Why Hempel's Paradox Isn't a Paradox (And What We Could Do If It Was)

Sara L Friedemann

20th May 2003

1 What is a paradox?

A paradox "arises when a set of apparently incontrovertible premises gives unacceptable or contradictory results" (Blackburn 1996, p. 276). Paradoxes are what result when apparently true and straightforwardly acceptable premises lead to apparently false and patently unacceptable conclusions. Paradoxes are the bane of formal systems; systems that are sound are ones where it is *provable*, beyond a shadow of a doubt, that no paradoxes will arise. Naturally, proofs of this fact (soundness proofs) easily separate out the sheep of the formal systems from the goats. Unsound systems are tossed to the side, and sound systems remain the sole focus of interest.

In informal systems, such as human reasoning, this differentiation is not so easy. Paradoxes, or 'puzzles', if you wish to be less antagonistic, have been the focus of philosophers since the dawn of time. When a system of analysis cannot be rigidly defined, it is much harder to *prove* whether or not no paradoxes can be derived. The only way to work toward the soundness of the system is to counter each so-called paradox as they arise, and determine whether or not they are in fact genuinely paradoxical, or whether they can be solved by "either showing that there is a hidden flaw in the premises, or that the reasoning is erroneous, or that the apparently unacceptable conclusion can, in fact, be tolerated" (*op cit*).

2 What is Hempel's paradox?

Hempel's paradox[1], in a nutshell, is that two equally plausible and straightforwardly acceptable premises concerning the nature of scientific confirmation lead to results that some people find extremely non-intuitive. The two premises are the Nicod-derived Condition, which says that a hypothesis of the form "All *F*s are *G*s" is confirmed by the existence of any object that is both *F* and *G*, and the Equivalence Condition, which says that whatever is confirmatory of a hypothesis is confirmatory of anything that is logically equivalent. Thus, the hypothesis "All ravens are black" is confirmed both by a black raven and a pink goldfish. Why? Because "All non-black objects are not ravens"

[1] First introduced in Hempel 1945.

is logically equivalent to "All ravens are black." By the Nicod-derived Condition, pink goldfish are objects which are both F and G (F and G in this case being "non-black" and "non-raven" respectively), and thus their existence is a confirmatory instance of the hypothesis. Since the two hypotheses are logically equivalent, both are confirmed in both cases, according to the Equivalence Condition.

This result can certainly be seen as anomalous, and, as Humberstone says, "it seems grossly counterintuitive to suppose that the observation of a red pencil (say) should provide any confirmation for the hypothesis that all ravens are black" (392). But as counterintuitive as it is, is this paradoxical? The conclusion is not *prima facie* contradictory. Yes, it may seem counterintuitive, but then the question should become, how far should we be guided by our intuitions? If one is wholly committed to the two conditions and willing to be hard-nosed about it, then one has no other choice than to say "my intuitions have failed me in the past, they must be failing me now; all right, this red pencil, that silver dress, that pink goldfish, these *are* confirmatory instances." It wouldn't be the first time that intuition has failed.[2] Following a Quinean lead, when presented with apparently contradictory beliefs one must give up that which is least central to the web of belief as a whole. In the case of Hempel's ravens, nothing else appears to ride on a belief that non-black non-ravens are not confirmatory of the hypothesis 'All ravens are black', and therefore it won't hurt us terribly to give up that intuition. When one considers that in order to ever be absolutely certain about the truth of any universal statement, one must look to every object in the domain of discourse to find any falsifying instances, accepting this conclusion begins to not only be more palatable but even correct. There is nothing paradoxical with granting the conclusion that a grey cat is confirmatory of the hypothesis "All ravens are black" but only to a very, very low degree.[3]

However, Hempel's paradox can be connected with a similar problem in such a way that people are much less desirous of giving up their cherished intuitions, for this problem concerns the rationality (or irrationality as the case may be) of human beings

[2]Maher (1999) gives a formulation of the problem wherein it *is* genuinely paradoxical:

> The following three principles regarding confirmation have all been regarded as plausible:
>
> **Principle 1 (Nicod's condition)** *In the absence of other evidence, the evidence that some object is both F and G confirms that all F are G.*
>
> **Principle 2 (Equivalence condition)** *If evidence confirms a proposition then it also confirms any proposition that is logically equivalent to that proposition.*
>
> **Principle 3** *In the absence of other evidence, non-black ravens [sic] do not confirm that all ravens are black.* (50)

Unless Maher is trying to state the obvious in Principle 3, I believe he actually meant that 'non-black *non*-ravens' are not confirmatory. Given that substitution of the Principle, these three principles are indeed inconsistent. However, this is easily resolvable by pointing out that Principle 3, of the three, is by far the most contentious, and thus if one should be given up, that is the one; in fact, a portion of my paper is to argue that it is not true, and thus taking it as a premise will certainly lead to an unsound argument.

[3]This is because there are so many more non-ravens than there are ravens. Should the distribution of objects be different, observations of grey cats might be extremely confirmatory. Consider the hypothesis "All ravens are self-identical." In this case, one would naturally want to look at all the non-self-identical objects first, to determine whether any of them are ravens, than to look at all ravens, and then determine whether or not they are self-identical. Leavitt (1996) discusses this point: When seeking to confirm the hypothesis "All ravens are black" we look at ravens first, because there are fewer ravens than non-black things.

such as ourselves, and how non-philosophers respond to questions similar to those brought up by Hempel's paradox.

3 Enter Wason

The Wason selection task comes in almost as many flavors are there are psychologists. The original description of the task is as follows:

> The subjects (students) were presented with an array of cards and told that every card had a letter on one side and a number on the other side, and that either would be face upwards. They were then instructed to decide which cards they would *need* to turn over in order to determine whether the experimenter was lying in uttering the following statement: *If a card has a vowel on one side, then it has an even number on the other side* (Nickerson, p.3)

The only disconfirming instance would be that of a card that has a vowel on one side but an odd number on the other side. Thus, the logical answer is to check cards whose visible side shows a vowel, and ones whose visible side which show an odd number. Take P as 'the card has a vowel on one side' and Q as 'the card has an even number on the other side', then the rule in question is $P \supset Q$ and the only time that this is false is in the presence of $P\& \sim Q$. Yet upon giving this test in numerous forms to numerous subjects, the number of people who get the 'right' answer is less than 10% in many cases. The most common answers, instead, where to pick just the P card, or the P and Q cards.[4]

Both Humberstone (1994) and Nickerson (1996) draw the connection between the Wason task and Hempel's ravens. The problems can be described virtually congruently. Consider this formulation of the raven's hypothesis: *If the card has a raven on one side, it is black on the other*. Now consider an amendment to the task where the cards show colors on one side and birds on the other. Then, to determine whether the modified hypothesis is true, the logically correct answer is to check all cards that are ravens and all cards that a not black. If an instance of $P\& \sim Q$ is found, then the rule or the hypothesis is disconfirmed. This means that if one finds a non-black card that has something other than a raven on the other side, this is a confirmatory instance. This is precisely the same 'paradoxical' result discussed above. However, even if we are willing to grant that that conclusion, especially when considered from the point of view of classical logic, is not paradoxical or even all that counter-intuitive, it is true, as Humberstone says that "we can think of the difficulty people have in seeing that the

[4]There is actually rather more variation in responses than I've indicated here, and these correspond to different content that can be given to the test, and different ways the test can be presented (*see* Nickerson 1996). (E.g., instead of concerning cards with numbers and letters, the test can be formulated to concern people in a bar, where the relevant rule is *If a person is drinking beer, then that person is over 21*. How do you determine whether this rule is being followed? Here, the correct answer (though still not the most common answer) is by checking all people drinking beer and all people who are not over 21. However, the presence of these variations in responses corresponding to variations in the content of the test will not concern us here. While it is interesting to ask why more people get the 'right' answer in some cases than in others, the question that interests me most is why so few people get the 'right' answer, period.

card showing '*White*' needs to be turned over as analogous to the counterintuitiveness of thinking that the observation of a white swan (or any other non-black non-raven) could count as a confirming the hypothesis that all ravens are black" (395). Given that only 10% or so of people given the Wason task reach what would be the logically correct answer, that means 90% or so *do not* answer in accordance with classical logic. Whether or not one is willing to accept that non-black non-ravens are confirmatory of the raven's hypothesis, many people will be much less willing to accept either the conclusion that human rationality is not logical or that the majority of people do not act rationally.[5]

Thus, if the two problems can be described in identical language, then if one is unwilling to accept the result of Wason's task, one should be unwilling to accept the conclusion of Hempel's paradox. Contrapositively, if we can show a way to perhaps resolve some of the lingering intuitions concerning Hempel, then perhaps these can be carried over to a consideration of understanding Wason. We shall come back to this at the end of the paper.

Exit Wason.

4 The Equivalence Condition

If one is to pick holes in Hempel's paradox, trying to undermine the Nicod-derived Condition does *not* appear to be the way to go. If the existence of a black raven is not confirmatory even of the hypothesis "There is a black raven," what will be? So let us turn our attention to the Equivalence Condition. What are some of the justifications given for this condition? Humberstone argues that it is "reasonable because whether or not a hypothesis is confirmed by an observation should depend on the content of the hypothesis and not on the way that it happens to be formulated" (391). The Equivalence Condition could be restated as saying that logically equivalent formulas have the same content. Thus, since $P \supset Q$ could just has easily, in logical terms, be formulated as $\sim Q \supset \sim P$, and vice versa (because the two are logically equivalent), whatever confirms the one must confirm the other. However, this is not apparent, in the same way that the Nicod-derived Condition is, and needs to be argued for, as one might conceivably react to this principle as follows: "But $P \supset Q$ doesn't have the same content that $\sim Q \supset \sim P$ has!" Certainly 'ravens' and 'black things' appear to be separate from 'non-ravens' and 'non-black things'. In fact, there is good reason to think that contrapositives of conditionals do not concern the same subject matter. Consider the observation of a purple Q-tip. It is confirmatory of both **(H1)** "All ravens are black" and **(H2)** "All ravens are fluorescent orange," because purple Q-tips are both non-orange and non-black.[6] Therefore, black ravens and purple Q-tips are confirma-

[5]Nickerson's paper is an attempt to show that, from a psychological point of view, "people's typical performance in the selection task can be explained, by consideration of what constitutes an effective strategy for seeking evidence of the tenability of universal or conditional claims in everyday life" (1). But even if this can be demonstrated, this still doesn't explain why it is that acting in a survivally fit manner doesn't coincide with acting in a truth-functionally sound manner. For me, *this* is, if anything is, paradoxical in the extreme. But the question "Why is rationality not logical?" is also a lot harder to answer.

[6]Nickerson makes the same point: "...[this observation] leads also to the conclusion that we should take the same observation as confirmation of the contradictory [sic] claim that all ravens are red, because the

tory instances of different hypotheses. The purple Q-tip confirms both H1 and H2, while black ravens confirm H1 and *disconfirm* H2. Because the confirmatory instances are not co-extensive in what hypotheses they confirm, it seems reasonable to believe that those hypotheses do not concern the same content. Thus, the justification (which concerns sameness of content) given for the Equivalence Condition (which concerns logical equivalence) does not seem adequate.

This brings up two possible threads to follow: Either to develop a specification of 'content' that is robust and not *ad hoc*, and demonstrate sufficiently that content does not coincide with logical equivalence, and thus the content argument for the Equivalence Condition fails, or to develop a system where the notion of content is not explicated fully, but logical equivalence is restricted in such a way that logically equivalent statements can, on a notion of content left rather undefined, coincide with sameness in content. This latter route was discussed by Sylvan and Nola (1991), but referenced only briefly in a footnote in Humberstone (401). It will occupy us for the rest of the paper.

While Sylvan and Nola make many claims about how switching from classical logic to a weak relevance logic (such as first-degree entailment, or FDE) will resolve all the paradoxes from Hempel to Goodman to others, very few arguments are actually given for these claims in their paper. Furthermore, they do not even specify which systems are sufficient to remove the paradoxes (whether this is all relevance logics, all weak logics, or perhaps a subset of both). Their main point is that if contraposition can be removed, then the problematic inferences which lead to the paradoxes can be removed. (Such a result can be reached with a Stalnaker/Lewis type of conditional theory, or a theory of conditionals that treats conditional statements as conditional probabilities, which do not validate contraposition, but these suffer from their own weaknesses and I will not discuss them here.) If a conditional and its contrapositive are no longer logically equivalent, then it is certainly not the case that they have identical content, and one can maintain the Equivalence Condition without trepidation. The task is then to determine a logic that does not validate contraposition but is still sufficiently strong that we can reconstruct other, non-problematic inferences. FDE is not, contrary to Sylvan & Nola, the right choice, as contraposition *is* still valid. Consider a relational four-valued semantics of the type given by Priest (2001), where a truth value assignment to the propositional parameters is not a function f from atomic formulas to truth values but rather a relation ρ between atomic formulas and truth values. The values are *1* and *0*, and an atomic formula can be related to one, both, or neither of these values. The truth values of complex formulas are then defined recursively as:

- $\sim A\rho 1$ iff $A\rho 0$.
- $A\&B\rho 1$ iff $A\rho 1$ and $B\rho 1$. $A\&B\rho 0$ iff either $A\rho 0$ or $B\rho 0$.
- $AvB\rho 1$ iff either $A\rho 1$ or $B\rho 1$. $AvB\rho 0$ iff $A\rho 0$ and $B\rho 0$.

$P \supset Q$ is defined as $\sim PvQ$. A set of sentences entails a conclusion iff when all it's premises related to *1* or both, the conclusion is also related to *1* or both. Given

contrapositive of the latter claim is that all nonred things are nonravens" (2). While he errs in saying that the two hypotheses are contradictory (they are only contrary: They can both be false but they can't both be true), the point is still germane.

this definition of logical consequence, it can be proved[7] that a set of formulas entails a sentence iff when the conjunction of the premises is reduced to Disjunctive Normal Form and the conclusion to Conjunctive Normal Form, each disjunct of the premise shares an atom (a propositional letter or the negation of one) with each conjunct of the conclusion. Since $P \supset Q$ is defined as $\sim PvQ$, then $P \supset Q$ and $\sim Q \supset \sim P$ reduced to DNF and CNF are identical.[8] There is no way that a conditional be related to *1* and it's contrapositive not be. Thus,what is needed is a relevance logic that is even weaker than FDE, and in the next section I will introduce a logic for which this is true.

5 Relevant Negation

Relevance logics take their name from the idea that in order for an inference to be correct, the premises have to in some way be relevant to the conclusion which is derived from them. In science, law-like statements such as "All copper conducts electricity" are generally taken to be asserting more than just that any object either is not copper or does conduct electricity; rather, the claim is meant to illustrate some type of relevant connection between being copper and being a conductor of electricity. Because the truth tables for the material conditional allow for classical conditionals having no connection between their antecedent and consequent, "relevance logicians often motivate their studies by suggesting that their logic, not the logic of material implication, is the logic of scientific discourse" (Waters 462).[9]

In his 1971 paper, Urquhart presents a semilattice semantics for relevant logics, including Church's weak theory of implication and systems of the type developed by Anderson and Belnap. A model in this semantics is a triple <*S, 0, u*>, where *S* is a set, 0 is an element of that set, and *u* is a binary operation on elements of *S* that satisfies the following constraints:

- $XuX = X$
- $(XuY)uZ = Xu(YuZ)$
- $XuY = YuX$
- $Xu0 = X$

A natural way to interpret such a model is to have *S* be the power set of some (finite or infinite) set, 0 be identified with the empty set, and *u* be set theoretic union. On these models, the valuation rule for the implicational connective $\rightarrow$ is as follows:

- $V(A \rightarrow B, X) = T$ iff for all Y, either $V(A, Y) = F$ or $V(B, XuY) = T$

[7]Mike Byrd did so in Philosophy 512 this semester.

[8]This rests partly on the fact that double negation and commutation are also acceptable; $\sim Q \supset \sim P$ is technically $\sim\sim Qv \sim P$, but it is easy to see how this is the same as $\sim PvQ$.

[9]Waters doesn't discuss Hempel's paradox in his paper so much as he treats Hypothetico-Deductivism in general, and he offers relevant logic up as a (partial) way to formulate H-D in an acceptable fashion. However, as he doesn't discuss negation at all, his account can't be taken to be adequate to discuss removal of contraposition as a way to block Hempel's paradox.

- $V(A \to B, X) = F$ otherwise.

Propositional variables are assigned either T or F at each element of the semilattice, and these values are then used to compute the values of complex implicational statements based on the rule above. A formula A is valid when $V(A, 0) = T$ for any valuation V in any semilattice S where the bottom of the lattice is 0.

Urquhart extends his semantics with valuation rules for conjunction, disjunction, and quantification, but these need not concern us here. What does concern us are the varying ways to deal with negation. As Urquhart says, relevance logics were "devised to exclude such 'paradoxical' formulas as $(p\& \sim p) \to q$ and $p \to (qv \sim q)$" (164). The logics must accommodate both truth value gluts (sentences that are both true and false) and truth value gaps (sentences that are neither true nor false), in order for these entailments to not hold. One way of treating negation is as follows:

- To the semilattice S add a function C under which S is closed, and which satisfies $CCX = X, C0 = 0$. Then, define the valuation rule for$\sim$as follows:
- $V(\sim A, X) = T$ iff $V(A, CX) = F$ and $V(\sim A, X) = F$ otherwise.

This variety of negation does not validate as theorems either *reductio ad absurdem* or contraposition, though these principles can still be taken as rules of inference.[10] This is hopefully precisely what we need: A logic that does not validate contraposition. If a conditional and it's contrapositive are no longer equivalent, then we can dodge the Equivalence Condition with gusto and Hempel's paradox is blocked before it reaches fruition.

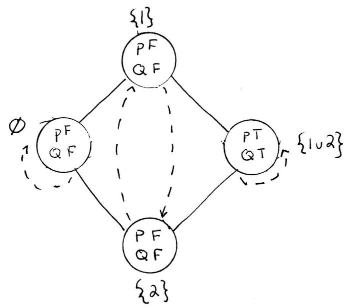
6 A Model Proof

If contraposition is not a theorem, then that means there is a model where $V((P \to Q) \to (\sim Q \to \sim P), 0) = F$.[11] This section introduces a model and proves that it makes the compounded conditional false.[12] Consider the following:

[10]This way of treating negation ultimately invalidates *reductio ad absurdem* and contraposition. The other possibility in the article validates these two, but does not validate DeMorgan's or the Law of Excluded Middle.

[11]It is not the case that any model that makes $V(P \to Q, 0) = T$ will make $V(\sim Q \to \sim P, 0) = F$ (though in fact the model discussed here does). This is because, as Urquhart notes, "both principles [contraposition and *reductio*] are valid, however, as inference rules...if $A \to B$ is valid, so is $\sim B \to \sim A$" (165).

[12]We need not consider cases where the antecedent and the consequent are themselves compound conditionals. The result is provable for all antecedents and consequents, and for our purposes proving the simpler case is sufficient: These are the only types of hypotheses that we generally consider anyway.



<S, 0, u, C> S = {0, {1}, {2}, {1,2}}, *0*=0, and *u* is set theoretic union, and C\{1\}=\{2\}, C{2}={1}, C0=0, and C{1,2}={1,2}. It is easy to check that this C meets the conditions specified above. As shown by the diagram, $V(P) = V(Q) = F$ at all elements except the top of the lattice, where $V(P) = V(Q) = T$. This model makes $V((P \rightarrow Q) \rightarrow (\sim Q \rightarrow \sim P), 0) = F$.

Proof: $V(\sim Q, \{2\}) = T$ because $V(Q, C\{2\}) = F$. $V(\sim P, \{1,2\}) = F$ because $V(P, C\{1,2\}) = T$. Thus $V(\ Q \rightarrow \sim P, \{1\}) = F$, because there is a Y (namely {2}), where the antecedent is true but the consequent is false at the union. $V(P \rightarrow Q, \{1\}) = T$, however, because for all Y, either $V(P, Y) = F$ (this is the case for 0, {1}, and {2}), or $V(Q, Yu\{1\}) = T$ (this is the case for {1,2}). Thus, $V((P \rightarrow Q) \rightarrow (\sim Q \rightarrow \sim P), 0) = F$, because there is a Y (namely {1}), where the antecedent is true and the consequent is false at the union of 0 and {1}.
Q.E.D.

The fact that contraposition is still acceptable as a rule of inference is not a hindrance here, because in general hypotheses are not tautologies of classical logic, and thus it is not the case that they will be valid on all models of the above sort.

In order to extend this result to hypotheses of the form "All ravens are black," quantification must be added to the model. This is easily enough done. To the semilattice structure, add a non-empty set *D*. This is the domain of quantification, and it does not vary from element to element in the semilattice (as the truth of the propositional parameters does or can vary). Then, extend the *V* as follows: To each variable x_1 assign $V(x_1) \in D$. To each n-ary predicate letter F assign a set $V(F) \subseteq D^n x S$, so that atomic statements are evaluation by the following rule:

- $V(Fx_1, ..., Fx_n, X) = T$ iff $<V(x_1), ..., V(x_n) \in V(F)$
- $V(Fx_1, ..., Fx_n, X) = F$ otherwise

Quantified sentences are evaluated in the expected way:

- $V((x)A, X) = T$ iff for every V' that differs from V by what it assigns to x, $V'(A, X) = T$ and false otherwise
- $V((\exists x)A, X) = T$ iff there is a V' that differs from V by what it assigns to x, $V'(A, X) = T$, and false otherwise

We now have everything we need to evaluate any universally quantified scientific hypothesis.

7 Enter Wason Again

It is one matter to find a logical solution to a puzzle, and quite another to motivate it philosophically. In the case of Hempel's paradox, the motivation behind finding a logic that does not validate contraposition was not so much so that red herrings no longer counted as empirical evidence for the claim "All ravens are black," but rather so that a possible alternative could be given that would make the results of the Wason tasks more palatable.[13] It is hard to see how the Urquhart semilattice semantics could be interpreted in any natural way. In fact, it is widely recognized that these semantics are opaque at best and *ad hoc* at worst.

However, because these semantics, and semantics for FDE on a broader scope, allow both truth value gaps and truth value gluts, this cannot be a logic of propositions, as propositions have only one truth value and have it regardless of the user's epistemic position with regard to that proposition. But this can be (and has been put forward as) the logic of *information*, where different sources may give different (and perhaps contradictory) truth value assignments to different propositions, depending on their epistemic position. Because it is impossible in classical logic, given the truth tables, to assign an atomic sentence letter and its negation both the value T, a contradiction will entail anything. However, in reality, people are often faced with conflicting information concerning the truth value of certain propositions without any knowledge of which is correct. People are often in a situation where they must rely on disparate sources for the information on which they base their decisions. Some sources may say one thing and other another; and what they say may be contradictory. But if one source says "$2 + 2 = 4$" and another one says "$2 + 2 = 5$", it is not rational, even if it would be logical, for the person receiving these pieces of information to conclude "I am king of the world." Truth value gluts allow for the failure of *reductio ad absurdem*, which is a plus when the framework within which we are working is a non-formal framework of information.

To see how the Urquhart semantics lend itself to an interpretation as modeling pieces of information, recall that a natural way to construct a semilattice is to identify it with the power set of some finite or infinite set. Let this finite or infinite set compose all the relevant pieces of information that are available (for the sake of ease, let us suppose that these are atomic pieces of information). Then the 0 element of the semilattice is either the set of no information or, as is sometimes the case, the set of background, already determined information. Then each atom of the element represents some new piece of information that someone would be given, e.g. "Object Y is black," "Object Y is a raven," "Object X is not a raven," "Object X is black," etc. Since the u operation results in the unioning of sets of pieces of information, if $\{1\}$ comprises the information

[13]Because Wason's task was studied primarily by psychologists and not logicians, more people have given psychological reasons why people answer the way that they do. These answers do not interest me because it is not apparent from these answers what (if anything) is the appropriate logic that could be underlying these psychological mechanisms.

“Object Y is a raven” and {2} the information “Object Y is black”, then {1,2} is the set of information “Object Y is a black raven.” It is important to note though that this is something of a simplification of the picture; it is actually the case that each atom can contains many different propositions, and different truth values for these propositions. Thus the information that is given at an element of the lattice is *not* the proposition, but rather an assertion about the truth or falsity of a certain set of propositions. It is just an easy short-hand for our purposes to take a statement “Object X is black” to stand for “The propostion ‘Object X is black’ is true.” Thus, at each stage in the process, in general we are not *adding* new propositions to our set, but just changing our judgement of their truth values, based on the new information from our “sources”.

Perhaps one of the strongest objections that can be raised to these semantics is that they are non-monotonic. If a set {1} tells us that piece of information *P* is true, there is no guarantee that it will remain true in any other set that contains 1 as a member of it, or vice versa.[14] This means that there is no way to predict what truth value an atomic letter will have at any given element in the lattice. Yet if such a monotonicity condition was added, then the resulting logic is provably equivalent to intuitionistic logic (Urquhart 166-167), which, while it doesn’t validate some forms of contraposition, does validate others, and is therefore unsuitable for our purposes. But it is precisely because the truth value can vary from source to source that a conditional can be true while it’s contrapositive is false. So what appears to be problematic is actually what is necessary in order to obtain the results that we desire.

8 Concluding Remarks

While the focus of this paper was the psychological test known as the Wason selection task, no part of this paper was devoted to showing that the logic considered is an accurate representation of the psychological mechanisms used by people in completing this task. Not only would this be extremely difficult to show, it is the purvue of psychologists and not logicians. The main goal of this paper was to introduce a possible alternative to classical logic that would, in some senses, at least, mimic how we function when doing scientific confirmation, while at the same time explaining how it is that people *could* be justified in their answers to the Wason task. Beyond just providing a neat way around the problems of contraposition, which are problematic not only in confirmation studies in philosophy of science but also in studies on causation (a causal conditional and its contrapositive are generally thought to not be equivalent, though I have not seen any study the purports to offer an acceptable logical alternative to classical logic, or one of its extensions, such as modal logic), relevance logic satisfies the desire for a requirement that there be a connection beyond what is necessary in classical logic between the antecedent and consequent of a conditional statement.

It is important to remember, though, that such a solution is required only if one finds

[14] This can be seen by considering a model just like the one discussed above except that $V(Q, 0) = T$. The proof could still be completed (this is easily enough checked), and yet the knowledge that *Q* would perforce not remain constant; sometimes we would have the knowledge that *Q* and sometimes the knowledge that *Not Q*; this corresponds to conflicting reports that you can get from different sources concerning different information.

the results of the Wason task explanable only by recourse to removing the paradox or changing the logic. There are certainly viable alternatives: Human 'rationality' is only so-called rationality, and the Wason task does not indicate that our logic is wrong but that our fellow people are in need of education. (The fact that different presentations of the task can lead to higher proportions of people getting the 'right' answer would support this claim. If people can do well when the questions concern underage drinking but poorly when the questions concern connections between letters and numbers, then perhaps we should be expending our energies to teach people how to recognize the tasks as variants of the same thing, rather than creating a possibly *ad hoc* explanation for their performances).

In sum, the moral of the story is: Hemple's paradox is only a paradox if you think it's one.

References

[1] Blackburn, S. (1996). *Oxford Dictionary of Philosophy*. Oxford: Oxford University Press.

[2] Hempel, C. G. (1945). "Studies in the logic of confirmation", *Mind* 54: 1-26.

[3] Humberstone, I.L. (1994). "Hempel Meets Wason", *Erkenntnis* 41: 391-402.

[4] Le Morvan, P. (1999). "The Converse Consequence Condition and Hempelian Qualitative Confirmation", *Philosophy of Science* 66: 448-454.

[5] Leavitt, F. (1996). "Resolving Hempel's Raven Paradox", *Philosophical Inquiry* 18, nos. 3/4: 116.

[6] Maher, P. (1999). "Inductive Logic and the Ravens Paradox", *Philosophy of Science* 66: 50-70.

[7] Nickerson, R. S. (1996). "Hempel's Paradox and Wason's Selection Task: Logical and Psychological Puzzles of Confirmation", *Thinking and Reasoning* 2 (1): 1-31.

[8] Priest, G. (2001). *An Introduction to Non-Classical Logic*. Cambridge: Cambridge University Press.

[9] Sylvan, R. & R. Nola. (1991). "Confirmation Without Paradoxes", *Advances in Scientific Philosophy: Essays in Honour of Paul Weingartner*, eds. G. Schurz and G.J.W. Dorn. Amsterdam: Rodopi: 5-44.

[10] Urquhart, A. (1971). "Semantics for Relevant Logics", *Journal of Symbolic Logic* 37 (1): 159-169.

[11] Waters, C. K. (1987). "Relevance Logic Brings Hope to Hypothetico-Deductivism", *Philosophy of Science* 54: 453-464.