cases may very well be output states of a transmission system, from an information theoretic view. The problem is not that the experiences are not output states of an information transmission system (although they may not be), but that the source of the information is not the scene before the eyes at all, or a former scene before the eyes, in these cases. One sees just in case their visual experiences are output states of a transmission system with the scene *currently* (when the experience happens) before the eyes as the source[25]. Lewis also restricts seeing to the current scene before the eyes. In both cases, any information content of the experience is not transmitted from the current scene before the eyes, because the processes that produce the experience operate independently of the current scene before the eyes. Thus, they are not cases of seeing.

B. Veridical Hallucinations

The next three examples are more difficult for Lewis, since in all three cases, something in the scene before the eyes *does* cause the visual experience. Two of these examples (Brain Before the Eyes and The Wizard) are what Lewis calls "veridical hallucinations": Hallucinatory experiences that happen to match the actual scene before the eyes. Furthermore, since something in the scene before the eyes is the cause of the experience in all three cases, Lewis must add his counterfactual analysis to adequately explain these examples. None of these three cases is a case of seeing, according to Lewis, because they are not cases in which the causal connection between scene and experience has the right counterfactual structure. Counterfactual scenes would not have produced correspondingly different experiences.

Similar results are obtained with an Information Theoretic account of seeing. Not just any causal connection between a source and an output state is enough to make a transmission system. It is built in to the concept of a transmission system that there be the kind of counterfactual dependence that Lewis requires for cases of seeing. The Brain Before the Eyes is not a case of seeing because, although the cause of the experience is the scene before the eyes, the process by which the scene before the eyes causes the experience is not a

25 We shall see shortly that things other than human eyeballs may count as 'the eyes' for purposes of determining the scene before the eyes. Watching a live TV broadcast probably should count as a case of seeing, where the scene before the 'eyes' is the scene before the TV camera lens.. Watching a videotape will not count as a case of seeing, for much the same reasons as Memory is not a case of seeing.

transmission system. It is not clear at all that the output states of this process are sensitive to the scene before the eyes in the way that a transmission system must be. Indeed, it is not clear that any change in the scene before the eyes would result in a different experience. If so, then there are no alternate output states possible and so no information carried by the output state. Similarly for the Wizard. The process by which the wizard causes the visual experience is not a process which transmits information about the scene before the eyes, since the process is not sensitive to the scene before the eyes. Suppose the wizard were casting his spell from behind the subject, or off to the side of the subject, or in front of a completely different field of background objects. There is no reason to think the visual experience would be any different, and so no reason to believe the subject's visual experience carries any information about the current scene before the eyes.

The light meter example is slightly more difficult. In this case, the visual experience of the grassy field is caused by the scene before the 'eyes', which in this case is the light meter responding to the scene. The process does appear to be one which transmits information about the scene. The light meter system is sensitive to the light levels of the scene before it, producing different output states when there are different light levels in the scene. I think perhaps Lewis dismisses this example a bit too quickly. It does seem to share many features of a genuine case of seeing. In particular, it could be viewed as a kind of prosthetic eye, and Lewis *does* think that one can see with a prosthetic eye. Lewis' restrictions on the counterfactual dependence, regarding the size of the class of alternative scenes and the number of subclasses that produce different visual experiences is deployed to handle this case, ruling it as a case of not seeing.

I find the Information Theory explanation to be better. I believe the correct answer to the Light Meter case is that it is a case of seeing, albeit extremely poor seeing. So poor that, for most purposes, it would be appropriate to say that a person who has visual experiences produced in this way is not seeing. Again I think this may be appropriate for typical conversation, but technically wrong. A person with such a visual apparatus does see, if only a little, about what is before their eyes.

Perhaps the temptation to classify this case as a case of not seeing is that the light meter only transmits information about the brightness of the scene before the eyes. It is insensitive to the color or shape of objects contained in the scene, since the light meter detector is not sensitive to those aspects of the scene. Thus, the experiences produced in the Light Meter case do not carry information about those aspects of the scene. The image of the grassy field, when interpreted 'in the normal way', would lead to a misrepresentation of the scene before the eyes most of the time.

However, Light Meter visual experience is sensitive to *some* aspects of the scene (the brightness). Because of this, the light meter should be viewed as a type of poor prosthetic eye which allows one to see just a little bit about the scene before the light meter. We might vary example, say using a dense grid of tiny light meters that each affected a small part of the overall visual experience. It is more tempting to say that the subject does see in these cases, because more information about the scene is transmitted with the proposed modifications. Again, I think this is just a linguistic convention. It should not be mistaken with a criteria for seeing. A dense grid of tiny light meters allows one to see more - to see better - than one can see with the single light meter. The difference seems to be only a matter of degree. Both light meter cases are cases of seeing, or neither is. Furthermore, arguing that neither is a case of seeing seems to rule out too much as seeing, including normal human vision[26].

C. Prosthetic Vision and Broken Prosthetic Eyes

26 The human eye is not much more than a dense mass of light meters, with a lens to allow focusing. Some receptors respond to specific frequency ranges in the 400nm to 700nm wavelengths, while others respond to all the radiation in these wavelengths. The difference between human vision and the light meter seems to be just a matter of degree as well.

The next two examples (The Minority and The Prosthetic Eye) are, as Lewis argues, genuine cases of seeing. They are, however, cases which do not involve the same kinds of causal mechanisms that go on in normal sighted human beings. Lewis argues that both these examples are cases of genuine seeing, by pointing out that they are processes by which the visual experience is causally connected to the scene before the eyes, and the connection has the right kind of counterfactual dependency. Information Theory, too, would include these as cases of seeing. So long as the causal process in question is a transmission system, whose detector is sensitive to the scene before it, both examples meet the Information Theory criteria for seeing. No big mysteries here.

The next example (The Deathbed Cure) is problematic, and the example after that (The Loose Wire) presents a great deal of difficulty. I group these two examples together because The Deathbed Cure is the extreme case of The Loose Wire example, where the wire contacts only once. I will focus my remarks on the Loose Wire example. In this case, the subject has a prosthetic eye, which is functionally equivalent to a genuine human eye [27]. However, there is a loose wire in the prosthetic eye. When this wire comes off the correct contact point, the eye malfunctions, producing visual experiences consisting of random splotches of color. When the wire touches, however, the eye functions as it is supposed to. Lewis suggests that "Whether [the subject] sees has nothing to do with whether the wire touches the contact often, or seldom, or only this once."[28] He takes this to be a case where the subject does see, at least when the wire makes contact.

However, it is not obvious at all from Lewis' remarks why he should say this. In fact, applying the counterfactual dependence criteria he outlines, one might conclude that he should say that the subject does not see if the wire contacts rarely or only once. For, it seems

[27] For our purposes, take 'functionally equivalent' to mean that if your genuine eyes were replaced with the prosthetic version without your knowledge, you would not notice anything different with your vision.
[28] p 281

plausible that probability comes in to the picture for this example. Rather than a strict dependence or independence of the experience on the scene before the eyes, it appears as though there is a probability of the experience being counterfactually dependent on the scene. And, this probability is influenced by whether the wire touches frequently, or seldom, or only this once. If the wire touches most of the time, then there is a high probability that the experience is counterfactually dependent on the scene. If the wire touches seldom, there is a low probability of that dependence holding. It is plausible to say that one sees only so long as there is a high enough probability of the counterfactual dependence holding[29].

Lewis discusses this in the Postscript to the article. He essentially argues the following. At any particular instant, the dependence either holds or does not hold. Probability only arises when considering the dependence through time. It is *not* the case that, at a particular instant in time, the counterfactual conditional is probabilistic. Only when considering spans of time can the probability in question be appealed to. He suggests that, since the dependence either holds or does not hold at any particular instant, the subject either sees or does not see at that instant. Sometimes they do see, and sometimes they do not see. However, they do see at any instant where the counterfactual dependence holds. The fact that, over a span of time, the subject does not see at some instants in the span, has no bearing on whether the subject sees or does not see at *other* instants in the span of time.

An information theoretic explanation of the example yields the same results as Lewis. Whether one sees or not is independent of how often the loose wire makes contact. It is only dependent on whether the wire actually does make contact or not at some instant in a span of time. The subject sees when the wire makes contact, and does not see when the wire comes loose. The temptation to classify this as a case in which one does not see when the wire rarely contacts has to do, not with whether or not information is transmitted from the scene to the

[29] It seems unlikely that there is a 100% chance of the correct dependency holding even in the case of normal

experience, but whether or not one may *rely* on their visual experience as a source of information about the scene before the eyes. All that is necessary for a system to transmit information at some instant is that, at that instant, the output states of the device are sensitive to the states of the source which it transmits information about. So long as at that instant of time, if the source were in a different state, the output state at that same instant would have been different, the system transmits information about the source. When the prosthetic eye is functioning as designed, it is sensitive to the scene in this way[30]. When the wire contacts, the system transmits information about the scene before the eye. When the wire comes loose, we are dealing with a *different* system. That system (the eye with the loose wire) has different functional properties from the other system (the eye with the wire in place). The former system is incapable of transmitting information about the scene before it, the latter system is perfectly capable of transmitting that information.

It is, however, a further question to ask *which system is currently in place.* If the broken contact wire is the current system, then one ought not to use visual experience as a source of information about the scene. If the good contact wire is the current system, then one ought to rely on their visual experiences as a source of information. However, without the additional information as to *which* of these systems is in place at any given instant, it can be difficult to decide whether to rely on one's experiences or not. When we lack this information about a particular instant, the probabilities come in to play. If the wire contacts frequently, then one can reasonably rely on their visual experiences as a source of information about the scene before the prosthetic eyes. If the wire contacts rarely, then one should not rely on those experiences as a source of information about the scene. And, of course, there are cases where it is unclear whether one can or cannot reasonably rely on their visual experiences for

---

human vision. Unless something below certainty of the dependence holding through time is acceptable, it may turn out that even normal human beings do not see! This would be an unacceptable result.

[30] Notice that a broken eye is different from an eye which, when functioning *normally*, fails to produce output states that carry information about the source. We shall look at bad prosthetic eyes later on, and why they may

information about the scene before the prosthetic eyes. The question of whether one ought to rely on their experiences as a source of infor mation about the scene before the eyes is, however, distinct from the question of whether the visual experiences *are* a source of information about the scene before the eyes. And, as mentioned earlier, they are if the good wire contact system produces them, and they are not if the broken wire contact system produces them.

D. Alternative Scenes and Discriminable Scenes

The next two examples (Laser Beam and Hypnotic Suggestion) are intended to show that the counterfactual dependence that Lewis proposes need not be complete. Not every possible alternate scene, regardless of how subtle or drastic the change, need produce a correspondingly different visual experience, or any visual experience at all. Ideally, this would be the case. However, one can see even if not every feature of the scene, or every possible alternate scene, before the eyes is accessible by way of visual experience. There can be invisible differences between scenes, such as in the Hypnotic Suggestion. There can be differences under which the visual system cannot function properly, as in the Laser Beam example. The number of subclasses within the class of alternative scenes needs to be 'many', but need not be a 1-to-1 correspondence of alternative scenes with visual experiences, nor need it exhaust the space of possible scenes before the eyes.

An information theoretic analysis provides similar answers to these examples. Even though there are possible circumstances under which the transmission system cannot operate, information transmission is possible so long as the actual circumstances are not one of these. Furthermore, the possibility of differences that the transmission system output states cannot discriminate between has no impact on whether or not information about the scene is

---

fail to be systems by which one can see.

transmitted by the system in other circumstances, unless no possible alternate scene would produce a different output state. All Information Theory can be used to show is that the less discriminating a system is, and the fewer scenes it responds to, the less information it transmits. Still, a system may transmit *some* information about *some* source states, even if it is incapable of transmitting all the available information about any source state.

The fact that, for example, a volt meter cannot measure an electric potential of 115,200 volts, without being destroyed, does not mean it cannot transmit information about voltage. The fact that a volt meter with a sensitivity of 0.01 volts cannot discriminate between a 2.0001 and 2.0002 volt current has no impact on whether it carries information about voltage or not. The volt meter cannot measure, or discriminate between, *those particular* voltages. It still transmits information about voltage as long as the voltages fall within the range of voltages the detector can respond to. It can discriminate between voltages that produce different output states. Of course, the more voltages within the measurable range, and the more output states for differentiating voltages, the *better* the volt meter (the more voltage information it transmits). These considerations only affect how much information the volt meter carries, and not whether it carries voltage information or not (unless how much information is carried turns out to be zero information).

A transmission system need not be perfect to be useful enough to gather information. Likewise, a visual system need not transmit all the information in the scene before the eyes to be a system that allows one to see. Visual systems, like volt meters, can be better or worse. A volt meter may be extremely bad, capable of measuring only whether there is 0 volts or not. But a bad volt meter is still a volt meter, and a bad visual system is still as visual system.

E. Backtracking Dependence

The next example exposes a weakness in Information Theory. I believe Lewis has the correct analysis of The Screen example, although Information Theory as I understand it would disagree. In this example, the normal causal connection between the scene before the eyes and the visual experience is reversed. Instead of the scene causing one to have visual experiences, the visual experiences one has (pseudorandomly in this example) cause the scene before the eyes to change in corresponding ways. A very fast computer monitors your brain state while you hallucinate. Based on the measurements it takes, the computer predicts what your next visual experience will be and projects an image on a screen before the eyes that matches that experience. Lewis argues that this is not a case of seeing, for although the counterfactual dependence between the scene and the visual experience is of the right kind, it nevertheless "goes the wrong way" to count as an instance of seeing. This is what Lewis calls a "backtracking" counterfactual dependence. The scene is counterfactually dependent on the visual experience. So, the experience is counterfactually dependent on the scene by "backtracking" this counterfactual dependence. He suggests this is not an appropriate way of establishing a counterfactual dependence between scene and experience for seeing to occur.

Unfortunately, Information Theory cannot say the same kind of thing in response to this example. The distinction between a source and an output state in a transmission system is essentially arbitrary. If the output states contain information about the source, then the source states contains information about the output as well. The source is sensitive to the output state in the right way for a functional transmission system. Any change in the output state would accompany a corresponding change in the source. Even worse, transmission systems need not even follow causal connections. In one instance of this, a transmission system may exist where the source is one of two or more effects of some phenomenon, and the output state is another of these effects. One example is the connection between a barometer reading and rain tomorrow. Neither barometer reading or rain tomorrow is a cause of the other; both are effects of low atmospheric pressure. Yet, barometers are often used to supply information about tomorrows weather.

As another example, there is a transmission system that exists between the television set in my house, and a television set tuned to the same station in your house. Any changes in the source (your television) results in corresponding changes in the output state (my television), even though your television is in no way causally connected to my television. Furthermore, one can not appeal to interference or a malfunction in one of the sets to show that no such information channel exists. For, such a change alters the *transmission system* in question. A different system may transmit completely different information, or none at all. However, this *particular* system does transmit information, even without a causal connection.

My intuition on how to handle these problems is unclear, however. The latter case - where transmission systems do not follow causal connections - seems to have import for a case that Lewis rules out in his analysis. The possibility of seeing some scene other than the one before the eyes appears to be related to this. Lewis' intuitions are that this is not a case of seeing, although even he admits that these cases are not clear one way or the other[31]. He ignores them in pursuing his analysis of seeing.

Backing up a bit, the case where the visual experience is the cause, rather than the effect, of the scene before the eyes is unclear to me as well. Some science-fiction scenario similar to The Screen could be imagined, where one's visual experience brought about corresponding changes in a physical world that the subject had to navigate, etc... Such an individual would have useful information about their environment, based on their visual experiences. Even though their experiences caused said environment to be the way it actually is, there is sufficient information transmission to make the information content of that person's visual experience usable in the same way that a truly sighted person would make use of them[32]. For now, I will simply stipulate that the "right kind" of transmission system must

31 p276

32 I assume that the subject whose visual experience changes the world has no conscious control over what their

be in place for seeing to occur, where the "right kind" of transmission system is a system where the source is a causal ancestor of the output state. This is an unsatisfactory way of resolving the difficulty, but exploring the problem here would take us too far afield.

F. Bad Prosthetic Eyes versus Broken Prosthetic Eyes and The Censor

We come to the final example presented by Lewis in the paper. This is the most difficult example to explain, and so Lewis saves it for last. This is the case of The Censor. In the case of The Censor, the censor wishes to ensure that the subject has only a limited range of visual experiences (in this case, only one experience). He has the capability of interfering with the subjects visual system, causing the subject to have whatever visual experience the censor desires them to have. However, he intervenes only when the scene before the eyes would produce an unacceptable visual experience. When the actua l scene before the eyes will produce an experience the censor permits, he does not interfere with the operation of the subjects visual system. In such cases, the subject's visual experience arises from the same kind of process that produces visual experience in normal sighted humans. As soon as the scene changes to something other than one which will produce an acceptable visual experience, the Censor intervenes to ensure the subject has such an acceptable experience. The question here is: Does the subject see while the Censor is *not* intervening (when the scene before the eyes is the cause of the visual experience)?

Lewis' answer to the question is no, but it is not immediately evident that he should respond this way. This example shares features in common with two previous examples. It is similar to The Loose Wire case, save for the fact that in the Loose Wire case, the intervention is essentially random. In the Censor case, intervention into the visual process is dependent on the scene before the eyes, and so may be different enough to warrant a different response. It

---

visual experiences are, just like normal sighted humans also bears a striking similarity to the Hypnotic Suggestion case. Even Lewis points out that this is "Example 11 [the Hypnotic Suggestion] carried to extremes".[33] The difference seems to lie only in the frequency of intervention; the Censor intervenes more often than the Hypnotic Suggestion does to ensure only acceptable visual experiences are produced. Unfortunately, we already discussed the issue of frequency of intervention in the Deathbed Cure and Loose Wire examples. There, we came to the conclusion that frequency of intervention (how often does the process work to produce visual experience with the correct counterfactual dependence on the scene) was irrelevant to whether or not the subject sees. Can the fact that the intervention in Censor-type cases is not random, but dependent on the scene before the eyes, make such a difference to the analysis?

I believe the answer is that it does make the difference. To see why, I again turn to Information Theory for an explanation. In the Loose Wire case, I argued that we were dealing with two distinct systems. One system was the prosthetic eye with the wire making good contact, the other was a prosthetic eye with the wire not making good contact. With the good contact system, the subject sees; with the bad contact system, the subject does not see. Whether or not the subject sees depends only on which system is currently producing the visual experiences; it does not depend on the frequency with which one system or the other is currently operating. The reason these are two separate systems is that the switch between good contact and bad contact is essentially random. It has no dependence on the scene before the eyes.

However, when the switch from one system to the other *is* dependent on the scene, the switch is part of the function of a single, larger system that incorporates these processes as subsystems. The composite system composed of a Censor, his intervention system and a human visualsystem is simply an alternative system for responding to those changes. This is

33 p285

more like the case of a *bad* prosthetic eye - a system which fails to transmit information about the source *as part of it's normal functioning.* In the Censor example, it is part of the function of the system that alternate visual experiences are produced when the scene before the eyes would not produce acceptable experiences by the other means (the means like those in normal sighted humans). In the extreme case, where only one specific visual experience is allowed by the Censor, the visual experiences of the individual affected carry no information about the scene. Since there is only one possible visual experience, the system cannot carry any information at all. There have to be at least two possible output states - visual experiences in this case - for the system to even be capable of transmitting information about a source.

However, I would describe a Censor case where there are two permissible visual experiences as a case of seeing, if only very little. With two permitted visual experiences, the system conveys some information about the scene before the eyes - at least the information that a scene which produces the *other* permitted visual experience is not before the eyes. This is not very much information at all, but it is more than no information. We should, in typical conversations about seeing, probably say that a person who only has two distinct visual experiences cannot see. But, as I have said before, I think this is technically incorrect. The person who only has two possible visual experiences does see, so long as those experiences are the output states of a system that transmits information about the scene before the eyes. They just see extremely poorly.

V. The Paradox of Seeing that One Cannot See

I wish to consider one additional example, briefly mentioned by Lewis but not pursued in any detail. It is not one of the numbered examples that appear in the article, but presents an opportunity to clarify the notion of seeing that is being used throughout this paper. The example I have in mind is the case of total darkness. Lewis mentions this near the beginning of Section VIII.[34] He suggests that, on his analysis, that we do not see in total darkness. Yet, he finds this to be paradoxical. After all, we do see *that it is totally dark* when in such situations. If true that we see that it is too dark to see, then it seems to follow that we see in the grammatically intransitive sense Lewis describes. "How else - by smell?"[35] It also can not be the case that we see that it is dark by the very fact that we do not see, for we also do not see in thick fog or dazzling bright light. These failures to see are qualitatively different than when we do not see in total darkness. Yet, if we see this, that contradicts the very proposition seen to be the case! It seems that we must see that we cannot see, but seeing that we cannot see is a flat-out contradiction. Lewis simply points out the problem, suggests there is an ambiguity in the word 'seeing', and moves on without further consideration. I wish to offer an explanation of what is going on when we 'see' that we can not see. More precisely, I shall try to offer a disambiguation of seeing, such that the first occurrence of 'see' in the paradoxical statement is not the sense of 'seeing' used throughout this paper and Lewis' paper.

Transmission systems, even ones designed to transmit information about certain kinds of sources, may be used to transmit other kinds of information. They may even be used self-diagnostically, to provide information about their own workings. This self-measuring ability of transmission systems is what goes on when one calibrates a system. For example, a scale is typically calibrated by placing an object with known mass on the scale. The indicator presents a value, which we compare to the known value for the object on the scale. In this

---

34 p283

way, the scale delivers not information about the mass of the object (we know that already), but about it's own operation. We test the accuracy of the scale, and make adjustments to improve that accuracy when needed. The scale, which normally delivers information about the mass of other objects, is used to transmit information about itself in these circumstances.

Another situation in which a device's output states carry information about itself, rather than about an external source, is when an error indicator is present. This may be an explicit indicator, such as a light marked "Error" or a digital display of the letters "ERR". An error state may also be implicit in the normal indicator of values, such as when a scale or meter is pegged[36]. When the needle or pointer pegs, the position of the needle or pointer does not indicate anything meaningful about the source it usually transmits information about. It merely indicates that something is not correct with the system itself. Frequently, this occurs when the value for the property being measured exceeds the range in which the device can operate. If you place a 5 1/3 pound weight on a postal scale intended to measures weights up to 1 pound, the pointer will peg. Of course, other things can cause pegging as well. The spring may be broken or badly deformed, so that even a ½ pound object pegs the scale. It may be improperly calibrated. Pegging a scale is *often* caused by exceeding the range of the scale, but not *always*. All a pegged scale indicates is that something has gone wrong with the device.

Furthermore, this kind of failure is distinguishable from other kinds of failures. The object may be too light to be measured by the scale. The mechanism may be jammed, preventing the needle from moving. The fact that the needle does not move off zero, even when an object is resting on the scale, also indicates that something has gone wrong. Furthermore, this failure is *distinguishable* from a failure in which the scale pegs. Something

---

35 p283

36 'Pegging' a meter occurs when the needle or pointer on the display quickly moves to the top or bottom of the range of values, usually slamming in to a small peg designed to stop the needle from moving further. This usually results in a clicking noise as the needle hits the peg.

has gone wrong, and something different from what goes wrong when the scale pegs has happened. Both indicate that something is wrong, but they indicate that different sorts of things have gone wrong. In either case, what is indicated is information about the apparatus used to transmit the weight information, not information about the weight of the object on the scale. Something is not right with the scale. One may infer from this that the object is too light or too heavy to be measured by the scale, since this is the explanation of the failure with the highest probability. However, this is *inferred* from the information presented. It is not the information presented by the scale's output state.

My explanation of what is going on when one 'sees that it is too dark to see' should be evident by now. When it is too dark to see, the visual system cannot transmit any information about the scene before the eyes. It may still transmit information, but this information is about the visual system itself and not the scene before the eyes. One can *infer* from the visual experience of total darkness that there is insufficient light in the scene before the eyes to allow proper visual functioning. This is, however, not what the visual information content *is*. What the experience of total darkness indicates is that something has gone wrong with the visual system. Furthermore, it indicates a particular way things have gone wrong. What has gone wrong is *different* from the kinds of things that go wrong when surrounded by thick fog or confronted with dazzling bright lights. In any event, these experiences provide information, not about the scene before the eyes, but about the visual system itself. One infers from this, based on previous occurrences, that it is too dark, or too bright, or too foggy to see, depending on the kind of experience. One infers from the fact that, in the past, failures of this particular kind were corrected by flipping on a light switch, lighting a candle, etc…, that this particular failure of the visual system is more than likely caused by a lack of light in the scene before the eyes. But the inference is based on information about the visual system.

Since the information present in such error experiences is not information about the scene before the eyes, but about the visual system itself, these experiences do not count as cases of seeing in the sense used throughout this paper, and by Lewis. They are visual experiences, and they are visual experiences that provide us with information. However, since seeing (in the sense we are interested in) only occurs when the source of information in visual experience is the scene currently before the eyes, experiences that provide no information about the scene before the eyes do not count as experiences by which one sees. One infers that it is too dark to see based on one's visual experiences, but does not 'see' in the sense that we are interested in. The sense of the word 'see' used first in the phrase 'seeing that one cannot see' implies a sense of grammatically intransitive seeing in which one's visual experience carries some information, regardless of the source. This is a perfectly fine sense of the word 'see'; it just is not the sense which interests us here. Furthermore, it is different than the sense used in the second occurrence of the word 'see' in that phrase. Thus the phrase appears paradoxical.

I think a similar explanation is open to Lewis, although he does not avail himself of them. When it is too dark, the scene before the eyes does not cause visual experiences. Light is required to cause responses in the visual system. Unless the absence of causes is itself a cause, the scene with no light cannot cause the visual system to respond. Since the scene before the eyes must *cause* the visual experience, with the right sort of counterfactual dependence, total darkness is not a case in which one sees. The experience is caused by the internal workings of the visual system, not the scene before the eyes. It therefore does not count as a genuine case of seeing, even though it may be useful for learning things about the scene before the eyes

VI. Conclusions

Lewis was on the right track in "Veridical Hallucination and Prosthetic Vision". Many of the ideas he puts forward fit nicely with an information theoretic account of what it means to see. However, I think Lewis made at least two mistakes. One was his refusal to explore some of the restrictions on seeing, such as the requirement that there be 'many' alternate scenes or that there must be 'suitably different' changes in the scene before the eyes, and so on. I suspect this is what led him to the second mistake; deciding that cases of extremely poor sight (The Light Meter) are cases of not seeing. This may accord with more common usage of the word in conversation, but I suspect that when pushed, we can see that this is technically correct. After all, if the person with the light meter can differentiate between bright and dim scenes based on his visual experiences, how else can he do this without seeing? By smell?

In spite of this, there is quite a bit of agreement between Lewis and the Information Theory account I have described. Indeed, Lewis' account of the Screen example still strikes me as the correct account. Information Theory alone is not adequate to explain why The Screen case is not a case of seeing, and merely stipulating that seeing requires a transmission system where the source is a causal ancestor of the output state is unsatisfactory at best. Unfortunately, I have no better suggestion for how to handle this kind of a case.

However, I do think Lewis rules out too much as genuine cases of seeing. His lack of a fine-grained account of this grammatically intransitive sense of 'seeing' requires him to give a yes or no answer to the question. This I think leads him to rule cases of seeing very badly as cases of non-seeing, such as the Light Meter example. I see no substantial difference between the Light Meter and normal human vision, save for a difference of degree. The individual who has only two possible visual experiences cannot see very much, but not seeing much is not the same as not seeing at all, and I think this is where Lewis goes wrong.

Bibliography

Dretske, Fred: Knowledge and the Flow of Information. MIT Press. ©1981

Lewis, David: Philosophical Papers, V 2. Oxford University Press. © 1986

---------------- "Veridical Hallucination and Prosthetic Vision" contained in above

Shannon, Claude E.: The Mathematical Theory of Communication (3rd printing)
Urbana, University of Illinois press. © 1949 (3rd printing ©1964)