Rationality and the Wason Selection Task: A logical Account*

Abstract

The main goal of the paper is to investigate the relation between indicative conditionals and rationality. We will do this by considering several interpretations of a very well-known example of reasoning involving conditionals, that is the Wason selection task, and showing how those interpretations have different bearings on the notion of rationality. In particular, in the first part of the paper, after having briefly presented the selection task, we will take a look at two pragmatic responses to the challenge posed by the task, through Wason’s notion of confirmation bias and Grice’s theory of conversational implicature. The second part will introduce Adams’ probabilistic view of indicative conditionals and will give reasons for preferring his account to those aforementioned. The conclusion will evaluate the question of human rationality in the light of the new standpoint acquired.

Duca Simone
University of Bristol

1 The Question

There is a venerable (philosophical) tradition which describes human beings as rational creatures. Famously, Aristotle said “Man is a rational animal”. It may be added that it is so unless we are very tired or angry or drunk, but generally we agree with that idea. However, during the last 40 years, 

*This work benefitted from the support of the European Science Foundation and the Arts and Humanities Research Council.
various psychologists have suggested the existence of much deeper problems than the above-mentioned exceptions. They pointed out that, given certain circumstances, people systematically fail to apply simple rules of (classical) logical inference. At this point, we could legitimately ask ourselves a very basic question. Let’s call it The Question: “Are people irrational?”.

2 The Plan

In the first part of this paper, we shall present where the psychologists’ worries about human rationality originated from, namely the Wason Selection Task or ‘The Task’, as we shall call it. Actually, we will see that, according to “traditional” cognitive psychology, The Task provided sufficient grounds to give a positive answer to The Question. The second part of the paper will take a broader perspective and will show that, in answering The Question, the interpretation of a certain type of conditional sentence, that is indicative conditionals, is crucially relevant. Firstly, we will take a brief look at a pragmatic defense of rationality through Wason’s notion of confirmation bias and Grice’s theory of conversational implicature. Secondly, we will introduce Adams’ probabilistic view of indicative conditionals and will give reasons for preferring his account to those aforementioned. The Conclusion will evaluate The Question in the light of the new standpoint acquired. In particular, all interpretations that we will be mentioning favour a negative answer to The Question, but Adams’, in virtue of its higher generality, seems to grasp the very essence of the problem.

1 Conditional sentences are traditionally classified in two main groups: indicative and subjunctive conditionals or counterfactuals. An example of the first kind is *If Brutus didn’t kill Caesar, then someone else did*. The corresponding counterfactual would be *If Brutus hadn’t kill Caesar, then someone else would have*. In general, indicative conditionals are of the form *did/did*, while counterfactuals has the *had/would have* form. Most importantly, we would tend to accept the former example, but not the latter. In fact, we would say that *If Brutus hadn’t kill Caesar, then someone else would have* is false, because, contrary to what the counterfactual seems to express, there is no necessity in the fact that Caesar was murdered.
3 The Task

Once a friend of mine came to me and said “I want to challenge your reasoning skills” and produced four cards. Each card had a number on one side and a letter on the other. Two cards were showing their number-side and the other two their letter-side. The cards were showing something like the following string: A, K, 2, 7. The Task was to determine which card/s should one turn in order to verify the following conditional sentence or ‘rule’: If there is vowel on one side of the card, then there is an even number on the other.

The Task, which was originally developed by the English psychologist Peter Wason, was meant to show that the majority of people do not reason consistently with the principles of classical logic, therefore the psychologists’ worries. But let us be more precise, starting with spelling out how The Task actually works. As a start, we can formalise the rule in the following straightforward way: \( p \supset q \), where \( p \) stands for There is a vowel on one side, \( q \) stands for There is an even number on the other and ‘\( \supset \)’ represents the material implication of classical logic. In this way, the string \( p, \neg p, q, \neg q \) (where \( \neg \) means “not”) would represent the previous arrangement of the cards. Let’s now take a look at a typical distribution of results regarding The Task (see Figure 1 below). It’s interesting to note that, after 40 years and innumerable iterations of the experiment, the distributions is still found to be the same.

Now, anyone who has some acquaintance with propositional logic should be quite startled by the data. In fact, the results show that only a small percentage of the participants choose the right cards, namely \( p \) and \( \neg q \). At this point, it should be pointed out that, in the last few decades the study of human reasoning had been dominated by roughly two tendencies. On the one hand, some researchers (such as Wason himself) considered classical logic as the ultimate criterion for rationality, possibly out of ignorance of
the alternatives; on the other, some others (such as Herbert Simon and Gerd Gigerenzer) were too quick in abandoning logic altogether. Having said this, it seems clear why cognitive psychology tended to interpret the previous data as undermining the idea that humans are rational. Nonetheless, the psychologists’ reaction is justified only in case “If... then...” is interpreted correctly as the classical Material Implication. Furthermore, classical logic cannot capture the notion of uncertainty, which is an essential feature of reasoning. Two questions may then arise: Does classical logic yield the best criterion of validity of inferential reasoning? What if we interpret “If... then...” differently? In this paper, we shall deal with the latter question. In particular, we will explore the idea that different interpretations may be able to give a better account of The Task outcome and, consequently, cast some light on the notion of rationality. However, let us conclude our analysis of The Task first.

We said that the correct answer to The Task was choosing $p$ and $\neg q$. Why? The key is given by the classical rule of inference called Contraposition, that is $p \supset q \vdash \neg q \supset \neg p$ and the corresponding axiom $(p \supset q) \equiv (\neg q \supset \neg p)$. Contraposition tells us that any argument with $p \supset q$ as premise and $\neg q \supset \neg p$ as conclusion is valid. Now, one way to make clear how this is relevant for solving The Task is saying that Material Implication puts a ‘condition’ on its antecedent. This means that, when we consider $p \supset q$, we should pick the card $p$, since that’s the relevant one. On the other hand, if we consider the logically equivalent $\neg q \supset \neg p$, the antecedent is $\neg q$; therefore, we should choose the corresponding card. Doing so, we have enough information on both components of the rule. This reasoning guarantees that, on our particular formulation of The Task, the correct answer consists in choosing A and 7. Another way to put this would be to look at another equivalence.

<table>
<thead>
<tr>
<th>Card/s selected</th>
<th># Persons (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$, $q$</td>
<td>46</td>
</tr>
<tr>
<td>$p$</td>
<td>33</td>
</tr>
<tr>
<td>$p$, $q$, not-$q$</td>
<td>7</td>
</tr>
<tr>
<td>$p$, not-$q$</td>
<td>4</td>
</tr>
<tr>
<td>Other combinations</td>
<td>10</td>
</tr>
</tbody>
</table>

Figure 1: A typical outcome regarding the Wason selection task (Wason & Johnson-Laird, 1972)
In fact, we know that the material conditional $p \supset q$ is also equivalent to $\neg(p \land \neg q)$. This means that the conditional is false when both $p$ and $\neg q$ are the case. Consequently, we should choose $A$ and $7$ to check for the falsity of the rule. So far, so good. It should now be clear how we should reason in order to give the right answer. However, we still have not explained of why the majority of people choose also irrelevant cards, notably $q$. The next two sections will be partly concerned with this problem.

4 Wason and Grice

A first attempt to explain why people generally perform The Task poorly came from Wason himself. Its basic idea is what in psychology and cognitive science is called Confirmation Bias. Roughly, this is a tendency to interpret new information in order to confirm one’s preconceptions, avoiding information and interpretations which contradict prior beliefs. Wason was among the first to investigate this phenomenon. His experiment consisted in presenting his subjects the following numerical series: 2, 4, 6, and telling them that the triple conforms to a particular rule. The participants were then asked to discover the rule by generating their own triples and use the feedback they received from the experimenter. Every time the subject generated a triple, the experimenter would indicate whether the triple conformed to a particular rule. The outcome of the experiment showed that, in spite of the fact that the rule was simply “Any ascending sequence”, the subjects often announced rules that were far more complex than the correct one. More interestingly, the subjects seemed to test only “positive” examples, that is triples that subjects believed would conform to their rule and confirm their hypothesis. The point is that they did not attempt to falsify their hypotheses by testing triples that they believed would not conform to their rule. Wason referred to this phenomenon as confirmation bias: Subjects systematically seek only evidence that confirms their hypotheses. In respect to The Task, this seems to imply that people would generally choose only those cards which could confirm the given conditional rather than refute it. However, it seems that this account does not fit with the empirical results. In fact, if confirmation bias was the whole story, then one would expect people to choose $\neg p$, along with the other “confirmatory” card, that is $q$, whereas they choose $p$, which is not “confirmatory”. Hence, the confirmation bias seems insufficient to
explain why people deviate from classical logic. It has been suggested that this may depend on whether we interpret the word ‘confirmation’ in a strict logical sense, as we did in our previous analysis, or in a broader psychological sense. In fact, confirmation bias has been advocated by several researchers in order to explain many heterogeneous phenomena, from The Task to people’s belief in pseudoscientific ideas, from governmental self-justification to medical diagnosis, from jury-deliberation processes to conservativism among scientists and even from witch-hunting to depression (See Nickerson, 1998, on distinguishing different notions of confirmation and for a very good survey of the interpretations of confirmation bias). If, on the one hand, this shows the pervasiveness of the notion of confirmation bias, then, on the other, investigators are left with further important questions: Is the confirmation bias a consequence of specific cognitive activities? Are those activities subject to any constraint? Does it persist because it has some functional value? Does it reflect a lack of understanding of logic? Is it more important to be able to make valid inferences from positive than from negative premises? Finally, the foregoing should be enough to allow us to maintain that the role of confirmation bias in the understanding of The Task is neither a central nor a privileged one, because of the presence of other basic issues regarding, for instance, logic and cognition. Let us now turn to another point of view on the matter coming from the philosophy of language. Grice’s famous account of conversation offers another possible way to interpret people’s general performance of The Task (Grice, 1975). Firstly, we need to introduce the concept of conversational implicature. The underlying idea is that what is literally said is not what is actually meant. For instance, if the teacher’s reply to the question “Is she a good student?” were “She always attends classes”, we would have some reason to believe that he implied she is not an outstanding student. Grice specifies this situation by means of some conversational maxims and a general principle of cooperation. Roughly, according to the maxims, one’s contribution to the conversation should be adequately informative, relevant, not believed to be false by its utterer and generally unambiguous and brief. The principle states that participants expect that each of them will make a conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose of the talk exchange. For instance, when a speaker makes an apparently uninformative remark such as “War is war”, the addressee assumes that the speaker is being cooperative and looks for the implicature
the speaker is making. We can see how the Gricean implicature is non-conventional, for it is drawn in accordance with pragmatic principles only, rather than involving the meaning of a linguistic expression. Finally, on this view, truth and assertibility crucially part ways. According to Grice, in fact, a proposition can be true, without being assertible: *If spaghetti grow on trees, then the pope is a German* is (vacuously) true, but not assertible, because the antecedent is false and a speaker will mislead her audience in uttering the corresponding sentence. This would contradict one of the maxims. Consequently, our intuition that that proposition is false is due, not to its falsity, but to its lack of assertibility. However, we could easily think about someone who would assert the sentence of the previous example and genuinely believe that the antecedent was true, without violating Grice’s maxim\(^2\). Be that as it may, since Grice’s theory of implicatures has been highly influential among both philosophers and psychologists interested in the study of natural language, it is beyond the scope of this paper to attempt any criticism of his ideas worth the name. Instead, we will just focus on how the foregoing is relevant to our particular discussion. So, how would a Gricean account of The Task look? Briefly, one might say that when The Task is asserted, its implicatures affect people’s reply to it. In particular, people seem to understand the “If...then...” of the conditional sentence to mean “…if and only if...”. In other words, that sentence might not come across as a material implication, but as a *biconditional*. On this account, in order to satisfy both the conditions put on the sentence by each “direction” of the biconditional, one would expect the majority of people to choose *all* the four cards. Nonetheless, that is contrary to what the data highlight, namely that the majority goes for *p* and *q* only. As a result, it seems that also Grice fails to provide a suitable justification of people’s response to the problem.

At this stage, before moving to the core of the paper, we should draw attention to the fact that since the literature that sprang from the Wason selection task is of gargantuan proportions and it is still growing, our exposition is by no means an attempt to give an even remotely comprehensive survey of it. However, it is worth just mentioning a few of different approaches that gained a lot of attention in more recent years.

\(^2\)Historically, this had been the case. When the BBC broadcasted a short documentary about the crop of spaghetti, hundreds of people contacted the production to ask where they could buy their own spaghetti trees. Needless to say, the show was a hoax and was transmitted on April fools’ day in 1957.
Even if our analysis will not be directly concerned with any of them, this brief synopsis will be useful later on in order to emphasize the novelty of Adams’ ideas. Let us begin our list with Sperber’s and Wilson’s attempt to develop Grice’s conversational implicature into a theory of relevance. Roughly, this means that the inferential component of communication rests very largely on the ability to work out what is and is not relevant, in terms of contextual effect, in what people are saying to you (Sperber, Cara, Girotto, 1995, for an application of this idea to the Wason selection task). Another very influential proposal was given in the framework of evolutionary psychology. In particular, Leda Cosmides claimed that people have not evolved in such a way that would allow them to perform The Task successfully. This idea seems to find strong support in experiments conducted with different versions of The Task involving “social exchange” scenarios. In these cases, people’s success in performing the tests highly increases (Cosmides, 1989). One more interesting project has been pursued by Oaksford and Chater. They have argued in various papers for an analysis of The Task based on the theory of optimal data selection in Bayesian statistics (Oaksford and Chater, 1994, for instance). By applying such standard, they try to justify the claim that the most frequent card selections are also the rational ones. However, so much for the alternatives. It is now time to start spelling out our suggestion by means of an analogy between psychology and the philosophy of language. From what has been said in this section, it seem that both a psychologist, Wason, and a philosopher, Grice, invoked pragmatic reasons, in order to clarify people’s behaviour towards (indicative) conditionals. In particular, pragmatic explanations have been given for the results of the Wason selection task and for the plausibility of the theory of material implication respectively. However, as the debate in the philosophy language regarding conditionals progressed beyond pragmatics, the material analysis of conditionals became less appealing, giving way to other explanations, such as, for instance, Adams’ probabilistic account. It may be argued then that what happened in philosophy of language could occur in psychology and thus that pragmatics would fall short of reasons also in the latter framework. In particular, we are interested in investigating whether Adams’ successful ideas in philosophy of language may cast some light on psychological issues as well. This will be the major concern of the next section. However, one should notice that there are several accounts based on probability available apart from Adams’. Roughly, one of the main reasons to prefer Adams’
view is the insight given by his new logical treatment of probability and its application to classical logic. More precisely, unlike the other accounts, e.g. Chater and Oaskford’s, Adams does not assign weights to objective probability measures in order to explain selection-task related data. Instead, his theory sets an alternative normative interpretation of rationality based on subjective probabilities of patterns of inferences. Let us now move on to the main section of the paper, where Adams’ account will be introduced and discussed in much more detail.

5 A Different Norm: Adams

In the previous section, we have seen how Wason and Grice maintained that the classical Material Implication provides a good rendering of the english “If... then...”. In other words, they endorsed a classical truth-functional interpretation of conditionals and focused their respective enquiries on related pragmatic issues. We will change our focus here to examine a different view originating from the philosophy of language, namely Ernest Adams’ probabilistic semantics for indicative conditionals (Adams, 1975; for a good overview of Adams’ ideas on this topic, see Bennett, 2003, ch. 9). Our main goal will be to show how this approach can provide a compelling normative account for rationality and to confront its explanatory power with respect to The Task with the one held by classical logic.

Among other things, Adams is credited with taking seriously the idea that conditionals can be accepted with different degrees of closeness to certainty and with developing that idea into a fully worked-out formal theory, which is known also as the Suppositional Theory of Conditionals. The reason behind this label is that, if we ask what it is to believe, or to be more or less certain, that q if p, e.g. that Ellen cooked the dinner if Lauren did not, that I recover if I sleep more, and so forth, we suppose (assume, hypothesize) that p, and make a hypothetical judgment about q, under the supposition that p, in the light of your other beliefs. That is how we make judgments such as the ones exemplified above. Originally, this idea was expressed by Frank Ramsey in a famous footnote to one of his paper:

“If two people are arguing ‘If p will q’ and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q; [...]they are fixing their degrees of belief in q, given p.” (Ramsey, 1929, p.129).
Notably, the first sentence of the previous quote is now known as the Ramsey Test for the acceptability conditions of conditionals and it has massively influenced the current debate on conditionals. One may say that if Ramsey originated this debate, Adams established some of the fundamental rules of it. Let us try to be more precise. It is important to notice that so far we have not mentioned truth conditions. This is not fortuitous. In fact, according to the Suppositional Theory, the connective ‘→’, commonly used to express the indicative mode of conditional sentences, is not truth functional and hence indicative conditionals do not have truth-values. In particular, the former tenet follows from the axioms of standard (Kolmogorov) probability calculus, while the latter and most interesting is proven by Adams (Adams 1975, p. 49-51 and for an illustration p.4). If we write the second claim formally as the following $u(A \rightarrow B) = \frac{u(A \supset B)}{p(A)}$, we may notice that this is equivalent to what Adams called Equation 1\(^3\), which says that the uncertainty of an indicative conditional is necessarily greater than the uncertainty of its material counterpart, unless either both uncertainties equal 0 or their antecedent $A$ has probability 1. This seems to highlight a connection between the equation expressed by the second idea and the standard ratio formula for conditional probabilities. Adams is aware of this and adds a very interesting remark involving the famous triviality result by Lewis. A possible way to put this is: “If the conditional probability measure for conditional’s probabilities is correct, and given other standard assumptions of probability theory, there is no way of attaching dichotomous truth values to conditionals in such a way that their probabilities will equal their probabilities of being true”\(^4\). As Adams points out, this gives good grounds to believe that conditional sentences lack truth-values. However, if this is so, then a major question arises: How can arguments using indicative conditionals be valid? Adams quick reply would be that ‘→’ expresses a high conditional (subjective) probability. Let us see what this means. One of Adams’ main insights on this topic is given by the notion of probabilistic validity. We will introduce this new concept by means of uncertainty. In probability theory, a proposition’s uncertainty amounts to its improbability, which equals 1 minus its probability. This leads to the

\(^3\)Ivi. p. 3.
\(^4\)Ivi. p. 5.
following identities:

\[ u(A) = p(\neg A) = 1 - p(A), \]

where \( u(\ ) \) and \( p(\ ) \) are uncertainty and probability functions respectively and \( A \) is a single proposition, that is \( A \) does not stand for an indicative conditional. In Adams’ terminology, \( A \) is a factual sentence, not a conditional one. We will be more precise about this distinction later. What is important now is that, according to Adams, indicative conditionals do not express propositions and, therefore, no probability (of being true) can be meaningfully assigned to them\(^5\). Now, to get to grips with the notion of probabilistic validity, one should keep in mind that, as classical validity does not allow falsity to enter along the way from premises to conclusion, in the same way, a high degree of uncertainty isn’t allowed to enter in any probabilistic valid forms of argument. In other words, in a probabilistically valid argument, if the probabilities of the premises tend to 1, then the probability of the conclusion also necessarily tends to 1. As a consequence, a probabilistically invalid argument is such that if the probabilities of the premises tend to 1, then the probability of the conclusion does not necessarily tend to 1. To be more precise, Adams has proven the following result:

**Theorem 1 (Adams’ probabilistic definition of validity)**

An inference \( Q_1, \ldots, Q_n \therefore Q \) is probabilistically valid (p-valid) if it satisfies the Uncertainty Sum Condition (USC):

\[ u(Q) \leq u(Q_1) + \cdots + u(Q_n), \]

for all uncertainty functions \( u(\ ) \)\(^6\).

\(^5\)Hannes Leitgeb offers an alternative to the standard debate on this point. Instead of assuming that either an indicative conditional expresses a unique proposition or does not express a proposition at all, he argues for the position that each indicative conditional corresponds to two propositions, which are the contents of so-called conditional beliefs. According to this view, this is how we believe in indicative conditionals. For more details, see Leitgeb, 2007.

\(^6\)In contrast with letter like \( A \) and \( E \) in italics, which are metalinguistic variables ranging over factual formulae of our language, capital letters, such as, \( Q \), with or without subscripts, are metalinguistic variables ranging over all formulae of the language, including factual and conditional ones. We adopt this convention from Adams. See Adams, 1998, p. 151.
An important remark is in order. USC is a basic theorem of probability logic and it is the key for justifying Adams’ application of probability to deductive logic or, in his own words, “the application of deductive logic to inferences with somewhat uncertain premises, so long as there are not too many of them” (Adams, 1998, p.38). Most importantly, the fact expressed by USC, that the uncertainty of the conclusions of an argument cannot be greater than the sum of the uncertainties of the premises, is a necessary and sufficient condition for inferences without conditionals to be classically valid. This entails that, as long as arguments with only factual sentences are concerned, the notions of probabilistic and classical validity are equivalent. With the following theorem, Adams established that this is true also for arguments with any number of premises:

**Theorem 2** If $A_1, \ldots, A_n$ entail $C$, then $u(C) \leq u(A_1) + \ldots + u(A_n)$, for all uncertainty functions $u(.)$.

This means that the uncertainty of the conclusion of any classically valid argument cannot be greater than the sum of the uncertainties of the premises. Moreover, this shows that any classically valid argument using factual sentences is also probabilistically valid.

At this point, it seems useful to concretely show that Theorem 2 holds by means of an example. In particular, we will construct a model, which illustrates that arguments of the classically valid form $A_1, \ldots, A_n \vdash A_1 \land \cdots \land A_n$ are also probabilistically valid. Let us consider the simplest instance of the argument, consisting of the two premises $A$ and $B$ and the conclusion $A \land B$. Our model consists in four possible worlds, which have been assigned a weight each, according to their respective probability of being the case, in the following way: $p(\{w_1\}) = 0.2$, $p(\{w_2\}) = 0.4$, $p(\{w_3\}) = 0.1$ and $p(\{w_4\}) = 0.3$ (See figure 2 below).

---

7It is important to keep in mind that the results discussed in this section are understood in terms of subjective probabilities of an idealised thinker who knows all the logical truths and makes no logical mistakes.

8Usually, mathematicians tend to assign probability or uncertainty functions to sets, while philosophers are inclined to assign them to formulae. However, the connection between the two practices seems quite straightforward. Precisely, when one refers to formulae, usually assume that each formula express a proposition. Now, since possible worlds may be understood as sets of propositions, it is easy to see why our notation should make mathematicians happy as well.
Now, it is obvious what the probabilities of the premises and the conclusion are. Remembering that, for factual sentences, ‘probability’ means ‘probability of being true’, premise $A$ has probability 0.6 of being true, premise $B$ has probability 0.7 and the conclusion $A \land B$ has probability 0.4. One should notice that, in spite of the fact that the probability of each premise is greater than the probability of the conclusion, the argument is valid. Let us spell this out in terms of uncertainty: premise $A$ has uncertainty 0.4 of being true, premise $B$ has uncertainty 0.3 and the conclusion $A \land B$ has uncertainty 0.6. This amounts to say that the uncertainty of the conclusion is less than the sum of the uncertainties of the premises. Therefore, the argument satisfies Adams’ probabilistic criterion of validity and, since it deals with factual sentences only, it is both classically and probabilistically valid. This reasoning shows another important consequence: “[...] p-validity differs from more traditional concepts of validity, e.g. that a conclusion must be ‘at least as true’ as its premises” (Adams, 1998, p.153). Another way of putting the foregoing appeals to the mathematical notion of limit. In particular, what Adams found is that it is impossible to make the probability of each premise tend to 1 at the limit, without their conjunction (in this case, the conclusion) tending to 1 as well.

After having made the acquaintance of p-validity relative to arguments with only factual sentences, it is time to have a look at how this notion allows us to analyze arguments with indicative conditionals, in spite of their lack of truth-values. Suppose we have fixed the semantics for factual sentences $A$, $B$, $C$, $D$, $E$, ... of a propositional language $\mathcal{L}$ in the standard classical way,
and for indicative conditionals of the form \( A \rightarrow B \), where \( A \) and \( B \) are members of \( \mathcal{L} \). Now, consider arguments of either of the following forms:

\[
\begin{array}{c}
A_1 \\
\vdots \\
A_m \\
B_1 \rightarrow C_1 \\
\vdots \\
B_n \rightarrow C_n \\
\hline
D
\end{array}
\quad
\begin{array}{c}
A_1 \\
\vdots \\
A_m \\
B_1 \rightarrow C_1 \\
\vdots \\
B_n \rightarrow C_n \\
\hline
E \rightarrow F
\end{array}
\]

According to Adams’ semantics such arguments are probabilistically valid if and only if, for all sequences \( p_1, p_2, p_3, \ldots p_i, \ldots \), where \( i \) tends to infinity, of subjective probability measures \( p_i \) on \( \mathcal{L} \), the following holds:

If

\[
\begin{align*}
p_i(A_1) & \quad \text{tends towards 1,} \\
p_i(A_m) & \quad \text{tends towards 1,} \\
p_i(C_1 | B_1) & \quad \text{tends towards 1,} \\
p_i(C_n | B_n) & \quad \text{tends towards 1,} \\
p_i(D) & \quad \text{tends towards 1}
\end{align*}
\]

Then

\[
\begin{align*}
p_i(F | E) & \quad \text{tends towards 1}
\end{align*}
\]

(where if \( p_i(\varphi) = 0 \), then the conditional probability \( p_i(\psi | \varphi) \) of the corresponding conditional, by stipulation, equals 1). One could encapsulate this idea in the following slogan: “The more certain the premises, the more certain the conclusion”.

In summary, Adams found that arguments using factual sentences are classically valid if and only if they are probabilistically valid. On the other hand, arguments using indicative conditionals can be evaluated only on the basis of probabilistic validity, because no truth-values can be meaningfully assigned to conditional sentences. In the next subsection, we shall give an
example of how that works, and we will show why this is particularly relevant for our general discussion.

5.1 $A \rightarrow B : \neg B \rightarrow \neg A$

As the above section title indicates, the main concern of this section will be the classically valid rule of inference called Contraposition. This already seems to require justification. As anyone familiar with the psychological literature on the Wason selection task would know, the reason why people generally fail (according to the classical view) to give the correct answer to the task is not clear. Aside from invoking pragmatic factors (see section 4), there are two possible explanations for this failure. We could say that the outcome is due to the fact that people fail to apply the classical valid form of Modus Tollens (MT) or we could say that they fail to employ the rule of contraposition. Let us spell out this distinction. Suppose we were to consider the upshot of The Task in terms of MT. Then our concern would be why people are not drawing an inference from the two premises $A \rightarrow B$ (the given rule of The Task) and $\neg B$ to the conclusion $\neg A$. Instead, if we see the issue in terms of contraposition, we should explain why people do not reason from $A \rightarrow B$ (the given rule) to $\neg B \rightarrow \neg A$ and then, by Modus Ponens (MP), from the latter and $\neg B$ to $\neg A$. This distinction could seem rather irrelevant to psychologists. The reason is that they usually endorse classical validity, according to which both contraposition and MT are valid and, therefore, it does not matter which inferential step people actually carry out. Consequently, they would conclude that logic is insufficient to explain the data, mainly because they understand ‘logic’ as ‘classical logic’, and therefore they would look for an explanation elsewhere, e.g. pragmatic assumptions. On the other hand, we can avoid that situation by considering, as it were, the right logic for the right kind of conditionals. In particular, it seems that Adams’ normative standard (which is usually overlooked by psychologists) can better accommodate the data without charge of irrationality. More precisely, since MT is probabilistically valid, but contraposition is not (as we will see later on), it does matter whether people use the former or the latter. This amounts to saying that, roughly speaking, if people use MT, they are rational, but if they use contraposition, they are irrational. From the foregoing, we can envisage two interrelated reasons for preferring the explanation in terms of contraposition. Firstly, since contraposition is not probabilistically valid,
the fact that people do not implement that pattern seems perfectly rational. So, perhaps, unless further empirical evidence is provided, we should prefer explanations according to which people turn out to be rational rather than irrational (following some sort of Principle of Charity). Thus, if the possible answers to “Why do people not turn the \( \neg q \)-card?” are (i) because they think MT is invalid, which is irrational, or (ii) because they think contraposition is invalid, which is rational, we should opt for the latter explanation. Secondly, in the reasoning involving contraposition, the only inference from a conditional and a factual premise that people need is MP. At this point, we may even speculate that it would not be surprising if research on the human brain uncovered evidence that MP is the most basic inference that we can use. Be that as it may, it is important to keep in mind that psychologists should not overlook the previous distinction, because of its bearing on questions about rationality. We can now go back to our analysis of contraposition.

What we really want to explore here is what happens to Contraposition when we plug indicative conditional signs into the places of material implication signs. In other words, how can we evaluate the argument \( A \rightarrow B \therefore \neg B \rightarrow \neg A \)? Needless to say, our main tool to spell that out will be probabilistic validity. Let us consider the following equivalent version of Contraposition:

\[
B \rightarrow \neg A \therefore A \rightarrow \neg B,
\]

where ‘\( \rightarrow \)’ is the non-truth-functional connective for indicative conditionals. Our next move is showing that Contraposition is a probabilistically invalid rule of inference. Let us look at an example of it. If we substitute the sentences Jane is drinking vodka and Jane is at least 16 years old for \( A \) and \( B \) respectively, we obtain

- **Premise:** If Jane is at least 16 years old, then she is not drinking vodka

\[ \therefore \]

- **Conclusion:** If Jane is drinking vodka, then she is less than 16 years old

Now, recalling that according to Adams’ view, ‘\( \rightarrow \)’ expresses a high conditional probability, it should be intuitive that it might be reasonable
not to draw the above inference. The reason is that the probability of the premise might be high, while at the same time the probability of the conclusion might be low. Let us assume that our example refers to a country where people are allowed to drink if and only if they are at least 16 years old, and where people usually abide the law. Then we would assign a high conditional probability to the premise, because most people might not drink vodka at all, but not to the conclusion, because in this country people very rarely break the law. This means that we have reasons to believe that the argument does not accomplish the definition of probabilistic validity. In conclusion, Contraposition is probabilistically invalid. Nonetheless, we may still feel some skepticism about this reasoning. Conveniently, Adams is in the position of giving us some more reason to buy his explanation. He does so through an ingenious application of Venn diagrams. Firstly, we want to characterize the already-mentioned distinction between factual and conditional sentences by means of Venn diagrams. Let us assume that all possible worlds are represented by points enclosed in a rectangle $D$. Factual sentences such as $A$ and $B$ are represented by the corresponding sub-regions of $D$ containing the possible worlds in which the propositions are true. This is shown in Figure 3 below:

![Venn Diagram](image)

Figure 3: The diagrammatic representation of a highly probable conditional $A \rightarrow B$.

In this case, the advantage of the Venn diagram is representing all the truth-functional combinations of $A$ and $B$, such as $\neg B$ and $A \land B$. On the other hand and most importantly, the diagram doesn’t give us a representation for conditional sentences in terms of region, which is in accordance with the assumption that they lack truth-values. However, we did not say anything about probabilities yet. Adams’ move is identifying
the areas of the sub-regions corresponding to factual propositions with the probabilities of the propositions. It’s should be noted that this makes sense if the area of $D$ is assumed to equal 1. For this reason, we can say that the bigger the region corresponding to a proposition, the bigger the area, and therefore the bigger the probability of the proposition. In Adams’ words: “The diagram represents a possible probabilistic state of affairs, where propositions corresponding to large regions are represented as probable while propositions corresponding to small regions are represented as improbable” (Adams 1975, p.10). The case of conditional sentences is unfortunately more complicated. In fact, neither their corresponding propositions are represented by regions of $D$, nor their probabilities by areas of $D$. Nonetheless, this seems to fit nicely with the fact that, unlike unconditional probabilities, conditionals’ probabilities are not probabilities of truth. The probability of a conditional $A \rightarrow B$ is instead identified in the diagram with the proportion of sub-region $A$ which lies inside sub-region $B$. But this is just the geometric analogue of saying that the probability of a conditional $A \rightarrow B$ is the ratio of the probability of the conjunction $A \land B$ to the probability of its antecedent $A$, that is the standard definition of \textit{conditional probability}: 

$$p(B|A) = \frac{p(A \land B)}{p(A)}$$

On this view, if most of region $A$ corresponding to the antecedent of a conditional lies inside region $B$ corresponding to the consequent, the probability of the conditional $A \rightarrow B$ should be understood as \textit{high}. Conversely, if most of region $A$ lies outside region $B$, the probability of the conditional $A \rightarrow B$ is \textit{low}. Thus, it should be easy to see how Figure 3 shows the diagrammatic representation of a highly probable conditional $A \rightarrow B$. In fact, most of region $A$ lies inside region $B$, therefore the probability of the corresponding conditional $A \rightarrow B$ is high. At this point, a couple of important limitations concerning probabilistically interpreted Venn diagrams should be noted. The first is that Adams’ view deliberately ignore the case in which the probability $p(A)$ of the antecedent of the conditional $A \rightarrow B$ equals 0. This is because, when it is taken into account, it may cause problems to the given picture of conditionals and their probabilities. In fact, whenever $p(A) = 0$, the probability of $A \rightarrow B$ is not defined, since the corresponding proportion is not defined. The second constraint is that probabilistic Venn diagrams can represent factual and \textit{simple conditional}
propositions, but not more complex constructions, such as conjunctions, disjunctions and negations of conditionals. This is the diagrammatic expression of the fact that it is highly problematic to assign probabilities to such constructions. However, dealing with this problem would be beyond the scope of this paper, so let’s go back to our main concern.

We have now the necessary informations to turn to our example involving Contraposition. With the help of Venn diagrams, we want to show Contraposition to be probabilistically invalid. Let us start with its premise $B \rightarrow \neg A$ and probabilities being given by Figure 4 below. Its probability is high, because most of region $B$ lies inside region $\neg A$. A concrete instance that fits with the previous diagram is our previous *If Jane is at least 16 years old, then she is not drinking vodka.*

![Figure 4: The diagrammatic representation of a highly probable conditional $B \rightarrow \neg A$.](image)

On the other hand, Figure 4 also shows that the probability of the corresponding conclusion $A \rightarrow \neg B$ is low. In fact, most of region $A$ lies inside region $\neg B$. In this case, a concrete example is *If Jane is drinking vodka, then she is less than 16 years old.* Finally, on the one hand the probability of the premise of our argument is highly probable, on the other the conclusion is assigned a low probability. But, if this informal reasoning in terms of Venn diagrams is turned into more formal reasoning in terms of sequences of probability measures, then one can see that it is possible to construct a counterexample to our definition of probabilistic validity. In particular, we could construct a model, similar in kind to the one of Figure 3 above, that shows that the uncertainty of the conclusion is actually greater than the sum of the uncertainties of the premises (which in our previous example consist of only one). Therefore we can conclude that
contraposition is a probabilistically invalid rule of inference. However, one might still feel somewhat baffled by this reasoning, maybe because the previous example was given for the less familiar (but nonetheless equivalent) version of contraposition, that is $B \rightarrow \neg A \therefore A \rightarrow \neg B$. Hence, at the cost of being redundant, I will provide one more example showing how Contraposition, in its standard formulation, that is $A \rightarrow B \therefore \neg B \rightarrow \neg A$ fails to be probabilistically valid. Here is a counterexample to Contraposition illustrated by means of Adams’ adaptation of Venn diagrams.

![Venn Diagram](image)

Figure 5: The probability of the conditional $A \rightarrow B$ is high, while the probability of $\neg B \rightarrow \neg A$ is low.

Now that we have hopefully cleared any doubt about Contraposition, we will try to gauge whether at this stage we are in any position to attempt an answer to our initial question: “Are people irrational?”.

### 6 The Conclusion

We started this paper presenting the following view: Since people, when confronted with simple selection tasks, fail to apply basic, classically valid rules of inference, such as contraposition, one needs to look for an explanation of the data outside logic. As examples of pursuing this route, we suggested Wason’s and Grice’s approaches. In our view, they both failed to recognize that logic can still tell us something important regarding the notion of rationality, provided that we use it in a suitable way. In the main section of the paper, we tried to show that Adams’ theory of conditionals provides such a way. The reason Adams’ point of view seems compelling is
that, by means of a general probabilistic perspective, it takes into account a fundamental feature of human rationality, that is subjective uncertainty. Furthermore, it provides a new normative interpretation of conditionals, according to which contraposition, whose failed application was held responsible by psychologists for people’s poor performance of The Task, is invalid. Hence, we gained a new perspective on how to define rationality. In doing so, Adams can justify why people fail to apply Contraposition on normative grounds. In other words, he can explain, unlike Wason and Grice, our failure to apply a logical principle of inference within logic, without endorsing any ‘merely’ pragmatic justification. In fact, we can infer that, since contraposition is probabilistically invalid, people’s failure of applying contraposition to the solution of The Task may well be considered as a rational response. However, if on the one hand, Adams’ account, unlike Wason’s and Grice’s, gives a purely logical explanation of why people don’t generally choose the card corresponding to \( \neg q \), on the other, it still does not justify people’s choice of the \( q \)-card. At this point, the best that we can say is that, according to Adams’ view, people are rational in not drawing the inference based on contraposition, but they still are irrational in not drawing the inference on *Modus Tollens*, which is probabilistically valid. On these grounds, we can conclude that even if our suggestion does not completely rescue people from irrationality, it seems to support at least the idea that rationality is not an “on-off” phenomenon, but a partial one. It is plausible that some help may come from psychology itself. In fact, recent studies in the field seem to give empirical evidence for Adams’ view (Pfeifer and Kleiter, 2005, 2007). Further research in both psychology and logic may corroborate the idea that people actually do what they *ought* to do when confronted with The Task. If this is the case, rationality would be an essential property of what we may tentatively call ‘probabilistic cognizers’.

---

Adams himself is not exactly crystal clear on the matter. For instance, when he raises the question about the status of every day inferences which are not probabilistically valid, he says that contraposition is rational, but “it is not rational in virtue of being of the contraposition form” (Adams 1975, p.15). According to him, in fact, further conditions need to be met for that reasoning to be rational. These conditions are not part of the meaning of propositions such as \( B \rightarrow \neg A \), but they obtain when people are *told* such propositions. Adams goes further and makes some comments worthy of being reported here. He calls the conditions under which a pattern is probabilistically valid its conditions of *partial rationality*. 
References


