Everyday Reasoning with Conditional Sequences

Ruth M. J. Byrne

MRC Applied Psychology Unit, Cambridge, U.K.

Everyday conditional reasoning commonly occurs in a sequence, such as:

... If she leaves home early today then she will get to work quickly.
If she gets to work quickly then she will get extra things done ...

Experimental studies of conditional reasoning assume that people make the same amount of inferences from sequences as they do from simple arguments that contain just a single conditional premise. In a series of four experiments this assumption is shown to be wrong. Our first experiment showed that people make more inferences from transitive sequences, and fewer inferences from certain attransitive sequences, than from corresponding simple arguments. We established in the second experiment that these differences arise because people can construct a single representation of a sequence. The third experiment replicated the result that people make fewer inferences from certain attransitive sequences. In the final experiment we found that people make more inferences from transitive sequences, regardless of the presence of negatives. We argue that the inferences people make from common everyday sequences depends on their representation of the state of affairs that sequences describe.

We frequently reason about the contingent relations between hypothetical events—this activity is central to our ability to plan and predict. Our reasoning commonly proceeds in sequences—for instance:

... If she leaves home early today then she will get to work quickly.
If she gets to work quickly then she will get extra things done ...

Requests for reprints should be sent to Ruth M. J. Byrne, MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge, CB2 2EF, U.K.

I want to thank Mark Keane and Phil Johnson-Laird for their instructive comments on many drafts of this paper; and also Angus Gealotly, Nick McDonald, Stephen Newstead, Howard Smith, and the anonymous referees for a critical reading of an earlier draft. Most of the experiments were conducted while I was in receipt of a postgraduate award from the Trinity Trust, of Trinity College, University of Dublin, Republic of Ireland. I am indebted to my parents Paul and Mary Byrne, and to all my family, especially Melly Tracey, Linda Deighan, and Cathly Fitzgerald, for their encouragement. Some of the results were presented at the Second Annual Conference of the BPS Cognitive Section at Oxford in 1985.

© 1989 The Experimental Psychology Society
In each sequence, the conditional sentences are related by a shared proposition. In this case, "getting to work quickly." We can make several types of inferences when a conditional sentence is accompanied by an assertion about its antecedent (the "if" clause) or its consequent (the "then" clause). For example, we can make a series of valid modus ponens inferences, when the first antecedent of a sequence is affirmed:

...If she leaves home early today then she will get to work quickly.
If she gets to work quickly then she will get extra things done...
She left home early today.

The first inference supports the intermediate conclusion:
Therefore, she got to work quickly.
and on the basis of this conclusion, the second inference yields:

Therefore, she got extra things done.

We can make a series of valid modus tollens inferences when the last consequent is denied:

...If she leaves home early today then she will get to work quickly.
If she gets to work quickly then she will get extra things done...
She did not get extra things done.

The first inference supports the intermediate conclusion:

Therefore, she did not get to work quickly.
and on the basis of this conclusion, the second inference yields:

Therefore, she did not leave home early today.

As well as these valid inferences, we sometimes make inferences that logicians have decided are invalid, i.e., e.g., Croll, (1982). For example, the denial of the antecedent gives rise to a series of fallacious inferences:

...If she leaves home early today then she will get to work quickly.
If she gets to work quickly then she will get extra things done...

She did not leave home early today.
Therefore, she did not get to work quickly.
Therefore, she did not get extra things done.

These inferences are fallacies because there might be alternatives, for instance, to leaving home early, which could also ensure getting to work quickly. Similarly, the affirmation of the consequent also gives rise to a series of fallacious inferences:

...If she leaves home early today then she will get to work quickly.
If she gets to work quickly then she will get extra things done...

She got extra things done.
Therefore, she got to work quickly.
Therefore, she left home early today.

Studies of the four inference types—modus ponens and modus tollens, denial of the antecedent and affirmation of the consequent—have concentrated on simple arguments. Simple arguments consist of a single conditional premise accompanied by an assertion about its antecedent or consequent. Typically people make the valid inferences; however, they also frequently make the fallacious inferences (for reviews, see Wasserman and Johnson-Laird, 1972; Evans, 1982). Conditional sequences have been used to investigate memory for arguments (Mackintosch, 1992) and in tasks analogous to the conditional syllogism (Croll and Butterfield, 1981, Experiment 5). But the inferences people make from conditional sequences have never been studied, despite the ubiquity of sequences in everyday planning. Instead, it has been assumed that reasoning performance on simple arguments can generalize directly to sequences. This assumption is made explicit in formal theories of reasoning, which posit that people make inferences by applying formal, syntactic rules (e.g., Garfield, 1972; Braine, 1984; Shoben, 1984; Croll & Croll, 1982). In formal theories, reasoning from sequences comprises a simple iteration of the inference rules that are applied singly in simple arguments. In sequences, the rules are applied repetitively both on the information given in the premises and to the conclusions of each premise.

However, there is at least one reason to suspect that people might make more inferences from sequences than from simple arguments: they can integrate the information in the two conditions of a sentence into a single representation. The premise:

If she leaves home early today then she will get to work quickly.
If she gets to work quickly then she will get extra things done.

suggests a model similar to:

Leaves early——get to work quickly——get things done

gets to work quickly——gets things done

where the first premise is represented as meaning that the first event leads to the second event. The second event might occur without the first event, and so an occurrence of the second event on its own is included in the model, tagged by an "O" to indicate that it is optional (cf. Johnson-Laird, 1983).

Indeed, such optional events might be mentally omitted from the model (Johnson-Laird & Byrne, 1984; Johnson-Laird & Byrne, 1988). The meaning of the second premise is then integrated into the model of the first premise.

Thus, the final model indicates that the first event leads to the second event,
which, in turn, leads to the third event. The premises of simple arguments obviously cannot be integrated into a single model. The integrated representation of a conditional sequence supports the same type of inference as the separate representations of each simple argument. But, we know that people find it easier to make inferences from an integrated representation of categorical syllogisms (e.g., Johnson-Laird & Breidenbach, 1978), three-term syllogisms (e.g., Potts, 1972; Stegeberg, 1953), and spatial relations (Elwyn & Johnson-Laird, 1980). Perhaps they will also find it easier to make inferences from the integrated representation of a conditional sequence. Our first aim was to establish whether or not people draw as many inferences from sequences as they do from simple arguments.

EXPERIMENT 1

The first experiment examined whether people make more, or fewer, inferences from conditional sequences than from simple arguments. First, we compared reasoning from transitive sequences, such as:

- If she rings her friend then she will go to visit her.
- If she goes to visit her friend then she will listen to music.
- She rang her friend.

with two comparable simple argument pairs:

- If she rings her friend then she will go to visit her.
- She rang her friend.

We expect that people can construct an integrated representation of the information in the sequence, but not of the information in the simple argument pair.

Although a transitive sequence yields an integrated representation, an associative sequence does not. Consider a sequence where the first occurrence of the shared proposition is negated:

- If she rings her friend then she will not go to visit her.
- She rang her friend.

It is difficult to assimilate the information in this associative sequence into an integrated representation. It supports the model:

*If she rings her friend, she will not go to visit her.*

As it is hard to represent associative sequences, perhaps people draw fewer inferences from them than from transitive sequences. They might also draw a different amount of inferences from associative sequences and from corresponding pairs of simple arguments. We compared such associative sequences with comparable simple argument pairs, such as:

- If she rings her friend then she will go to visit her.
- She rang her friend.

and:

- If she goes to visit her friend then she will listen to some music.
- She did not go to visit her.

One kind of associative sequence results from negating the first occurrence of the shared proposition, but another kind results from negating the second occurrence of the shared proposition:

- If she rings her friend then she will go to visit her.
- If she goes to visit her friend then she will listen to some music.
- She rang her friend.

It is also hard to construct an integrated representation of this associative sequence:

*If she rings her friend, she will go to visit her.*

We examined this negative-antecedent, associative sequence, because people make different inferences when the presence and location of a negative is varied in simple arguments (e.g., Adams, 1968; Evans, 1992, 1972; Pollard and Evans, 1980; Reber, 1983). We compared it with a corresponding pair of simple arguments:

- If she rings her friend then she will go to visit her.
- She rang her friend.

and:

- If she does not go to visit her friend then she will listen to some music.
- She went to visit her friend.

*If she rings her friend, she will go to visit her.*

We thank Simon Crookall for suggesting this explanation.
Method

Design: There were six groups of subjects in this experiment. Each group received a single type of argument: affirmative transitive sequences, negative-consequent intransitive sequences, negative-antecedent intransitive sequences, or the simple arguments corresponding to one of these types of sequences. The arguments received making inferences corresponding to medius pinnus, medius tollens, denial of the antecedent, and affirmation of the consequent.

For each of the three types of sequence, we constructed six types of argument, by manipulating whether the accompanying assertion entailed or denied the first, middle, or last term. People could make two inferences from each type of argument. We assigned two lexical contexts to each of the six argument types and the resulting 12 problems yield 24 inferences (12 intermediate conclusions and 12 final conclusions). But we included two of the argument types(those constructed by accompanying the sequence with an assertion about the middle term—as did problems only, and the remaining four argument types yielded 3 problems of interest. There were 24 simple arguments, as there were two simple arguments corresponding to each sequence. We presented the arguments appropriate to each group in a different random order to each subject of that group, and the simple arguments were presented, not as pairs, but also in a random order.

Materials: We generated 12 pairs of lexical materials for the 3x conditions. The materials involved realistic mundane events (examples of the domain and see the Appendix). They referred to the actions of a single person (identified as “she” for female subjects and as “he” for male subjects). The nature of the actions and events was such that they could be either transitive or reflexive, assigned at either the middles or terms. The actions were arranged in a temporally sequential order. We gave subjects in all conditions the same lexical materials.

Procedure: The subjects were tested individually. Their instructions were illustrated with an example, which for the sequence groups was:

If she gets up early then she will go into town.

If she gets up early then she will go into town.

Then, they were instructed to imagine that, having been given the information in the premises, some time later they heard the further information.

She goes up early.

Their task was to say what they considered to follow from this information. The uncontrolled construction of their own conclusions allowed subjects to formulate any type of conclusion and to make however many conclusions they wished to make. For each sequence, if a subject made an inference about only one of the two possible events that they could refer to in their conclusions, they were given the groups, “anything else?” They were given examples of several responses, which for the sequence groups were

“you might think it follows that she went into town, or that she did not go into town, or that you can't say for definite whether she went into town or not, and you might think it follows that she spent all her money, or that she did not spend all her money, or that you can't say for sure whether she spent all her money or not”.

For the simple arguments groups, the examples responses referred to the first conditional only.

Subjects: Sixty undergraduates from Trinity College, University of Dublin, participated in this experiment. They were randomly assigned to one of the six conditions (m=10). Six subjects were replaced during the experiment, prior to any data analysis, because they had received false instructions.

Scores: Subjects could make at least two inferences from each sequence, or from each pair of simple arguments. They considered the hypothetical case where subjects made the first inference, which led to an intermediate conclusion in a sequence, affirming both from sequences and from simple arguments. For the sequences, this conclusion increases the frequency of the second inference, because the second inference depends on the intermediate conclusion. Hence, we considered the case, in which subjects could make both inferences, either in a sequence or in a pair of simple arguments. Consequently, we assigned a value of 1 to a subjects’ conclusions from a sequence only if they made both inferences from

*See Gordon J.S. & T. Evans for pointing out that some of the context could exist.
that sequence. That is, if subjects responded to the sequence:

If she works hard then she will leave work early.
If she leaves work early then she will eat out.
She works hard.
with the two conclusions:
She leaves work early.
She eats out.

then they were given a score of 1 for that sequence. Similarly, we assigned a
value of 1 to the subjects’ conclusions from a pair of simple arguments only if
they made both inferences from that pair. That is, if subjects responded to
the simple argument:

If she works hard then she will leave work early.
She works hard.

with the conclusion:

She leaves work early.

and they also responded to the simple argument:

If she leaves work early then she will eat out.
She leaves work early.

with the conclusion:

She eats out.

then they received a score of 1 for that pair of simple arguments. A subject’s
responses were given a score of 0 when they did not make the first inference
(or made it probabilistically, as indicated by such modal qualifiers as
“possibly”, “probably”, “maybe”), or when they made the first inference
but did not make the second inference (or made it probabilistically), or
when they made neither inference (or made them both probabilistically).

An important feature of this scoring procedure is that it allows us to
remain neutral with respect to the logical accuracy of the inferences, because
they are assigned a score regardless of whether they are logically valid or
invalid. In this respect, the scoring follows the recent trend (e.g., Evans, 1982)
to examine the overall frequency of inferences that people make, rather than
the frequency of logically correct inferences that they make. This scoring
convention is best suited to our purpose of investigating the possibility that
constructing an integrated representation of a sequence can affect the
frequency of any inference that people make.

## Results and Discussion

The results corroborated our expectations. There were considerable dif-
ferences in the frequency of inferences that subjects made from sequences
and from corresponding simple arguments, and there were also differences
between the transitive and intransitive sequences. The percentages of in-
ferences that the subjects made from the transitive sequences and their
corresponding simple arguments are presented in Table 1 and the percent-
ages of inferences from the two intransitive sequences and their corresponding
simple arguments are shown in Table 2.

The size and direction of the differences between sequences and simple
arguments varied with the type of sequence, as shown by an analysis of
variance, F(2, 44) = 1.1, p > .05. (7) Subjects made more inferences, both

---

**Table 1**
The Percentages of Inferences Made from the Transitive Sequences and
Simple Arguments in Experiment 1

<table>
<thead>
<tr>
<th>Argument Type</th>
<th>MP</th>
<th>DP</th>
<th>AC</th>
<th>MT</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affirmative transitive</td>
<td>15</td>
<td>15</td>
<td>65</td>
<td>60</td>
<td>66</td>
</tr>
<tr>
<td>Corresponding simple</td>
<td>38</td>
<td>25</td>
<td>45</td>
<td>50</td>
<td>45</td>
</tr>
</tbody>
</table>

*Each percentage is based on the responses of ten subjects in two items.

**Table 2**
The Percentages of Inferences Made from the Intransitive Sequences and
Simple Arguments in Experiment 1

<table>
<thead>
<tr>
<th>Argument Type</th>
<th>MP</th>
<th>DP</th>
<th>AC</th>
<th>MT</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affirmative transitive</td>
<td>25</td>
<td>30</td>
<td>40</td>
<td>35</td>
<td>33</td>
</tr>
<tr>
<td>Corresponding simple</td>
<td>68</td>
<td>50</td>
<td>60</td>
<td>70</td>
<td>66</td>
</tr>
<tr>
<td>Affirmative transitive</td>
<td>55</td>
<td>45</td>
<td>60</td>
<td>65</td>
<td>50</td>
</tr>
<tr>
<td>Corresponding simple</td>
<td>65</td>
<td>73</td>
<td>53</td>
<td>55</td>
<td>58</td>
</tr>
</tbody>
</table>

*Each percentage is based on the responses of ten subjects in two items.
valid and invalid, from transitive sequences (66%) than from corresponding simple arguments (59%), although this difference reaches statistical significance only at the 0.05 level, by planned comparisons (t(16) = 1.62, p = .10). (2) Subjects made fewer inferences, both valid and invalid, from the negative-consequent, transitive sequences (35%) than from their corresponding simple arguments (59%), and this difference is reliable, t(16) = -2.26, p < .05. Although subjects made fewer inferences from the negative-antecedent, transitive sequences (30%) than from the corresponding simple arguments (59%), this difference does not approach significance, t(16) = 0.5, p > .25.

There were also differences between the two sequences. Subjects made more inferences from transitive sequences (66%) than from negative-consequent, transitive sequences (35%), t(16) = 2.26, p < .05. They made more inferences from transitive sequences (66%) than from negative-antecedent, transitive sequences (30%), although this difference is not reliable, t(16) = 1.1, p > .20. They also made more inferences from transitive sequences (66%) than from the negative-antecedent, transitive sequences (30%), although the difference is not reliable, t(16) = 1.2, p > .20. Thus, subjects made more inferences from transitive sequences than from attitudinal sequences, presumably because of the relative ease of integrating the transitive information into a single representation. There were no reliable differences between the three types of simple arguments, a finding that is also consistent with previous studies; although subjects make different inferences depending on the location of negative in simple arguments with attitudinal material, this effect disappears when the simple arguments contain explicit location material (see Evans, 1971).

There were no reliable differences as a function of whether an argument was presented by an assertion affirming its truth or deniability, denying its truth, and so on. Inspection of the tables shows that there are no major trends in the data across the different truth values of the arguments; in contrast to previous findings, this departure might be because of the type of temporal-cause material used, which people frequently make all four inferences at a similar rate (e.g., Fischbein, 1975; Marcus and Rips, 1979). An unexplained factor in the results is that subjects made only 56% of the inferences from the attitudinal sequences, as compared to the 66% of the inferences from the transitive sequences. This is considerably lower than the usual 80% of the inferences from the transitive sequences. Subjects frequently made probabilistic responses to these arguments, which might be attributed, again, to the temporal-cause material used. Despite this anomaly, subjects made more inferences from the transitive sequences than from the simple arguments, not just in the case of valid inferences, but also in the case of the other three types of inferences, as the table shows.

When subjects did not make both inferences, they made either the first or the second, or both inferences, with some probabilistic qualification, such as "probable." For responses were of the inappropriate for simple arguments, an inference (e.g., "It is raining that she will get wet," to the conclusion "Therefore, she will get wet.") Consequently, their probabilistic responses are a mirror image of the attitudinal responses. As these appeared to be valid inferences in the frequency of first inferences above or of second inferences above, these data were not subjected to analysis.

The pattern of results from Experiment I raises two questions: (1) Why are more inferences made from transitive sequences than from corresponding attitudinal sequences? (2) Why are fewer inferences made from attitudinal sequences than from corresponding simple arguments? We can eliminate several of the potential explanations for their effects simultaneously. As subjects made more inferences that were both valid and invalid from transitive sequences, it is not simply a combination of two valid inferences or of two invalid inferences that creates the observed differences. Conversely, the differences between the transitive sequences and the corresponding simple arguments do not arise from the combination of a valid and invalid sequence with an invalid one; although we have found differences between positive-consequent, transitive sequences and corresponding simple arguments, the difference between negative-antecedent, transitive sequences and their corresponding simple arguments is not reliable. We can conclude that the differences between transitive sequences and negative-consequent, transitive sequences do not arise from the combination of a valid sequence with an invalid one, as against the combination for attitudinal sequences, of valid and invalid inferences, as the difference between transitive sequences and negative-attitudinal sequences is not reliable.

Therefore, based on formal inferences we do not help us to understand the disparity between the inferences made from certain sequences and their corresponding simple arguments, as reasoning from sequences should consist of an analysis of the reasoning scheme applied to simple arguments. Nor do theories based on domain-specific rules (e.g., Cheng & Holyoak, 1986; Cheng, Holyoak, Nisbett, & Oliver, 1986) provide any insight into these findings, because the pragmatic schemes that either of these, as Chinese inference schemes, or the (binary) inference schemes, or the (binary) inference schemes, seem unlikely to be forms of conditional sequences. A transitive sequence, should access one scheme—or a persistence one—and another form, a negativeconsequent a negative sequence, access a different kind of schema—one, a causal one. None of these possibilities suggests that there should be differences between sequences and simple arguments. Finally, explanations of reasoning behavior based on "bias", such as a bias in favor of negative conclusions (e.g., Evans, 1971, 1977; Pollick & Evans, 1980), cannot account
for there could be either, because the sequence and their corresponding simple arguments do not differ in the form of the inferences to be made.

Instead, it seems that the differences arise because the information in the conditional sequences is integrated into a single representation, or model (e.g., Johnson-Laird, 1983; 1984). We have suggested that subjects represent transitive sequences, such as:

If A happens then B will happen next.

If B leaves early then C will not catch it.

In an integrated way, and that their model affects the frequency of inferences that they make. When they know, for example, that the first event is true:

She arrived late,

eye can infer that the second event is also true:

She left work early,

and so they can infer that the third event is true:

She is out.

Subjects would not infer this final conclusion unless they had also established that both of the intermediate events were true; if subjects suspect that all the conditions for a consequence have not been satisfied, the logical premises and modus tollens inference are suppressed (Byrne, 1983). Hence, from an integrated representation, subjects explicitly make both inferences and not the frequency of inferences from transitive sequences is greater than from simple arguments.

In contrast, it seems likely that transitive sequences are represented in a relatively non-integrated way. We have suggested that subjects represent negative-consequent sequences in a disjointed model. They can make fewer inferences from the disjointed model and so they make fewer inferences from those transitive sequences than from transitive sequences, and also fewer than from corresponding simple arguments. The third kind of sequence is also represented in a non-integrated model, and so people may have fewer inferences from the negative-consequent, transitive sequences than from corresponding simple arguments. As the models of the two transitive sequences are not the same, perhaps it is unsurprising that we have observed differences between these transitive sequences.

Our next aim is to test this explanation. We will examine whether the differences between sequences and simple arguments arise because a single representation can be constructed from sequences.

EXPERIMENT 2

Our proposal is that different sequences are made from sequences and simple arguments because subjects represent the two conditions in a single model. We will test this explanation against a corresponding one in this experiment. The alternative explanation is that the differences arise because the sequences only half of the inferences are based on assertions that have been given to the subjects as premises. The other half are based on the intermediate conclusions reached from the subjects' own first inference. Subjects might find uncertain about using their self-derived conclusions. This uncertainty could be mitigated by the form of the intermediate conclusions, because subjects frequently find negative assertions harder to deal with than affirmative ones (e.g., Clark & Clark, 1981). Whereas the effects of uncertainty on reasoning have begun to receive attention (e.g., Rips, 1981), the use of prior conclusions has not, despite both being employed in everyday reasoning (Collins, 1979).

The form of the intermediate conclusions for the negative-consequent, transitive sequences, for which the most dramatic differences have been observed, are negative; in the case of the valid modus ponens and modus tollens inferences, and affirmative in the case of the denial of the antecedent and the affirmation of the consequent fallacies. For example, the modus ponens inferences from the premises:

If she goes to the family gathering then she will not enjoy herself.

If she enjoys herself then she will go to bad school.

She went to the family gathering.
yields the negative conclusion:

She did not enjoy herself.

Subjects might have difficulty in using this negative conclusion to make a simultaneous inference because of their uncertainty about its truth.

Consequently, the next experiment examines whether fewer inferences are made from negative-consequent, transitive sequences than from simple arguments because subjects can only draw the single model for the two conditions in a single model, or because they are uncertain about their self-derived conclusions. The experiment is designed to eliminate systematically the two conflicting features from sequences, and to test the two conclusions in parallel. First, we need to eliminate the possibility of representing the two premises in a connected way. But at the same time maintain whether uncertainty subjects might feel in using their intermediate conclusions. We can achieve this by examining subjects' performance with a modified version of the simple argument pairs, where they cannot represent the conditional in a single representation, but where they must use their own conclusion, derived from the first argument, to make an inference from the second condition. If
representation is at the heart of the observed differences, and uncertainty does not matter; thus subjects should make fewer inferences from the original sequence than from the uncertainty-maintained sequence. Thus, we can give subjects the first argument in a simple argument pair:

If she goes to the family gathering then she will not enjoy herself. She went to the family gathering from which they will infer:

She did not enjoy herself. We can then give them the second conditional:

If she enjoys herself then she will go to bed tired. and remind them that they had inferred:

She did not enjoy herself.

from the first argument, and ask them what follows from their conclusion and the current conditional. Thus, these subjects cannot represent the two conditionals in a combined representation, but it is their own intermediate conclusion on which their second inference is based.

Second, we need to eliminate any uncertainty that subjects might have in their intermediate conclusions, but maintain the possibility of representing the two conditionals in a single rule. We can examine subjects' performance on a modified version of the sequences, where they are given assurance that their intermediate conclusion is correct. If uncertainty is at the root of the difference and representation does not matter, then subjects should make fewer inferences from the original sequence than from the representation-maintained sequence. Hence, we can give subjects the sequence:

If she goes to the family gathering then she will not enjoy herself. She went to the family gathering, and when they make their first inference on the intermediate conclusion:

She did not enjoy herself.

we can inform them that this conclusion is correct. Thus they need feel no uncertainty about using their intermediate conclusion to make a second inference, yet they can combine the conditionals into a single representation.

Method

Design. We gave four groups of subjects the conditionals used in the negative-consequent, inerrant sequences of Experiment 1. We gave one group of subjects arguments that were designed to eliminate the possibility of representing the two conditionals in a single rule, but these subjects still had to use their own intermediate conclusions to make half of the inferences. These subjects received simple arguments, but the second simple argument in each pair was not accompanied by a premise asserting information about one of the clauses in the conditional. Instead, the subjects had to use their own conclusion, inferred from the first simple argument to make an inference from the second simple argument. We gave subjects in the second group arguments designed to eliminate the uncertainty of using intermediate conclusions, but these subjects could still assimilate the conditionals into a single representation. They were given sequences, but they were told, after they had made their first inference, that their conclusion was correct (regardless of its actual source). The third group of subjects received negative-consequent, inerrant sequences, and so presumably they could assimilate the conditionals into a single representation and they also had to use their intermediate conclusions to make a second inference from each sequence. We gave the final group corresponding simple arguments; presumably they could not represent the conditionals in a single representation, and they did not have to use their own conclusions to make any inferences.

We constructed four arguments by accompanying the sequences with assertions about the first antecedent or the last consequent, for six of the lexical materials used in Experiment 1 (see the Appendix). The 24 sequences and corresponding 24 simple argument pairs required 48 inferences. We gave each group of subjects arguments corresponding to the four inferential modal patterns and modal tokens, directed to the antecedent and antecedent or the consequent. The design allowed for the assessment of the influence of the direction of the inference, either forwards, moving from the antecedent to the consequent (the modus ponens and denial of the antecedent), or backwards, moving from the consequent to the antecedent (the affirmation of the consequent and modus tollens inferences). The second characteristic is whether the assertion affirms the antecedent or consequent (the modus ponens and affirmation of the consequent inferences) or denies it (the denial of the antecedent and modus tollens inferences). The arguments for each group were presented in a different random order to each subject of that group.

Materials. The arguments for the uncertainty-maintained group eliminated the possibility of assimilating the information into a single representation. They were similar to simple arguments, after subjects had made their first inference, they were given the conditional premise of the next simple argument, without an accompanying assertion. They were reminded of their previous conclusion and asked to say what followed from that conclusion, together with the current conditional. The argument for the uncertainty-
measured group were seances. After making an inference based on the
relevant premises, the subjects were informed that their conclusion was
correct, before they proceeded to use this intermediate conclusion to make
the second inference.

Procedure. The subjects were tested individually. They read and listened to
the items being read aloud. Their instructions were identical to those for
Experiment 1. The procedure involved two additional manipulations for
subjects in the uncertainty-maintained conditions: they were asked to imagine
the first stimulus argument of a pair, so that it could apply to the next
stimulus. Subjects in the repetition-maintained conditions were informed
that their intermediate conclusions were correct. The subjects’ verbal
responses were recorded on a cassette recorder.

Subjects. Thirty-two undergraduates from Trinity College, University
of Dublin, participated in the experiment. They were randomly assigned in
double-blind pairs to one of the four conditions (n = 16). Two subjects were subsequently elimin-
arized, prior to any data analysis, one was a non-native English speaker, and
the responses of the other had been handwriting destroyed. None of the
subjects had received training in logic.

Results and Discussion
The results indicated that the differences we observed between negative-
consequent, affirmative sequences, and affirmative and negative sequences in the first experi-
ment were not significant because the conditions of a sequence can be reanalyzed into a
single sequence. The scoring procedure used in the first experiment was
employed, again, and the responses are displayed in Table 3.

<table>
<thead>
<tr>
<th>The Percentages of Inference Stems in the Conditions of Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence Type</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Uncertainty-maintained condition</td>
</tr>
<tr>
<td>Negative-consequent, affirmative sequence</td>
</tr>
<tr>
<td>Representation-maintained condition</td>
</tr>
<tr>
<td>Corresponding simple argument</td>
</tr>
</tbody>
</table>

*Each percentage is based on the responses of eight subjects in the uncertainty-maintained and representation-maintained conditions or seven subjects in the extraneous and simple argument conditions, respectively.

Subjects made more inferences in the uncertainty-maintained condition (79%) than in the sequence (38%), and the difference is significant by
planned comparisons, (65) = 2.95, p < 0.01, on the interaction of the manipulations of representation and uncertainty, F(1, 24) = 5.16, p < 0.05,
found in an analysis of variance. There were no reliable differences in the
number of inferences that subjects made in the representation-maintained conditions.
Subjects made more inferences in the uncertainty-maintained condition (79%) and the sequence (38%), G = 1.36, p < 0.10. These findings suggest that it is the representation of the conditions in a single
model that leads to the difference between sequences and simple arguments.

The analysis also showed that fewer inferences were made in a forward direction (43%) than in a backward direction (57%), and this difference was
significant, F(1, 24) = 4.2, p < 0.05. This supports the suggestion that the directionality of an inference is influenced (e.g., Braine, 1978). More inferences
were made from affirming assertions (43%) than from denying assertions (46.5%), and this difference is also significant, F(1, 24) = 4.2, p < 0.05.

One surprising feature of the results is that subjects make very few inferences from simple arguments (40%), and consequently the difference
between these simple arguments and the sequences (38%) does not reach significance. This failure to replicate the differences observed in Experiment 1
will be given further attention in the next experiment. It may be a spurious failure to replicate, as inspection of each group revealed that an unknown-
certainly high proportion of recently modeled undergraduates had been
randomly assigned to the simple argument conditions; the uncertainty of their
responses may reflect extraneous factors.

EXPERIMENT 3

The experiment aimed to replicate the differences observed in Experiment 1
between the negative-consequent, affirmative sequences and their corres-
ponding simple arguments. A successful replication should ensure any
similarities the reliability of the differences.

Method

Materials and Design. There were two groups: one group the negative-consequent, affirmative sequences, and the other group corresponding to simple arguments. The materials from Experiment 1 for these two conditions were used.

Procedure. The procedure and instructions were similar to that of the
previous experiment. Subjects read the items aloud before giving their verbal
response. The experimenters recorded their responses, the time in writing.

Subjects. Sixteen subjects participated in this experiment. They were
members of the subject pool of the MRC Applied Psychology Unit,
Results and Discussion

The results constitute a successful replication of the differences that we observed in Experiment 1. Once again, the scoring of the responses was the same as in the first and second experiments. The responses are displayed in Table 4.

Subjects made more inferences from the negative-precedent, transitive sequence (279%) than from corresponding simple arguments (25%), and this difference was significant, χ²(1) = 8.27, p < .001. Consequently, this replication supports the reliability of the finding that subjects make fewer inferences from negative sequences, even negative sequences than from simple arguments.

From the results of the second experiment we can attribute the differences between the negative-precedent, transitive sequence and corresponding simple arguments to the construction of a single representation of the information in a sequence. The transitivity of the sequence ensures that subjects form a non-integrated representation of the two conditions, and from this disintegrated representation they make few inferences. But perhaps it is instead the presence of negatives in the conditions sequences that results in a representation from which subjects are unwilling to make inferences.

If subjects make few inferences from transitive sequences because they contain negatives, then they should make few inferences even from transitive sequences that contain negatives. In the next experiment we will turn our attention to transitive sequences. If subjects make few inferences from transitive sequences because they contain negatives, then they should make few inferences even from transitive sequences that contain negatives.

<table>
<thead>
<tr>
<th>Sequence Type</th>
<th>N</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative-precedent sequence</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Corresponding simple arguments</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

Such percentage is based on the responses of nine subjects for the sequence condition of seven subjects in the simple arguments condition in two tests.

Experiments 1 and 2 addressed the possibility that negatives do not matter as long as they do not affect the transitivity of the argument. We compared sequences where the first proposition was negated.

If she does not work hard then she will leave the library before dinner.

If she leaves the library before dinner then she will go to the battery.

We also compared sequences where both occurrences of the shared proposition are negated.

If she does not leave the library before dinner then she will go to the battery.

### Table 4

<table>
<thead>
<tr>
<th>Inference Type</th>
<th>NP</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP-NP</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>NP-NP</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

### Experiment 4

People make fewer inferences from negative-precedent, transitive sequences than from corresponding simple arguments; they make more inferences from affirmative, transitive sequences than from corresponding simple arguments. Can we say simply that people make more inferences from transitive sequences, and fewer inferences from affirmative sequences, regardless of the presence of negatives? The next experiment addresses the possibility that negatives do not matter as long as they do not affect the transitivity of the argument. We compared sequences where the first proposition was negated.

If she does not work hard then she will leave the library before dinner.

If she leaves the library before dinner then she will go to the battery.

We also compared sequences where both occurrences of the shared proposition are negated.

If she does not leave the library before dinner then she will go to the battery.

### Method

**Design and Materials:** There were four groups of subjects. We gave one group negative-precedent, transitive sequences and the second group...
A table with data on the percentage of both-sequences and single-sequences of the forms in Experiment 1:

<table>
<thead>
<tr>
<th>Sequence Type</th>
<th>NP</th>
<th>DP</th>
<th>AC</th>
<th>MT</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative-sequent mixed</td>
<td>59</td>
<td>20</td>
<td>11</td>
<td>54</td>
<td>89</td>
</tr>
<tr>
<td>Corresponding mixed-sequent</td>
<td>54</td>
<td>26</td>
<td>48</td>
<td>46</td>
<td>86</td>
</tr>
<tr>
<td>Negative-sequent neutral</td>
<td>52</td>
<td>36</td>
<td>41</td>
<td>53</td>
<td>80</td>
</tr>
<tr>
<td>Corresponding neutral-sequent</td>
<td>66</td>
<td>31</td>
<td>22</td>
<td>41</td>
<td>80</td>
</tr>
</tbody>
</table>

Each percentage is based on the number of subjects in the negative-sequent sequences and single-sequences in the negative-sequent sequence and single-sequent sequence in the form.

There was a modest, but unreliable, difference between the sequences and the single-sequent arguments, $t(25) = 1.24, p = .21$. (1) Subjects made more inferences from negative-sequent, transitive sequences (47%) than from corresponding simple arguments (45%), although this difference reached significance only at the .013 level, $t(25) = 1.27$. (2) Subjects made more inferences from negative-sequent, transitive sequences (45%) than from comparable simple arguments (47%), although this difference reached significance only at the .01 level, by planned comparisons, $t(25) = 1.27$. Despite the lack of statistical significance of these results, they do suggest, nonetheless, that subjects make more inferences from the sequences because of the presence of negation. It is clear that the presence of negation did not lead subjects to make fewer inferences from those sequences. Consequently, it is of no surprise that the presence of negation in a sequence leads subjects to make fewer inferences from that sequence than from those sequences. Instead, it is the transitivity of the transitive arguments that alters the inferences people make. Similar effects of transitivity have been found for three-term syllogisms and attributed to the relative ease of representing transitive sequences in an integrated sequence (Holtz, 1972). Another effect has been found for categorical syllogisms (e.g., Johnson-Laird & Wason, 1977). The "spatial effect" describes differences in the conclusions people make from premises with an adjacent middle term, such as:

All of the ants are beetles.
All of the beetles are chickens.

and premises with a distal middle term:

The results suggest that the transitivity of a sequence leads subjects to make more inferences, regardless of the presence of negation. Once again the scoring procedure was the same one used in the previous experiments. The responses were presented in Table 5.
GENERAL DISCUSSION

The four experiments that we have discussed show that people make different frequencies of inferences from sequences than from corresponding simple arguments. The pattern of these differences is that people make more inferences from transitive sequences than from corresponding simple arguments, whether the transitive sequences are affirmative, as they were in the first experiment, or whether they contain negatives, as they did in the fourth experiment. This is indicated by Table 5, which displays the overall frequency of inferences that the subjects made in the five sequences we have examined.

People make fewer inferences from affirmative sequences, particularly when the negated subject appears in the first condition. We would expect this difference to occur regardless of the location of the negated subject in the sequence, and so it is not clear why the size of the difference between the negative-plus-affirmative sequence and corresponding simple arguments was so small. Nonetheless, subjects made fewer inferences from the two affirmative sequences than from corresponding simple arguments.

We have proposed that the differences between the sequences and the simple arguments arise because subjects can form a single model of the two conditionals in sequence. We have also proposed that the differences between the transitive sequences and the affirmative sequences arise because the representation of transitive sequences is a more integrated one, whereas the representation of affirmative sequences is a non-integrated one. From an integrated representation, subjects make more inferences because the intermediate and final conclusions seem to be interdependent. In contrast, from a non-integrated representation subjects make fewer inferences because the inferences do not seem to be interdependent. The representation of a

sequence in an integrative or non-integrative fashion does not depend on the presence of negatives in the sequence: even negative transitive sequences are represented in an integrated way.

We have considered several other possible interpretations of these results. As we discussed earlier, the differences between the various kinds of sequences cannot be attributed to the mixture of valid inferences and inferences. We have also seen that the differences between the sequences and their corresponding simple arguments cannot be attributed to the form of the conclusions derived from the different sequences. Third, the differences do not arise because subjects are uncertain about using their own intermediate conclusions. The second experiment eliminated this possibility and established instead that the representation of sequences in a single model was responsible for the effect. Finally, the mere presence of negatives in the sequences is not responsible for the differences. Experiment 4 eliminated this possibility and indicated instead that it is the integrality or interactivity of a sequence that is responsible for differences between the sequences and their simple argument counterparts.

Related theories cannot easily accommodate these results. Theories based on formal rules make the wrong predictions for performance on sequences and simple arguments, because they specify that reasoning from sequences should consist in a single iteration of the rules applied to simple arguments. For a theory based on domain-specific rules, such as pragmatic schemata, to extend to sequences, the schemata must be either amalgamated or iterated. For such a theory to encompass conditions of different forms would require that different negative conditions are transformed and encoded into an affirmative schema, or a schema containing
negative conditionals is included in the repertoire. It is not clear whether the theory predicts that any of these possibilities will lead to more inference from stimulus and fewer from antecedent sequences. Finally, however, such as a basis in favour of making negation conditionals, cannot account for the facts of other, as the sequence and their corresponding simple arguments do not differ in the form of the inference to be made. Instead, it seems that these results are best accounted for by the theories that give the representation of information a crucial role in reasoning, such as the theory of reasoning based on causal models. Some support for the importance of the representation of information is provided by the fact that non-antecedent forms are especially by the second argument work.

In conclusion, reasoning from conditional sequences displays similar effects as reasoning that depends on specific items—for example, in categorical syllogisms—and reasoning about relations, such as three-valued series problems and spatial relations. But, it is not possible to generalize from reasoning with simple arguments to reasoning with sequences. As sequences are common in everyday reasoning, we need to gain an understanding of how people reason from them.

REFERENCES


APPENDIX

Examples of the Type of Materials Used in the Experiments

<table>
<thead>
<tr>
<th>Used in Experiment:</th>
<th>Conditionals</th>
</tr>
</thead>
</table>
| 1 and 3 only        | If she works hard then she will leave work early.  
                     | If she leaves work early then she will eat out.  
                     | If she has the morning off then she will go for a swim.  
                     | If she goes for a swim then she will have a late breakfast.  
                     | If she gets a copy of the textbook then she will read it today.  
                     | If she reads it today then she will do her essay at the weekend.  
                     | If she goes shopping then she will get home late.  
                     | If she gets home late then she will cook herself dinner.  |
| 1, 2 and 3 only     | If she drives to work then she will go by the main route.  
                     | If she goes by the main route then she will get to work on time.  
                     | If she meets her friend then she will go to see the play.  
                     | If she goes to see the play then she will get home late.  |
| 1, 2, 3, and 4      | If she goes to the gathering then she will drink a lot.  
                     | If she drinks a lot then she will go to bed early.  
                     | If she rings her friend then she will go to visit her.  
                     | If she goes to visit her then she will listen to some music.  |
| 2 and 4 only        | If she meets her friends then she will go down the country.  
                     | If she goes down the country then she will have a pleasant weekend.  
                     | If she works hard then she will leave the library before dinner.  
                     | If she leaves the library before dinner then she will go to the battery.  |