The Interaction between Interpretation and Reasoning

An Application of the Closed World Assumption to Children's Conditional Reasoning

Universiteit van Amsterdam
Department of Philosophy

Arjan Berkeljon

The Interaction between Interpretation and Reasoning

An Application of the Closed World Assumption to Children's Conditional Reasoning

by

Arjan Berkeljon
Studentnumber 9906509

A Master's thesis (*doctoraalscriptie*) submitted in partial fulfillment of the requirements of the degree

Master of Arts in Philosophy
(*Doctorandus in de Filosofie*)

Supervised by
prof. dr. M. van Lambalgen

Second reading by
prof. dr. F. Veltman

Universiteit van Amsterdam
Faculty of Humantities
Department of Philosophy

Zaandam, June 2006

Contents

1

Introduction

Man is the Reasoning Animal. Such is the claim.
I think it is open to dispute.

"The Lowest Animal", Letters From the Earth.
Mark Twain

For the interest conditionals have received since the beginnings of philosophy it is somewhat remarkable so much dispute should still arise over them. Their abundance in everyday language and their seemingly similar form make them a much discussed and debated topic.

This thesis aims to contribute to the on-going debate by discussing children's conditional reasoning in the common ground between philosophy and psychology. The analytical methods of philosophy are combined with the experimental methods of psychology in an attempt to develop a clearer picture of conditional reasoning. An outline of what follows is given below:

[1] Form and semantics of conditional sentences are introduced in a discussion of philosophical and psychological theories of the conditional and conditional reasoning. Prominent ideas in philosophy on the nature of the conditional are described and psychological research into conditional reasoning is reviewed.

[2] The philosophical background thus introduced is used to explain the necessity of an approach to conditional reasoning, suggested by Stenning and Van Lambalgen [2005], that distinguishes two kinds of logical reasoning: reasoning *to* and interpretation and reasoning *from* an interpretation. Reasoning *to* an interpretation is the nontrivial process of assigning a logical form to the task at hand on the basis of content and context. Reasoning *from* an interpretation is employing the assigned logical form to derive conclusions based on premises and assumptions.

[3] Drawing on Stenning and Van Lambalgen [2005] the implementation of this approach, logic programming with negation as failure, is introduced and applied to a paradigmatic case in the psychology of reasoning, an experiment by Byrne [1989]. It is shown that several reasoning patterns not conforming to classical logic can be accounted for by the proposed framework.

[4] To extend the proposed approach a conditional reasoning experiment performed with children is described and analyzed. The results are reviewed quantitatively and qualitatively. A statistical analysis is described to gain insight into the overall performance of children on the conditional reasoning task. An analysis of transcriptions of experimenter-subject dialogue is described to gain insight into the interpretative reasoning process underlying the performance on the task.

[5] Findings from developmental psychology are discussed to help explain the pattern of results and interpretative processes obtained in the conditional reasoning experiment.

[6] The concluding chapter of this thesis contains an overview of the reported findings and a discussion of possible future topics of research.

1.1 The conditional

A conditional, in its simplest form, is an 'if,... then' sentence. It has as its body at least one if-clause (also known as the antecedent) and a main clause (also know as the consequent). Of course this very broad definition characterizes many sentences that, although they appear to be very similar, are in fact very different. This is exactly what makes conditionals so fascinating, their apparent similarity and simplicity hides a great deal of diversity and complexity.

Two main types of conditionals are distinguished, *indicative* and *subjunctive* conditionals. Consider for example, the indicative conditional 'If we want to be there on time then we will have to leave by 10 am'. Contrast this with the subjunctive conditional 'If you hadn't dawdled then we would have been on time'. Both express a certain state of affairs which is conditional on another state of afairs. Typically, though not always, the if-clause of subjunctive conditional expresses something that is currently false, hence they are also called counterfactuals.

A conditional, indicative or subjunctive, expresses a dependence of some sort between one sentence and another. For example, '*If it's Tuesday this must be Belgium*' expresses a dependence between '*It's Tuesday*' and ' *This must be*

Belgium'. Exactly because it is Tuesday do we know we must be in Belgium. This trait of conditionals can be used in a number of ways. For example:

> **Informative:** If it is raining tomorrow then we will go to a museum instead of the beach
>
> **As a promise:** If you hurry up then I'll give you some candy
>
> **As a threat:** If you do that again then I'll hit you!
>
> **As a law:** If you are under 18 years of age then you cannot drink

Of course we can also use 'and' and 'or' to imply a dependence such as in '*Do that again and I'll hit you!*' or '*Don't do that again or I'll hit you!*' but one cannot use a conditional (honestly) and yet deny a dependence between antecedent and consequent. It would simply defy the purpose of using a conditional. Consider the rather awkward ' *If people are drunk they shouldn't drive but I'm drunk and yet I can still drive!*' This makes a weak point even if you are as a great a driver as you claim to be.

Many conditionals are used to express acts of hypothetical reasoning, i.e. reasoning from hypotheses or assumptions. They are either explicitly but most often implicitly part of some argument. These arguments, presented here in minimal form, contain at least two premises, a conditional premise of the form 'If A is the case then B is the case' and a categorical premise of the form (for example) 'A is the case'.

Because of the great diversity of conditionals (and conditional arguments) it is useful to employ some kind of classification. One way to do this is to abstract from the specific terms contained in the conditional and focus on its form instead. A formal way to do this is by using variables, say p and q, for the terms. We shall see later on that this is but a first step in properly formalizing a natural language conditional (or any sentence for that matter). In section 1.1.3 we will discuss what is involved in translating a natural language conditional (or sentence) into a formal language. What follows here will suffice for now.

The advantage of this approach is that we can represent different conditional arguments in a similar fashion and examine their differences and similarities. Consider a simple conditional 'If p then q' which we write as $p \rightarrow q$[1]. Adding a categorical premise, p or q we can construct four common conditional argument forms:[2]

[1] Let the arrow, $\rightarrow$, express the conditional relation between p and q.

[2] We use '/' to separate our premises from our conclusion; to the left of the '/' we write our premises and to the right our conclusions. We use $\neg$ to signify 'not' and like above $\rightarrow$ means 'if,... then'.

- Modus ponens (MP): $p \rightarrow q,\ p\ /\ q$
- Modus tollens (MT): $p \rightarrow q,\ \neg q\ /\ \neg p$
- Denial of the antecedent (DA): $p \rightarrow q,\ \neg p\ /\ \neg q$
- Affirmation of the consequent (AC): $p \rightarrow q,\ q\ /\ p$

Precisely two of these arguments are valid within classical logic, i.e. when its premises are true, its conclusion *must* also be true: Modus ponens and Modus tollens. The other two, Denial of the antecedent and Affirmation of the consequent are invalid within classical logic.

That the above is correct can be grasped quite easily for MP: '*If it rains then I will get wet; it rains, therefore I will get wet*'. MT may be a little harder but most people seem to get it right anyway. Say I tell you "*If I get a large tax return then I will buy a new bike*". Now if you see me a few months from now without a new shiny bike you can safely and rightly assume my tax return turned out rather low this year.

The fallacies are not always easy to notice but a sharp ear or eye and some practice can do much good. Consider the following argument: '*If this necklace is golden then it is very valuable; it is not golden, however, so then it is not very valuable.*' Although golden necklaces are valuable they are certainly not the only things of value in this world. The DA argument makes a fair assumption but its conclusion is too narrow.

Lastly, an even less obvious one. You may have encountered the following argument: ' *If God created the Earth we must see order everywhere*', '*We see order everywhere, therefore God must have created the Earth*'. This argument, of the form AC, is considered invalid because even if its premises are true the conclusion does not follow necessarily. Perhaps Genesis is correct, but then again perhaps our planet owes its existence to the Big Bang, gravity and plate tectonics; this argument certainly does not settle it. The difference between the AC and MP argument is in the latter case no reservations of the above form have to be made; the rain will make me wet and no amount of Biblical or scientific argumentation can change that.

Of course even this very informal approach to validity needs to be based on an understanding of what counts as a valid argument and how to determine an argument's form. In the next section I will describe several accounts dealing with the validity and form of conditional sentences that have been put forth in the last two millennia. Evidently doing justice to 2000 years of logic in a mere section is an impossible task. I hope instead to give a concise overview of some of the major issues and offer a suitable background for understanding what will be presented in the following sections of this paper.[3]

1.1.1 The philosophy of the conditional

The Stoic account of the conditional

The first systematic treatment of the conditional in philosophy can be traced back to the Stoics (a ancient Greek school of philosophy founded in 310 BC which centered around freeing oneself from the influences of the passions). Three distinct views of the conditional have survived from this period, those of Philo, Chrysippus and Diodorus.

Philo On Philo's account a conditional is a truth-functional[4] sentence which is true if it does not have a true antecedent and a false consequent. At first glance this is a very acceptable view since the conditional 'If it rains then I will get wet' is true if it is raining and I do indeed get wet but not if it is raining and I stay dry.

Stated otherwise, Philo's truth-functional interpretation states that any conditional with a false consequent and a true antecedent is false. Any conditional not false for this reason is true and so any conditional with a true consequent is true and any conditional with a false antecedent is true. This is not a view everyone necessarily agrees with. An objection to Philo's interpretation is that the truth or falsity of an entire conditional cannot depend solely on the truth or falsity of one of its compound clauses.

Chrysippus It is exactly this objection, which is central in Chrysippus' account of the conditional. In Chrysippus' view, as opposed to Philo, the truth or falsity of a conditional does not depend merely on the truth of the consequent or the falsity of the antecedent. For Chrysippus a connection between antecedent and consequent is relevant. He believed a conditional is true whenever the denial of its consequent is incompatible with its antecedent. In other words, for a conditional 'If p then q' to be true, it must necessarily be so that p and $\neg q$ not occur together.

To make clear the difference between Philo's and Chrysipuss' views consider the following two conditionals:

[1] If atomic elements of things do not exist, then atomic elements of things do exist.

[3]This overview derives largely from Sanford [1989].

[4]A truth-functional sentence is one which truth value is a function of the truth value of its compounds.

> If atomic elements of things do not exist, then nothing exists that is an atomic element of a thing.
>
> [Sanford, 1989, p. 24]

Let us assume that Philo and Chrysippus, like most ancient Greeks, believed there were atomic elements of things. The antecedent of both conditionals is thus false. On a Philionian interpretation this means that both *conditionals* are therefore true since neither has a true antecedent and a false consequent.

The Chrysippean account, however, says that the first conditional is false since the denial of its consequent is not incompatible with its antecedent (the denial is in fact the if-clause so it is perfectly compatible[5]). This is of course exactly what we would want. Clearly a conditional such as the first ought not be deemed valid since it is contradictory.

Diodorus Much like Chrysippus, Diodorus gives a modal[6] account of the conditional. In Diodorus' view this modality is connected with temporality. Diodorus defines a true conditional as one which neither is nor ever was capable of having a true antecedent and a false consequent. In other words a Diodorian conditional holds if the corresponding Philonian conditional holds *at all times.*

At first glance this may seem to advocate a view of conditionals as universal laws, constant over time. Such conditionals are perfectly reasonable of course but what about 'normal' conditionals? For this to work we need merely realize that many conditionals contain an implicit temporal component. A simple conditional such as *If it is raining then I will get wet* may be rephrased by making explicit its temporal component:

[2] If it is raining at time t, (and I am in the rain at time t) then I will get wet at time t.

In the above example we see an almost Humean idea of temporal proximity implying causality or at least a connection between antecedent and consequent.

These three accounts each emphasize a different aspect of conditional sentences. Philo's account is primarily concerned with the truth-functional relation between antecedent and consequent. Chrysippus and Diodorus provide more sophisticated accounts which consider more context-dependent aspects of con-

[5] Of course in intuitionistic logic (developed by Brouwer in the early 1900s), where Aristotle's *Law of the excluded middle* (P or not-P) does not hold, not-(not-P) is not the same as P so this example breaks down. I think it is safe to assume Philo and Chrysipuss were not intuitionists however.

[6] Modal here refers the so-called modal operators of necessity and possibility. Modal logic is the logic of necessity and possibility.

ditionals such as relevance and temporality. We will now turn to the treatment of conditionals in medieval philosophy where yet other aspects of conditionals were described.

Form and content: the medieval philosophers

Where in ancient philosophy the conditional was treated — albeit using different interpretations — as a fairly unambiguous thing, medieval philosophers recognized several different types of conditionals. I will describe several important distinctions that can be made by briefly reviewing the work of two medieval philosophers, Abelard and Ockham.

Abelard Peter Abelard (1079-1142) drew various important distinctions concerning the structure of conditionals. His most notable contribution is having recognized that some conditionals are true by virtue of their structure or form, the so-called perfect conditionals. For example, 'If no As are Bs, then no Bs are As'.

On the other hand, while perfect conditionals are always true, true conditionals need not be perfect. For example the conditional 'If Socrates is a human, then Socrates is an animal' is true, but not so by virtue of its form. Its form 'If x is A, then x is B' (where x is an individual and A and B are characteristics of this individual) does not have only valid instances. If x, A and B are taken to be variables we can give them values so that the conditional does not hold. We can, however, conclude if we add that 'All A are B' (e.g. in the above case we add 'All humans are animals'). Now the relation between antecedent and consequent is true no matter what values the variables H and A may take provided they obey the rule that all H are A.

Abelard's distinction shows the importance of background knowledge for dealing with certain conditional problems. It will not always been self-evident that a conditional holds. In fact solving conditionals problems assumes quite a lot of knowledge is already present. This treatment of conditionals focusing more on world knowledge in relation to conditionals predates various modern approaches which I will discuss later on.

Ockham Related distinctions between the form and content of a conditional were posited by William Ockham (ca. 1295-1349). Firstly he drew a distinction between conditionals that hold by extrinsic versus intrinsic means (relevant to the subject matter). A conditional that holds purely in virtue of its form does not hold by anything intrinsically connected to the subject mater. A conditional that holds by intrinsic means, by virtue of its content, can only be shown to hold by adding a second, necessarily true, statement. Much like an imperfect conditional can only be shown to hold by adding a requirement that all H are A.

A second distinction Ockham makes is between formal and material conditionals.[7] Ockham defines a material conditional as one that is true because it either has a necessary consequent or an impossible antecedent but is *not* true by intrinsic or extrinsic means. This is so because Ockham believed two maxims (which are not uncontroversial): 1) the necessary follows from anything; 2) anything follows from the impossible.

This distinction may seem rather contrived in modern eyes. More than likely this is because of Ockham's peculiar sense of necessary statements. To Ockham, for example, the statement 'God exists' is not a matter of empirical investigation, nor of faith; it is simply necessarily true. This allows him to make a distinction between conditionals that are true by virtue of their form (formal extrinsic conditionals) and conditionals that are true because of their content where this content is necessarily true (material conditionals). This in contrast to conditionals that are also true because of their content but this content can only be shown to be true by reference to an additional, necessarily true, statement (formal intrinsic conditionals).

Ockham reasoned that if an argument,*P, therefore Q*, is valid if is impossible that *P and not Q*. That is, it is valid if and only if it is necessary that *not both P and not Q*. This means that the necessity of Q or the impossibility of P alone is enough to establish the validity of the argument. If it is impossible that P then it is impossible that *P and not Q*, irregardless of Q. Likewise if it is necessary that Q then it is necessary that *not both P and not Q* regardless of P. Thus a material conditional is true for one of the above reasons, the impossibility of the antecedent or the necessity of the consequent. They are not true by intrinsic means because no additional truth is required to explain their truth. They are not true by extrinsic means because the impossibility of the antecedent or the necessity of the consequent is not a consequence of their form but rather of their content.

The third distinction central in Ockham's treatment of conditionals is of that between simple conditionals and those that are *ut nunc*, 'as of now'. When a simple conditional is true its antecedent can never be true while its consequent is false. When a conditional is *ut nunc* however it is possible that the antecedent is true while the main clause is false, but this is currently not the case. Thus an *ut nunc* conditional resembles a conditional that is true by intrinsic means however the additional statement needed to determine its truth is contingently true (happens to be true now) instead of necessarily true.

[7] Not to be confused with the contemporary material implication often used to describe the conditional relation denoted by $p \rightarrow q$.

An *ut nunc* interpretation of conditionals is in fact quite common. Consider the following situation:

[3] If I turn the ignition key then my car will start.

In deciding if this conditional is true or not we might realize that the car is indeed likely to start when I turn the key provided that the battery is in good working order. Dead batteries being the main cause of cars failing to start this is a reasonable reservation to make. On this interpretation, provided that the statement 'The battery is working properly' is true (contingently so), we can conclude the conditional is true. The experiment described later on in this paper uses this principle and describes how extra information affects children's conditional reasoning.

Applying the theory The distinctions between form and content presented above are not merely academic. Reasoning with conditionals involves deciding on their domain, i.e. is a conditional concerned merely with the specific terms it describes; the terms in a more general sense; or perhaps the terms are not important and only its form should be considered? This decision will affect subsequent interpretation of the conditional and what follows from it. Separating form and content (like Ockham does but also many logicians after him) is instructive because it can shed light on the contributions of both form and content to deciding the domain of the conditional at hand.

An excellent example of this is provided by Stenning and Van Lambalgen [2004] when describing the results of a variation of the Wason 4-card task [Wason, 1968]. In the original Wason task subjects are presented with four cards labeled A, K, 4 and 7. They are told that each card has a letter on one side and a number on the other. They are then asked to decide which cards they *must* turn to decide whether the following rule is true for these four cards:

[4] If there is a vowel on one side, then there is an even number on the other side.

Most subjects choose to turn A and 4 (i.e. check if the rule can be confirmed). Wason on the other maintained the only correct response is to turn A and 7 (i.e. check the rule against the A card and then try to falsify it by looking at the 7). Maintaining this commits Wason to a particular interpretation of this task — an interpretation according classical logic, more on this in section 1.1.3 — but who is to say this is the only correct interpretation of the task? There are different interpretations of this task that subjects could, or are likely, to take.

Rather than simply assume one interpretation it would be more instructive to judge a subject's answer according to the interpretation of the task *they* adopt.

These differences in interpretation, i.e. checking if the rule can be confirmed versus seeing if it can be falsified thus have profound effects on the final answer given in response to this problem. Because Wason recorded only responses and not the reasons underlying them little can be concluded on why there is a difference in interpretation. It is because of the contribution of Stenning and Van Lambalgen that we now know why this phenomenon occurs.

To understand the various interpretations subjects may use and their reasons for doing so Stenning and Van Lambalgen devised a variation on the standard Wason task, the two-rule task. They presented four cards labeled U, I, 8 and 3 with the added information that one side contains either U or I and the other side contains either 3 or 8. They asked the subject to read two rules, one of which was true and the other false, and asked them to decide which is which. The rules were:

[5]
1. If there is a U on one side, then there is an 8 on the other side
2. If there is an I on one side, then there is an 8 on the other side

The cards were real cards and subjects were first asked to select cards, then asked to imagine what could be on the other side, and finally to turn all the cards. Subjects were then given the opportunity to revise their previous selection. In their experiment the U and I cards had an 8 on the other side, the 8 card had an I, and the 3 card a U.

Reasoning according to classical logic, as Wason would have subjects do, would lead to selection of the 3 card. As in the original Wason task this is not what subjects do however. Stenning and Van Lambalgen provide many examples of interpretations subjects give to the task which, although perfectly reasonable, do not accord with classical logic. The crux of the matter is concerned with subjects notions of truth and falsity as compared to the meaning these notions have in classical logic. Not only do many subjects have an entirely different interpretation of what constitutes the truth or falsity of a rule (which will be discussed in section 1.1.4, but many subjects also choose a domain of interpretation for this task that stands in sharp contrast to what Wason would have expected.

In the original Wason task the intended domain of the supplied conditional consists of the four cards shown. Stenning and Van Lambalgen report subjects giving interpretations that are at odds with this. Some subjects appear to treat each card as a separate domain, i.e. they think each card should be checked against the rule independently of the others. The truth of the rule can thus be determined by one card, if the rule is true for *this* card, then the rule is true. As the authors suggest this is of course true for deontic conditionals, i.e. conditionals related to obligation. Consider the deontic conditional *If you are under 18, then you cannot not drink alcohol.* Surely finding one drunk 16-year old does not make this rule false, it simply puts the 16-year old in violation of the rule. If the card rule is interpreted as a deontic conditional rather than a descriptive one this would explain the particular answer in this case.

Another possible explanation for this interpretation is an *ut nunc* or as-of-now reading of the conditional (mentioned above). If the domain is restricted to one card at a time an *ut nunc* reading of the conditional leads to the conclusion that the conditional holds given that the card accords with the rule. The main difference between these two interpretations, deontic and *ut nunc*, lies in the treatment of exceptions. In the deontic case the conditional is assumed to hold and any exception to it does not falsify the conditional but rather is in violation of it. In the *ut nunc* case the conditional can either hold or not hold which in turn depends on the information supplied by the domain. An exception here means the conditional does not hold as of now (but may hold some at time in the future should new information become available).

Opposite to interpreting each card as its own separate domain are subjects who take the four cards to be a sample of a much larger domain. The term 'card' in the conditional is understood to point at any card from a large population of cards instead of to the four cards or even one card specifically (as was the case in the previous interpretation). This is a very reasonable assumption to make since most conditionals apply to an open-ended domain.

In the above examples we saw how an interaction of form (deontic or *ut nunc*, content and context (here the experimental instructions and background rules) influence the interpretation of the domain of the conditional and thus influence the interpretation of what follows from the conditional. In section 1.1.4 below a systematic and formal way to incorporate the important role of interpretation into conditional reasoning will be discussed.

Logical formalism

In the late 19th and early 20th century ideas on the conditional, and logic in general, took a more mathematical turn with the publication of Frege's *Begriffsschrift* in 1879 and Whitehead's and Russell's *Principia Mathematica* (three volumes) in 1910–1913.

In these works a symbolic (mathematical) and systematic approach is taken to logic in general and the conditional in particular. In addition to many valuable and detailed considerations on logic, inference, validity, etc. the formalism expressed in these works has done much to aid the systematic study of the conditional.

In his *Begriffsschrift* Frege reintroduces the conditional as a Philonian, truth-functional conditional. This was made more explicit in the *Principia* where Whitehead and Russell contended that 'a false proposition implies any proposition' and 'a true proposition is implied by any proposition'. Although undoubtedly useful in the development of formal logic the paradoxical nature of these statements leaves one with a feeling of confusion as to the nature of the conditionals described here.

A clearer account of the nature of the conditional and the connection between it and its compound statements can be given by means of a truth table. The truth table method was introduced by Wittgenstein in his *Tractaus Logico-philosophicus* first published in 1921.

Truth tables In such a table the truth value of a proposition (sentence) is tabulated along with the truth value all of its compound statements. In order to do this we first need a proper way to dissect sentences into their compound statements; we need a formal way of analyzing sentences. Consider again the example used at the beginning of this section. Here we took a sentence say 'If A is the case then B is the case' and rewrote it in formal terms like so: $p \to q$. We call this conditional a material implication and it is used in *classical* or *propositional* logic. Propositional logic is a simple logic in which we can analyze sentences from natural language in a formal manner. This language contains several formal symbols to indicate the relationships between the terms of the propositions in the language. Among these are $\wedge$ for 'and', $\vee$ for 'or', $\neg$ for 'not', $\to$ for 'if,...then' and $\leftrightarrow$ for 'if and only if'. Now that we have dissected are 'if,...then' sentence (which is called a material conditional or implication in propositional logic) we can tabulate it in a truth table:

p	q	$(p \to q)$
1	1	1[a]
1	0	0
0	1	1
0	0	1

[a] Let 1 stand for 'true' and 0 for 'false'.

This shows how all possible truth values of the compound statements combine to give a truth value for the entire conditional. The first and second lines will come as no surprise. The third and fourth lines may appear counter-intuitive, however. The trouble is that if the antecedent is untrue we are faced with an awkward problem, namely what truth value to assign to the conditional as a whole. Consider the sentence '*If John has just bumped his head, then he is now crying*'. Surely the conditional is true when John has just bumped his head and is now crying and it is false if he has just bumped his head but is not crying. What if he has not just bumped his head though? Certainly one would not wish to say that the conditional must always be false in this case but it also seems less than attractive to say that it must always be true. Since we wish to be able to assign a truth value to all elements in our truth table (that is the point of making one after all) we choose the least unattractive alternative and say that the material conditional is true if their antecedent is false.

Now that we know how the truth values of the compound statements of the conditional relate to the truth value of the conditional as a whole we can check for the validity of entire conditional arguments. Recall from the beginning of this section that a conditional argument contains at least two premises, one conditional and categorical premise. We defined an argument as valid if, if all its premises are true its conclusion *must* also be true.

I will demonstrate how this works for two conditional arguments, MP and AC.

MP '*If it rains then I will get wet; it rains, therefore I will get wet*'. We translate this formally into $p \rightarrow q$, p / q. Using a truth table we can easily show why this inference is valid:

		1st premise	2nd premise	Conclusion
		$p \rightarrow q$	p	q
p	q			
1	1	1	1	1
1	0	0	1	0
0	1	1	0	1
0	0	1	0	0

We see that if both premises are true (if $p \rightarrow q$ and p are 1) the conclusion is also true (q is also 1). This is precisely what we require so MP is valid.

AC *If God created the Earth we would see order everywhere; we see order everywhere, therefore God must have created the Earth*'. We translate this formally into $p \rightarrow q$, q / p. Using a truth table we can show why this is inference is *invalid*:

		1st premise	2nd premise	Conclusion
		$p \to q$	q	p
p	q			
1	1	1	1	1
1	0	0	0	1
0	1	1	1	0
0	0	1	0	0

Here we see that the premises of the argument are true on the first and third line. On the third line however, the conclusion is false. This violates our definition of validity which says that whenever an argument's premises are true, its conclusion must also be true. This then demonstrates why AC is invalid.[8]

The elegance of such a simple system is not difficult to appreciate but the reader may already have some intuitions why this propositional (or classical) account of conditionals cannot adequately deal with all conditional sentences. More precisely its syntax and semantics are not equipped to deal with the various forms of conditionals and the interpretations people give to them. In section 1.1.3 we will look at reasons why propositional logic is insufficient and an alternative will be suggested.

Recent accounts of the conditional

Recent logical accounts of the conditional focus more on the semantics (e.g. meaning) rather than on form as discussed in the previous section. I will discuss two of the most notable theories for indicative conditionals. *Possible world semantics* [Stalnaker, 1968] and *situation semantics* [Barwise, 1986].

Possible world semantics Consider the following conditional "If Iran produces nuclear weapons, Israel will declare war on Iran." Is this conditional true or false and how would one decide? Stalnaker [1968] suggests three possible answers.

The first is based on the Philonian truth-functional approach outlined in the section on Philo and in the previous section. To check the truth of the conditional two questions need to be answered: 1) 'Will Iran produce nuclear weapons?'; and 2) 'Will Israel declare war on Iran?' If the answer to the first question is no, or if the answer to the second is yes then the truth values of the compound statements of the conditional are such that it is true.

[8] For more information about propositional logic and the truth table method the reader can consult Gamut [1991] (in English) and van Benthem et al. [2003] (in Dutch).

On this interpretation of the conditional, however, the sentence "I firmly believe that Iran will not produce nuclear weapons *therefore* the above conditional is true." must follow, which is odd, to say the least. As we saw above an account in terms of a simple Philonian (truth-functional) conditional is insufficient.

The second answer then is based on the idea of a connection or relation implicit in the conditional. Although Stalnaker does not call it by name, this account is known as *relevance logic*.

In brief, relevance logic is a system of logic developed primarily to deal with the paradoxes of the implication (e.g., the material implication $p \rightarrow (q \rightarrow p)$ or the strict implication $p \rightarrow (q \vee \neg q)$). The problem relevance logicians allude to is that in such paradoxes the antecedent seems irrelevant to the consequent.[9] To solve this proposed problem the semantics of the implication is changed in such a way that the antecedent and consequent must share some relation for a conditional to be valid, i.e. they must be relevant to each other.

Applied to the Iran-Israel example above, we would have to establish whether a 'connection' exists between antecedent and consequent in this example. If the relation between the propositions expressed in this conditional holds, the conditional is said to hold.

Stalnaker's criticism of this account is not that he does not see the sense in a connection between antecedent and consequent. It is more that he does not believe such a connection is always relevant for the evaluation of the conditional. Using the above example one could wonder, analogously to Stalnaker, what would happen if one believed that Israel declaring war on Iran in the near future is inevitable given the relations between the two nations. If you think Iran's aspirations to produce nuclear weapons will not make much of a difference to Israel's intentions to declare war on Iran you may well affirm the original conditional while believing the antecedent and consequent to be causally independent.

Obviously none of this should be taken to mean that relevance logic is trivially dismissed. This is not the case. Stalnaker's dismissal of relevance logic in this case is merely instrumental in presenting his theory which emphasizes different aspects of conditionals. He claims that in determining if a conditional is true we must make use of the so-called Ramsey test.[10] Stalnaker's formulation of it is as follows:

[9]For example, take $p \rightarrow (q \vee \neg q)$. That q is true or false exhausts all possibilities in a two-valued logic such as classical logic. The antecedent p thus seems completely irrelevant to the consequent.

[10]So named after Frank Ramsey who described a process to do this in a footnote in one of his papers.

> This is how to evaluate a conditional: First, add the antecedent (hypothetically) to your stock of beliefs; second, make whatever adjustments are required to maintain consistency (without modifying the hypothetical belief in the antecedent); finally, consider whether or not the consequent is then true.
>
> [Stalnaker, 1968, p. 44]

The transition from belief conditions to truth conditions ought then to proceed by reference to a *possible world*. Stalnaker writes:

> Consider a possible world in which A is true, and which otherwise differs minimally from the actual world. *'If A, then B' is true (false) just in case B is true (false) in that possible world.*
>
> [Stalnaker, 1968, p. 45]

For obvious reasons this idea and its continuations have come to be known as *possible world semantics*. Stalnaker's 1968 article was mainly concerned with subjunctive (counterfactual) conditionals. In a 1975 article [Stalnaker, 1975] he extends his possible world semantics to include indicative conditionals as well.

His idea is this, a conditional *if A then B*, states that the consequent is true, not necessarily in this world, but in the world as it would be if the antecedent were true. Formally this is expressed by a function f, which takes the value of that possible world, most similar to the actual world, in which the antecedent is true: a conditional *if A then B* is true in a possible world i when B is true in possible world $f(A, i)$. [Stalnaker, 1975, p. 198]

To get this to work properly Stalnaker makes two assumptions:

> - Limit assumption: for every possible world i and non-empty proposition (one that is true in some possible world) A, there is at least one A-world minimally different from i.
> - Uniqueness assumption: for every possible world i and proposition A there is at most one A-world minimally different from i.
>
> [Stalnaker, 1981, p. 89]

Given these two assumptions and the selection function f a conditional can then said to be true in a world picked out by this selection function.

It appears as if the possible worlds approach is a rather coarse and general theory of conditionals. There is no accounting for the content of a conditional. Surely in most, if not all, cases what a conditional is about affects its evaluation. This content or rather its interpretation is often affected by the context in which the conditionals is embedded.

Situation semantics An approach more sensitive to the content and context of conditionals is given by Barwise [1986] employing the framework of *situation semantics*. According to situation semantics the meaning of a sentence can be found in relations between the different situations described in this sentence. A situation is an event or fact with a specified temporal and locational component (e.g. 'George's jacket is blue at the Hilton Hotel lobby in New York, March 20th 2005 at 4pm').

A situation can hold information because of certain constraints that exist between types of situations. Let S, S', ... denote types of situations and $s{:}S$ denote if a situation, s is of type S. A situation of type S is said to be realized if there is a real situation such that $s{:}S$. A constraint C is a relation between types of situations: $S \Rightarrow S'$. If this relation holds and S is realized then so is S' (in other words, if there is a situation of type S, then there is a situation of type S').

Using Barwise's example [Barwise, 1986, p. 35], consider the following constraint that if Claire is rubbing her eyes,then she is sleepy. This constraint is a relation between two types of situations, S and S', which holds at $\boldsymbol{l}$ (a space-time parameter):

[6] S = the type of situation where at $\boldsymbol{l}$, Claire is rubbing here eyes

which we write as:

[7] [s | in s: at $\boldsymbol{l}$: rubbing, Claire's eyes, Claire; 1]

and, using the same notation:

[8] S' = [s | in s: at $\boldsymbol{l}$: sleepy, Claire; 1]

So if S involves S', and if at some specific space-time location l it is the case that $s : S(l)$, then there is a situation $s' : S'(l)$. Thus, at that location l is in fact sleepy in s'. Therefore, the proposition that $S'(l)$ is realized entails the proposition that at l, Claire is sleepy.

An additional point is to be made regarding background conditions. Consider, in terms of the above example, that Claire has hay fever and rubs her eyes whenever a certain pollen, say pollen X is in the air. What is needed is a way to incorporate this background condition into the constraint. This is done by interpreting the constraint C (e.g., $S \to S'$) as a parametric constraint where a parameter $\boldsymbol{B}$ expresses the background conditions. In other words we have $C \mid \boldsymbol{B}$ instead of just C and $\boldsymbol{B}$ is anchored to the background conditions, in this case $B = [s \mid \text{in } s\text{: at } \boldsymbol{l}\text{: pollen } X; 0]$.

The above process of evaluating conditionals based on situations and their background conditions differs greatly from possible world semantics. In contrast to possible world semantics the problem of evaluating conditionals in situation semantics is not hypothesizing in which possible world, most similar to our own, the conditional holds but rather whether the background conditions in *this* world are such that the conditional holds. Strictly speaking in possible world semantics we consider a total model of the world, which is a complete yet minimally different model of the actual world where the conditional holds, and base our evaluation of the conditional on that. In situation semantics on the other hand we consider a partial model of the world, a model of the relevant constraints affecting the conditional, and base our evaluation on *that*.

The intuition behind both accounts is that reasoning with conditionals involves creating a model of the relationship between the clauses of the conditional. On the one hand we want this model to be terse enough so we do not consider *all* possible events relating to our conditional (e.g. while considering 'If I want to get to work on time then I will need to leave by 8am' I do not consider an outrageous event such as an asteroid hitting the Earth which may well make me really, really late). On the other hand we also want our model to be complete enough so that relevant events are not ignored.

Closed world reasoning The approach suggested in this paper is that people engage in *closed world reasoning* when reasoning with conditionals. In closed world reasoning one constructs a *minimal model* of the premises of a conditional argument which assumes only those facts contained in the premises; everything else is assumed to be false. The formal details of this approach will be discussed in section 1.1.3 but intuitively this idea can be grasped quite easily. Consider the following problem: I want to travel from Amsterdam central station to Schiphol airport and have to be there by 11:00. In order to get there by 11:00 I will have to take the 10:35 train from Amsterdam central. In the timetable I see no other train after 10:35 that will get me to Schiphol on time so I assume there is no such train. I have now constructed a minimal model of my original problem, namely how to get to Schiphol by 11am. This minimal model contains what needs to be true on the basis of the information I have: where I am traveling from and to; what time I will have to be there; and the information deduced from my timetable. In my minimal model I do not consider there may be trains not listed in the timetable, I assume the timetable is complete. This model is both terse and complete: it contains enough information for me to solve my problem and nothing more.

In section 1.1.3 closed world reasoning will be applied to conditional reasoning and it will be shown how it can account for various phenomena observed in psychological studies of conditional reasoning. The results of these psychological studies are the topic of the next section.

1.1.2 Psychological theories of conditional reasoning

Philosophers have not been the only ones to lay claim to the conditional. Psychologists too have investigated the conditional by studying people's performance on various types of conditional arguments.

The most replicated and robust studies are those dealing with the four most common types of conditional argument, Modus ponens (MP), Modus tollens (MT), Denial of the antecedent (DA) and Affirmation of the consequent (AC). These studies show that people do not find them equally easy. Evans et al. [1993] report frequencies of endorsement of these modes across several studies. MP is consistently high across all studies, MT however, is less frequently endorsed. Figures as high as 81% but also as low as 41% are reported. The rates for the classical fallacies DA and AC have a rather broad range, from below 20% to slightly above 70%.

For children the results are slightly different. The rate of endorsement for MP is lower for younger children as compared to adults (a rate 62% for 6-year olds is reported). DA and AC are endorsed with high frequency by younger children. The rates for both MP and the fallacies develop towards the adult rate as the age of the children increases.

The rates for MT are interesting because they do not conform to the above pattern. Some studies suggest the rate of MT actually increases for children around 10-12 years old and then declining towards the adult rate. This seems to support the idea that children around this age are better at MT than adults. The reason for this is unclear.

Several theories have been suggested to explain human reasoning in general and the above results in particular. Two of the most notable, which I will discuss here, are 'mental logic' and 'mental models'.[11]

[11] There are others such as a an heuristic approach in which people are supposed to use non-logical heuristics in reasoning.

Mental logic

The 'mental logic' (or rules) approach proposes that people use formal inference rules to solve reasoning problems [Rips, 1983, for example]. This is envisioned to be a process of abstraction where first the logical form of the problem is laid bare (is translated into formal logic). Next a formal proof is constructed which is finally translated back into the original context of the problem. The proof construction can be seen as generating a natural deduction proof tree in working memory by using several formal rules which govern where and how certain inferences can be made.

Natural deduction is a formal system that lets one generate a proof for an conclusion based on certain premises and assumptions. Generating a proof is done step by step by applying inference rules to a set of premises until the desired conclusion is reached. These inference rules are rules that govern what conclusions can be drawn on the basis of certain premises. For example the Modus ponens inference rule states that if we have a premise 'IF p THEN q' and we have a premise 'p' we may conclude 'q'. Another inference rule, 'and introduction' states that if we find p in our set of premises somewhere and we also find q in there (or perhaps we have been able to derive them) we may conclude that 'p AND q' holds. Another useful inference rule is *reductio ad absurdum*. A *reductio ad absurdum* argument is one where we assume a claim, derive a contradiction (an absurdity) and then conclude our original claim must have been false. Using these and other inference rules can lead us from premises to a conclusion in an argument (having studied this I would hardly call this process 'natural however').

In mental logic the letters p and q stand for propositions in working memory. Thus a letter here matches an event, such as 'John is in Paris' or 'Mary is Madrid', contained in working memory. The uppercase words, IF,... THEN and AND are believed to be logical constants which must match those words in a working memory proposition [Rips, 1983].

A simple inference such as 'If John is in Paris, then Mary is in Madrid, John is in Paris therefore Mary is in Madrid' should proceed as follows according to mental logic:

- Translation:
  - John is in Paris becomes p
  - Mary is in Madrid becomes q
  - If,... then matches the IF,... THEN modus ponens inference rule

- Construct proof:
  1. premise: IF p THEN q
  2. premise: p
  3. conclusion: q (Modus ponens on 1 and 2)

Thus we have proved that q, i.e. 'Mary is in Madrid'.

An MT inference is not as easy however. Consider the same first premise but now the second premise is 'Mary is not in Madrid':

- Translation:
  - John is in Paris becomes p
  - Mary is not in Madrid becomes $\neg q$
  - If,...then matches the IF,...THEN modus ponens inference rule
- Construct proof:
  1. assumption: p
  2. premise: IF p THEN q
  3. derivation: q (Modus ponens on 1 and 2)
  4. premise: $\neg q$
  5. derivation: q AND $\neg q$ (AND introduction on 3 and 4)
  6. conclusion: $\neg p$ (*reductio ad absurdum* on 5)

Thus we have proved that $\neg p$, i.e. 'John is not in Paris'.

Mental models

The 'mental models' approach on the other hand proposes that people do not use formal inference rules at all when reasoning. Instead they manipulate a model of the premises and 'read off' the conclusion that follows from this [Evans et al., 1993].

Reasoning is though to proceed in three stages: first subjects form a mental model of the world consistent with the information given by the premises; second subjects draw a putative conclusion that is more informative than the information supplied in the premise (i.e. an inference is made); third this putative conclusion is tested by trying to construct counterexamples in which the premises are true but the conclusion is false. If no counterexample is found the conclusion is inferred to be valid. For example for an MP inference this is thought to proceed as follows:

Consider the premises 'If the letter is an A then the number is a 3' and 'The letter is an A'. The model of the first premise would be:

The first line represents the A and 3 contained in the premises. The brackets indicate that As are represented exhaustively with respect to 3s, i.e. As cannot occur any other model unless 3s occur in that model too. Now given the second premise that the letter is an A, it follows that there must be a 3 since A is exhaustively represented with respect to 3.

A modus tollens problem (so if the second premise reads 'The number is not a 3') is more difficult because there ¬3 (not 3) is not present in our model. The solution in the mental models approach is to flesh out the model we have constructed so far, i.e. we will have to include additional models:

The premise 'The number is not a 3' does not accord with the models on first and third line so the second line is the only possible model. This warrants the MT inference 'The letter is not an A'

Debating rules and models

We saw that both the mental logic and the mental models aim to explain conditional reasoning within the framework of classical logic. The former by positing people use mental inference rules and the later by positing that people construct mental models. A theory of reasoning cannot not just be about what people do right, though, it must also account for what they do wrong; the fallacies (AC and DA). According to the rule-theorists fallacious inferences can be explained by comprehension errors and thus misapplication of a rule. Accordingly, people do not posses inference rules for the fallacies.

Rumain et al. [1983] showed that fallacious inferences can be suppressed by supplying a so-called *alternative premise*. For example:

[9] If she has an essay to write then she will study late in library.
If she has some textbooks to read the she will study late in the library.
She does not have an essay to write.

Given these premises people are not likely conclude 'She will not study late in the library' and therefore not commit the fallacy Denial of the Antecedent. A similar argument goes for AC. That the fallacies can be suppressed in this manner is taken as evidence by the rule-theorists that we indeed do not posses rules for these fallacies.

Byrne [1989] (from the mental models camp) put this argument on its head and claimed that if not only the fallacies but also the valid inferences can be suppressed this would show that we do not have rules for the valid inferences either.

By supplying an additional premise,[12]:

[10] If she has an essay to write then she will study late in library.
If the library remains open then she will study late in the library.
She has an essay to write.

Byrne found that the valid inferences, MP and MT, can suppressed. In other words, given the additional premise subjects are less likely to make an MP or MT inference (in the example above most people will not want to conclude that she has an essay to write thus suppressing a classical MP inference). Using the argumentation of the rule-theorists Byrne concluded that we do not posses inference rules for the valid inferences either. This was taken as evidence against the mental rules approach to conditional reasoning and as evidence for the mental models approach.

1.1.3 Theories of conditional reasoning revisited

Putting aside the debate between mental logic and mental models the experiments with additional and alternative premises present interesting results for any theory about conditional reasoning. How and why do these premises affect people's conclusions?

Byrne [1989] argues that it is the context that makes the interpretation different. Undoubtedly so, but *how* does this happen? Later on she writes:

> The moral of these experiments is that in order to explain how people reason, we need to explain how premises of the same apparent logical form can be interpreted in quite different ways. The process of interpretation has been relatively neglected in the inferential machinery proposed by current theories based on formal rules. It plays a more central part, however, in theories based on mental models. [Byrne, 1989, p. 79]

Here lies the crux of the matter. Stenning and Van Lambalgen [2005] remark that whereas Byrne takes logical form to be given, in fact logical form is assigned by the interpretative process involved in reading the premises. In reading the sentences above, for example, logical form is not 'read off' from them, as Byrne would claim, but rather it is assigned to them. Logical form is not, as is the naive view, simply translating the surface structure of a given sentence into some formal language, a lot more is involved. Stenning and Van Lambalgen provide the following:

[12]To clarify the terminology: additional premises make salient a possible obstacle in the way of achieving the consequent while alternative premises present another motive for wanting to achieve the consequent

[11] Let $\mathcal{N}$ be (a fragment of) natural language. A more complete list of what is involved in assigning logical form to expressions in $\mathcal{N}$ is given by:

1. $\mathcal{L}$ a formal language into which $\mathcal{N}$ is translated
2. the expression in $\mathcal{L}$ which translates an expression in $\mathcal{N}$
3. the semantics $\mathcal{S}$ for $\mathcal{L}$
4. the definition of validity of arguments $\psi_1, \ldots, \psi_n/\varphi$, with premises ψ_i and conclusion φ.

This brings us full circle. We started out by stating that it would be useful to formalize conditionals to study their differences and similarities. An initial attempt was made to formalize a simple conditionals $p \rightarrow q$. In discussing several accounts of conditionals we saw the conditional relation itself can be subject of debate (Chrysippus and Diodorus). We saw that form, content and domain of the conditional are important in deciding when it holds (Ockham and the results of Stenning and Van Lambalgen [2004]). We also saw that it is useful to construct a model of the relationship between the clauses of conditional (as in possible world semantics, situation semantics and closed world reasoning).

Considering these remarks and the four items on the list turns reading and interpreting a conditional into a process of setting parameters. These parameters then guide our reasoning with the conditional. Stenning and Van Lambalgen [2005] call the former process reasoning *to* an interpretation and the latter process reasoning *from* an interpretation. While reasoning *from* an interpretation is usually assumed to be what reasoning is all about, it appears reasoning *to* an interpretation is at least as important and plays a central role in any reasoning task.

1.1.4 Conditional reasoning: reasoning with exceptions

Having posited that reasoning with conditionals involves setting certain parameters such as choosing a formal language, an expression for the conditional in this language and the semantics for this language we need to establish what the values of these parameters are. With these parameters set we can then reinterpret the results obtained in experiments with simple, additional an alternative conditionals such as:

[12] If she has an essay to write then she will study late in library.[13]
(If the library remains open then she will study late in the library.)[14]

[13] Simple premise.
[14] Additional premise.

(If she has some textbooks to read then she will study late in the library.)[15]
She has an essay to write.

[For example see studies by Rumain et al., 1983; Byrne, 1989; Dieussaert et al., 2000]

Since there is a lot of data on these types of conditionals it is useful to see if the approach of parameter-setting can shed new light on the results obtained in these experiments.

Formalizing the conditional

In section 1.1.1, in discussing the Wason task, we saw that there is ample evidence that subjects do not reason according to classical logic in this task [Stenning and Van Lambalgen, 2004]. In section 1.1.1 we saw how their interpretation of the domain of the conditional affects their reasoning. It was also noted how nonclassical notions of truth and falsity influence the conclusions subjects infer.

Recall that in the two-rule task there were four cards, U, I, 3 and 8 and two rules of which one is said to be true:

[13] 1. If there is a U on one side, then there is an 8 on the other side
2. If there is an I on one side, then there is an 8 on the other side

One subject says the following:

[14] I wouldn't look at this one [3] because it wouldn't give me appropriate information about the rules; it would only tell me if those rules are wrong, and I am being asked which of those rules is the correct one. Does that make sense? [Stenning and Van Lambalgen, 2004, p. 498]

In classical logic 'not-false' is the same as 'true' but apparently some subjects disagree.

Similarly for 'false'. Stenning and Van Lambalgen [2004] report an effect they call *strong falsity* where subjects interpret a conditional $p \rightarrow q$ being false as $p \rightarrow$ not q! An example of this is given also:

[15] *E.* OK turn them.
S. [turns U, finds 8] So rule one is true.

[15] Alternative premise; additional and alternative premises are supplied with a simple premise, they are never presented together.

E. OK for completeness' sake let's turn the other cards as well.

S. OK so in this instance if I had turned that one [I] first then rule two would be true and rule one would be disproven. Either of these is different. [U or I]

E. What does that actually mean, because we said that only one of the rules could be true. Exactly one is true.

S. These cards are not consistent with these statements here.

[Stenning and Van Lambalgen, 2004, p. 501]

Together with the problems of domain interpretation mentioned in section 1.1.1 the findings Stenning and Van Lambalgen report suggest that, for at least some subjects, a nonclassical framework is better suited to capture their reasoning process than a classical one.

The original Wason task conditionals were intended to be true without exceptions. Stenning and Van Lambalgen make the interesting observation that this is not how conditionals are interpreted in real life. In fact they report that some of their subjects found it difficult to accept a notion of truth for the conditionals that knows no exceptions. For example a subject performing the original Wason task (with A, K, 4 and 7 cards):

[16] *S.* If I just looked at that one on its own [7/A] I would say that it didn't fit the rule, and that I'd have to turn that one [A] over, and if that was different [i.e. if there wasn't an even number] then I would say the rule didn't hold.

E. So say you looked at the 7 and you turned it over and you found an A, then?

S. I would have to turn the other cards over . . .well it could be just an exception to the rule so I would have to turn over the A.

[Stenning and Van Lambalgen, 2004, p. 502]

A logical form for a conditional 'If A then B' that includes the above considerations can now be given:

[17] If A, *and nothing abnormal is the case*, then B.

This form explicitly allows exceptions, although if none are the case it simply reduces to 'If A then B'. Of course we do not always know if something abnormal is the case or not; in real life we usually reason with incomplete information. Formally, reasoning on the basis of incomplete information is known as nonmonotonic reasoning [Brewka et al., 1997]. A time-tested example of how nonmonotonic reasoning works is the following. Consider these sentences:

[18] If Tweety is a bird then Tweety can fly.
Tweety is a bird.

Everyone should be quite happy to conclude that Tweety can indeed fly. Now what if we add an abnormality? What if additional information shows that Tweety is a penguin? Anyone with a even some modest zoological knowledge should now want to retract their previous conclusion.

This example thus shows that the semantics we will use to to embed our new formalization of the conditional in has to allow nonmonotonic inference relations in order to make the handling of abnormalities useful.[16]

Of course this is not to say that conditionals were never thought of as allowing exceptions. Consider Ockham's *ut nunc* conditional introduced in section 1.1.1. Recall that this is a conditional which is true 'as of now', but not necessarily true. In other words, the antecedent of the conditional can be true while its consequent is false, for example when a peculiar contingent and confounding fact is the case. Such ideas are not disputed here.

Nor is it disputed that researchers, such as Wason, could believe that people allow conditionals or rules to have exceptions. The point made here is that according to Wason and those who support him, interpreting the 4-card task as containing rules that allow exceptions, is 'obviously' incorrect. This is so because he believes, apparently, that the natural and obvious interpretation of the rules in this task is given by (some form of) classical logic with conditionals formalized as material implications. Since most subjects do not follow a classical interpretation, this task is seen as proof that (classical) logic is a poor model of human reasoning. For some this has discounted any possible contribution of logic and formal analysis to the psychology of reasoning (e.g. the mental models approach introduced in section 1.1.2 actively denies the involvement of any formal rules in human reasoning).

The next section provides, following the introduction given in the current section, an account of conditional reasoning thoroughly grounded in logic, that formalizes the inclusion of exceptions in such reasoning. Thus, apart from good task-related and context-related reasons for treating conditionals as allowing exceptions, there can also be sound logical reasons for doing so. Even, and perhaps especially, in tasks where traditionally (i.e. in psychological research concerning reasoning) this has not been considered the case. What follows

[16]This is yet another reason why a classical semantics will not do. Classical logic is *monotonic*. In a monotonic logic the following holds:

$$\text{If } A \vdash p \text{ then } A, B \vdash p$$

In other words, if it is known that p follows from A, p must also also follow from A *and* B.

should therefore be considered an argument *supporting* the merits of a logical and formal analysis in the study of human reasoning.

One final note regarding the status of different logics and different interpretations. Of course, as the authors acknowledge, the above remarks do not imply that subjects (perhaps aspiring logicians) who wish employ a classical interpretation of a conditional are now forbidden to do so. The approach presented here should be seen in two ways: (1) suggesting a new interpretation of conditionals which is not the only possible one; and (2) suggesting that the interpretation people give to a task itself suggest the model to be used to judge their performance. This leaves subjects free to choose an interpretation they deem best while acknowledging that by doing this they pin themselves to certain correct and incorrect answers. Obviously logic is not a mere pick-and-choose business, one cannot coherently maintain a classical, truth-functional, interpretation of the conditional while at the same time allowing exceptions. A useful analogy might be made using games: this approach does not tell you which game you *must* play, it does tell you that, once you have decided, what the rules for your chosen game are.

1.2 Logic programming

In the previous section we have committed to modeling conditionals in a non-classical semantics in such a way that it allows exceptions. We formalized the conditional as follows:

[19] If A, *and nothing abnormal is the case*, then B.

At the end of section (5) we briefly mentioned closed world reasoning as a means to reason with conditionals using minimal models. We now need to integrate the above form of the conditional in a process of constructing minimal models by closed world reasoning.

The framework Stenning and Van Lambalgen [2005] propose to use for this process is logic programming, a fragment of propositional (or predicate[17] logic with a nonclassical semantics. In its most general sense logic programming can be described as a means by which a logic program (a collection of clauses) is run through an execution procedure to allow deductions to be made from this program in order to derive a an answer to queries [for an introduction see van Benthem et al., 2003]. Below I will discuss the definitions Stenning and Van Lambalgen use in their implementation of logic programming. Firstly however

[17] Predicate logic, developed by Frege is an extension of propositional logic which includes a way to express predicates (characteristics of things and individuals such as red, tall etc.) and quantification.

I will characterize this implementation in an a rather informal manner because the definitions are rather technical. Hopefully an informal characterization will aid in the understanding of the proper definitions.

1.2.1 A characterization

Stenning and Van Lambalgen [2005] see reasoning as a process of considering logical models[18] of the premises given (this is the logic program). Because out of a set of premises quite a few models can be constructed a systematic way to choose which model to use is needed (the so-called preferred model). To do this the *closed world assumption* is used. This simply says that every nonprovable proposition is assumed to be false. For example, if there is no train scheduled to leave for Amsterdam at 7 pm in my timetable I assume there is no such train. As this example shows the use of this assumption is can be justified by considering how much we make use of it in daily life. Implementing the closed world assumption into the semantics of this approach is done by what is known as 'negation as failure'. Simply put, this means that if an attempt to derive a clause fails, its negation is held to be true. This ensures that our framework can not only deal with positive clauses but also negative ones.

To find out if a particular clause follows (nonmonotonically) from our program we compute the completion of our logic program. We have to do this because up until now our program contains information not written out explicitly. In order to derive what follows we will need to 'complete' our program. This process of completion gives us the 'closed world' we were after. To reuse a previous example of me wanting to go from Amsterdam to Schiphol by train. Initially my program contains the information that if I want to be at Schiphol airport by 11:00 I will have to take the 10:35 train. Completing this program consists of making this information explicit: if I want to be at Schiphol airport by 11:00 I will have to take the 10:35 train *and* if I take the 10:35 train I will be at Schiphol airport by 11:00. Thus I exclude irrelevant and unlikely events from model and thus from my reasoning. Looking ahead to the a more formal description of this process we can already see that the completion involves recasting a conditional relation into a biconditional relation (from 'if' to 'if and only if'.

The above process of completion is also called minimization and hence the model we find by computing the completion is also referred to as the minimal model. This minimal model will now allows us to state our conclusion based on what follows from it.

[18] To contrast this approach with the mental models field: models here are theoretical constructs, not things people actually hold in their heads. Later on in their article the authors show how nonmonotonic inferences can be derived in a neural network model.

1.2.2 Formal definitions

We will now proceed with a formal characterization of this process. The definitions used come from Stenning and Van Lambalgen [2005].

In standard logic programming, a (positive) logic program consists of a finite number of positive program clauses of the form:

[20] $p_1, \ldots, p_n \rightarrow q.$

This is read as $(p_1, \ldots, p_n)$ implies q. In this formula q is called the head and $p_1, \ldots, p_n$ the body.

Initially we define a model $\mathcal{M}$ of our logic program P to be a simple assignment of truth values to the proposition letters L contained in P. This considerations give us the following two definitions:

Definition 1 *A* (positive) clause *is a formula of the form* $p_1, \ldots p_n \rightarrow q$*, where the* q*,* p_i *are propositional variables; the antecedent may be empty.*

Definition 2 *Let* P *be a positive program on a finite set* L *of proposition letters. An assignment* $\mathcal{M}$ *of truth values* $\{0, 1\}$ *to* L *(i.e. a function* $\mathcal{M} : L \longrightarrow \{0.1\}$*) is a* model *of* P *if for* $q \in L$*,*

1. $\mathcal{M}(q) = 1$ *if there is a clause* $p_1, \ldots p_n \rightarrow q$ *in* P *such that for all* i*,* $\mathcal{M}(p_i) = 1$
2. $\mathcal{M}(q) = 0$ *if for all clauses* $p_1, \ldots p_n \rightarrow q$ *in* P *there is some* p_i *for which* $\mathcal{M}(p_i) = 0$.

These definitions allow only positive clauses. In order to also allow negative clauses we extend our semantics by what is known as 'negation as failure'. This simply means that $\neg\varphi$ is taken to be true if an attempt to derive φ from our program fails. This gives us:

Definition 3 *A* (definite) clause *is a formula of the form* $(\neg)p_1 \wedge \ldots \wedge (\neg)p_n \rightarrow q$*, where the* p_i *are either propositional variables,* $\top$ *or* $\bot$[19]*, and* q *is a propositional variable.* Facts *are clauses of the form* $\top \rightarrow q$*, which will usually be abbreviated to* q*. Empty antecedents are no longer allowed. A* definite logic program *is a finite conjunction of definite clauses.*

Finally the completion of our definite program will have to be computed. This is given by the following:

Definition 4 *(a) The* completion *of a program* P *is given by the following procedure:*

[19] Let $\top$ denote an arbitrary tautology, and $\bot$ an arbitrary contradiction

1. *take all clauses $\varphi_i \rightarrow q$ whose head is q and form the expression $\bigvee_i \varphi_i \rightarrow q$*

2. *replace the $\rightarrow$s by $\leftrightarrow$s (here, $\leftrightarrow$ has a* classical *interpretation given by: $\psi \leftrightarrow \varphi$ is true if ψ, φ have the same truth value, and false otherwise).*

3. *this gives the completion of P, which will be denoted by $comp(P)$.*

(b) If P is a logic program, define the nonmonotonic consequence relation $\approx$ by

$$P \models\!\!\approx \varphi \text{ iff } comp(P) \models \varphi.$$

If $P \models\!\!\approx \varphi$, we say that φ follows from P by negation as failure*, or by* closed world reasoning*. The process of completion is also referred to as* minimization.

Using the completion of our logic program it is now possible to infer an answer to a given query. An application of this is discussed in section 1.3.

1.2.3 Closed world reasoning for facts and rules

The above analysis is adequate for the so-called forward inferences, MP and DA. It does not work for the backward inferences, MT and AC however. This is because for the backward inferences our factual information consists of the negated or affirmed consequent of our conditional ($\neg q$ in the case of MT and q in the case of AC). Since the above process applies closed world reasoning to *facts* only our final (minimal) will inevitably lack information about p.

Consider for example the AC case in which we have $p \wedge \neg ab \rightarrow q$ and fact q. Closed world reasoning for facts gives the completion $\{((p \wedge \neg ab)\top) \leftrightarrow q; ab \leftrightarrow\}$, from which nothing can be concluded about p.

Stenning and Van Lambalgen [2005] suggest that the closed world approach be extended in this case by also allowing closed world reasoning for *rules*. In other words we assume that only those rules are true, which are known to be true (and just the facts).

This extension of the closed world assumption from atomic formulas to also allow program clauses is implemented in logic programming by so-called *integrity constraints*. On this view the consequent should be seen as a constraint our model must satisfy if the antecedent holds (comparable to situation semantics where when interpreting conditionals the relation between the antecedent situation and the consequent situation is also expressed as a constraint).

Closed world reasoning for rules proceeds differently than closed world reasoning for facts. If the problem is deciding whether a formula φ follows from a program, φ is used as a query denoted $?\varphi$. This query is said to succeed with respect to a program P if $comp(P) \models\!\!\approx \varphi$, i.e. if φ follows nonmonotonically from the completion of φ. Similarly the query $?\varphi$ fails with respect to a program P if $comp(p) \models \varphi$. A concrete application of this to MT and AC is discussed in section 1.3.

With these definitions established the closed world reasoning approach for facts and rules can now be applied to specific reasoning problems and an explanation can be suggested of why reasoners give the answers they do. In a way this approach bridges the gap between philosophical accounts of the conditional on the one hand and psychological accounts on the other. It has a solid logical base, which has been lacking in psychological theories of conditional reasoning so far, but does not ignore the rich sources of data provided by investigations of these psychological theories, which has been underutilized in philosophy up and until recently.

1.3 Applying logic programming to conditional reasoning

Applying this new framework to Byrne's [Byrne, 1989] results (which are a paradigmatic case of much of the research done in this area) gives the following hypothesis: a) the conditionals used by Byrne are best captured by a logic programming clause of the form $p \wedge \neg ab \rightarrow q$ where ab is a proposition letter signifying something abnormal is the case; and b) that when people make interpretations they do not consider all models of the premises but only the minimal models (given by the completion).

Using these formal definitions we are now able to predict specifically how subjects should solve MP, MT, DA and AC inferences given that they reason to the interpretation specified here, namely that conditionals are to be formalized as compound statements of the form $p \wedge \neg ab \rightarrow q$ and subsequently that closed world reasoning should be applied to them. Below I will state these predictions and show how a derivation in the present model is thought to proceed. I will first present an analysis (due to Stenning and Van Lambalgen [2005]) of forward inferences MP and DA and subsequently I will present the backward inferences MT and AC.

1.3.1 The forward inferences: MP and DA

MP problems with a single conditional premise

Suppose that for an MP inference we are given the following:

[21] If she has an essay to write then she will study late in the library.
She has an essay to write.

Using our rules for formalizing conditionals we represent the first premise as $p \wedge \neg ab \rightarrow q$ and the second as p. Remember that $\neg ab$ denotes that nothing abnormal is the case. Since there is no evidence to support a claim that something abnormal is the case we conclude all is normal, i.e. we conclude that the clause 'there is no abnormality' is false. Our problem can now be represented as the following logic program: $\{p; p \wedge \neg ab \rightarrow q; \bot \rightarrow ab\}$. The completion of this program is $\{p; p \wedge \neg ab \leftrightarrow q; \bot \leftrightarrow ab\}$ which in turn reduces to $\{p; p \leftrightarrow q\}$ from which q obviously follows.

Thus we have derived q from p in a manner that might seem superficially similar to a standard classical derivation, but is not. The crux of the problem is applying the definition of $p \wedge \neg ab \rightarrow q$ to the conditional instead of the simple $p \rightarrow q$ (this is what Stenning and Van Lambalgen call reasoning *to* an interpretation). In reasoning *from* an interpretation we then derive q, which is trivial in the case of MP.

DA problems with a single conditional premise

In the case of MP the present model predicts the same result as if a classical interpretation of the task was used, namely the derivation of q. For DA inferences the predictions diverge however. Consider:

[22] If she has an essay to write then she will study late in the library.
She does not have an essay to write.

To conclude that she will not go to the library, as many subjects do, is a fallacy under the classical view because it is possible for both premises to be true while the conclusion is false. If we assume subjects interpret this problem nonclassically and use the interpretation described above we see why (some) subjects may draw this particular conclusion.

Again we formalize the conditional as $p \wedge \neg ab \rightarrow q$ and the second premise as $\neg p$. Using closed world reasoning we take $\neg p$ to signify the absence of p in this model. As the example with the timetable given above we assume the negation of p implies its absence (negation as failure). As with the simple MP inference there is no information about a possible abnormality so we assume nothing is out of the ordinary. The program for this problem then becomes

$\{\neg p; p \wedge \neg ab \rightarrow q\}$. By the same reasoning as above the computed completion leaves us with $\{\neg p; p \leftrightarrow q\}$ from which $\neg q$ follows and so we have a DA inference.

Additional and alternative premises

This application can be extended to inferences with additional and alternative premises (recall that additional premises present a possible obstacle relevant to achieving the consequent while alternative premises present another motive or reason for achieving the consequent). The data for problems with additional and alternative premises have different effects on the four inferences types. In problems with an additional premise subjects make fewer MP and MT inferences compared to single premise problems while the number for DA and AC is comparable to that of single premise problems. Conversely, for problems with an alternative premise subjects make fewer DA and AC inferences compared to single premise problems while the number for MP and MT is comparable to that of single premise problems.

In problems with an additional or alternative premise subjects are presented with not one but two conditional premises. We saw that in the case of a single conditional premise, $p \wedge \neg ab \rightarrow q$ and absence of any further information minimization sets $ab \leftrightarrow \bot$ so that $p \wedge \neg ab \rightarrow q$ finally reduces to $p \leftrightarrow q$. If an additional premises is given, i.e. a premise that makes salient a possible obstacle relating to the first conditional, the situation changes however.

MP and DA problems with an additional premise The additional premise, formalized as $r \wedge \neg ab' \rightarrow q$, highlights a possible obstacle $\neg r \rightarrow ab$ for obtaining q from the first conditional; in other words, absence of r is an abnormality with the regards to the first conditional which is reflected in ab being set to true. Analogously the first conditional might now be taken to highlight a possible obstacle relating to the second conditional such that $\neg p \rightarrow ab'$ needs to be added to our program. Thus, our complete program for problems with an additional premise becomes

[23] $\{p; p \wedge \neg ab \rightarrow q; r \wedge \neg ab' \rightarrow q; \bot \rightarrow ab; \bot \rightarrow ab'; \neg r \rightarrow ab; \neg p \rightarrow ab'\}.$

The completion of this program is

[24] $\{p; (p \wedge \neg ab) \vee (r \wedge \neg ab') \leftrightarrow q; (\bot \vee \neg r) \leftrightarrow ab; (\bot \vee \neg p) \leftrightarrow ab'\},$

which reduces to $\{p; (p \wedge r) \leftrightarrow q\}$. For MP problems this leads to the interesting situation that since there is no information regarding r the truth of q cannot be established, i.e. a classical MP inference is suppressed.

In the case of DA, so with premise $\neg p$ instead of p as with MP we get $\neg(p \wedge r)$ from which $\neg q$ follows by negation as failure. DA then is not suppressed, as the data shows.

MP and DA problems with an alternative premise In the case of alternative premises (recall that alternative premises present another motive for achieving a goal stated in the first premise) we have two conditionals: $p \wedge \neg ab \rightarrow q$ and $r \wedge \neg ab' \rightarrow q$. Alternative premises do not highlight any abnormalities so for MP our logic program is the set

[25] $\{p; p \wedge \neg ab \rightarrow q; r \wedge \neg ab'; \bot \rightarrow ab; \bot \rightarrow ab'\}$,

of which the completion is

[26] $\{p; (p \wedge \neg ab) \vee (r \wedge \neg ab') \leftrightarrow q; \bot \leftrightarrow ab; \bot \leftrightarrow ab'\}$.

In the case of MP q is easily derived from this completion. As the data shows, MP inference are not suppressed by alternative premises.

For DA inferences we again take $\neg p$ to follow from the absence of p by negation as failure. The completion then is $\{(p \vee r) \leftrightarrow q; p \leftrightarrow \bot\}$ from which nothing follows. So indeed, as Byrne's [1989] original study showed, the DA fallacy can be suppressed by supplying an alternative premise. Subjects often do not draw this conclusion but instead some apply DA, i.e. they infer $\neg q$.

1.3.2 The backward inferences: MT and AC

MT problems with a single conditional premise

For a single premise MT problem consider:

[27] If she has an essay to write then she will study late in the library.
She does not study late in the library.

We formalize the first premise as $p \wedge \neg ab \rightarrow q$ but now, unlike with the integration strategy used for MP and DA, we interpret the second premise as a query $?q$ which must fail on our model. The completion of our first premise is $\{((p \wedge \neg ab) \vee \top \leftrightarrow q; ab \leftrightarrow \bot\}$ so the only way for query $?q$ to fail is if one of p and $\neg ab$ is false. We know by closed world reasoning that $\neg ab$ is true (there are no abnormalities) so we conclude p must be false.

AC problems with a single conditional premise

For a single premise AC problem suppose we are given:

[28] If she has an essay to write then she will study late in the library.
She does study late in the library.

As with a DA inferences, the obvious conclusion is not classically valid. Many subjects draw it anyway and infer that she does have an essay to write. The update strategy makes it possible to account for this.

Our model is the same as for MT but now we know our query $?q$ must succeed on this model. The truth of q can only be guaranteed if both p and $\neg ab$ are true. By closed world reasoning we know that this is the case for $\neg ab$ but not for p. We therefore posit its truth and we have AC.

Additional and alternative premises

As the integration strategy, the update strategy can also be extended to incorporate additional and alternative premises. In the case of an additional premise the program is made up by $p \wedge \neg ab \rightarrow q$, $r \wedge \neg ab' \rightarrow q$, $\neg p \rightarrow ab'$, and $\neg r \rightarrow ab$.

MT and AC problems with an additional premise For MT we have have the query $?\neg q$ which must succeed, or rather the query $?q$ must fail. The completion of P entailing $\neg q$ must either satisfy $\neg p$ or $\neg r$ but which one is unknown. This gives suppression of MT.

For AC one starts with the query $?q$. The completion of the program reduces to $\{(p \wedge r) \leftrightarrow q\}$ so a model that satisfies q must also satisfy p and r. If such a strategy were to be used by a particular subject this would explain an AC answer.

MT and AC problems with an alternative premise For a conditional and an alternative, $p \wedge \neg ab \rightarrow q$ and $r \wedge \neg ab' \rightarrow q$ the completion becomes $\{(p \vee r) \leftrightarrow q\}$ since nothing abnormal is the case the conditionals reduce to $p \rightarrow q$ and $r \rightarrow q$. For AC this gives that either p or r must be true but we do not know which. For MT the query $?q$ must fail so we know that both p and r must be false.

1.3.3 A comparison of semantics

Now that we have established the proposed model it is useful to look back and compare this to the approaches sketched in section 1.1.1, possible world semantics and situation semantics.

Possible world semantics

Recall from section 1.1.1 that possible world semantics [Stalnaker, 1968] is a description of conditionals in terms of a possible world, which differs minimally from the actual work, in which 'If A then B' is true (false) in case B is true (false) in this possible world.

An application of this principle to simple conditionals is essentially given by the above. The most complicated issue is deciding upon which possible world to choose; how do we decide which world, minimally different from the actual world, is needed to evaluate our conditional?

Presumably this decision should be based on careful considerations so that we select not just any world but a world in which precisely that much is true about the antecedent so that we can infer something meaningful about the consequent.

Consider the premises Byrne [1989] used:

[29] If she has an essay to write then she will study late in library.
If the library remains open then she will study late in the library.

To judge if the first conditional is true we consider a world different from our own only in that she has to write an essay; it then follows that she will go to the library. To judge if the second conditional is true we consider a world different only from our own in that the library remains open; we then conclude she goes to the library.

What the suppression task shows that at least some participants consider these conditionals, not in isolation, but as expressing a dependency. That is, the second conditional highlights an abnormality with respect to the first.

The problem for possible world semantics then lies in accounting for this abnormality handling. Obviously a world which is minimally different from the actual world is one where no abnormalities hold; it is different only in that the antecedent holds in it (and possibly matters need to hold in order to maintain consistency). The distinction of when an abnormality is relevant or when to disregard it in considering the minimally different possible world is not conveniently made within possible world semantics.

Situation semantics

In the last part of section 1.1.1 we saw that situation semantics [Barwise, 1986], contrary to possible world semantics, entails partial models of the world. In situation semantics meaning is given by relations between situations expressed by constraints.

A general conditional sentence is interpreted as a constraint between situations: if there exists a constraint C between a type of situation S and a type of situation S' and S is realized (there is a situation s of type S denoted $s : S$) then S' is also realized (denoted $s' : S'$).These constraints are parametric, a parameter anchored to the prevailing background conditions is part of our interpretation of the constraint.

Applying this to the conditionals we have been discussing so far we can write a conditional

[30] If she has an essay to write then she will study late in library.

as a constraint C which between two types of situations S and S' which are given by:

[31] $S = [s \mid$ in s: at $\boldsymbol{l}$: has an essay to write, she; 1]

[32] $S' = [s \mid$ in s: at $\boldsymbol{l}$: study late in the library, she; 1]

If we then take the additional conditional to express relevant background information with respect to the first conditional (as is done by a fair amount of people) we will read

[33] If the library remains open then she will study late in the library.

as signifying the possibility of being closed. So we introduce a background condition $\boldsymbol{B}$ expressing this:

[34] $B = [s \mid$ in s: at $\boldsymbol{l}$: library closed; 0]

and interpret our constraint C as $C \mid \boldsymbol{B}$, i.e. C holds given the specified background condition $\boldsymbol{B}$. Of course this is slightly contrived, we could just as easily posit

[35] $B = [s \mid$ in s: at $\boldsymbol{l}$: library open; 1]

but this merely begs the question what other relevant background conditions have to hold in order for the conditional to hold. Computationally it is far more efficient to consider certain abnormalities as the need arises than to consider all possible relevant background conditions which have to hold.

This is not to say specifying parametric constraints is radically different from specifying abnormalities, they are of course related. The difference between them is that a defining a background condition is a much less specific approach to parameterizing a conditional than specifying abnormalities. This is because the possibility of an abnormality occurring results from considering the relevant background conditions. To use the above example, the information that the library is open makes one consider the library's opening hours as a relevant background condition to embed the conditional in. An abnormality relating to the conditional and its relevant background is then the library being closed while its being open is not an abnormality but rather something which facilitates the first conditional.

1.3.4 Conclusion

With the above considerations in mind Stenning and Van Lambalgen [2005] reinterpret Byrne's [1989] work in a more appropriate logical framework. They show that with proper focus on the interpretation subjects give to a logical task a lot more can be said about their performance. Of course subjects need not reason according to the interpretation specified here. Subjects might instead employ a purely classical interpretation in addressing these inferences. The strength of this model then is, that it offers a different way to look at how and why subjects might make an inference. It offers a second competence model by which subject's performance can be judged. Subjects can get the problems right for different reasons and they can get them wrong for different reasons. This model provides several logical reasons to account for subjects' performance that go beyond the classical approach. As an extension of their work I will apply their model to children's conditional reasoning. More specifically I aim to:

[a] replicate the quantitative results obtained in the suppression studies by Byrne [1989] and most notably by Dieussaert et al. [2000] who replicated Byrne's experiment with more statistical power.[20] This is especially interesting since suppression has not been studied extensively in children (Evans et al. [1993] do cite conditional reasoning experiments done on children but these included only single premises and no additional or alternative premises).

[b] give a qualitative account of the reasoning processes behind these results. This will be done by applying the closed world reasoning framework described in sections 1.2 and 1.3 to the answers and reasons subjects give.

The experiment done to achieve these goals and the results thereof are discussed in the next two sections.

[20] The results to be replicated are only those reported in the first experiment described in their article. The second and third experiment defined answer sets people could use to respond (instead of yes, no or maybe). Unfortunately the required their subjects to choose only *one* answer from an answer set that includes answers that are always true (e.g. $p \wedge \neg p$ and answers that are dependent on each other (e.g. the set $\{p \vee q, p \wedge q, p, q\}$). Sadly this renders the statistics produced by these experiments uninterpretable.

2

Conditional Reasoning Experiment

2.1 Participants

A total of 40 children ages 8–12 ($M = 10.28$ years, $SD = 1.23$) participated in this study. They were recruited from two local primary schools in an urban area of the Netherlands with consent of their primary caregivers. One participant had to be excluded from analysis because she got almost none of the problems right, even after thorough prompting. This left 39 children for the final analysis.

2.2 Material

Subjects were presented with a total of twelve sets of conditional arguments consisting of (1) one or two initial premises, (2) a categorical premise and (3) a question as to what the conclusion might be.

For example:

[1] *(1) If Jeroen has an essay to write then he will go to the library.*
(2) Jeroen has an essay to write.
(3) Will Jeroen go to the library?

The material was presented in Dutch and is supplied as an appendix.

A within-subjects design was used in which every subject completed three conditions, each consisting of four conditional arguments. The three conditions differed in the number and nature of the initial premise(s) given as follows: 1) a simple conditional was given; 2) a conditional and an additional premise were given; and 3) and a conditional and an alternative premise were given. Each of these conditions was presented in all four modes MP, MT, DA and AC.

It should be noted that most studies done in this particular field have used a between-subjects design. A within-subjects design was used here instead for two reasons: 1) to maximize the amount of data that could be gathered given the small number of participants; 2) to study correlations between performance on the inferences across the three conditions.

Three different story lines were used to vary the content used across the three conditions. Which storyline corresponded to which condition was randomized across subjects. The order in which the subjects were presented with the three conditions (simple, additional or alternative) was also randomized across subjects. Finally, the order in which the four modes were presented was randomized across conditions.

2.3 Procedure

Testing was done in an empty classroom, one subject at a time. The material was presented on laptop and the experimenter-subject dialogue was recorded to the hard disk using a microphone.[1]

After the necessary demographical data were recorded (age, sex and year in school) the session was started. No time limit was imposed on each individual problem but nearly all sessions were completed in 10-15 minutes.

A semi-structured procedure was used to test the children on the problems. All children were given the same written and oral instruction to try and answer the question about the conclusion with yes, no or maybe. They were also asked to clarify their answer.

After reading this general instruction subjects were asked to read each of the premises out loud and then answer to the best of their ability. Following their initial answer some subjects were asked to clarify this answer or reread the sentences in case they made a mistake in the first reading. Subsequently subjects were prompted to consider other possibilities where appropriate (e.g. after giving the classically correct answer to an MP problem with an additional premise subjects were asked to consider the additional premise again. Successive prompting was then used to determine if subjects would make a nonmonotonic inference after all).

2.4 Data analysis

The recorded experimenter-subject dialogue was transcribed and analyzed in a two-stage process. Firstly, the answers of the children were categorized as either classically correct or classically incorrect. Statistical analysis of the changes in

[1] The source code of the program used in this experiment is available upon request

answers on the various inference types in the three different conditions allows for a comparison with earlier studies, most notably the one by Byrne [1989].

Secondly, the transcripts of the experimental sessions were also analyzed in a more qualitative manner. This was done by relating the subject's answers to the framework of logic programming by Stenning and Van Lambalgen [2005] described above.

3

Results

3.1 Percentages of endorsement

The percentages of endorsement of the various inference types across the three premise conditions are shown in table 3.1.

	Inference type			
	MP	MT	DA	AC
Single premise	94.9%	84.6%	71.8%	59.0%
With an additional premise	61.5%*	69.2%*	48.7%*	61.5%
With an alternative premise	92.3%†	89.7%†	41.0%*	30.8%*†

* There is a statistically significant difference ($\alpha = .05$) between the number of endorsements in this condition compared to the number of endorsements in the *single* premise condition.
† There is a statistically significant difference ($\alpha = .05$) between the number of endorsements in this condition and and the number of endorsements in the *additional* premise condition.

Table 3.1: Percentages of endorsement

As table 3.1 suggests there is a distinct effect of inference type (MP, MT, DA, and AC) and premise type (Simple, Additional, and Alternative) on the number of inferences subjects endorse. An analysis of variance for repeated measures reveals a significant main effect for inference type, $F(3, 114) = 18.481$, $p < .001$; subjects more frequently endorsed the classically valid inferences MP and MT than they did the classically invalid inferences DA and AC. A main effect for premise type was also found, $F(2, 76) = 6.090$, $p < .005$; subjects more frequently endorsed inferences in the simple premise condition than in either the additional or alternative premise condition. Finally, an interaction effect of inference type and premise type was found, $F(6, 228) = 5.220$, $p < 0.001$). This result is indicative of the suppression effect as reported by Byrne [1989], i.e. an additional premise should suppress MP and MT inferences while an alternative premise should suppress DA and AC inferences.

Planned comparisons (one-tailed) show that subjects indeed less frequently endorsed MP in the additional premise condition than in the simple premise condition (61.5% vs. 94.4%, $t(38) = 3.929$, $p < .001$). Subjects also made fewer MT inferences in the additional premise condition than in the simple premise condition (69.2% vs. 84.6%, $t(38) = 1.707$, $p < .05$). Subjects made fewer DA inferences in the alternative premise condition than in the simple premise condition (41.0% vs. 71.8%, $t(38) = 2.993$, $p < .01$). Subjects also made fewer AC inferences in the alternative premise condition than in the simple premise condition (30.8% vs. 59.0%, $t(38) = 2.320$, $p < .05$).

Subjects also made significantly more MP and MT inferences in the alternative premise condition than in the additional premise condition (92.3 vs. 61.5%, and 89.7% vs. 69.2%, $t(38) = 3.376$ and $t(38) = 2.454$, respectively, $p < .005$ for both comparisons). This is to be expected since an alternative premise should have no effect on the valid inferences while an additional premise is likely to suppress them.

AC inferences were made more often in the alternative premise condition than in the additional premise condition (30.8% vs. 61.5%, $t(38) = 2.933$, $p < .005$) but no significant difference was found for DA inferences between these conditions. An alternative premise should suppress AC and DA while an additional premise should leave them untouched. The absence of a significant difference between DA in the additional and alternative conditions may therefore seem unexpected. It is highly probable however, that this difference is solely attributable to the low number of endorsements in the additional condition (48.7%). The real unexpected result then, is the low number of DA endorsements in the additional premise condition compared to the number of DA inferences in the single premise condition (48.7% vs. 71.8%, $t(38) = 2.042$, $p < .05$). It is unclear why this is the case.

3.1.1 An extended study of conditional reasoning

In addition to the analysis of the results of the current study, results from a study by Schmittmann and Raijmakers were also analyzed. In this study three tests of cogntive ability were adminstered to 227 subjects ages 7–16. The tasks included a discrimination learning task, a working memory capacity task, a reaction time/inhibition task, and a conditional reasoning task resembling the one used in the current study.

Included here is an analysis, done by the present author, of the data gathered with the conditional reasoning task in Schmittmann and Raijmakers[1]. In this version of the task subjects were presented with 8 different conditional sentences of the type used in this study (examples can be found in the appendix): four sentences with a single premise, one sentence for each of MP, MT, DA, and AC; two sentences with an additional premise, one sentence for MP and one for MT; and two sentences with an alternative premise, one sentence for DA and one for AC. To solve the inference problems subjects had to choose from one of three possible answers: "Certainly yes", "Certainly no", and "Maybe". Three subjects had to be excluded from the analysis because they did not complete all items. This leaves 224 subjects, 109 males and 115 females, for analysis.

	Inference type			
	MP	MT	DA	AC
Single premise	78%	44%	58%	69.0%
With an additional premise	57%*	49%	NA	NA
With an alternative premise	NA	NA	25%*	44%*

* There is a statistically significant difference ($\alpha = .05$) between the number of endorsements in this condition compared to the number of endorsements in the *single* premise condition.
NA: data was not available for this inference type since it was not tested in this experiment.

Table 3.2: Percentages of endorsement in Schmittmann and Raijmakers

The three answer categories were recoded into two categories, "Inference made" and "Inference not made", in order to analyze and compare the results to the current study. This entails that subjects who answered "Maybe" in the original study were always classified as not having made the inference in question. This may introduce a somewhat conservative bias into the results of the analysis.

Table 3.2 shows the percentages of endorsement in the study by Schmittmann and Raijmakers. As with the current study, the suppression

[1] I would like to thank Verena Schmittmann for generously making available her data for this analysis.

effect is evident for MP, DA, and AC. Subjects make significantly fewer MP inferences in the additional premise condition than in the simple premise condition, 57% vs. 78%, $t(223) = 5.178$, $p < .001$. Subjects also made significantly fewer DA and AC inference in the alternative premise condition than in the simple premise condition (58% vs. 25% and 69% vs. 44%, $t(223) = 9.466$ and $t(223) = 7.096$, respectively, $p < .001$ for both comparisons).

For MT a suppression effect is not found. MT inferences were endorsed by rather low number of subjects. It is commonly found that MT is less frequently endorsed than MP (e.g. see the results of the current study in table 3.1) but the difference found in the study by Schmittmann and Raijmakers, 44% for MT vs. 78% for MP, is rather large. This deviance is not due to the recoding of the data but quite simply to the low number of subjects who give the classically correct answer to MT ($p \rightarrow q$, $\neg q$, hence $\neg p$). Only 44% of subjects give this answer. Only 17 % were unwilling to commit to either making or not making an MT inference which seems to exclude the possibility that the deviance can be attributed to the complexity of MT causing more answers in the "Maybe" category. It thus remains unclear why subjects less frequently endorsed MT than is commonly found, especially since the results for the other inferences are in line with reported empirical findings.

Because the number of subjects in the study by Schmittmann and Raijmakers was far greater than the number of subjects in the current study the influence of age on conditional reasoning could be tested. Subjects, ages 7–16 ($M = 11.68$ years, $SD = 2.70$) were divided into three age groups: (1) ages 7–9 (61 subjects); (2) ages 10-13 (83 subjects); and (3) ages 14–16 (80 subjects).

Age-related effects

An analysis of variance of inference type by age group reveals significant effects for the classical fallacies in the simple as well as the alternative premise condition (DA^{simple}: $F(2, 221) = 4.891$, $p < .01$; DA^{alternative}: $F(2, 221) = 19.859$, $p < .001$; AC^{simple}: $F(2, 221) = 10.708$, $p < .001$; and AC^{alternative} $F(2, 221) = 11.042$, $p < .001$). Post-hoc Tukey tests reveal the differences presented in table 3.3 (p. 48).

These data show that younger children more frequently endorse the classical fallacies than older children (table 3.3 a and b). Table 3.3b shows that, at least where the classical fallacies are concerned, the suppression effect (by an alternative premise) varies in magnitude according to age. As can be seen in table 3.3b, DA and AC inferences are suppressed in all age groups. However, the suppression effect is far greater for older children (ages 10 and up) than it is for younger children (under 10 years of age). Note that the age-related differences in percentages of endorsement are substantially larger in the alternative premise condition than in the simple premise condition. This shows that the alternative premise is indeed the relevant factor in this effect. Evaluating the differences between age groups by calculating effect sizes supports this claim.[2] The effect sizes for the age-related effects on simple premise DA and AC inferences reported in table 3.3a are .54, .37, and .40, respectively ($M = .44$). By contrast, the effect sizes for the age-related effects on alternative premise DA and AC inferences reported in table 3.3b are much larger, namely .80, .94, .43, and .82 respectively ($M = .75$).

An age-related difference was also found for single premise MT problems (MT^{simple}: $F(2, 221) = 7.270$, $p < .005$). It was noted before that endorsement of single premise MT inferences was rather low across all age groups. The difference reported in table 3.3 (c) provides a possible partial explanation. The rate of endorsement is especially low in the first age group (7–9 years) but significantly higher in the third age group (14–16 years). Still, even in this group, the rate of endorsement is not as high as, for example, in the current study. It is possible the method of experimentation has been a relevant influence here; in the study by Schmittmann and Raijmakers this was paper-/computer-based while in the current study tutorial sessions were used. The tutorial sessions probably provided subjects with more time and feedback which helped them solve the reasoning problems. This is especially relevant for MT since it MT problems are considered to be among the most difficult conditional reasoning problems.

The absence of an age-related effect for single premise MP problems is more than likely the result of the relative ease with which reasoners of all ages endorse MP inferences (75% of subjects in the first two age groups endorse single premise MP inferences; 84% of subjects in the third group do so). Such inferences are unproblematic and commonplace in all age groups and thus an age-related effect is not expected.

The absence of an age-related effect for MP and MT problems with an additional premise is caused by a rather constant percentage of endorsement of such problems across age groups (61% and 41% for MP and MT in the first group; 57% and 47% in the second; and 56% and 58% in the third). For DA and AC problems with an alternative premise the difference in percentage of endorsement between the first and third age group is substantial (see table 3.3b).

[2] The effect size (Cohen's d), or magnitude of the experimental manipulation, is calculated by dividing the difference between the means of experimental group and the control group by either of their standard deviations (if the variance is homogenous) or by their pooled standard deviation. Effect sizes are commonly defined as "small" if $d = .2$, "medium" if $d = .5$, and "large" if $d = .8$ [Cohen, 1988].

No such difference is apparent for MP and MT problems with an additional premise. An interpretation of this finding is given in section 3.3.1.

	Inference type	Age group (% endorsements)			Tukey
(a)					
	DA^{simple}	7–9 (72%)	—	14–16 (46%)	$p < .01$
	AC^{simple}	7–9 (87%)	—	14–16 (53%)	$p < .001$
		10–13 (72%)	—	14–16 (53%)	$p < .05$
(b)					
	$DA^{alternative}$	7–9 (52%)	—	10–13 (17%)	$p < .001$
		7–9 (52%)	—	14–16 (13%)	$p < .001$
	$AC^{alternative}$	7–9 (66%)	—	10–13 (45%)	$p < .05$
		7–9 (66%)	—	14–16 (28%)	$p < .001$
(c)					
	MT^{simple}	7–9 (28%)	—	14–16 (59%)	$p < .005$

Age group 1: 7–9 ($N = 61$)
Age group 2: 10-13 ($N = 83$)
Age group 3: 14–16 ($N = 80$)
Only statistically significant differences ($\alpha = .05$) are reported

Table 3.3: Percentages of endorsement: age related differences

3.1.2 Comparison of studies

So far, the results of three studies of conditional reasoning have been discussed, Byrne's [1989] suppression research, the study by Schmittmann and Raijmakers, and the current study. Not mentioned so far is a study by Dieussaert, Schaeken, Schroyen, and d'Ydewalle [2000] which sought to replicate and refine the results obtained by Byrne [1989].

Taken together these studies represent a solid body of research of conditional reasoning. Some consistent effects are observed, and some difference are notable. It is interesting to compare the results of these studies and to notice differences and similarities. To do this results of comparable experiments done by Dieussaert et al. [2000, experiment 1] and Byrne [1989, experiment 1], and the results from Schmittmann and Raijmakers and the current study are presented in figure 3.1 (p. 49; for the exact percentages see table 3.1 above and for the studies by Dieussaert et al. and Byrne see tables A.1 and A.2 in the appendix). These figures show the percentages of endorsement in each study for each of the four inference types across the conditions, single, additional and alternative.

Please note that it is notoriously difficult to quantitatively compare studies with very different sample sizes. In this study 39 subjects participated in the task. In the study by Dieussaert et al. there were 70 subjects. Byrne's study had only eight participants while the study by Schmittmann and Raijmakers had 224 subjects. Because the effect of one particular response is much greater if there are only eight subjects than if there are 39, 70 let alone 224 subjects, Byrne's results should be interpreted with care.

With regards to participants it is worth mentioning that college students participated in Byrne's study while Dieussaert et al. used children in their early teens as particpants. In the study by Schmittmann and Raijmakers and in the current study several age groups were tested, ranging from 7–16 in Schmittmann and Raijmakers, and 8–12 in the current study.

Figure 3.1: Comparison of results from four studies

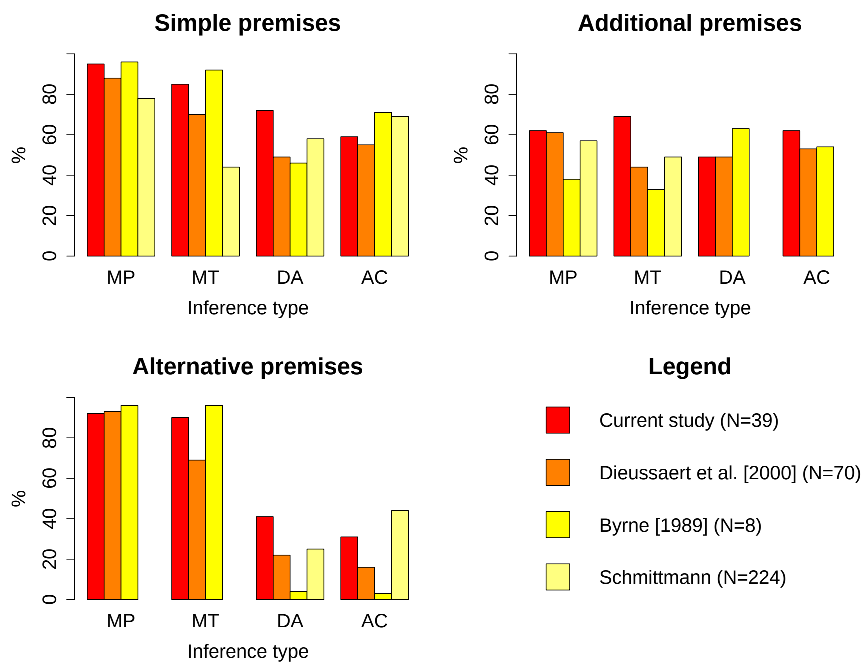

Figure 3.1 Percentages of endorsement of four argument patterns are shown across conditions with simple, additional and alternative premises.

The most striking result of this experiment the excellent performance of the children on the problems, i.e. compared with adult performance their results are not terribly different (with the exception of MT perhaps, which will be discussed in detail in section (3) below). Young children, of ages 8–12, as used in this study but also in the study by Schmittmann and Raijmakers, already appear to be sophisticated reasoners able to cope with abnormalities and reason with exceptions. Of course a still developing reading ability and lesser working memory capacity inevitably leads to a somewhat worse performance than that of adults on the same task. Issues relating to development will be discussed in chapter 4. Nevertheless their conditional reasoning ability is quite impressive and is certainly worth noting.

In the single premise condition (figure 3.1, top-left chart) the results are fairly similar although, as remarked earlier, the high number of DA endorsements in this study is high considering the results of the other studies. For a possible explanation of this see section (9). Also, as remarked above, the number of MT endorsements is low in the study by Schmittmann and Raijmakers

In in the second chart of figure 3.1 the suppression effect is clearly visible for MP and MT, they are much less endorsed compared to single premise condition. Notice the high number of MT inferences compared to the other two studies. It is sometimes observed that children ages 10–12 are quite good at MT [Evans et al., 1993]. The high number of inferences here may well be related to this observation. This is discussed in further detail in section (3) below.

In the alternative premise condition again clear suppression can be seen but this time for the classical fallacies, DA and AC. The number of endorsements in the current study as well as in the study by Schmittmann and Raijmakers is larger than in the other two studies, in other words, the suppression effect appears smaller. The most likely explanation for this is that the current study and the study by Schmittmann and Raijmakers both tested young participants. There is ample evidence that younger children endorse the fallacies more than older children and adults. Support for this claim can be found in table 3.3 in section 3.1.1 above, where it was observed that younger children more frequently endorse the classical fallacies than older children.

3.2 Two competence models

Explaining the frequencies in table 3.1 has traditionally, but unsatisfactory, been done by using classical logic and pointing to frequent errors subjects make. If subjects reason according to the rules of purely classical logic, they should endorse MP and MT and refuse to endorse DA and AC (for single conditional problems). As can been seen from table 3.1 this is clearly not the case. A better approach is that of closed world reasoning. Here interpretative processes have a much more prominent role to play in the explanation of why certain inferences are endorses while others are not. Both accounts are discussed in more detail below.

3.2.1 The classical model

The classical competence model would predict no change in these frequencies of endorsement for problems with additional or alternative premises. Since classical logic is monotonic, the extra information supplied by either the additional or alternative premise should not affect the inference made based on the first conditional premise.

As table 3.1 (for a single premise) shows, subjects perform well on the valid inferences MP and MT. Both are endorsed by a high percentage of subjects. Then there are the fallacies, DA and AC. Although not as frequently endorsed as MP and MT subjects do endorse these fallacies quite frequently. Interpretation of these results in terms of the classical competence model is thus in conflict.

On problems with additional and alternative premises there is even more confusion for the classical logician. On problems with an additional premise subjects suddenly perform much worse on the valid inference while on problems with an alternative premise they appear to endorse the fallacies less than in the other conditions.

In an attempt to take away some of this confusion (but undoubtedly add more in other areas) I will reinterpret the results above by applying the framework of nonmonotonic, closed world reasoning to the subject's reasoning (as apparent from the recorded experimenter - subject dialogue) in section 3.3.2 below.

3.2.2 Closed world reasoning

Recall from section 1.1.4 that in the closed world reasoning approach by Stenning and Van Lambalgen [2005] conditionals are of the form $p \wedge \neg ab \rightarrow q$ where ab denotes a possible abnormality.

In the analysis in section 3.3.2 below the formal derivations in logic programming as described in section 1.3 will given in the appropriate context of the particular inference. To contrast the predictions of the closed world approach with those of classical logic these predictions will be described here in brief.

Closed world reasoning for facts is applied to the forward inferences MP and DA. For single premise problems this means that both MP and DA can be endorsed as is observed.

For additional problems MP will be suppressed while DA is not. This is because of the abnormality the additional premise makes salient. This affects MP because a positive conclusion has to be derived while sufficient information is lacking. For DA this problem does not arise since in this case a negative conclusion is to be derived.

For problems with an alternative premise, i.e. a premise that states a possible alternative to achieve the consequent, no suppression takes place for MP while it does for DA. This is because in the case of MP there is enough information license the inference on the basis of the first conditional alone. For DA this is not the case. There is only negative information available for the first conditional (by closed world reasoning) and no information about the alternative conditional. Hence the reasoner is stuck: suppression.

To the backward inferences, MT and AC, closed world reasoning for rules is applied. Recall from section 1.2 that closed world reasoning for rules interprets the problems as integrity constraints that must be satisfied.

For single premise MT problems we know that there are no abnormalities (by closed world reasoning for facts) and hence the only reason why our consequent could be false is if our antecedent is false as well. Similarly for AC the absence of an abnormality lets us infer that the antecedent is true given that the consequent is.

In the case of additional premise again the saliency of an abnormality affects MT but not AC. Since we have two conditionals with different antecedents but the same consequent. We know from the premises that the consequent is false so we reason that at least one of the antecedents must be false, but we do not know which. Hence MT is suppressed. For AC the converse applies, again we have two conditionals with a different antecedent and the same consequent but now we know that this consequent is true. Therefore we conclude that both antecedents must be true and AC is not suppressed.

Alternative premises on the other hand lead to suppression of AC but not of MT. The constraint that the consequent is false here means that the antecedent of both conditionals (the first and the alternative one) must be false. For AC we are constrained by the fact that the consequent must be true. We do not know which of the antecedents must be true however. It could be the antecedent of the first conditional or the antecedent of the second (alternative premise).

3.3 Closed world reasoning and suppression

Although interesting, the aggregate data presented in section 3.1 does not provide much insight in the reasoning proces underlying the reported percentages of endorsement, and more specifically, the reported suppression effect. In order to gain insight into these matters, and thus assess the viability of the closed world reasoning approach, is is useful to examine the data obtained in the conditional reasoning experiment more closely. This is done in two ways:

(1) the pattern of suppression for each of the four inferences MP, MT, DA, and AC is examined per subject; and (2) key excerpts from the transcripts of experimenter-subject dialogue recorded during the conditional reasoning experiment are presented and discussed in terms of the closed world reasoning approach.

3.3.1 Patterns of suppression

Recall from section 3.1 that MP and MT inferences were suppressed by additional premises while DA and AC inferences were suppressed by alternative premises. Recall also that an additional premise makes salient a possible abnormality or exception regarding the consequent given in the first conditional while an alternative premise makes salient alternate cause for the occurrence of the consequent given in the first conditional.

This suggests two distinct closed world reasoning processes: (1) closed world reasoning applied to *abnormalities* or *exceptions*, which accounts for suppression of MP and MT in the case of an additional premise; and (2) closed world reasoning applied to *rules* (e.g. conditionals) which accounts for the suppression of DA and AC in the case of an alternative premise [se chapter 6 of Stenning and Van Lambalgen].

Stenning and Van Lambalgen report empirical data which suggests these two processes can be dissociated. It is reported that autists are quite good at the second reasoning proces, closed world reasoning for rules, but bad at the first, closed world reasoning for exceptions. This is supported by results from a conditional reasoning experiment, such as the one performed in the current study, in which suppression was observed for DA and AC inferences and an alternative premise, but much less so for MP and MT and an additional premise [see chapter 8 of Stenning and Van Lambalgen].

To examine the dual process and dissociation hypotheses, data gathered in the conditional reasoning experiment was classified according to suppressed inference (MP, MT, DA, or AC). Table 3.4 (p. 56) shows the outcome of this classification. The columns list, per subject ($N = 39$), which inferences were suppressed (denoted by a '1') and which were not (denoted by a '0'). Suppression was defined as the refusal to endorse an MP or MT inference when an additional premise is supplied, or the refusal to endorse a DA or AC inference when an alternative premise is supplied, provided that the respective inference *was* endorsed when only single premise was provided. This is a somewhat conservative measure because it excludes those subjects who, for one reason or another, do not endorse a single premise conditional and yet show suppression on conditionals with additional or alternative premises by applying closed world reasoning. Nonetheless the measure used is deemed adequate because it includes only those subjects who appear to exhibit the suppression effect proper.

Based on the dual proces hypothesis suggested above the data can be presented in two ways: (1) in table 3.4a rows are sorted by number of suppressed inferences starting with MP and MT, i.e. subjects are sorted according to how well they apply closed world reasoning to *exceptions*; (2) in table 3.4b rows are sorted by number of suppressed inferences starting with DA and AC, i.e. subjects are sorted according to how well they apply closed world reasoning to *rules*.

Two distinct groups of reasoners can be distinguished in table 3.4a: (1) subjects who apply closed world reasoning to exceptions (i.e. either MP, MT or both are suppressed); (2) subjects who do not apply closed world reasoning to exceptions (i.e. neither MP, nor MT is suppressed). In 18 (46%) subjects either MP, MT, or both is suppressed. In 21 (54%) subjects no suppression is found for these inferences.

The data in table 3.4b reveal two distinct groups of reasoners as well: (1) 25 (64%) subjects who apply closed world reasoning to rules (i.e. either DA, AC, or both are suppressed); and (2) 14 (36%) subjects who do not apply closed world reasoning to rules (i.e. neither DA, nor AC is suppressed).

Note that although there is overlap between these two groups, i.e. some subjects apply closed world reasoning to exceptions and to rules, these data also indicate the two processes can be dissociated. Of the subjects who apply closed world reasoning to exceptions (group 1 table 3.4a) three (17% of this group of 18) fail to apply closed world reasoning for rules (i.e. neither DA nor AC is suppressed). Of the subjects who apply closed world reasoning to rules (group 1 table 3.4b) 10 (40% of this group of 25) fail to apply closed world reasoning to exceptions (i.e. neither MP nor MT is suppressed). This suggests that either process can be present while the other is absent and is thus indicative of a double dissociation. This result is compatible with the reported result that in autists suppression of DA and AC is similar to that observed in normal subjects while suppression of MP and MT is virtually absent.

Concerning possible developmental factors relevant to the apparent dissociation between closed world reasoning for exceptions and for rules it is interesting to consider the results reported in section 3.1.1 once more. Recall from this section that there was an age-related increase of suppression of DA and AC by an alternative premise but no such effect for the suppression of MP and MT by an additional premise. Table 3.3b shows a steady decline of the percentage of endorsements (and thus an increase of suppression) of DA and AC inferences with an alternative premise across age groups, i.e. the tendency to suppress these inferences appears to increase with age. No corresponding increase of suppression associated with age is observed for MP and MT inferences with an additional premise. The absence of an age-related effect for closed world reasoning applied to exceptions (suppression of MP and MT) is plausibly explained by noting that reasoning with exceptions may be more difficult than closed world reasoning with rules. This is compatible with the finding that patients with autism, a developmental disorder, tend to show suppression on inferences that require closed world reasoning for rules but not on inferences that require closed world reasoning for exceptions.

The present study thus offers support for the dual process hypothesis of closed world reasoning applied to exceptions and to rules discussed in Stenning and Van Lambalgen. Support was found for two distinct groups of reasoners in the current data set; there are subjects who apply closed world reasoning only to exceptions and those who apply it only to rules. This suggests the two reasoning processes are dissociated. In qualifying these finding it was noted that they demonstrate a double dissociation of these two processes. That is, there are subjects who apply closed world reasoning to exceptions and fail to apply closed world reasoning to rules, and there are subjects who apply closed world reasoning to rules and fail to apply closed world reasoning to exceptions. This qualification was further extended by noting compatible results have been obtained with autistic subjects, who appear to apply closed world reasoning to rules but not to exceptions, and by studying age-related effects in reasoning, which show the tendency to apply closed world reasoning to rules increases with age although no such effect is found for closed world reasoning applied to exceptions. Implications of these findings relevant to a developmental discussion of children's conditional reasoning are presented in chapter 4. More information on the specific processes that guide children through reasoning with conditionals is presented in the next section, where transcripts of experimenter-subject dialogue are discussed.

3.3.2 Transcript analysis

This section describes an analysis of the experimenter-subject dialogue as recorded during each experimental session. Relevant parts of these transcripts were translated from Dutch into English for inclusion in this section. Also the data used in the statistical analysis is presented here in crosstabulated form to make clear the pattern of change for each inference type across the three conditions.

Great care has been taken to include, for each of the four inference types, a representative amount of transcribed dialogue to provide a meaningful frame-

a.						b.					
		Suppression of				Suppression of					
Nr.	**Age**	**MP**	**MT**	**DA**	**AC**	**DA**	**AC**	**MP**	**MT**	**Nr.**	**Age**
32	10	1	1	1	1	1	1	1	1	32	10
8	10	1	1	1	0	1	1	1	0	9	10
28	11	1	1	0	0	1	1	1	0	18	10
9	10	1	0	1	1	1	1	1	0	23	9
18	10	1	0	1	1	1	1	1	0	35	11
23	9	1	0	1	1	1	1	0	1	31	11
35	11	1	0	1	1	1	1	0	0	7	10
11	11	1	0	1	0	1	1	0	0	13	12
17	11	1	0	1	0	1	1	0	0	36	11
24	9	1	0	1	0	1	0	1	1	8	10
10	10	1	0	0	1	1	0	1	0	11	11
15	11	1	0	0	1	1	0	1	0	17	11
20	11	1	0	0	0	1	0	1	0	24	9
38	8	1	0	0	0	1	0	0	1	29	11
31	11	0	1	1	1	1	0	0	0	3	8
29	11	0	1	1	0	1	0	0	0	14	11
25	12	0	1	0	1	0	1	1	0	10	10
33	11	0	1	0	1	0	1	1	0	15	11
7	10	0	0	1	1	0	1	0	1	25	12
13	12	0	0	1	1	0	1	0	1	33	11
36	11	0	0	1	1	0	1	0	0	1	8
3	8	0	0	1	0	0	1	0	0	5	8
14	11	0	0	1	0	0	1	0	0	12	10
1	8	0	0	0	1	0	1	0	0	26	11
5	8	0	0	0	1	0	1	0	0	37	9
12	10	0	0	0	1	0	0	1	1	28	11
26	11	0	0	0	1	0	0	1	0	20	11
37	9	0	0	0	1	0	0	1	0	38	8
2	9	0	0	0	0	0	0	0	0	2	9
4	9	0	0	0	0	0	0	0	0	4	9
6	11	0	0	0	0	0	0	0	0	6	11
16	12	0	0	0	0	0	0	0	0	16	12
19	11	0	0	0	0	0	0	0	0	19	11
21	9	0	0	0	0	0	0	0	0	21	9
22	10	0	0	0	0	0	0	0	0	22	10
27	12	0	0	0	0	0	0	0	0	27	12
30	12	0	0	0	0	0	0	0	0	30	12
34	12	0	0	0	0	0	0	0	0	34	12
39	9	0	0	0	0	0	0	0	0	39	9

Nr. denotes the number assigned to each individual subject; subjects are sorted first by number of suppressed inferences (starting with MP and MT in part a. and with DA and AC in part b.) and then by subject number. Age denotes subject age.
Suppression is defined as the refusal to endorse an MP or MT inference when an additional premise is supplied, or the refusal to endorse a DA or AC inference when an alternative premise is supplied, provided that the respective inference *was* endorsed when only single premise was provided. Suppression of any one of the inferences, MP, MT, DA, and AC, is indicated by '1'; a '0' denotes no suppression.

Table 3.4: Suppression of inferences per subjects

work in which to understand the results described in section 3.1. A selection was made based on the idea that those pieces of text that give a good indication of the common response to a problem and those that reflect a more rare interpretation needed to be included in this analysis.

Modus ponens

Single and alternative premises As can been seen in table 3.5 (p. 59) children are quite happy to make an MP inference when supplied with this type of problem; 37 of 39 subjects make this inference.

In the single premise and alternative premise conditions (i.e. when no abnormality is highlighted) almost all children draw this inference readily and provide the (classically) correct reason for doing so. For example:

[1] *Example* Subject 52, a 9-year old girl

If Jeroen has to write a paper then he will go to the library.
Jeroen has to write a paper.
Will Jeroen go to library?

S: "Yes because he has to write a paper."

Additional premise Fourteen subjects make an MP inference in the single premise condition but not in the additional premise condition. Since Byrne Byrne [1989] this effect has been know as the 'suppression effect'. What is happening here is that subjects when confronted with a problem like $p \wedge \neg ab \rightarrow q$, $r \wedge \neg ab' \rightarrow q$ and fact p do not infer that q is the case because information about r is lacking. As discussed above minimizing both conditionals reduces the problem to $p \wedge r \leftrightarrow q$ and since only the premise p is given there is no information about r, hence no conclusion can be drawn. In the previous section this was referred to as applying closed world reasoning to exceptions.

Of this group of 14 subjects 12 answer in line with this interpretation, i.e. they will affirm q but cite r as a necessary condition.

[2] *Example* Subject 36, a 10-year old girl

If Erik has to go shopping for his mother then he will go to the store.
If the store is open then Erik will go to the store.
Erik has to go shopping for his mother.
Will Erik go to the store?

> *S:* "He will go to the store and then the second sentence says [reads second premise]. So he will go to the store if it is open. If it is not open he will not go."

The two remaining subjects think that q holds because r holds. For example:

[3] *Example* Subject 66, an 8-year old boy

> If Erik has to go shopping for his mother then he will go to the store.
> If the store is open then Erik will go to the store.
> Erik has to go shopping for his mother.
> Will Erik go to the store?
>
> *S:* "Yes because if the store is open Erik goes to the store. And you can also see this from the first sentence because [reads the first premise].

Although this subject does believe Erik will go to the store his answer suggests he does so because of a categorical (as opposed to conditional) reading of the conditionals. Thus rather than licensing the inference q providing that p holds, as in the previous cases, here premise r *alone* appears to be enough to infer q. A similar answer was given by the other subject. This particular interpretation of the problem was also observed for certain DA inferences and I will return to it when discussing those types of inferences below.

The groups of subjects that make an MP inference in the alternative premise condition but not in the additional premise condition largely overlaps with the group that makes an MP inference in the single premise condition but not in the additional premise condition. Here the problem is $p \wedge \neg ab \rightarrow q$, $r \wedge \neg ab' \rightarrow q$ and fact p. Since no abnormalities are made salient ab and ab' are minimized to $\bot$. The completion then becomes $(p \vee r \leftrightarrow q) \wedge p \wedge (r \rightarrow \bot)$, from which q follows.

Only a small number of subjects did not make an MP inference in the single premise (two subjects) or alternative premise condition (three subjects). Their answers do not show any systematic pattern however.

Modus tollens

Single premise MT inferences are also frequently made by subjects in the single premise condition (see table 3.6 p. 59). Although not as frequently endorsed as MP, 84.6% (33 out of 39) is quite a high figure. Furthermore the suppression of MT inferences when additional premises are presented is less pronounced than it is for MP.

Table 3.5: MP inferences

	MP additional	
MP single	Inference not made	Inference made
Inference not made	1	1
Inference made	14	23

	MP alternative	
MP single	Inference not made	Inference made
Inference not made	0	2
Inference made	3	34

	MP alternative	
MP additional	Inference not made	Inference made
Inference not made	1	14
Inference made	2	22

Table 3.6: MT inferences

	MT additional	
MT single	Inference not made	Inference made
Inference not made	5	1
Inference made	7	26

	MT alternative	
MT single	Inference not made	Inference made
Inference not made	1	5
Inference made	3	30

	MT alternative	
MT additional	Inference not made	Inference made
Inference not made	2	10
Inference made	2	25

Of the 33 children who give the classical correct answer $\neg p$ to the single premise MT problems 18 explicitly mention $\neg q$ as the reason for their answer. For example:

[4] *Example* Subject 36, a 10-year old girl

If it is Sanne's birthday then she will go to the zoo.
Sanne does not go to the zoo.
Is it Sanne's birthday?

S: No.
E: Can you say some more about that; why do you think that?
S: No, er yes. It is not her birthday so she does not go to the zoo. It says here that only if it is her birthday she will go to the zoo but here it says that she does not go to the zoo so it is not her birthday.

Another 13 children gave the same answer but did not give an explicit reason. It was clear however that their reasoning was similar to the 17 children above. Using closed world reasoning for rules, described in the introduction and in Stenning and Van Lambalgen [2005], the completion of this particular problem becomes $\{p \wedge \neg ab \leftrightarrow q, ab \leftrightarrow \bot\}$. Starting with the query $?\neg q$ we can conclude $\neg p$.

Additional premise Of interest is the group of seven children who *do* make an MT inference in the single premise condition but do *not* make an MT inference in the additional premise condition. This is the group affected by the so-called suppression hypothesis; added information in the form of an additional premise should suppress MT (and MP as we saw above) inferences. In other words, closed world reasoning is applied to the additional premise highlighting a possible abnormality. Recall that for MT in the additional premise condition from the constraint that query $?q$ must fail. This means that at least one of p or r must fail but we do not know which. Subject 60 is a good example:

[5] *Example* Subject 60, a 10-year old girl

If it is Sanne's birthday then she will go to the zoo.
If the zoo is open then Sanne will go to the zoo.
Sanne does not go to the zoo.
Is it Sanne's birthday?

S: No. . . yes. . . well. . . maybe.

> *E:* Why maybe?
>
> *S:* It could be that the zoo is closed but it could also be that it is not her birthday.
>
> *E:* So you mean it could be that it is her birthday but that the zoo is closed so she does not go. Is that what you mean?
>
> *S:* Yes.

Apart from this girl there was only one other subject in this group of seven to explicitly use the additional premise in his answer. His approach is quite different though:

[6] *Example* Subject 56, an 11-year old boy

> If Jeroen has to write a paper then he will go to the library.
> If the library is open then Jeroen will go to the library.
> Jeroen does not go to the library
> Does Jeroen have to write a paper?
>
> *S:* Yes because on closed days he cannot go to the library.
>
> *E:* No.
>
> *S:* If it is open he can.
>
> *E:* So does he have to write a paper or not?
>
> *S:* Yes he does.
>
> *E:* He has to write a paper?
>
> *S:* Yes.
>
> *E:* But he does not go and that is because the library is closed?
>
> *S:* Yes.

So although this subject also refuses to make an MT inference his reasoning is not as complete as that of the previous subject. He simply realizes that the library being closed is enough reason for Jeroen to not have to go there. He does not infer from this that Jeroen does not have an essay to write. Closed world reasoning for rules indeed concludes that *at least one* of the antecedents must be false. Where subject 60 concluded that either could be false but could not say which this subject is convinced only the antecedent of additional premise is false.

The other subjects had a more difficult time solving MT additional problems. Only three of them did not make an MT inference because of the additional premise. Striking is the subject from a previous example (who did not affirm an MP inference because of an additional premise):

[7] *Example* Subject 36, a 10-year old girl

If Erik has to go shopping for his mother then he will go to the store.
If the store is open then Erik will go to the store.
Erik does not go to the store.
Does Erik have to go shopping for his mother?

S: Yes, because his mother tells him to go shopping.
E: But he does not go to the store...
S: No, because the store is not open or something.
E: So he does not go, but why not? Does he not go because the store is closed or because his mother has not told him to go? Or do you not know?
S: It says here "If Erik has to go shopping" so no.
E: So it is about this sentence [the first premise]?
S: [Reads the first premise]. It says 'if' but you do not know if he has to do it or not.
E: But it says here that he [Erik] does not go to the store. So then you think that she [Erik's mother] has not told him to go?
S: Yes.

After minor prompting she seems to think that the second premise is indeed relevant to her answer. Upon asking her for the specific reason why Erik does not go to the store she doubts again and after more questions arrives at a standard MT inference. Nevertheless, I categorize this subject as not wanting to make an MT inference in the additional premise condition because her initial answer suggests exactly this.

The other four children did not infer $\neg q$ because they read the additional premise as alternative information rather than additional. For example:

[8] *Example* Subject 57, an 11-year old boy

If it is Sanne's birthday then she will go to the zoo.
If the zoo is open then Sanne will go to the zoo.
Sanne does not go to the zoo.
Is it Sanne's birthday?

S: I think [the answer is] maybe.
E: Why maybe? Why not yes or no?
S: If it is her birthday then she will go to the zoo perhaps and if it is open then she will go to the zoo perhaps.

> *S:* But there is no 'maybe' in those sentences right? It says [reads premise 1 and 2]. So what do you mean by those 'maybes'?
> *S:* If it is her birthday then she will go to the zoo and if it is open then she will also go to the zoo, perhaps for fun.

This boy, apparently does not read the second premise as an additional premise but as alternative one, i.e. Sanne goes to the zoo on her birthday and she goes if the zoo is open. The strange interpretation this leads to, namely that the zoo being open is a by itself a good enough reason for Sanne to go, apparently occurs to him as well because he inserts a reason for her to go to the zoo when it is open, namely for fun. Similar solutions were given by the other children.

Of the five children who did not make an MT inference in the additional premise condition and also not in the single premise condition only one used the second premise as we would expect:

[9] *Example:* Subject 51, a 9-year old girl

> If Erik has to go shopping for his mother then he will go to the store.
> If the store is open then Erik will go to the store.
> Erik does not go the store
> Does Erik have to go shopping for his mother?
>
> *S:* Yes, only if the store is open.
> *E:* But is says here, he does not go and you think this could be because. . .
> *S:* The store is closed.

Thus out of the group of 12 children who do not make an MT inference in the additional premise condition only five explicitly mention the additional premise in their answer. Recall that this was 12 out of 14 for MP so children can and do in fact use additional premises as is to be expected, just not for MT. Whatever reason they have for ignoring these premises for MT problems must therefore be something rather specific to MT problems instead of a more general aspect of conditional reasoning.

The major difference between MP and MT problems with an additional premise is that with MP problems the additional premise may modify the conclusion but information about this is lacking (q holds unless r holds in which q does not hold, hence the dependence). For MT however, there *is* information about the additional premise, namely $\neg q$ (hence our query that $?q$ fail). So where for MP we know that p holds but not if r holds, for MT we know that at least one of p or r does not hold but not which one.

As we saw above only subject 60 handled this added complexity with no trouble at all. The other subjects appeared more confused than aided by the additional premise however. This confusion quite possibly leads them to consider partial solutions (subject 56) or even an alternative reading of the additional premise (subject 57) in an attempt to arrive at an answer. Since with an additional premise it is tempting to view the premises as dependent on each other (the situation expressed in the second premise may reflect an abnormality with regards to the first) these problems may trigger an increased working memory load.

Alternative premise In the alternative premise condition there was much less confusion. As expected almost all of the children make an MT inference in this case. Their reasoning is similar to that in the single premise condition except that the completion of the problem now becomes $\{(p \vee r) \leftrightarrow q\}$ from which $\neg p$ follows. As with the difference in the number of inferences between the single and additional premise condition, the difference in the number of inferences between the additional and alternative premise condition is also smaller than for MP problems.

The strange case of MT The results for MT are thus quite different from the results for MP problems. Although children endorse MT less frequently than MP for single premises they seem less affected by additional premises on MT problems than on MP problems. Interestingly enough it has been reported that children around ages 10-12 are quite good at MT, i.e. the rate of endorsement is quite high compared to the common adult rate [Evans et al., 1993]. This is strange because MT problems are commonly believed to be more difficult than MP problems if approached classically.

Rumain et al. [1983, p. 475] suggest children might be good at MT for non-classical reasons. They performed an experiment in which they investigated the relevance of invited inferences on conditional reasoning problems. An invited inference is one that is not explicitly entailed by the main inference but can reasonably be said to hold (e.g. $p \rightarrow q$ invites $\neg p \rightarrow \neg q$). An invited inference then would be an 'easy' way to solve and MT problem. They contrasted two hypotheses: 1) invited inferences are necessary and thus a part of children's lexical understanding of *if* which is read as a biconditional; 2) invited inferences are caused by conversational comprehension strategies.

Their method consisted of two experiments performed on adults as well as children. In their experiment subjects were presented with extra information that countermanded the major conditional premise i.e. '$p \rightarrow q$ (major premise), but $\neg p \rightarrow (q \vee \neg q)$ and $q \rightarrow (p \vee \neg p)$' (countermanding information).

If the first hypothesis is correct then children should only be confused by this information. In the second experiment countermanding the invited inferences was made implicit by presenting another conditional in a manner similar to that of Byrne's [1989] alternative premises.

Their results do not support a biconditional reading of *if* but rather lend support to the idea that invited inferences can indeed follow from *if, . . . then* sentences if the context accommodates such inferences. Subject's reasoning is affected by explicitly as well as implicitly countermanding information but they are not very much confused by it. This leads the authors to believe that invited inferences are possibly but not necessarily linked to conditional sentences. In other words, people use invited inferences because of pragmatic assumptions (such as processing difficulty) and context of discourse. The authors suggest that differences between age groups in application of invited inferences may be due to an ability to separate ordinary discourse assumptions from particular logical assumptions required in performing certain reasoning experiments. It is possible that children appear better at MT precisely because they are more likely to apply the invited inferences to their reasoning.

These results are in line with the closed world reasoning approach advocated here. Here too, extra information (be that in the form of additional or alternative premises) is proposed to affect the interpretation of the reasoning problem in a way that influences if and how conclusions are drawn. It is exactly this, and similar considerations, that lie at the heart of the approach advocated here: closed world reasoning implemented in logic programming.

Denial of the antecedent

Single premise The high proportion of subjects who make a DA inference (71.8%) is not unusual. Especially in the single premise condition it is all too easy to forget that something else might be going on; some other event might still warrant the conclusion q. For example:

[10] *Example* Subject 38, a 10-year old boy

If Erik has to go shopping for his mother then he will go to the store.
Erik does not have to go shopping for his mother.
Will Erik go to the store?

S: No.
E: And why doesn't he go to the store?
S: Ehm, to. . . because if he has to, he will go.
E: Yes but could it be that he will go to the store for some other reason? For

example because he has to go shopping for someone else?

S: No, because we are talking about his mother, not about someone else.

An explanation of the reasoning behind this fallacy is straightforward. Given the premise $p \wedge \neg ab \rightarrow q$ and fact $\neg p$. The clause $\neg p$ is taken as the absence of clause p by closed world reasoning. This in turn sets $p \leftrightarrow \perp$. There is no information about ab so we set $ab \leftrightarrow \perp$. This gives $p \leftrightarrow q$ and therefore $\neg q$.

The DA fallacy is common not only to children. Adults too quite often commit this fallacy so this result is to be expected. The steep decline of DA inferences in the additional premise condition already described above is unexpected though.

Additional premise As table 3.7 (p. 69) shows, 15 subjects make the DA inference in the single premise condition but not in the additional premise condition. This rather steep decline is unexpected since an additional premise is not thought to have any effect on a DA inference. As said above, the trick to solving a DA problem correctly (in the classical sense) is to realize an alternative fact might be relevant to one's answer.

Analysis of the transcripts of these subject's experimental sessions reveals that 7 out of 15 subjects who do not make the DA inference in the additional condition but do in the single premise condition do so because of a their reading of the additional premise as an alternative premise. Instead of considering the additional premise as extra information regarding the first conditional premise, these subjects regard it as expressing a separate conditional relation (something similar was seen for a few MT inferences, see above). For example:

[11] *Example* Subject 45, an 11-year old girl

If it is Sanne's birthday then she will go to the zoo.
If the zoo is open then she will go to the zoo.
It is not Sanne's birthday.
Will Sanne go to the zoo?

S: "Maybe, if the zoo is open then she will go."

The DA answer would have been 'no, she will not go to the zoo.' Instead this subject is unsure. Quite possibly she is unsure because she thinks the zoo being open is a good enough reason for Sanne to go and of course it is unknown whether the zoo is open or not, hence her answer. Where this subject remains unsure subject 33 has no problem taking this one step further:

[12] *Example* Subject 33, an 8-year old girl

If it is Sanne's birthday then she will go to the zoo.
If the zoo is open then she will go to the zoo.
It is not Sanne's birthday.
Will Sanne go to the zoo?

S: "Yes."
E: "Yes? And why yes?
S: "Yes, if the zoo is open then she will go to the zoo."

To clarify why these children might be interpreting these premises as alternative rather then additional premises the order of the conditions was examined. 9 of the 15 children who do not make a DA inference here saw problems with alternative premises before they saw problems with additional premises. Of this group of nine, six children explicitly mention an alternative in their answer showing that this is clearly an important factor. It could very well be that the order of the problems, making known the possibility of alternatives, and the subsequent switching of problem type triggered a perserveration error in these children. Findings discussed in section 4 show that it is not uncommon for children to stick to an earlier rule or mode of responding even after a switch should have taken place (these are known as perseveration errors).

Alternative premise As can been seen from table 3.1 (p. 43) 23 subjects indeed do not make a DA inference in the alternative premise condition. All but three subjects mention the alternative premise in their answer confirming that children respond as expected to alternative premises.

Formalization of a DA problem gives $p \wedge \neg ab \rightarrow q$, $r \wedge \neg ab' \rightarrow q$ and fact $\neg p$. Minimization sets ab and ab' to $\perp$ and the clause $\neg p$ is interpreted as the absence of clause p by closed world reasoning. The completion of this is, $\{(p \vee r \leftrightarrow q) \wedge (p \leftrightarrow \perp)\}$ explaining why the answers in this condition were mostly of the form 'although p does not hold, q might still hold if r holds' (of which there is no information).

Note the difference with suppression of MP and MT by an additional premise described above. There closed world reasoning was applied to the additional premise, highlighting a possible abnormality. Here, instead, closed world reasoning is applied to rules, i.e. the alternative premise, and thus leads to the consideration of r as an alternative to p in order to attain q.

The 16 children that do make a DA inference ignore the alternative premise until I point it out to them. For example:

[13] *Example* Subject 54, an 11-year old boy

If it is Sanne's birthday then she will go to the zoo.
If it is her girlfriend's birthday then she will go to the zoo.
It is not Sanne's birthday.
Will Sanne go to the zoo?

S: No, then she will not go to the zoo.
E: But what does the second sentence say?
S: [Reads the second premise] Oh!
E: And you know that it is Sanne's birthday but do you know if it is her girlfriend's birthday? *S:* Maybe. *E:* So if I then ask you if Sanne will go to the zoo, what can you tell me then?
S: Maybe.

As is apparent from this transcript this boy ignored the second premise at first. Not considering something else might motivate Sanne to go the zoo he answers no. After I tell him to consider the second sentence once more he has a 'Eureka moment' and now thinks the answer is 'maybe' since it is unknown whether or not it is the girlfriend's birthday (by the same reasoning as above).

Affirmation of the consequent

Single and additional premises As can be seen from table 3.1 (p. 43) the number of AC inferences is quite low in the single premise as well as in the additional premise condition. In the alternative condition there were few AC inferences made. As with the DA inferences above, the smaller number of AC inferences in the alternative premise condition is to be expected since alternative premises make salient precisely what subjects need to consider to realize making this inference would be fallacious.

The low number of AC inferences in both the single premise and additional premise condition (remember that in these conditions no clues are provided that help recognize the fallacy) is surprising. Usually children frequently endorse both DA and AC inferences.

As can been seen from table 3.8 (p. 69) 16 subjects do not make an AC inference in the single premise condition and 15 subjects refuse to make to this inference in the additional premise condition.[3] In both groups five subjects refuse to make this inference because a possible alternative has occurred to them (given is $p \wedge \neg ab \rightarrow q$ and these children also consider $r \wedge \neg ab' \rightarrow q$).

[3]Actually these groups partially overlap so 25 unique subjects do not make either a single premise AC inference, an additional premise AC inference or neither of these inferences.

Table 3.7: DA inferences

	DA additional	
DA single	Inference not made	Inference made
Inference not made	5	6
Inference made	15	13

	DA alternative	
DA single	Inference not made	Inference made
Inference not made	7	4
Inference made	16	12

	DA alternative	
DA additional	Inference not made	Inference made
Inference not made	14	6
Inference made	9	10

Table 3.8: AC inferences

	AC additional	
AC single	Inference not made	Inference made
Inference not made	6	10
Inference made	9	14

	AC alternative	
AC single	Inference not made	Inference made
Inference not made	9	7
Inference made	18	5

	AC alternative	
AC additional	Inference not made	Inference made
Inference not made	11	4
Inference made	16	8

As with the low number of DA inferences in the additional premise condition, almost all of these children saw problems with an alternative premise before they saw single premise or additional premise problems.

Alternative premise The expected decline of AC inferences in the alternative premise condition compared to the single premise condition should be because of the alternative premise. 18 children who make an AC inference in the single premise condition do not in the alternative premise condition. Of this group 10 children explicitly mention the alternative premise in their answer and two more give their own alternative. In previous sections this has been referred to as applying closed world reasoning to rules. In addition four children refused to make an AC inference because the sentence contained an 'if'. For example:

[14] *Example* Subject 54, an 11-year old boy

If it is Sanne's birthday then she will go to the zoo.
If it is her girlfriend's birthday then she will go to the zoo.
Sanne goes to the zoo.
Is it her birthday?

S: No because there it says 'If it is Sanne's birthday...' but it is not her birthday so no.
E: But how do you know it is not?
S: Because there it says [reads the first premise] but it is not her birthday because it says 'if'.
E: So 'if' means it is not so?
S: Yes like a 'for example' or 'maybe'.
E: Yes but 'if' does not mean that it could be that it is her birthday or 'say it is her birthday, then she will go to the zoo'? If means that it is not so?
S: Yes but there it says [reads the third premise]
E: Yes but according to me it does not say anywhere that it is not her birthday. It only says what would happen if it was her birthday, because then she will go to the zoo. It also says what would happen if it was her girlfriend's birthday, then Sanne will also go to the zoo. And now you know she goes to the zoo. Then the question is, is it her birthday? So does she go to the zoo because it is her birthday?
S: But she could also go with her girlfriend perhaps. So it is maybe.

His first reaction of saying no because there is an if-clause, was seen several times across different problems in different subjects. This reaction seems to contradict earlier correct performances.

One reason why a fallacy is more difficult than a valid inference (in the classical sense) is that for the fallacies one must consider situations which are not explicitly mentioned but might still apply in this context. For example in a single premise DA problem the trick is to realize there might be a situation, $r \wedge \neg ab' \rightarrow q$ so that q could still hold. As can be seen here—and in several other transcripts— this boy does not consider the second premise in his answer at first. His initial answer then, seems to be inspired by a pragmatic argument that the 'if' signifies a lack of certainty, i.e. if it was her birthday it would have been stated. It might simply be an attempt to comply with the task at hand while being unsure what to answer. Reading the premises once more and thinking about my questions he realizes there is a clue in the premises given. In his final answer he uses this clue, namely the alternative premise, and refuses to endorse the inference.

4

Conditional Reasoning and Cognitive Development

In the previous chapter several aspects of children's conditional reasoning were discussed. An analysis of the percentages of endorsement revealed that children's performance on conditional reasoning problems is, especially for single premise problems, quite similar to that of adults. Thus children's conditional reasoning abilities appear quite sophisticated, even at a quite young age. Nonetheless some striking findings are worth noting. These findings concern children's ability to reason with complex conditional problems, i.e. those containing an additional or alternative premise.

As shown in the first section of chapter 3 children's reasoning is clearly affected by such premises in a way quantitatively similar to adults. That is, additional premises tend to suppress the endorsement of MP and MT inferences while alternative premise tend to suppress the endorsement of DA and AC inferences. Recall from section 1.3 reasoning with additional and alternative premises consists of computing the completion of the set of premises. The so-called minimal model that is computed in this manner is thus a combination of the information contained in the premises and the closed world assumption. This model of conditional reasoning predicts that suppression will occur for MP and MT problems if an additional premise is present, and for DA and AC problems if an alternative premise is present, but for different reasons. Suppression of MP and MT by an additional premise occurs because this premise, through closed world reasoning, highlights a possible exception pertaining to the first conditional. Thus we have closed world reasoning applied to exceptions. Suppression of DA and AC by an alternative premise occurs because this premise, through closed world reasoning, highlights a possible alternative to the first conditional. Thus we have closed world reasoning applied to rules (the alternative conditional).

In the second section of chapter 3 it was reported that these two means of suppression appear to be independent. They also appear to be doubly dissociated in the sense that certain subjects apply closed world reasoning to exceptions but not to rules, and other subjects apply closed world reasoning to rules but not to exceptions. Additional support for this dissociation was found in autistic patients. As reported by Stenning and Van Lambalgen, in autists suppression of DA and AC by an alternative premise occurs as it does in normal subjects. Suppression of MP and MT by an additional premise is virtually absent however. Furthermore, in the current data set an age-related increase was found for suppression of DA and AC by an alternative premise, but no such increase was found for the suppression of MP and MT by an additional premise (see section 3.1.1).

Finally, in the third section of 3, it was shown that a substantial number of children indeed vocalize the predictions made by the closed world reasoning approach (see section 1.3). Most importantly we saw that on MP or MT problems with an additional premise, suppression was often the result of noticing the abnormality or exception made salient by the additional premise. Also, on DA or AC problems with an alternative premise, suppression often followed the realization that the alternative premise expresses a reason for affirming the consequent of the first conditional which is independent of the antecedent of this first conditional.

In interpreting the above results it may be useful to incorporate certain relevant aspects from developmental psychology in the discussion of children's conditional reasoning abilities. As the children tested in this study are at a very interesting age regarding cognitive development, and since findings relating the extent of suppression of conditional inferences to age an autism have been reported, it would seem this could be a fruitful exercise.

Perhaps one of the most important developments, from a cognitive point of view, that occurs around the ages of the children tested in the current study, is the structural and functional maturation of the prefrontal cortex. This area of the brain, located directly behind the forehead (see figure 4.1), is thought to be functionally associated with cognitive abilities such as working memory capacity, behavioral inhibition, planning, and the ability to integrate information. These abilities, usually referred to as (a subset of) the so-called *executive functions*, have been studied for numerous years and are thought to be especially important regarding higher cognitive functions including problem-solving and reasoning. The next section provides a brief overview of some of the psychological literature on the development of the prefrontal cortex and the executive functions associated with it.[1]

Figure 4.1: The human brain

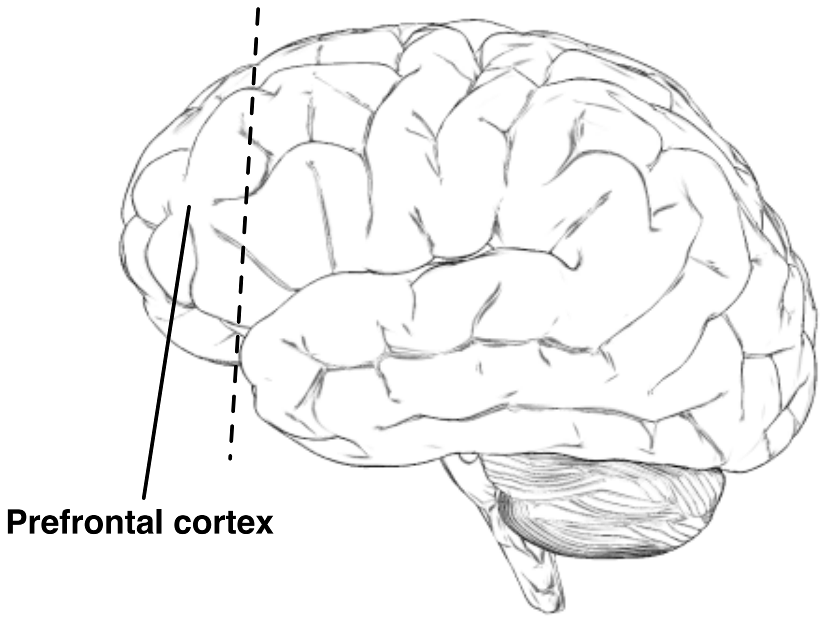

Figure 4.1 The human brain with the location of the prefrontal cortex indicated. [Adapted from: `http://www.cs.princeton.edu/gfx/proj/sugcon/models/`]

4.1 Executive functioning

Executive functions describe higher cognitive functionality such as working memory, planning, inhibition, integration of information, and task switching/attention [Fuster, 1997; Luciana, 2003]. Mainly through the work of Fuster [1997] these higher cognitive functions have come to be associated with the prefrontal cortex (PFC), the anterior area of the frontal lobe. This interpretation of broadly defined functions and crude localizations has not gone without criticism but it is generally accepted as a correct, although general, model. That is,

[1] The next section is adapted from *ADHD and Executive Control* [Berkeljon, 2006], which is a literature study of the merits of ADHD theories in which executive functioning (usually called executive control in this domain) is given a central role. The general conclusion is that while executive functioning deficits are readily observed in ADHD patients, disturbances of executive functioning do not reliably distinguish ADHD patients from patients with other disorders, some of them often comorbid with ADHD (such as Conduct Disorder or a learning disorder). Thus, while executive functioning is almost always impaired in patients with ADHD in one way or another, it is not a unique identifier of ADHD and thus theories which primarily emphasize executive functioning deficits as the core problems in ADHD are inadequate. Nonetheless, the executive functioning impairments observed in ADHD patients, even if not unique to this disorder, do make this population an interesting one to investigate with respect to conditional reasoning.

most everyone acknowledges the broader implications and predictions of this account while at the same time it is understood it is a coarse approximation not without caveats.

4.1.1 Development of executive functioning

On the basis of measurements through experimental tasks hypothesized to tap into executive functions and neuroanatomical development we can say several things about the development of executive functioning and the PFC. Broadly speaking the PFC and the executive functions correlated with it gradually develop during childhood until mature levels of development and function are reached in adolescence [Fuster, 1997; Luciana, 2003].

Cell-development in the PFC appears to be completed in the early postnatal period [Luciana, 2003]. Neurogenesis is achieved during pregnancy (rather later in development compared to other developmental trends), as is the formation of synapses that follows neurogenesis. This formation continues in infancy and levels off during later childhood and into puberty. After puberty gradual synaptic elimination begins which is thought to consolidate functionally stable neural circuits that have been acquired through behavioral experiences.

Behavioral evidence of the development of the PFC and executive functioning can be inferred from various experimental tasks, the one most often used being the spatial delayed response task. In such tasks, which have been tested on a variety of subjects (human and animal), the subject is presented with a stimulus in a particular location of the visual field. After a delay the subject must indicate the location of the stimulus without any aid from recognition memory. Thus the subject is thought to rely solely on an internal representation of the stimulus location kept in working memory. In non-human test subjects (primates) as well as human subjects the capacity to perform this task correctly emerges during late infancy, i.e. around the time that independent motor ability is achieved.

Performance on increasingly complex tasks gradually develops throughout childhood. For example in sorting tasks, such as the Wisconsin Card Sorting Task (WCST), children around ages 3–4 can reliably sort stimuli (cards with a varying number of geometric shapes in different colors in the case of the WCST) according to one criterion of stimulus attributes (e.g. shape only). As they get older children can switch between sorting according to one criterion to sorting according to another (between ages 4 and 5). Around age 6 children are able to sort according to three different criteria.

It has been observed that children make mistakes on the WCST similar to those frontal patients make. These so-called perseveration errors, which manifest themselves as the inability to switch to using a new sorting rule and continuing to sort according to the previous rule, are common to both groups. It has been suggested a lack of inhibition of the previously learned sorting rule (a prime example of an executive function) underlies these errors [Crone, Ridderinkhof, Worm, Somsen, and van der Molen, 2004].

Similarly, on tasks of behavioral inhibition tasks such as the Stroop task. In this task subjects are required to read aloud color names. Difficulty is created by printing the color names in colored fonts that are different from the name printed (e.g. BLUE is printed in green). For young children a variant of this task is used called the Day-Night Stroop task. In this task children must respond to to two cards, one showing a picture of the sun and the other of the moon. Difficulty is created by requiring the children respond by saying “Night” in response to a picture of the sun, and “Day” in response to a picture of the moon. This task is difficult for 3–4-year-olds but fairly easy for 6–7-year-olds. Thus performances on frontally guided tasks improves considerably during the preschool and early school period. Adequate performance on complex frontally guided tasks and behavioral patterns is not reached until after puberty however.

To follow the progression of development and performance across age groups a battery of test measuring frontal functions was developed that could be administered to subjects in age groups ranging from pre-schoolers to adults. This battery of tests, the Cambridge Neuropsychological Testing Automated Battery (CANTAB), consist of of subtasks in the following three domains: (1) working memory/planning; (2) visual memory; and (3) visual attention. Luciana [2003] mentions several studies using the CANTAB to assess the development of frontal lobe function. Results of these studies corroborate the earlier mentioned finding that performance on frontally guided tasks increases during childhood but does not reach a adult performance levels until after puberty.

4.1.2 Executive function and dysfunction

Findings of executive functioning deficits related to disturbances of the frontal areas of brain further substantiates the correlation between these areas and executive functioning. As the development of executive functions follows the maturation of the frontal cortex and thus provides a positive indication of the involvement of the frontal cortex in executive functioning, the pattern of dysfunctional behavior following frontal lesions provides a reverse indication of the role of the frontal cortex.

Fuster [1997] describes seven broad functional disorders related to PFC dysfunction: (1) attention and perception; (2) motility; (3) Memory; (4) planning; (5) intelligence; (6) temporal integration; and (7) affect and emotion. These dysfunctions are observed in frontal patients in varying degrees of severity and are therefore hypothesized to be functionally correlated with this area of the brain. Below is a general characterization of these dysfunctions.

Attention and perception

Deficits in these areas result from abnormalities of attention. Fuster distinguishes seven subtypes: (1) low alertness in which a patient appears generally less alert and aware of the world around him; (2) sensory neglect in which a patient lacks full awareness of one side of the body including any stimulation of this area; (3) distractibility in which a patient is abnormally attracted to irrelevant stimuli and finds it hard to not let these stimuli interfere in ongoing task performance; (4) disorders of visual search and gaze control manifested by an inability to direct attention and unsystematic visual scanning; (5) difficulty sustaining attention manifested by an inability to maintain concentration for any given period of time; (6) internal inference manifested by a sensitivity to interfering internal impulses; and (7) defective motor attention manifested by a sensitivity to internal interference on motor tasks (lack of motor inhibition).

Fuster also notes that frontal lesions are know to cause deficits of perception that cannot be related to deficits in attention. These are mostly disorders in perceiving spatial relationships in one's environment or performing tasks that require guidance by visual information.

Motility

Disorders of motor behavior, not resulting from lesions in the (pre)motor cortex, can result from damage to the PFC. Two kinds of disorders can be distinguished: (1) disorders of general spontaneous motor behavior; and (2) disorders of goal-directed motor behavior.

Disorders of spontaneous motor behavior fall into two categories, hypokinesia and hyperkinesia. Hypokinesia is diminished spontaneous motor activity which can vary greatly in degree, ranging from mild aspontaneity to severe apathy. Hyperkinesia on the other hand is characterized by excessive and aimless motor behavior (hyperactivity).

Disorders of goal-directed motor behavior often exist alongside the disturbed motor behavior just mentioned, but are not caused by them. They are thought to result from cognitive deficits in initiation, planning, and organization of action which in turn are manifestations of PFC lesions in their own right.

Memory

PFC patients are capable of forming and retrieving long-term memory episodes. Any problems associated with long-term memory are most often attributable to deficits in organization and monitoring of the material to be committed to memory. There thus appears to be a deficit in applying so-called mnemonic strategies most probably because of a lack of attention and drive.

A typical, and more structural problem is that of a working memory deficit. Frontal patients exhibit problems on tasks that require keeping in memory certain operations required for adequate task performance. These effects appear to be related to deficits in sustained attention and susceptibility to interference but are completely explained by them. There are patients who fail on tests of working memory even if attention and interference are controlled for.

Planning

Fuster notes that while deficits in memory lead to problems with hindsight, forethought is hindered by planning difficulties. This lack of "prospective memory" is closely related to difficulties in formulating and applying plans. As with working memory deficits, here too inattentiveness and lack of drive appear to play a role in this aspect of the disorder. Also the heightened sensitivity to internal interference (competing plans or ideas) most frontal patients exhibit may pose especially disruptive in executing plans.

Intelligence

Setting aside the semantic mess that plagues intelligence research a few general remarks about intelligence in relation to the PFC can be made. One striking finding is that most frontal patients have normal IQ scores and are capable of marked intellectual achievements. However, as the previous suggest, certain aspects of cognitive function are most definitely impaired. Upon closer inspection it appears that so-called 'fluid intelligence', or the capacity to solve novel problems, is indeed impaired in frontal patients. This is not necessarily completely detrimental to cognitive functioning but it is a definite impairment nonetheless. It speaks for the versatility and persistence of those patients that are able to achieve quite normal performance and behavior in spite of their cognitive impairments.

Temporal integration

Temporal integration concerns the integration of (novel) behavioral routines. Frontal patients have no trouble carrying out previously learned behavioral routines but learning new ones is an almost impossible challenge. Fuster remarks that the synthesis of new behavioral routines, acts or plans is perhaps the most quintessential aspect of 'executive function'. This deficit of temporal integration then, appears to be the cumulative effect of the previously mentioned functional disorders.

As with the deficits in intelligence the deficits in temporal integration do not preclude normal functioning. They are only evident in challenging situations and not necessarily in simple, daily activities. Thus frontal patients can lead relatively normal lives, albeit with much routine and habit and without much creativity or imagination.

Affect and emotion

These are among the most elusive of disorders associated with PFC lesions since it is difficult to disentangle these emotional problems from the cognitive disorders they may be secondary too. Be this as it may, two emotional manifestations are observed, often dissociated: apathy and euphoria.

Apathy is manifested by the same symptoms as inattention and hypokinesia (diminished motor activity). Patients show low awareness and lack of initiative and motility. Affect and emotional responses are most aptly described as blunted. Note that this is not the same as clinical depression although Fuster reports that PFC lesion can result in clinical depression. In both cases it is possible that the changes in affective behavior follows cognitive deficits as a reaction to the deterioration of cognitive functions.

Euphoria, as it occurs in frontal patients, is characterized by a sporadic but recurrent elevation of mood and resembles pathological mania. It is often associated with two symptoms of frontal dysfunction discussed earlier, distractibility and hyperactivity.

Social effects

As should be evident from the discussion of major symptoms of frontal dysfunction above, a lesion in the frontal lobes can have varying effects on cognitive functioning and behavior. As such, quite often lesions in this area of the brain lead to changes, be that subtle or dramatic, in the way lesioned patients interact socially. Some cognitive effects will be less noticeable than others. The effects on affect and emotion are usually among the most dramatic changes. Such changes can and do have adverse effects on the frontal patient's personality and social behavior.

4.2 Conditional reasoning

It is obvious that not all aforementioned aspects of executive functioning are relevant to conditional reasoning. Taken together the reported dysfunctions describe a broad range of behaviors, none of which need to be impaired to the extent described here for frontal patients to have implications for conditional reasoning. The above does however, provide a good indication of the functional role of the prefrontal cortex and the cognitive capacities it is associated with.

The findings of the functional role of the prefrontal cortex, together with the observed late development of this area of the brain in childhood, suggests a definite and direct impact on the nature and functional development of children's cognitive abilities. Especially relevant to the current discussion of conditional reasoning are response control (inhibition), working memory, and (fluid) intelligence. Directing attention to focus on relevant information contained in the premises of conditional reasoning problems, keeping in mind this relevant information and integrating it to from a conclusion, are all essential in solving such problems adequately.

Precisely which information is considered and in what way, is specified by the closed-world account of conditional reasoning put forth in Stenning and Van Lambalgen [2005] and advocated in the current paper. In this account conditional reasoning is the computation of a (minimal) model of the problem based on the information contained in its premises. The so-called 'suppression effect', which entails that MP and MT inference are suppressed by supplying an additional premise, and that DA and AC are suppressed by supplying an alternative premise, has proved especially useful in piecing apart how this information is used.

It was suggested that the different information contained in the additional and alternative premises would reveal two distinct closed world reasoning processes or strategies. For additional premises this is closed world reasoning applied to exceptions, and for alternative premises this is closed world reasoning applied to rules. An analysis of the suppression effect across subjects indeed supports a dissociation based on this hypothesis. There are subjects who apply closed world reasoning to exceptions and those who do not, and there are subjects who apply closed world reasoning to rules and those who do not. Most revealing was the finding that there are subjects who apply one type of closed world reasoning exclusively, i.e. there are subjects who apply closed world reasoning to exceptions but not to rules and subjects who apply closed world reasoning to rules and not exceptions.

These results are compatible with findings relating to possible developmentally relevant factors, namely the developmental disorder autism and age. As reported [see section 3.3.1 and chapter 8 of Stenning and Van Lambalgen] autists appear to apply closed world reasoning to rules but not exceptions (i.e. DA and AC are suppressed by an alternative premise but MP and MT are not suppressed by an additional premise). In the sample of normal school-age children tested in the current study an age-related increase of suppression of DA and AC by an alternative premise was found. No such age-related increase was found for the suppression of MP and MT by an additional premise.

The absence of an age-related increase in suppression of MP and MT was explained earlier by noting that the exception-handling involved—as hypothesized by the closed world reasoning approach and subsequently supported by empirical findings in the current study— is a difficult process. The fairly equal percentages of endorsement across age groups (see section 3.1.1) suggest subjects find these problems equally difficult regardless of age. It thus appears that the two independent types of closed world reasoning, for exceptions and for rules are under separate developmental control. Closed world reasoning applied to rules, as evidenced by suppression of DA and AC by an alternative premise, appears to be positively correlated with age and thus cognitive maturity. Closed world reasoning applied to exceptions seems much less directly correlated with age, at least for the age range the current results are based on.

Considering the theoretical model of conditional reasoning that has been proposed, namely closed world reasoning, the following conclusion seems plausible. Closed world reasoning is a cognitive process dependent on cognitive mechanisms which involve a bracketing of information that is considered in solving a problem, resulting in separating relevant from irrelevant, and likely from unlikely confounding influences. This bracketing can be applied to different types of information. Relevant to the present discussion of conditional reasoning are: (1) closed world reasoning applied to exceptions (to conditionals), and (2) closed world reasoning applied to rules (alternatives to conditionals). It is suggested that while closed world reasoning as a general ability is present and developing at age 7 or 8 (based on the data reported in the current study), the application of closed world reasoning to different types of information develop independently and at a different pace.

At first glance closed world reasoning applied to exceptions and closed world reasoning applied to rules might appear similar processes as far use of cognitive means are concerned. Both processes rely on the consideration of two premises, an additional one or an alternative one, in attempting to solve a problem. The essential difference between them then, is not the amount of information considered, but rather how this information is considered. This is best illustrated by considering the formalization of reasoning problems with additional and alternative premises introduced in section 1.3.

Recall that solving a (complex) conditional reasoning problem was proposed to involve computing a minimal model of its premises: the first conditional $p \wedge \neg ab \rightarrow q$; the second conditional $r \wedge \neg ab' \rightarrow q$; and some fact (e.g. p or q etc.). For any given conditional inference with an additional premise this means the minimal program consists of $(p \wedge r) \rightarrow q$ and some fact. By contrast, for any given conditional inference with an alternative premise, the minimal program consist of $(p \vee r) \rightarrow q$ and some fact. Based on this it is not difficult to see how two different cognitive mechanisms may be involved in these types of reasoning. One mechanism which processes an outcome/response based on one or more criteria to be satisfied (as in reasoning with an alternative premise), and one which processes an outcome/response based on an inhibiting or limiting exception imposed on a rule currently in use (as in reasoning with an additional premise). It is important to note that deficits in the latter mechanism would not manifest themselves as deficits of inhibition across the board. It is the proposed deficit in exception-handling which triggers the lack of inhibition on a rule currently in use. This is indicative of a cognitive inflexibility rather than a general deficit in inhibition of behavior/responses. This distinction will be important when considering precisely where a deficit with respect to exception-handling originates in cognitive functioning and is thus important to bear in mind.

The cognitive mechanisms relevant to the present discussion thus, at the very least, involve the capacity to retain and use multiple items of information (working memory) but also the capacity to inhibit/alter upcoming responses using these items of information (behavioral inhibition and planning). It is interesting in this regard that on the Wisconsin Card Sorting Task the number of criteria children are able to sort to correctly increases with age while the number of perseveration errors (responding according to the previous, instead of the current, sorting rule) decreases with age. This seems to reflect both an increasing capacity to handle multiple items of information, and an increasing capacity to inhibit/alter previously learned responses in cases where novel information requires this.

There thus appear to be cognitive mechanisms, related to those proposed to be involved in conditional reasoning, which develop during childhood and adolescence, and are associated with development of the prefrontal cortex. The connection between the closed world reasoning approach to conditional reasoning, including the defined subprocesses of closed world reasoning applied to exceptions and to rules, and the development of specific cognitive capacities associated with the prefrontal cortex therefore seems adequately established. These proposals, although incomplete, hopefully provide sufficient theoretical and empirical constraints in order to guide future research in this area.

5

Conclusion

In the conditional reasoning experiments reviewed in this paper by Byrne [1989], Dieussaert et al. [2000], Schmittmann and Raijmakers, as well as in the results of the current study, a suppression effect was observed. That is, supplying subjects with extra premises in conditional reasoning problems often decreases their willingness to endorse particular inferences they had no problem endorsing without the extra information. This highlights a systematicity in the cognitive processes underlying conditional reasoning. In accounting for this systematicity Stenning and Van Lambalgen [2005] have proposed a two-step reasoning process where a problem is solved by first reasoning *to* an interpretation of the problem based on content and context, and subsequently reasoning *from* this interpretation to solve the problem. This last step is taken to be a kind of model construction in which a preferred or minimal model is selected from all possible models of the premises given in the problem. A formal framework for this model construction is provided by logic programming with negation as failure.

This approach has proved useful in accounting not only for existing psychological data [Byrne, 1989] but also for data gathered in the current study. That is, Stenning and Van Lambalgen [2005] give a theoretical account of the reasoning processes guiding the subjects in Byrne's study in terms of logic programming with negation as failure and are able to give a explanation for why subjects make the inferences they do. In the current study data was collected explicitly to test Stenning and Van Lambalgen's formal approach on children's conditional reasoning. Here it proved useful as well in accounting for inferences and mistakes made by the subjects.

Central to this approach is the *suppression effect*, as reported by Byrne [1989]. This effect, which entails a decrease in the number of endorsements of the classical inferences MP and MT if an additional premise is supplied and a decrease in the number of classical fallacies DA and AC if an alternative premise is supplied, has proved robust yet difficult to explain. Some conclude from the existence of the suppression effect that subjects do not reason logically, at least not according to the rules of classical logic. Stenning and Van Lambalgen [2005] argue, however, that it is entirely possible and indeed likely that (most) subjects do not reason according to classical logic but instead adopt a nonmonotonic logic, i.e. a logic tailored for reasoning on the basis of incomplete information. Elsewhere [Stenning and Van Lambalgen, 2004] it has been observed that subjects tend to interpret conditionals as allowing exceptions, i.e. they assume they are reasoning on the basis of incomplete information and that further information might present an exception to the conditional statement under consideration. Nonetheless, subjects arrive at a conclusion. They do not get stuck considering *all* possible exceptions. The implementation of the reasoning process should thus allow us to derive a stable conclusion yet be flexible enough to allow for new information to update this conclusion.

The chosen implementation, logic programming with negation as failure, does just that. From the information obtained through the premises a model (of the world or rather the world which the premises describe) is constructed. This model is demarcated by the the so-called *closed world assumption* which is semantically implemented through *negation as failure*. Simply put this means that a model is constructed that assumes the information contained in the premises and nothing more. In fact, it assumes any information *not* contained in the premises is false. In this sense the model constructed is *minimal*. This minimal model provides us with exactly what we need to draw a conclusion (if the premises contain enough information to derive a conclusion, that is). This model construction is a computational process and is therefore sensitive to the computational abilities and constraints of the system in which it is implemented. The assumption is, quite justifiably, that in humans this system is working memory within the domain of executive functions. With regards to this assumption some notable results from this study can be reported.

An interesting result obtained in this study, and partially through the analysis of the data from Schmittmann and Raijmakers, is that children are quite good at simple conditional reasoning problems. For example, the rates of endorsements for MP problems are on par with those of adults (see figure 3.1). An analysis of age-related effects revealed a distinct difference between younger and older children in willingness to endorse the classical fallacies DA and AC. Overall performance on these inferences does not appear to differ substantially from that of adults (figure 3.1). If age is taken into account, however, a marked difference between younger children (under age 10) and older children (above age 10) appears (table 3.3). Younger children are much more likely to endorse the classical fallacies than older children are. For DA and AC, suppression by an alternative premise was found to increase with age. This increase is not solely attributable to the difference in endorsement of DA and AC between younger and older children as evidenced by a substantial difference in the mean effects sizes of these two findings. No such increase was observed for suppression of MP and MT by an additional premise. More than likely this finding, or rather the lack thereof, reflects the relative difficulty of MP and MT problems with an additional premise. The fairly equal percentages of endorsement across all ages seems to indicate subjects of all ages find these problems equally difficult.

Following the hypothesis, proposed by Stenning and Van Lambalgen, that closed world reasoning as applied to conditionals consists of two separate processes, namely closed world reasoning applied to exceptions (which accounts for suppression of MP and MT by an additional premise), and closed world reasoning applied to rules (which accounts for suppression of DA and AC by an alternative premise), an analysis of responses from individual subjects in sequence was conducted. Through this analysis support for the dual process hypothesis was indeed found. More importantly, these process were shown to be dissociated, i.e. there are subjects who apply closed world reasoning to exceptions but not to rules and vice versa.

Key excerpts from the experimenter-subject dialogue recorded during the reasoning task further establish these results and show that a substantial number of subjects are indeed led by the considerations the closed world approach proposes. It is encouraging to see such direct support for the predictions of the theoretical model. This is evidence of its potential not just as a general explanation of conditional reasoning, but also of its worth in explaining reasoning mechanisms and considerations in individuals.

In the discussion of topics from developmental psychology maturation of the prefrontal cortex and associated executive functions was shown to be relevant to conditional reasoning. Especially cognitive processes such as working memory, behavioral inhibition, and planning, perhaps the most quintessential executive functions, were suggested to be involved in closed world reasoning applied to exceptions and to rules. Note that this is in no way a trivial conclusion. The possible involvement of these functions suggest there may exist relevant similarities between the nature of, and performance on conditional reasoning tasks and tasks which are thought to measure working memory, behavioral inhibition and planning such as the Wisconsin Card Sorting task, the Stroop Task and Tower of London/Hanoi type problems. Also, the possible involvement of these functions suggests samples from populations known to be impaired on measures of working memory, behavioral inhibition, or planning may exhibit very specific impairments on conditional reasoning tasks. This thus suggests several possible avenues of research. For an example of relevant similarities in performance on measures of conditional reasoning, working memory, and inhibition, see Schmittmann and Raijmakers [in preperation]. For an example of specific impairments on conditional reasoning related to impaired executive functions see chapter 8 of Stenning and Van Lambalgen [forthcoming].

In conclusion it can be remarked that he suppression effect and associated phenomena now appear to be sufficiently well-established so as to serve as a paradigm case in the field of conditional reasoning. The studies reviewed in the current paper, and experimental results analyzed, consistently find similar results. The interpretation of these results in terms of closed world reasoning is not that well-established and fairly new. It does, however, appear to be very well supported, not only by data previously gathered and reinterpreted in terms of this approach (such as Byrne's [1989] results), but also by data explicitly gathered with this approach in mind (such as the current study). It also provides interesting suggestions regarding relevant cognitive mechanisms underlying conditional reasoning and the suppression effect. As far as these have been tested or assessed, it appears the presently advocated approach and its suggestions are supported. It will be interesting to see what the future brings. It certainly looks promising.

A

Results Previous Studies

Table A.1: Percentages of endorsement in Dieussaert et al. [2000]

	Inference type			
	MP	MT	DA	AC
Single premise	88.3%*	69.6%	49.3%	55.1%
With an additional premise	60.6%	43.9%	49.2%	53.0%
With an alternative premise	93.3%	69.3%	22.0%	16.0%

* $N = 70$; $n_{single} = 23$, $n_{additional} = 22$ and $n_{alternative} = 25$. Two subjects were excluded from the analysis. Each subject solved three problems.

Table A.2: Percentages of endorsement in Byrne [1989]

	Inference type			
	MP	MT	DA	AC
Single premise	96.0%*	92.0%	46.0%	71.0%
With an additional premise	38.0%	33.0%	63.0%	54.0%
With an alternative premise	96.0%	96.0%	4.0%	13.0%

* $N = 8$. Each percentage is based on the responses of 8 subjects to three problems.

B

Conditional Arguments

B.1 Single conditionals

MP
Als Jeroen een werkstuk moet maken dan gaat hij naar de leeszaal.
Jeroen moet een werkstuk maken.
Gaat Jeroen naar de leeszaal?

MT
Als Jeroen een werkstuk moet maken dan gaat hij naar de leeszaal.
Jeroen gaat niet naar de leeszaal.
Moet Jeroen een werkstuk maken?

DA
Als Jeroen een werkstuk moet maken dan gaat hij naar de leeszaal.
Jeroen moet geen werkstuk maken.
Gaat Jeroen naar de leeszaal?

AC
Als Jeroen een werkstuk moet maken dan gaat hij naar de leeszaal.
Jeroen gaat naar de leeszaal.
Moet Jeroen een werkstuk maken?

MP
Als Sanne jarig is dan gaat zij naar de dierentuin.
Sanne is jarig.
Gaat Sanne naar de dierentuin?

MT
Als Sanne jarig is dan gaat zij naar de dierentuin.
Sanne gaat niet naar de dierentuin.
Is Sanne jarig?

DA
Als Sanne jarig is dan gaat zij naar de dierentuin.
Sanne is niet jarig.
Gaat Sanne naar de dierentuin?

AC
Als Sanne jarig is dan gaat zij naar de dierentuin.
Sanne gaat naar de dierentuin.
Is Sanne jarig?

MP
Als Erik boodschappen moet doen voor zijn moeder dan gaat hij naar de winkel.
Erik moet boodschappen doen voor zijn moeder.
Gaat Erik naar de winkel?

MT
Als Erik boodschappen moet doen voor zijn moeder dan gaat hij naar de winkel.
Erik gaat niet naar de winkel.
Moet Erik boodschappen doen voor zijn moeder?

DA
Als Erik boodschappen moet doen voor zijn moeder dan gaat hij naar de winkel.
Erik moet geen boodschappen doen voor zijn moeder.
Gaat Erik naar de winkel?

AC
Als Erik boodschappen moet doen voor zijn moeder dan gaat hij naar de winkel.
Erik gaat naar de winkel.
Moet Erik boodschappen doen voor zijn moeder?

B.2 Additional conditionals

MP

Als Jeroen een werkstuk moet maken dan gaat hij naar de leeszaal.
Als de leeszaal open is dan gaat Jeroen naar de leeszaal.
Jeroen moet een werkstuk maken.
Gaat Jeroen naar de leeszaal?

MT

Als Jeroen een werkstuk moet maken dan gaat hij naar de leeszaal.
Als de leeszaal open is dan gaat Jeroen naar de leeszaal.
Jeroen gaat niet naar de leeszaal.
Moet Jeroen een werkstuk maken?

DA

Als Jeroen een werkstuk moet maken dan gaat hij naar de leeszaal.
Als de leeszaal open is dan gaat Jeroen naar de leeszaal.
Jeroen moet geen werkstuk maken.
Gaat Jeroen naar de leeszaal?

AC

Als Jeroen een werkstuk moet maken dan gaat hij naar de leeszaal.
Als de leeszaal open is dan gaat Jeroen naar de leeszaal.
Jeroen gaat naar de leeszaal.
Moet Jeroen een werkstuk maken?

MP

Als Sanne jarig is dan gaat zij naar de dierentuin.
Als de dierentuin open is dan gaat Sanne naar de dierentuin.
Sanne is jarig.
Gaat Sanne naar de dierentuin?

MT

Als Sanne jarig is dan gaat zij naar de dierentuin.
Als de dierentuin open is dan gaat Sanne naar de dierentuin.
Sanne gaat niet naar de dierentuin.
Is Sanne jarig?

DA
Als Sanne jarig is dan gaat zij naar de dierentuin.
Als de dierentuin open is dan gaat Sanne naar de dierentuin.
Sanne is niet jarig.
Gaat Sanne naar de dierentuin?

AC
Als Sanne jarig is dan gaat zij naar de dierentuin.
Als de dierentuin open is dan gaat Sanne naar de dierentuin.
Sanne gaat naar de dierentuin.
Is Sanne jarig?

MP
Als Erik boodschappen moet doen voor zijn moeder dan gaat hij naar de winkel.
Als de winkel open is dan gaat Erik naar de winkel.
Erik moet boodschappen doen voor zijn moeder.
Gaat Erik naar de winkel?

MT
Als Erik boodschappen moet doen voor zijn moeder dan gaat hij naar de winkel.
Als de winkel open is dan gaat Erik naar de winkel.
Erik gaat niet naar de winkel.
Moet Erik boodschappen doen voor zijn moeder?

DA
Als Erik boodschappen moet doen voor zijn moeder dan gaat hij naar de winkel.
Als de winkel open is dan gaat Erik naar de winkel.
Erik moet geen boodschappen doen voor zijn moeder.
Gaat Erik naar de winkel?

AC
Als Erik boodschappen moet doen voor zijn moeder dan gaat hij naar de winkel.
Als de winkel open is dan gaat Erik naar de winkel.
Erik gaat naar de winkel.
Moet Erik boodschappen doen voor zijn moeder?

B.3 Alternative conditionals

MP

Als Jeroen een werkstuk moet maken dan gaat hij naar de leeszaal.
Als Jeroen een boek moet lenen dan gaat hij naar de leeszaal.
Jeroen moet een werkstuk maken.
Gaat Jeroen naar de leeszaal?

MT

Als Jeroen een werkstuk moet maken dan gaat hij naar de leeszaal.
Als Jeroen een boek moet lenen dan gaat hij naar de leeszaal.
Jeroen gaat niet naar de leeszaal.
Moet Jeroen een werkstuk maken?

DA

Als Jeroen een werkstuk moet maken dan gaat hij naar de leeszaal.
Als Jeroen een boek moet lenen dan gaat hij naar de leeszaal.
Jeroen moet geen werkstuk maken.
Gaat Jeroen naar de leeszaal?

AC

Als Jeroen een werkstuk moet maken dan gaat hij naar de leeszaal.
Als Jeroen een boek moet lenen dan gaat hij naar de leeszaal.
Jeroen gaat naar de leeszaal.
Moet Jeroen een werkstuk maken?

MP

Als Sanne jarig is dan gaat zij naar de dierentuin.
Als haar vriendin jarig is dan gaat Sanne naar de dierentuin.
Sanne is jarig.
Gaat Sanne naar de dierentuin?

MT

Als Sanne jarig is dan gaat zij naar de dierentuin.
Als haar vriendin jarig is dan gaat Sanne naar de dierentuin.
Sanne gaat niet naar de dierentuin.
Is Sanne jarig?

DA

Als Sanne jarig is dan gaat zij naar de dierentuin.
Als haar vriendin jarig is dan gaat Sanne naar de dierentuin.
Sanne is niet jarig.
Gaat Sanne naar de dierentuin?

AC

Als Sanne jarig is dan gaat zij naar de dierentuin.
Als haar vriendin jarig is dan gaat Sanne naar de dierentuin.
Sanne gaat naar de dierentuin.
Is Sanne jarig?

MP

Als Erik boodschappen moet doen voor zijn moeder dan gaat hij naar de winkel.
Als Erik boodschappen moet doen voor zijn vader dan gaat hij naar de winkel.
Erik moet boodschappen doen voor zijn moeder.
Gaat Erik naar de winkel?

MT

Als Erik boodschappen moet doen voor zijn moeder dan gaat hij naar de winkel.
Als Erik boodschappen moet doen voor zijn vader dan gaat hij naar de winkel.
Erik gaat niet naar de winkel.
Moet Erik boodschappen doen voor zijn moeder?

DA

Als Erik boodschappen moet doen voor zijn moeder dan gaat hij naar de winkel.
Als Erik boodschappen moet doen voor zijn vader dan gaat hij naar de winkel.
Erik moet geen boodschappen doen voor zijn moeder.
Gaat Erik naar de winkel?

AC

Als Erik boodschappen moet doen voor zijn moeder dan gaat hij naar de winkel.
Als Erik boodschappen moet doen voor zijn vader dan gaat hij naar de winkel.
Erik gaat naar de winkel.
Moet Erik boodschappen doen voor zijn moeder?

References

Barwise, J. [1986]. Conditionals and conditional information. In Traugott, E.C., A. ter Meulen, Reilly, J.S., & Ferguson, C.A., editors, *On conditionals*, pages 21–54. University Press, Cambridge.

Berkeljon, A. [2006]. ADHD and Executive Control: On the Merits of an Executive Control Theory of Attention-Deficit/Hyperactivity Disorder. Master's thesis, Universiteit van Amsterdam. In preparation.

Brewka, G., Dix, J., & Konolige, K. [1997]. *Nonmonotonic reasoning: an overview*. CSLI Publications, Leland Stanford Junior University: United States.

Byrne, R.M.J. [1989]. Suppressing valid inferences with conditionals. *Cognition*, **31**:61–83.

Cohen, J. [1988]. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Hillsdale, NJ, 2nd edition.

Crone, E.A., Ridderinkhof, K.R., Worm, M., Somsen, R.J.M, & van der Molen, M.W. [2004]. Switching between spatial stimulus-response mappings: a developmental study of cognitive flexibility. *Developmental Science*, **7**(4):443–455.

Dieussaert, K., Schaeken, W., Schroyen, W., & d'Ydewalle, G. [2000]. Strategies during complex conditional inferences. *Thinking and Reasoning*, **6**(2):125–161.

Evans, J.St.B.T., Newstead, S.L., & Byrne, R.M.J. [1993]. *Human reasoning: the psychology of deduction*. Lawerence Erlbaum Associates, Hove, Sussex.

Fuster, J.M. [1997]. *The prefrontal cortex: anatomy, physiology and neuropsychology of the frontal lobe*. Lippincott-Raven Publishers, Philedephia, third edition.

Gamut, L.T.F. [1991]. *Logic, language and meaning*, volume 1: Introduction to logic. The University of Chicago Press, Chicago.

Luciana, M. [2003]. The Neural and Functional Development of the Human Prefrontal Cortex. In de Haan, M., & Johnson, M.H., editors, *The Cognitive Neuroscience of Development*. Psychology Press, New York, NY.

Rips, L.J. [1983]. Cognitive processes in propositional reasoning. *Psychological Review*, **90**(1):38–71.

Rumain, B., Connell, J., & Braine, M.D.S [1983]. Conversational comprehension process are responsible for reasoning fallacies in children as well as adults: If is not the biconditional. *Developmental Psychology*, **19**:471–481.

Sanford, D.H. [1989]. *If P, then Q: conditionals and the foundations of reasoning*. Routledge, London.

Schmittmann, V.D., & Raijmakers, M.E.J. Hypothesis testing and conditional reasoning styles. In preperation.

Stalnaker, R.C. [1968]. A theory of conditionals. Reprinted in Harper, Stalnaker & Pearce, 1981, pp. 41–55.

Stalnaker, R.C. [1975]. Indicative conditionals. Reprinted in Harper, Stalnaker & Pearce, 1981, pp. 139–210.

Stalnaker, R.C. [1981]. A defense of conditional excluded middle. In Harper, W.L., Stalnaker, R.C., & Pearce, G., editors, *IFS*, pages 87–104. D. Reidel Publishing Company, Dordrecht: Holland.

Stenning, K., & Van Lambalgen, M. Human reasoning and cognitive science. Forthcoming.

Stenning, K., & Van Lambalgen, M. [2004]. A little logic goes a long way: basing experiment on semantic theory in the cognitive science of conditional reasoning. *Cognitive Science*, **28**(4):481–530.

Stenning, K., & Van Lambalgen, M. [2005]. A working memory model of relations between interpretation and reasoning. *Cognitive Science*. Submitted.

van Benthem, J.F.A.K., van Ditmarsch, H.P., Ketting, J., Lodder, J.S., & Meyer-Viol, W.P.M. [2003]. *Logica voor informatici*. Pearson Education Benelux.

Wason, P. [1968]. Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, **20**:273–281.