Article scientifique   Article   2009   Published version   Open Access

––

Heuristic and analytic processes in mental models for conditionals: An integrative developmental theory

––

Gauffroy, Caroline; Barrouillet, Pierre Noël

Review

Heuristic and analytic processes in mental models for conditionals: An integrative developmental theory

Caroline Gauffroy *, Pierre Barrouillet *

Université de Genève, Faculté de Psychologie et des Sciences de l'Education, 40, bd du pont d'Arve, 1211 Genève 4, Switzerland

A B S T R A C T

This article presents a developmental dual-process theory of the understanding of conditionals that integrates Evans' heuristic–analytic theory within the revised mental model theory of conditional proposed by Barrouillet, Gauffroy, and Lecas (2008). According to this theory, the interpretation of a conditional sentence is driven by unconscious and implicit heuristic processes that provide individuals with an initial representation that captures its meaning by representing the cases that make it true. This initial model can be enriched with additional models (a process named fleshing out within the mental model theory) through the intervention of conscious and demanding analytic processes. Being optional, these processes construct representations of cases that are only compatible with the conditional, leaving its truth-value indeterminate when they occur. Because heuristic processes are relatively immune to developmental changes, while analytic processes strongly develop with age, the initial model remains stable through development whereas the number of additional models that can be constructed increases steadily. Thus, the dual-process mental model theory predicts in which cases conditionals will be deemed true, indeterminate, or false and how these cases evolve with age. These predictions were verified in children, adolescents and adults who were asked to evaluate the truth value and the probability of several types of conditionals. The results reveal a variety of developmental trajectories in the way different conditionals are interpreted, which can all be accounted for by our revised mental model theory.

© 2009 Elsevier Inc. All rights reserved.

* Corresponding authors. Fax: +41 22 379 90 20 (P. Barrouillet).
E-mail addresses: Caroline.Gauffroy@unige.ch (C. Gauffroy), Pierre.Barrouillet@unige.ch (P. Barrouillet).

As stressed by Evans and Over (2004), "If" is one of the most interesting and important words in the language, because it expresses hypothetical thought which is essential by permitting humans to imagine and analyze possible states of the world. Accordingly, the way individuals interpret and reason from conditional statements of the form *If p then q* has been the subject of many theoretical proposals (Evans, Newstead, & Byrne, 1993; Evans, Over, & Handley, 2005; Johnson-Laird & Byrne, 2002). How children and adolescents understand conditional relations is of particular relevance not only for the development of logic (Inhelder & Piaget, 1958), but also for scientific reasoning and how hypotheses are supported or contradicted by evidence (Kuhn, Amsel, & O'Loughlin, 1988), for causal learning (Kushnir & Gopnik, 2007), or for understanding of social rules and conventions (Harris & Nunez, 1996; Light, Blaye, Gilly, & Girotto, 1989). Thus, several studies have been devoted to the development of conditional reasoning (see for recent examples Artman, Cahan, & Avni-Babad, 2006; Barrouillet & Lecas, 2002; Daniel & Klaczynski, 2006; Klaczynski, Schuneman, & Daniel, 2004). In the present article, we explore the implications of the integration, within the developmental theory of conditional proposed by Markovits and Barrouillet (2002), of recent theoretical advances in the domain of reasoning, namely Evans' (2006) heuristic–analytic theory and the revised mental model theory of conditional proposed by Barrouillet et al. (2008).

Klaczynski et al. (2004) distinguished between three approaches to conditional reasoning in the developmental literature. A traditional approach, championed by Piaget (1972), Inhelder and Piaget (1958), Overton (1990), and Moshman (1990), assumes that conditional reasoning is underpinned by general-domain operators, available only at the late adolescence, permitting to comprehend and coordinate the concepts of necessity (given *If p then q*, *q* is a logically necessary conclusion when *p* is true) and indeterminacy (the determinate conclusion *p* does not necessarily follow from *If p then q* and the premise *q is true*). By contrast, other theories reject the idea of such domain-general operators, suggesting that conditional reasoning is driven by inferential rules either acquired through experience within particular domains or innate. For example, deontic reasoning would be an instance of these domain-specific structures differing from reasoning processes in other domains (Cheng & Holyoak, 1985; Cosmides, 1989; Cummins, 1996). A third approach, proposed by Markovits and Barrouillet (2002), suggests that reasoning relies on both general processes implemented in existing cognitive architecture and on domain specific knowledge. When reasoning from *If p then q* sentences, reasoners would retrieve relevant knowledge from long-term memory related to the content of the conditional in order to construct transient mental models in working memory. Development would result from age-related changes in both the range of available knowledge and the capacity of working memory, with increased capacity permitting the construction and maintenance of a higher number of models and a better access to stored knowledge.

Among the assumptions of Markovits and Barrouillet's (2002) theory is the idea that reasoning by mental models is underpinned by the automatic activation of experience-based knowledge. This echoes the opposition put forward by the dual-processing accounts of reasoning between, on the one hand, automatic, implicit, and unconscious mechanisms that would constitute a System 1 and, on the other hand, deliberative and conscious processes constituting a System 2 (Evans, 2006; Evans & Over, 1996; Stanovich, 1999). However, Markovits and Barrouillet (2002) did not explicitly introduce this distinction in their account of development of conditional. This article fills this gap by presenting a first step towards a developmental theory of conditional which integrates the main assumptions of Evans' (2006) heuristic–analytic theory of reasoning within the mental model framework proposed by Markovits and Barrouillet (2002). This integration is made possible by proposals about the nature of mental models for conditional we recently put forward (Barrouillet et al., 2008).

Mental models and the development of conditional

According to the mental model theory, people reason from conditional sentences of the form *If p then q* (e.g., *If the piece is square, then it is red,* in which *p*, the antecedent, stands for the proposition *the piece is square*, and *q*, the consequent, stands for the proposition *the piece is red*) by constructing and manipulating mental models of the states of affairs these sentences refer to (Johnson-Laird & Byrne, 1991). These models would represent possibilities that can occur when the sentence is true.

Because constructing and maintaining transient models in working memory is particularly demanding and goes against cognitive economy, Johnson-Laird and Byrne (2002) proposed a *principle of implicit models* stating that basic conditionals[1] have mental models representing the possibilities in which their antecedents are satisfied, but only implicit models for the possibilities in which their antecedents are not satisfied. Thus, individuals would construct an initial model representing both *p* and *q* satisfied, which corresponds in our example to representing the co-occurrence of the properties *square* and *red*, the other possibilities remaining implicit (the three dots in the following diagram):

square   red
          . . .

Implicit information can be made explicit through a demanding and time-consuming fleshing out process adding two additional possibilities, one representing pieces that are not square and not red, another representing pieces that are not square but red. This fleshing out process leads to the following complete three-model representation in which the sign ¬ refers to a negation tag (Johnson-Laird & Byrne, 1991; Johnson-Laird & Byrne, 2002):

  square   red
¬ square   ¬red
¬ square   red

In the following, we will refer to these models as the *p q*, the ¬*p*¬*q* and the ¬*p q* models respectively. It is worth noting that, within this account, there is no psychological difference between the explicit mental model pertaining to the initial representation (i.e., *p q*) and those added through the fleshing out process. Thus, the content of the initial representation is more postulated than justified in Johnson-Laird and Byrne's theory as well as in its developmental revision proposed by Markovits and Barrouillet (2002).

A revised mental model theory

We have recently proposed to abandon the notion of cognitive economy in accounting for the distinction between initial and fleshed out models and hypothesized that the initial model represents explicitly those states of affairs that make the sentence psychologically true when they occur (Barrouillet et al., 2008). In the case of conditionals with no semantic, referential or knowledge-based relation between the antecedent and the consequent (i.e., basic conditionals), this corresponds to the sole case in which both *p* and *q* are satisfied (e.g., a red square makes true the conditional *If the piece is square, then it is red*). The role of the fleshing out process would be to enrich this initial model by representing those states of affairs that do not make the sentence true when they occur, but that are nonetheless compatible with it (e.g., a green triangle or a red triangle). Thus, when presented with states of affairs and asked to evaluate the truth of a basic conditional, adults who are able to construct the complete three-model representation would deem the conditional true for the *p q* case, its truth value remaining neither true nor false but indeterminate for cases matching the models issued from fleshing out (i.e., ¬*p* ¬*q* and ¬*p q* cases). By contrast, those states of affairs that do not correspond to any model, even when the fleshing out process has been completed (i.e., *p* ¬*q*, in our example a blue square), are considered as incompatible with the sentence and as falsifying it.

These hypotheses were verified by taking advantage of the developmental changes that occur in the fleshing out process. In many previous studies, we have described the development of the interpretation of conditional by children and adolescents. In a first level predominant in primary school children (about 8 or 9 years of age), children are only able to construct the initial model *p q* but fail to envisage other possibilities, thus adopting a conjunctive interpretation in which the sole *p q* case is considered as compatible with the conditional, all the other cases being incompatible. In a second level often observed around age 12, young adolescents are able to flesh out the initial representation

---

[1] According to Johnson-Laird and Byrne (2002), basic conditionals are conditionals in which the antecedent and the consequent have no semantic or referential relations, or relations based on knowledge.

by adding a $\neg p \ \neg q$ model, favoring a biconditional interpretation. At this level, both the $p \ q$ and the $\neg p \ \neg q$ cases are compatible with the conditional and the others incompatible. Finally, older adolescents are able to construct the complete three-model representation postulated by Johnson-Laird and Byrne (2002, i.e., $p \ q$, $\neg p \ \neg q$, and $\neg p \ q$). These three developmental levels were observed in several experiments involving a variety of tasks (Barrouillet, Grosset, & Lecas, 2000; Barrouillet & Lecas, 1998; Barrouillet & Lecas, 2002), and we established that the number of models that can be constructed is linked to the age-related increase in working memory capacities (Barrouillet & Lecas, 1999).

According to Barrouillet et al.'s (2008) proposal, states of affairs matching the initial model make the conditional true while those matching fleshed out models are considered as irrelevant, leaving the truth value of the conditional indeterminate. This hypothesis leads to a series of predictions that were tested using a truth evaluation task. In this task, also called truth table task, participants are presented with a conditional sentence, and asked to say for each of the four possible logical cases (i.e., $p \ q$, $\neg p \ q$, $p \ \neg q$, and $\neg p \ \neg q$) if it makes the conditional true, false, or if it leaves its truth-value indeterminate. First, the children who favor a conjunctive interpretation should consider the conditional as true for the $p \ q$ case but false for all the other cases because they do not flesh out the initial representation. Second, adolescents who favor a biconditional interpretation should consider the $p \ q$ case as making the conditional true, the $\neg p \ \neg q$ case constructed through fleshing out as leaving its truth-value indeterminate, and the $\neg p \ q$ and $p \ \neg q$ cases as falsifying it, resulting in a defective biconditional interpretation. This interpretation is said defective because the connective has no truth value in all the possible cases, this truth value remaining indeterminate for $\neg p \ \neg q$. Finally, older adolescents and adults who have access to the conditional interpretation should deem the conditional true for $p \ q$, indeterminate for $\neg p \ \neg q$ and $\neg p \ q$, and false for $p \ \neg q$, thus exhibiting the so-called defective truth table of the conditional. It should be noted that these predictions differ from those that could be drawn from the standard mental model theory according to which the states of affairs matching the mental models of a sentence would make this sentence true. In this case, the three developmental levels would correspond in the truth evaluation task to "true" responses to the $p \ q$ case for the conjunctive level, to the $\neg p \ \neg q$ and the $p \ q$ cases for the biconditional level, and finally to the $p \ q$, $\neg p \ \neg q$, and $\neg p \ q$ cases for the conditional level. Actually, Barrouillet et al. (2008) observed the predicted developmental trend from conjunctive to *defective* biconditional and then *defective* conditional interpretations, leading strong support to the hypothesis that the mental models pertaining to the initial representation refer to state affairs that make the conditional true whereas those states of affairs that are only compatible with the conditional are represented by additional models through the fleshing out process. However, though empirical evidence supported our predictions, there was no rationale for the hypothesis of different epistemic status of initial and fleshed out mental models. We assume that the distinction made between heuristic and analytic processes by dual-process theories can provide such a rationale.

Heuristic and analytic processes in mental models construction

In his recent review of the dual-process theories of reasoning, judgment, and social cognition, Evans (2008) noted that although the writings of mental model theorists include little explicit discussion of a distinction between two systems of reasoning, this distinction is implicitly present. Indeed, he remarked that the mental model theory describes the formation of the initial model as a relatively automatic and effortless process whereas the fleshing out of initially implicit models is effortful, constrained by working memory capacities and as a consequence error prone. These two kinds of processes correspond to the heuristic and analytic systems described by Evans (2006) in his revised heuristic–analytic theory of reasoning. If we draw a parallel between the two theories, the formation of the initial model would be underpinned by heuristic, that is fast, tacit and unconscious processes. Following the so-called relevance principle of the heuristic–analytic theory, when interpreting a conditional statement, the heuristic system would produce an epistemic model which is the most plausible and believable with reference to prior knowledge elicited by context and the current goals. The heuristic system cues default mental models that imply default responses, inferences, or decisions. This is exactly what the mental model theory considers as the initial model.

Contrary to previous versions of the theory, Evans' (2006) extended heuristic–analytic theory assumes that the analytic system may or may not intervene depending on instructional set, time

available, and general intelligence. Thus, the age-related increase in intellectual abilities leads to assume that development should be one of the factors facilitating the intervention of this system. Analytic processes can revise or replace default mental models and inhibit default heuristic responding. They are controlled rather than automatic, conscious, slow, and sequential in nature. They require access to a single, capacity-limited working memory resource. As suggested by Evans (2008), this corresponds to the process of fleshing out described by the mental model theory.

This conception substantiates Barrouillet et al's. (2008) proposal about a difference in epistemic status between initial and fleshed out models. Because the initial model comes spontaneously to mind through fast, impenetrable, and unconscious processes when listening a sentence, individuals assimilate this representation to the meaning of this sentence.[2] Thus, when confronted with a situation matching this default mental model, they deem the sentence true. By contrast, fleshed out models are constructed through analytic processes that enrich the representation but are optional in nature. These analytic processes add models representing states of affairs that are compatible with the conditional sentence (i.e., they do not contradict the initial model), but that do not pertain to its basic meaning. Thus, when confronted to these states of affairs, individuals consider the conditional neither true, because the corresponding models did not come spontaneously to mind, nor false because these situations are nonetheless compatible with the conditional and do not contradict it. Thus, the heuristic/analytic distinction can provide the rationale our hypothesis was missing, but it leads also to new developmental hypotheses resulting from differences in the developmental trends of the two kind of processes.

The development of the two systems of reasoning

Because the analytic system, System 2, relies on general intelligence and working memory, it can be expected to develop with age until the late adolescence. Indeed, developmental studies have shown that working memory capacity evolves slowly and progressively with age, probably until adulthood (Barrouillet, Gavens, Vergauwe, Gaillard, & Camos, 2009; Case, Kurland, & Goldberg, 1982; Gathercole, Pickering, Ambridge, & Wearing, 2004). Accordingly, a host of studies have demonstrated System 2 development, children's reasoning becoming more analytical, complex, and abstract with age (Barra, Bucciarelli, & Johnson-Laird, 1995; Barrouillet & Lecas, 1998; Daniel & Klaczynski, 2006; Jenveau-Brennan & Markovits, 1999; Kokis, MacPherson, Toplak, West, & Stanovich, 2002; Moshman & Franks, 1986). Thus, as we suggested above, development should be added to the factors susceptible to influence both the likelihood and the efficiency of the intervention of the analytic system within Evans' theory.

By contrast, there is no consensus about the way cognitive development affects System 1. Theoretically, this heuristic system would be relatively independent of general intelligence and age (Reber, 1993; Stanovich, 1999), as the unconscious and implicit processes on which it relies (Vinter & Perruchet, 2000). However, many studies using tasks issued from the "heuristic and biases" literature have provided evidence that the heuristic system increases in efficiency with age. For example, Jacobs and Potenza (1991) found that, in the domain of social decision, the normative use of base-rate decreased with age while judgments based on a representativeness heuristic became more and more frequent. Davidson (1995) observed that the susceptibility to the conjunction fallacy increases during elementary school years. In their fuzzy-trace theory, Brainerd and Reyna (2001) assume that both the analytic and the heuristic systems develop with age. Accordingly, Reyna and Ellis (1994) observed that the framing effect in risky decision making, which relies on a heuristic bias, was more pronounced as children grew older. Markovits and Dumas (1999) reported that inferences about nonlinear relations concerning friendship that are nonlogically deducible increase with age. More recently, Morsanyi and Handley (2008) reported a marked increase in heuristic responding with age that was related to the developmental increase in cognitive capacity in children from 6 to 10 years of age. Thus, it could be concluded from these findings that the heuristic processes responsible for the construction of the initial mental models probably develop with age during adolescence.

---

[2] We would say its *core meaning*, but this expression has already been used in a different sense by Johnson-Laird and Byrne (2002) who name core meaning the complete three-model representation of the basic conditionals.

However, though this list may seem impressive and shows that the heuristic system can become more predominant with age in some circumstances, there are at least three reasons to doubt of such a developmental trend concerning the basic processes underpinning the production of the initial model of conditionals. First, the results evoked above proved to be particularly volatile. For example, whereas Morsanyi and Handley (2008) report an increase between ages 6 and 10 in the sunk cost fallacy based on applying the heuristic "do not waste", Klaczynski and Cottrell (2004) report the *reverse* trend between 7 and 14 years of age with an increase in normative decisions. Though Jacobs and Potenza (1991) found a developmental increase in the use of the representativeness heuristic from first grade to adulthood, Kokis et al. (2002) found the *reverse* developmental trend between the fifth and the eighth grade and suggested that Jacobs and Potenza's findings probably result from confounding the use of the representativeness heuristic with the increasing knowledge of social stereotypes. The susceptibility to the conjunction fallacy that is described as increasing in school years by Davidson (1995) is found to *decrease* between early and middle adolescence by Klaczynski (2001). Among the 17 measures of normative and non-normative responses taken by this author from a battery of reasoning and decision making tasks, none of them indicated an age-related increase of non-normative responses while the developmental trend for normative responses was significant on 14 measures. Thus, many of the most often cited studies reporting developmental increases in heuristic proved to be difficult to replicate.

Second, it is generally assumed in dual-process theories that the heuristic system is an older evolutionary product that is ontogenetically earlier developing whereas the analytic system is both a phylogenetically and ontogenetically later developing system (Kokis et al., 2002). The studies reporting a developmental increase in heuristic tendencies broadly confirm this conjecture. Though Jacob and Potenza's (1991) study is often referred to as having demonstrated that adult's judgments are inferior to those of children (e.g., Reyna & Ellis, 1994), a careful reading of the article reveals that the developmental change occurred mainly between the first and the third grade (with a decline in normative scores from .31 to .23), but that third graders did practically not differ from sixth graders and college students whose normative score was .21. In the same way, the studies of Davidson (1995), Reyna and Ellis (1994), Morsanyi and Handley (2008), and Markovits and Dumas (1999) that report development of heuristics are limited to preschool and school years. Thus, as we noted above, significant changes in heuristic use have more often been described during childhood than during adolescence.

Third, most of the developmental trends in heuristic use are related to social contents and seem to rely on the acquisition of social knowledge and stereotypes (e.g., Davidson, 1995; Jacobs & Potenza, 1991; Markovits & Dumas, 1999). These developmental trends are less often observed in non-social context (e.g., Jacobs & Potenza, 1991), suggesting that they depend on the acquisition of social norms and stereotypes rather than on changes at the level of the heuristic processes themselves. This could explain the surprising relationship observed by Morsanyi and Handley (2008) between cognitive capacities and heuristic responding, a relationship that is at odds with the dual-process theories that all agree that heuristics are independent of cognitive capacity. Indeed, how could cognitive capacities have an impact on the tacit, implicit, automatic, and unconscious processes constituting System 1? We will here endorse the explanation given by Morsanyi and Handley. They assume that the heuristics they studied (the conjunction fallacy, the if-only fallacy, the sunk cost fallacy, or the belief bias) at least partly depend on conscious capacity-demanding processes in the case of children, such as retrieving and integrating relevant knowledge into the representation of tasks. They conclude that, instead of challenging the idea of a heuristic system immune from developmental and individual differences, their results challenge the idea of non-normative responses being necessarily the result of System 1.

However, contrary to the complex processes involved in responding to the tasks used by Morsanyi and Handley (2008), the heuristic processes described by Evans (2008) as responsible for the construction of the default representation constituting the initial model are simpler basic processes relying on automatic activation and retrieval of knowledge from memory. These processes operate without any need of conscious and effortful integration by providing a representation that comes spontaneously to mind.

In summary, though a tendency exists for a developmental increase in heuristic responding in reasoning tasks, it is mainly observed in childhood and on tasks requiring the retrieval and integration of social knowledge. By contrast, several studies have shown that analytic processes develop through childhood but also during adolescence (Klaczynski, 2001) and until adulthood (Barrouillet et al., 2000; Barrouillet et al., 2008). Thus, there should be a sharp contrast in the way the two systems develop during adolescence, which is the period studied in this article. Those processes described by Evans (2006) as heuristic in his heuristic–analytic model, which are implicit, automatic, and largely unconscious, should remain largely unaffected by developmental changes whereas the analytic processes should strongly develop.

It is worth to note that this approach does not assume that during development, the analytic system progressively overcomes and supersedes the heuristic system. Following Evans (2007a), we adopt his default-interventionist model according to which the analytic system may or may not modify the default representation provided by the heuristic system. Of course, the development of the analytic system implies that reasoning should more often rely on analytic processes in older than in younger subjects, but the default interventionist conception also implies that the reasoning should never be on the exclusive reliance of the analytic system, even in adults. On the contrary, heuristic processes provide the conscious contents on which analytic processes apply and, as a consequence, heuristic processes influence and modulate reasoning at any age (Evans, 2006). As we will see, there are many cases in which these heuristic processes strongly constrain the representations constructed by children, adolescents and even adults, resulting in the disappearance of the developmental trends usually observed.

Developmental implications for conditionals understanding

This heuristic–analytic distinction and the differences in the developmental course of the two systems permit new developmental predictions. First, if the initial model results from heuristic processes, its constitution should remain stable through development, at least in the range of ages we previously studied (i.e., from 8 years onwards). As a consequence, following Barrouillet et al.'s (2008) assumption about the epistemic status of the different models, there should be no evolution with age of the cases considered as making the conditionals true. By contrast, and this is our second hypothesis, mental models constructed through fleshing out by the analytic system should evolve with age as well as, consequently, those cases deemed as leaving indeterminate the truth value of the conditional or making it false.

Though several studies, including from our research group, have investigated the way the interpretation of conditionals develops with age, we are not aware of a developmental study examining how the cases considered as making various conditionals true or false evolve with age from childhood to adulthood.[3] The three following experiments filled this gap by presenting children, adolescents, and adults with a truth table task in which they were asked to evaluate the truth value of basic, causal and inducement conditionals for the four logically possible cases. In a following part, we examine how development affects the evaluation of the probability of conditionals. This question is of particular interest because the responses given by adults to the probability task (Evans, Handley, & Over, 2003; Oberauer & Wilhelm, 2003) have been considered as the main empirical evidence supporting the suppositional theory put forward by Evans (2007b) and Evans et al. (2005) and contradicting the standard mental model theory of conditional proposed by Johnson-Laird and Byrne (2002). Before studying the responses given by children and adolescents to the probability task, we focus on the interpretation of basic conditionals revealed by the truth table task.

---

[3] Byrnes and Overton (1988) ran such a study in, 3rd, 5th, 8th graders and college students, but using exclusively familiar conditional relations ("If it has rained, then the grass is wet"). The problem raised by this kind of task is that participants know that the conditional is true before studying the cases (actually, it is true that when it has rained, the grass is wet). As a consequence, there is a risk that children, adolescents, and even adults do not think about the truth of the conditional, but about the possibility of the different cases from their own experience. Accordingly, children and adolescents favored a biconditional response pattern, responding true for *p q* (rain – wet grass) and ¬*p* ¬*q* (no rain – dry grass) and false for ¬*p q* (no rain – wet grass) and *p* ¬*q* (rain – dry grass). Adults often responded true for the ¬*p q* case (no rain – wet grass), being able to evoke alternative causes. These patterns are very rare with unfamiliar relations.

Basic conditionals

The aim of this study was to verify that the heuristic processes provide an initial model remaining stable across age, whereas analytic processes responsible for the fleshing out of this initial representation develop with age, while being still modulated by the heuristic system. For this purpose, 8-year-old children, 11- and 15-year-old adolescents and adults were asked to evaluate basic conditionals in a truth table task by deciding if given states of affairs made the conditional either true, false, or if its truth value was indeterminate. From the hypothesis that states of affairs matching the initial model make the conditional true, whereas those matching the models issued from the fleshing out leave the truth value of the conditional indeterminate, Barrouillet et al. (2008) predicted and observed that the responses to a truth table task evolve from a conjunctive pattern, favored by children, to a defective biconditional pattern in adolescents and a defective conditional pattern in adults. The first aim of this experiment was to replicate this developmental trend.

The second aim of this experiment was to explore how the heuristic system can constrain the process of fleshing out and the output of the analytic system, and how this interaction between the two systems evolves with age. For this purpose, we took advantage of a phenomenon reported by Barrouillet and Lecas (1998) and Barrouillet and Lecas (2002) who observed that basic conditionals involving binary terms as antecedent and consequent elicit a biconditional interpretation. For example, a sentence like "If the bird is a female then it has a light plumage" involves a biconditional interpretation in adolescents as well as in adults who reject both $p \neg q$ and $\neg p q$ cases as incompatible with the rule and produce a pattern of conditional inferences based on only two models, $p q$ and $\neg p \neg q$ (Barrouillet & Lecas, 1998; Barrouillet & Lecas, 2002).

However, what these previous studies left indeterminate is the nature of this biconditional interpretation, because the tasks they used did not involve to evaluate the truth of the conditionals. On the one hand, it could be assumed that conditionals with binary terms are understood as biconditionals because they pragmatically induce the inverse implicature, which would be added to the initial representation by the heuristic system ("If the bird is a female then it has a light plumage, hence if it is a male then it has a dark plumage"). In this case, binary terms would induce an equivalence reading in which both $p q$ and $\neg p \neg q$ cases make the conditional true, whereas $p \neg q$ and $\neg p q$ cases would make it false. On the other hand, and this was the explanation favored by Barrouillet and Lecas (1998), the structure of semantic memory could affect the fleshing out process. Because both "female" and "light" allow only one alternative, the first step of the fleshing out would lead to a complete representation in which each value of the antecedent is associated with a corresponding value of the consequent (*female light* and *male dark*). This would result in a one-to-one correspondence (p associated with q and $\neg p$ with $\neg q$) that blocks any further fleshing out and the production of the $\neg p q$ model (Barrouillet & Lecas, 1998). If this is the case, binary terms would not affect the output of the heuristic but of the analytic system. As a consequence, they should not involve equivalence but defective biconditional readings because $\neg p \neg q$ cases would match a model constructed through fleshing out, thus leaving the truth value of the conditional indeterminate, instead of making it true.

To decide between these two possibilities, the present experiment contrasted conditionals involving binary terms in both the antecedent and the consequent (BB conditionals) with conditionals in which the antecedent and the consequent do not have opposite terms (NN conditionals). According to the hypothesis that binary terms only affect the output of the analytic system, the initial model constructed by the heuristic system should remain unaffected by both the age and the type of conditional. Thus, in all age groups and for both BB and NN conditionals, $p q$ should be the sole case eliciting "true" responses. By contrast, because the development of the analytic system leads to an increase in the production of the $\neg p \neg q$ model through fleshing out, there should be an age-related increase in the frequency of responses of indeterminacy to $\neg p \neg q$ cases, and a corollary decrease in "false" responses. This effect should be observed with both BB and NN conditionals. The same developmental trend, although delayed, should be observed for $\neg p q$ cases, but only with NN conditionals. Indeed, binary terms should block the production of the $\neg p q$ model and abolish the developmental increase in responses of indeterminacy with BB conditionals, $\neg p q$ cases being considered as making the conditional false. Thus, we expected an interaction between age and type of conditional for the responses of inde-

Table 1
Predicted patterns of responses to BB (with binary terms in both the antecedent and the consequent) and NN (with nonbinary terms) conditionals at each of the three developmental levels in the truth evaluation task in Experiment 1 (Def.: Defective).

Type of rule	Logical case	Developmental level		
		Conjunctive	Def. biconditional	Def. conditional
NN	$p\ q$	True	True	True
	$\neg p\ \neg q$	False	Indeterminate	Indeterminate
	$\neg p\ q$	False	False	Indeterminate
	$p\ \neg q$	False	False	False
BB	$p\ q$	True	True	True
	$\neg p\ \neg q$	False	Indeterminate	Indeterminate
	$\neg p\ q$	False	False	False
	$p\ \neg q$	False	False	False

terminacy on these cases. Finally, in all the age groups and for both types of conditional, participants should deem the sentences false for $p\ \neg q$ cases. Because the locus of the effect of binary terms is in the fleshing out process, younger children who do not flesh out the initial model should remain unaffected. These predictions are summarized in Table 1.

Concerning individual responses, for the NN conditionals, a first developmental level should correspond to the predominance of conjunctive interpretations, followed by the predominance of defective biconditional and then defective conditional interpretations. Thus, the rate of conjunctive patterns should decrease with age, the rate of defective biconditional patterns should increase and then decrease, following a quadratic trend, whereas the rate of defective conditional patterns should increase. The BB conditionals should elicit conjunctive interpretations in younger participants and then a predominant defective biconditional interpretation even in adults. It is worth to note that these predictions are ordinal and concern developmental trends. Thus, each developmental level should not be considered as characterizing a precise age group.

Experiment 1

These predictions were tested in 20 third graders (mean age = 8.6 years, SD = 0.3, 14 females), 24 sixth graders (mean age = 11.6 years, SD = 0.6, 15 females), 22 ninth graders (mean age = 15.7 years, SD = 0.4, 12 females), and 26 students from the University of Geneva (mean age = 24.5 years, SD = 1.41, 19 females) who performed a truth table task containing both BB and NN conditionals (see Appendix). Four BB conditionals and four NN conditionals were associated with the four logical cases $p\ q$, $\neg p\ \neg q$, $\neg p\ q$ and $p\ \neg q$, resulting in 32 trials. The task was administered using a video projector. For each trial, the conditional statement displayed on the top of the screen described a pictured scene representing one of the four logical cases. For example, for the BB rule "If the door is open then the light is switched on", the presented cases took the form of "open door – light switched on" (i.e., p q), "closed door – light switched off" (i.e., $\neg p\ \neg q$), "closed door – light switched on" (i.e., $\neg p\ q$) or "open door – light switched off" (i.e., $p\ \neg q$). The experimenter read the conditional statement aloud and asked the participants to pay attention to it. After 7 s, a first picture representing either the p or the *not-p* instance appeared in the left side of the screen (e.g., an open door), followed 2 s after by a second picture representing either the q or the *not-q* instance on the right (e.g., a light switched on). The experimenter described aloud both pictures, and three boxes corresponding to the three response possibilities (i.e., "true", "one cannot know", and "false") appeared under the pictures. The same responses possibilities were printed on individual response sheets. The following instructions were given: "By looking at the pictures, you must decide if the sentence is true, or if it is false, or if the picture does not permit to know if the sentence is true or false. Thus, you notch the box "true" if the picture makes the sentence true, the box "one cannot know" if the picture does not permit to know if the sentence is true or false, and the box "false" if the picture makes the sentence false. Do not speed up, take your time, and think about it before responding". Presentation order of the 32 trials was counterbalanced in such a way that two successive trials never presented the same logical case, nor the same rule. The experimenter controlled when the next trial was initiated. This procedure ensured that all the participants had enough time to answer.

Table 2

Percent of responses "true" (T), "one cannot know" (I) and "false" (F) for each type of cases as a function of grades and type of conditionals, either BB (with binary terms in both the antecedent and the consequent) or NN (with nonbinary terms) conditionals, in Experiment 1.

| | | Cases | | | | | | | | | | | |
| | | $p\,q$ | | | $\neg p\,\neg q$ | | | $\neg p\,q$ | | | $p\,\neg q$ | | |
Grades	Conditionals	T	I	F	T	I	F	T	I	F	T	I	F
3	NN	100	0	0	0	21	79	0	14	86	0	14	86
$n = 20$	BB	98	0	2	2	19	79	0	19	81	2	19	79
6	NN	100	0	0	0	46	54	1	18	81	0	16	84
$n = 24$	BB	99	1	0	1	59	40	0	17	83	1	18	81
9	NN	99	1	0	1	82	17	1	30	69	0	0	100
$n = 22$	BB	100	0	0	0	93	7	0	22	78	0	5	95
Adults	NN	99	1	0	1	98	1	1	86	13	1	1	98
$n = 26$	BB	96	2	2	1	94	5	1	27	72	0	2	98

Analysis of responses

The age-related evolution in the rate of "true", "false" and "one cannot know" responses for each logical case on the two types of conditionals can be seen in Table 2. As we predicted, the $p\,q$ cases elicited a high rate of "true" responses that did not evolve with age and was not affected by the type of conditional, whereas the $p\,\neg q$ cases involved a high rate of "false" responses. A 4 (Grades: third, sixth, ninth and adults) × 2 (Conditionals: NN and BB) ANOVA with the last factor as within-subject variable on the number of the "false" responses on $p\,\neg q$ cases did not reveal any effect of conditionals, $F(1, 88) = 1.78$, $p = .26$, $\eta_p^2 = .06$, and no interaction between grades and conditionals, $F < 1$. However, there was an unexpected increase in "false" responses with age, $F(3, 88) = 3.70$, $p < .01$, $\eta_p^2 = .11$, which will be explained by the response patterns analysis. As it was expected, the rate of indeterminate responses on $\neg p\,\neg q$ cases increased with age for NN , $F(3, 88) = 27.35$, $p < .001$, $\eta_p^2 = .48$, and BB conditionals, $F(3, 88) = 30.50$, $p < .001$, $\eta_p^2 = .51$, with no effect of conditionals, $F(1, 88) = 1.98$, $p = .32$, $\eta_p^2 = .02$, and no interaction, $F(3, 88) = 1.01$, $p = .38$, $\eta_p^2 = .09$. Post hoc comparisons (Newman–Keuls test) revealed for both conditions a significant increase in indeterminate responses from grade 3 to 6, and 6 to 9 ($ps < .01$), whereas the increase from grade 9 to adulthood did not reach significance. Exactly as we predicted, this age-related increase in indeterminate responses on $\neg p\,q$ cases was delayed with NN conditionals, $F(3, 88) = 29.33$, $p < .001$, $\eta_p^2 = .50$, and was only significant from grade 9 to adults ($p < .001$), whereas there was no age effect with BB conditionals, $F < 1$, $\eta_p^2 = .02$. This resulted in a significant interaction between age and type of conditional, $F(3, 88) = 26.12$, $p < .001$, $\eta_p^2 = .47$, which was mainly due to the strong increase in indeterminate responses with NN conditionals exhibited by adults; this interaction was no longer significant when this group was discarded ($p = .10$).

Response patterns analysis

These results are clearer when expressed in terms of response patterns. We distinguished three expected interpretations: *conjunctive* (true, false, false, false responses to $p\,q$, $\neg p\,\neg q$, $\neg p\,q$, and $p\,\neg q$ cases respectively), *defective biconditional* (true, indeterminate, false, false) and *defective conditional* (true, indeterminate, indeterminate, false). One additional interpretation observed in third and sixth graders consisted in responding true for $p\,q$, false for $\neg p\,\neg q$ and indeterminate for $\neg p\,q$ and $p\,\neg q$ cases. Barrouillet et al. (2008) had already observed this pattern named *matching pattern* in younger participants (third and sixth graders). They suggested that it reflects a matching strategy in which the conditional is judged true when the two pictures verify both the antecedent and the consequent evoked in the sentence, indeterminate when only one of the two propositions is verified, and false when there is no match at all. The occurrence of this matching pattern in younger participants explains the age-related increase in false responses to the $p\,\neg q$ case. The interpretation associated with the response pattern for each of the eight conditional rules presented (i.e., four BB and four NN conditionals) was
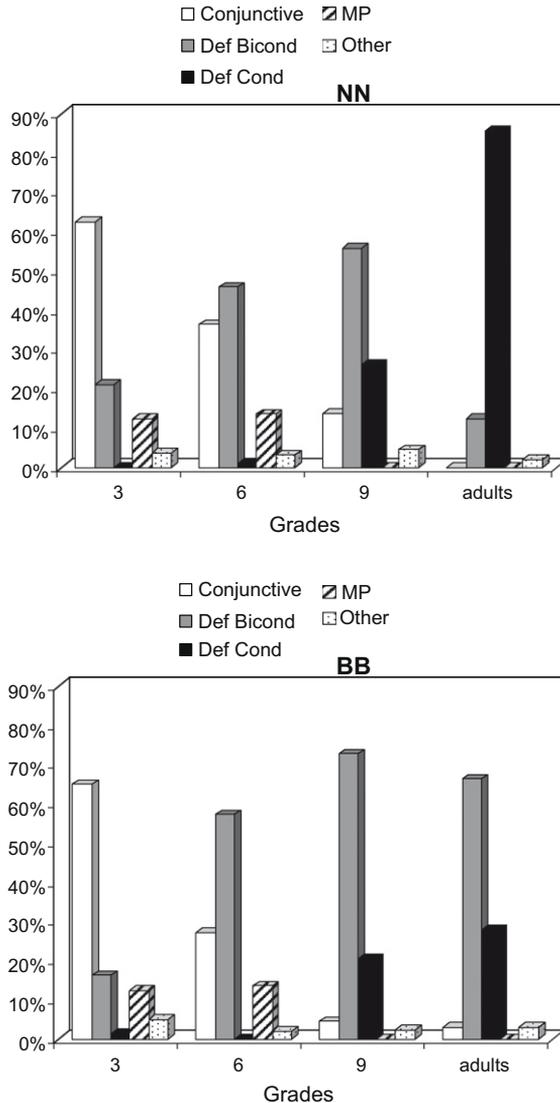
Fig. 1. Percent of response patterns categorized as conjunctive, defective biconditional (Def Bicond), defective conditional (Def Cond), matching (MP) and others as a function of grades for NN and BB conditionals in Experiment 1.

determined for each participant. As shown in Fig. 1, for each age and experimental condition, almost all the response patterns corresponded to one of the four interpretations described above.

First of all, as we predicted, conjunctive patterns were predominant in younger participants and their rate decreased with age for both types of conditionals, with significant linear trends in both cases, $F(1, 88) = 41.87$, $p < .001$, $\eta_p^2 = .32$, and $F(1, 88) = 54.96$, $p < .001$, $\eta_p^2 = .38$ for NN and BB conditionals respectively (Fig. 1). In line with our hypotheses, for NN conditionals, the defective biconditional interpretation represented an intermediate developmental step testified by a significant quadratic trend, $F(1, 88) = 17.51$, $p < .001$, $\eta_p^2 = .21$, and was followed by the defective conditional responses that increased with age and became predominant in adults with a significant linear trend, $F(1, 88) = 50.23$, $p < .001$, $\eta_p^2 = .51$. Concerning BB conditionals, our main prediction was that binary terms should affect

the fleshing out process and induce a predominant *defective* biconditional interpretation in adolescents and adults. Accordingly, these response patterns were rarer in sixth graders than in the older groups, $F(1, 88) = 22.47$, $p < .001$, $\eta_p^2 = .17$, in which they did not vary with age, $F(2, 88) = 1.14$, $p = .27$, $\eta_p^2 = .02$. In accordance with the idea that binary terms block the fleshing out process, defective conditional interpretations were rarer with BB conditionals and did not reach 30% in adults.

It is worth noting that, although the four BB conditionals did not differ in the rate of biconditional responses they elicited (51%, 53%, 55%, and 54% respectively), the occurrence of defective conditional patterns in older participants suggests that the effect of binary terms is not universal. In others words, the BB conditionals were sometimes treated as NN conditionals. When there was no effect of binary terms, the standard developmental trend from a conjunctive, to a defective biconditional and finally a defective conditional interpretation was observed (Fig. 1).

Discussion

Two main findings arose from this experiment. First, in line with our predictions, "true" responses remained stable across age and concentrated on *p q* cases only. As a consequence, participants who fleshed out this initial representation exhibited defective patterns, either biconditional in adolescents or conditional in adults, whereas children in whom analytic processes are not effective remained at a conjunctive level. Thus, the developmental pattern observed in NN conditionals replicated Barrouillet et al. (2008) results. Second, the biconditional interpretation induced by binary terms turned out to be a defective biconditional reading that was predominant in all age groups, except at the ages in which the analytic system does not intervene. In this latter case, binary terms did not have any effect, the interpretation remaining conjunctive. To our knowledge, this developmental pattern had never been reported before.

These results lent strong support to the hypothesis that the binary terms do not affect the heuristic processes responsible for the production of the initial model, but those processes providing the analytic system with material to construct additional models. These additional models do not pertain to the meaning of basic conditionals and, as a consequence, the cases matching them do not make the conditionals true but leave their truth-value indeterminate, hence the defective readings. When the output of the memory search is strongly constrained by the structure of the semantic spaces explored, the developmental differences resulting from differences in efficiency of the analytic system are reduced, and adults tend to adopt the same reading of conditional as adolescents. Although being in line with our predictions, these findings could be restricted to the basic conditionals involving unfamiliar relations between antecedent and consequent. Thus, the following experiment aimed at extending them using causal conditionals.

Causal conditionals

Causal conditionals are conditionals in which the antecedent refers to a cause and the consequent to its effect, such as *If a piece of metal is heated, then it expands*. Using causal conditionals is particularly appropriate to test our hypotheses. Indeed, there is a great deal of evidence that the structure of the general knowledge and the structure of the knowledge related to the causal relation under study have a strong impact on the way it is understood and people reason from it. Causal relations for whom people do not have access to potential alternative causes (i.e., strong causations) tend to elicit a biconditional reading, whereas weak causations for which alternative causes easily come to mind result in more frequent conditional interpretations (Barrouillet, Markovits, & Quinn, 2001; Cummins, 1995; Cummins et al., 1991; Markovits, 1984; Quinn & Markovits, 1998). What is the nature of the initial representation produced by heuristic processes for each type of causal conditionals and how does their interpretation evolve with age?

According to the mental model theory (Goldvarg & Johnson-Laird, 2001), both strong and weak causal conditionals would involve the same single initial representation in which the cause and the effect occur (Johnson-Laird, 2006). Following the principle of truth, naive individuals would represent each of the causal relations using the mental models:

$$p \qquad\qquad q$$
$$\qquad \cdots$$

with the ellipsis representing implicit possibilities in which the antecedent does not hold. Representing these possibilities would require a fleshing out process calling for the analytic system. Thus, the strength of the causal relation should not have any impact on the representation delivered by heuristic processes and $p\ q$ should be the sole case making both weak and strong causal conditionals true. The mental model theory assumes that the retrieval of alternative causes through fleshing out of the initial model results in the construction of possibilities of the form $\neg p\ q$ that lead to the construction of a complete representation with the $p\ q$, $\neg p\ \neg q$, and $\neg p\ q$ models, whereas, when no alternative cause is retrieved, this representation is restricted to the two models $p\ q$ and $\neg p\ \neg q$. It should be noted that the availability of alternative causes in long-term memory has the same effect on the understanding of causal conditionals as the structure of semantic memory explored in Experiment 1 has on the understanding of basic conditionals. Causal relations with no or few alternative causes constrain the construction of additional models as binary terms constrained the fleshing out process for BB conditionals. Thus, the developmental patterns for causal conditionals with either few or many alternative causes should mimic those observed for BB and NN conditionals respectively in Experiment 1. Strong causations should elicit defective biconditional rather than equivalence readings, but only on the ages at which the analytic process is efficient and the conjunctive interpretation overcome, whereas weak causations should lead to the conjunctive – defective biconditional – defective conditional developmental pattern.

Experiment 2

These hypotheses were tested in 40 third graders (mean age = 8.5 years, SD = 0.3, 29 females), 42 sixth graders (mean age = 11.9 years, SD = 0.4, 28 females), 44 ninth graders (mean age = 15.7 years, SD = 0.3, 27 females), and 42 students from the University of Geneva (mean age = 22.8 years, SD = 2.86, 28 females) who performed a truth table task containing both strong and weak causal relations. In a pretest, 15 additional adult participants were presented with 12 causal conditionals and asked to decide for each of them whether many causes could, or not, produce the effect evoked in the statement under study ("Are there many causes that can produce this effect?"). Four conditionals for which all the participants considered that there was no possible alternative cause, and four conditionals for which 14 participants judged that there were many alternative causes were selected as strong and weak causal relations respectively and used in the present experiment (see Appendix).

The procedure was similar as the previous experiment. The eight causal conditionals (four strong and four weak relations) were presented with each of the four logical cases, resulting in 32 trials. Contrary to the previous experiment, we also manipulated the nature of the false antecedent in $\neg p$ cases by representing the absence of the cause evoked in the conditional statement in two different ways. In half of the trials, the absence of the cause was represented along with another event that could be considered as a potential alternative cause of the effect (we named this condition the *other event* condition), whereas in the other half this other event did not occur (we named this condition: *no other event*). For example, for the causal conditional "If the lever 2 is down, then the rabbit's cage is open", the antecedent was represented by a picture of three levers. In the other event condition, the picture represented the lever 2 up but the lever 3 down, whereas in the no other event condition, none of the three levers was down (Fig. 2). This was done because, in the real life, a putative cause can be absent but another phenomenon can occur that can affect our analysis of the phenomenon. In the above example of the levers, the mere absence of the cause (no other event condition) in presence of the effect can lead to a "one can not know" response (participants thinking "if lever 2 were down, the door might be open too"), whereas observing lever 3 down and the door open could lead participants to think: "the rule is false, it is not lever 2 but lever 3 that makes the door opening". In each age group, half of the four weak and of the four strong causal conditionals were presented in the other event condition for $\neg p$ cases and the other half in the no other event condition, the nature of the causal rules affected to each condition being counterbalanced across participants. As in the previous experiment, the conditional statement was displayed on screen, followed by pictures representing successively p
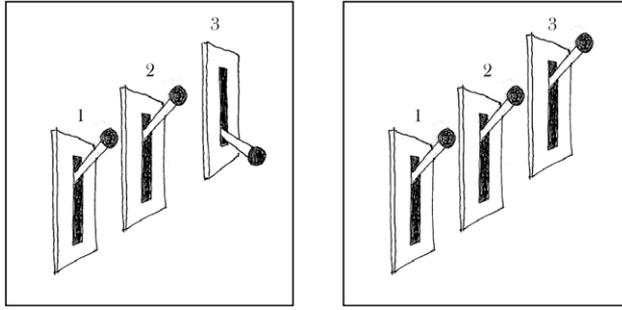
Fig. 2. Example of pictures representing the *other event* condition (in the left panel) and the *no other event* condition (in the right panel) for the ¬*p* cases in Experiment 2 (here for the causal conditional "If the lever 2 is down, then the rabbit's cage is open").

or ¬*p* instances, *q* or ¬*q* instances and finally the three possibilities of response, and the participants received the same instructions.

Analysis of responses

The presence or absence of another event along with the absence of the evoked cause had no effect on the interpretations of both strong and weak causal conditionals. The average percentages of conjunctive, defective biconditional, defective conditional, and matching responses across ages and types of causal relations were exactly the same in both conditions, and among the 40 comparisons permitted by 5 interpretations × 4 ages × 2 types of relations, the largest difference observed was 5% for defective conditional responses given by adults on strong causal relations. Thus, the analyses combined responses from the two conditions. As we expected, the *p q* cases elicited a high rate of "true" responses from 97% to 99% that did not vary with age (Table 3) whereas *p* ¬*q* cases mainly elicited "false" responses. As for NN and BB conditionals, their rate increased with age, $F(3, 164) = 14.96$, $p < .001$, $\eta_p^2 = .17$, an effect explained by the production of matching responses in third grade. As we predicted, the rate of indeterminate responses on ¬*p* ¬*q* cases increased with age for both weak, $F(3, 164) = 47.76$, $p < .001$, $\eta_p^2 = .47$, and strong relations, $F(3, 164) = 48.75$, $p < .001$, $\eta_p^2 = .41$. In both conditions, post hoc comparisons (Newman–Keuls test) revealed significant increases from grade 3 to 6, and 6 to 9 ($ps < .01$), but not from grade 9 to adults. There was no effect of the type of relation and no interaction with age, $Fs < 1$. By contrast, and in line with our predictions, the increase with age in indeterminate responses on ¬*p q* was delayed for weak relations, $F(3, 164) = 14.32$, $p < .001$, $\eta_p^2 = .21$, failing to reach significance between grades 6 and 9 ($p = .08$) and being significant only be-

Table 3
Percent of responses "true" (T), "one cannot know" (I) and "false" (F) for each type of cases as a function of grades and type of causal relations, either weak or strong, in Experiment 2.

Grades	Causal	Cases											
		p q			¬*p* ¬*q*			¬*p q*			*p* ¬*q*		
		T	I	F	T	I	F	T	I	F	T	I	F
3	Weak	98	1	1	1	16	83	0	24	76	0	24	76
n = 40	Strong	99	0	1	4	16	80	0	28	72	2	27	71
6	Weak	99	1	0	3	55	42	0	14	86	0	6	94
n = 42	Strong	99	0	1	1	66	33	1	11	88	2	7	91
9	Weak	98	2	0	1	89	10	1	34	65	1	0	99
n = 44	Strong	97	1	2	2	89	9	0	27	73	0	2	98
Adults	Weak	99	1	0	0	95	5	1	70	29	0	1	99
n = 42	Strong	98	0	2	0	94	6	2	33	65	1	1	98

tween grade 9 and adults ($p < .001$). This age-related effect was strongly reduced for the strong relations, $F(3, 164) = 2.38$, $p = .07$, $\eta_p^2 = .04$, resulting in a significant age × type of relation interaction, $F(3, 164) = 7.17$, $p < .001$, $\eta_p^2 = .11$. As in Experiment 1, this interaction was no longer significant when adults were discarded from the analysis.

Response patterns analysis

We categorized the response patterns for the eight causal conditionals on the basis of the four interpretations already observed in Experiment 1, that is conjunctive, defective biconditional, defective conditional and matching interpretations (Fig. 3). The matching pattern was produced only by younger participants (third graders), explaining the age-related increase in "false" responses on the $p \ \neg q$ case. First of all, as we predicted, conjunctive response patterns predominated in younger
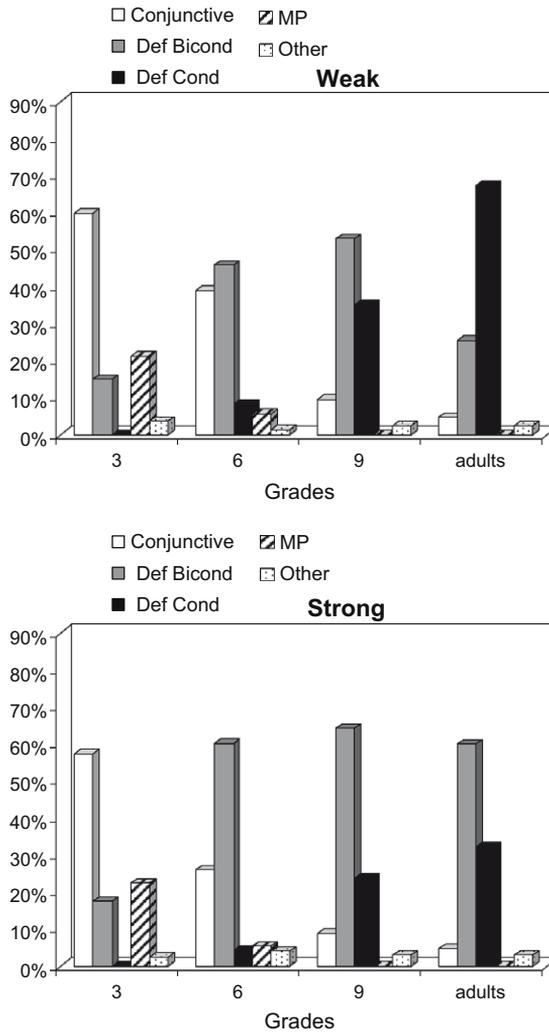
Fig. 3. Percent of response patterns categorized as conjunctive, defective biconditional (Def Bicond), defective conditional (Def Cond), matching (MP), and others as a function of grades for strong and weak causal conditionals in Experiment 2.

participants and tended to disappear with age for weak and for strong causal relations with significant linear trends in both cases, $F(1, 164) = 60.07$, $p < .001$, $\eta_p^2 = .27$, and $F(1, 164) = 47.19$, $p < .001$, $\eta_p^2 = .22$, respectively. In accordance with our main prediction, the defective biconditional responses were more frequent for strong than for weak relations (51% and 35% respectively), and their evolution with age differed from one type of relation to the other. As we predicted, for weak causal relations, defective biconditional interpretations represented an intermediary developmental step revealed by a significant quadratic trend, $F(1, 164) = 20.99$, $p < .001$, $\eta_p^2 = .11$, whereas for strong relations, these interpretations increased from third to sixth grade, $F(1, 164) = 37.70$, $p < .001$, $\eta_p^2 = .19$, and were no longer affected by age from sixth grade to adulthood, $F < 1$. The frequency of defective conditional interpretations steadily increased with age for weak relations, $F(1, 164) = 83.60$, $p < .001$, $\eta_p^2 = .3$, for the linear trend. In accordance with the hypothesis that few alternative causes block the construction of the $\neg p\ q$ model, defective conditional interpretations were rarer for strong than for weak causal relations (15% and 28% respectively), $F(1, 164) = 19.5$, $p < .001$, $\eta_p^2 = .13$.

Finally, as for BB and NN, though the four strong causal relations we used elicited comparable rates of defective biconditional readings (50%, 51%, 51%, and 48% respectively), their effect was not universal and some participants treated them as weak relations. In this case, the standard developmental trend was observed with an evolution from conjunctive to defective biconditional and then defective conditional interpretations.

Discussion

The present experiment revealed a striking developmental parallel between strong causal relations and BB conditionals on the one hand, and between weak causal relations and NN conditionals on the other. For both basic and causal conditionals, heuristic processes deliver a single-model initial representation and the development of the analytic system leads to conjunctive, defective biconditional and then defective conditional interpretations from childhood to adulthood. However, for both types of conditional, the heuristic system has still an impact on the output of this analytic system and the structure of knowledge available from long-term memory can modify this developmental trend and abolish the developmental differences existing between adolescents and adults.

These facts suggest that, though causal conditionals have been described as eliciting biconditional interpretations (Lecas & Barrouillet, 1999; Marcus & Rips, 1979), actually they do not differ from basic conditionals and, when a biconditional reading is adopted, it corresponds to a defective biconditional.[4] This suggests in turn that the biconditional interpretations of causal conditionals often adopted by adolescents do not result from implicatures and pragmatic principles. If this was the case, both $p\ q$ and $\neg p\ \neg q$ possibilities would be produced by the heuristic system and would appear as pertaining to the basic meaning of the conditional, making it psychologically true when they occur. Actually they do not differ from the biconditional interpretations resulting from strong causations or BB conditionals. Thus, defective biconditional interpretations are better described as resulting from failures in the fleshing out process due either to a limitation in the efficiency of the analytic processes or to the structure of knowledge from which additional models are constructed, a phenomenon described as modulation by Johnson-Laird and Byrne (2002).

Though basic and causal conditionals have the same initial model and reveal the same developmental trends, this does not mean that all the conditionals evolve in the same way. Assuming that the structure of the initial model remains stable across age does not mean that it is constituted by the sole $p\ q$ model for

---

[4] It should be noted that Newstead, Ellis, Evans, and Dennis (1997) reported a significant rate of interpretations of equivalence for causal conditionals in which $p\ q$ and $\neg p\ \neg q$ cases are considered as making the conditional true and the other two cases as making it false, something we never observed in the present study. However, differences in the experimental designs could explain this discrepancy in results between the two studies. First, Newstead et al. (1997) used verbal instead of pictorial material for the four cases under study, the $\neg p$ cases being verbal descriptions involving explicit negations. Schroyens (1997) has observed that verbal descriptions for the $\neg p\ \neg q$ case often lead to "true" instead of indeterminate responses. Second, causal conditionals were presented along with many other types of conditionals, some of them, such as promises and threats, eliciting very frequent equivalence readings as we will see in the following experiment. Thus, contamination effects can not be excluded. Third, Newstead et al. (1997) presented the four cases simultaneously. It is possible that this procedure prompts a search for coherence in participants who avoid to give different responses to cases that they judge as compatible with the sentence.

all kinds of conditional. Evans (2006) suggests that the mental models he refers to can include added pragmatic implicatures that would require more than one semantic type of mental model to capture. The developmental implications of this phenomenon are addressed in the following section.

Inducements

Our theory assumes that those states of affairs that make a sentence true are represented in its initial model, which construction depends on heuristic processes that are not sensitive to individual or developmental differences. A direct consequence of these proposals is that there should be no difference between younger and older children in the number of models that can be produced by the heuristic system. Thus, whereas an inefficient analytic system limits young children to a one-model representation for the basic and the causal conditionals, this limitation might be overcome for conditionals eliciting multi-model initial representations. This seems to be the case of promises and threats. Indeed, Newstead et al. (1997) had adult participants evaluating several types of conditionals with realistic material and observed that promises and threats elicited equivalence readings in which both $p\,q$ and $\neg p\,\neg q$ cases are considered as making these conditionals true, whereas $\neg p\,q$ and $p\,\neg q$ cases make them false.

According to our theory, this suggests that promises and threats involve a two-model initial representation. As suggested by Evans and Over (2004), in the case of inducements, this could result from pragmatic implicatures that add the inverse conditional *If not p then not q*. Indeed, the goal of conditional inducements is to persuade a person to do, or refrain from doing, some action and, as noted by Ohm and Thompson (2006), inferences are inherent to the meaning of these statements (see also Evans & Over, 2004; Fillenbaum, 1975; Fillenbaum, 1976). For example, a promise like "if you mow the lawn, then I'll give you five euros" will only be effective if the listener understands that the promised reward will not be forthcoming if the action is not performed. Thus, the pragmatic implicature "if you don't, I'll give you nothing" is not optional but inherent to the core meaning of this kind of speech act and should be represented in the initial model, resulting in the following two-model representation:

Lawn mown   5 euros
Lawn not mown   nothing

This initial representation would block in turn any further construction of models by the analytic system. The original promise "if you mow the lawn, then I'll give you five euros" leads to dismiss the $p\,\neg q$ possibility of mowing the lawn without receiving any reward, while the implicature "if you do not mow the lawn, then I wont give you five euros" dismisses the possibility to be rewarded without mowing the lawn that would contradict it. Thus, any intervention of the analytic system would remain fruitless. The same applies to threats like "If you break the vase, then I'll take your ball away". This speech act is intended to refrain the action but would loose any effectiveness if it was not clear that a good way to avoid the threatened outcome is to not break the vase, which invites the pragmatic implicature "If you don't, I don't take your ball away". Thus, a two-model initial representation is needed to convey the core meaning of inducements.

Our theoretical framework leads to two main predictions. First, the two-model initial representation of promises and threats should elicit an equivalence reading (i.e., true, true, false, false for $p\,q$, $\neg p$ $\neg q$, $\neg p\,q$, and $p\,\neg q$ respectively) that was never observed with basic or causal conditionals that have a one-model initial representation. Second, this reading should be predominant whatever the age of the participants from childhood to adulthood.

Experiment 3

Twenty third graders (mean age = 8.9 years, SD = 0.7, 11 females), 22 sixth graders (mean age = 12.1 years, SD = 0.3, 14 females), 19 ninth graders (mean age = 15.3 years, SD = 0.5, 10 females), and 27 students from the University of Geneva (mean age = 21.2 years, SD = 0.6, 18 females) performed a truth table task containing both promise and threat conditionals. As in Experiment 2,

Table 4
Percent of responses "true" (T), "one cannot know" (I) and "false" (F) for each type of cases as a function of grades and type of conditionals, either promise or threat, in Experiment 3.

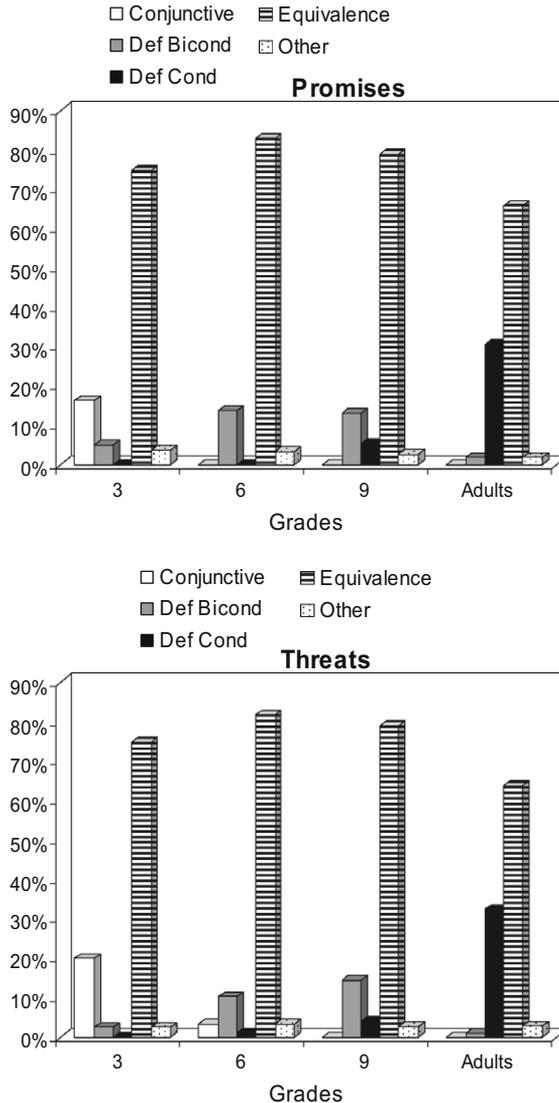| Grades	Conditionals	Cases											
		$p\ q$			$\neg p\ \neg q$			$\neg p\ q$			$p\ \neg q$		
		T	I	F	T	I	F	T	I	F	T	I	F
3	Promise	100	0	0	78	5	17	1	1	98	2	3	95
$n = 20$	Threat	100	0	0	76	4	20	3	1	96	1	1	98
6	Promise	97	2	1	85	14	1	1	2	97	1	1	98
$n = 22$	Threat	99	0	1	84	11	5	1	2	97	1	1	98
9	Promise	99	1	0	80	20	0	0	5	95	0	3	97
$n = 19$	Threat	99	0	1	80	20	0	1	4	95	0	0	100
Adults	Promise	99	0	1	71	29	0	0	32	68	0	1	99
$n = 27$	Threat	99	1	0	69	31	0	1	33	66	0	1	99

experimental sentences were selected from a pretest. An additional group of 15 adult participants studied six promises, whereas another group of the same size studied six threats. They were asked to judge the nature of each sentence on a 7-point scale with 7 corresponding to "This sentence perfectly conveys a promise (a threat)" and 1 to "This sentence does not at all convey a promise (a threat)". The four promises used in this experiment elicited a mean score of 6.1 and for the four threats a mean score of 6.4 (see Appendix).

The material was constructed and presented in the same way and the instructions given to participants were the same as in the previous experiments. For example, for the promise "If you score a goal then I name you captain", the presented cases took the form of "a ball in the goal – a boy with a captain's shirt" (i.e., *p q*), or "a ball not in the goal – a boy with a normal shirt" (i.e., ¬*p* ¬*q*), or "a ball not in the goal – a boy with a captain's shirt" (i.e., ¬*p q*) or "a ball in the goal – a boy with a normal shirt" (i.e., *p* ¬*q*).

Analysis of responses

We hypothesized that the heuristic system would produce a two-model initial representation (*p q* and ¬*p* ¬*q*) that would block any further fleshing out. Hence, for all grades, we expected "true" responses for *p q* and ¬*p* ¬*q* cases and "false" responses for ¬*p q* and *p* ¬*q* cases. This is what we observed (Table 4). The rate of "true" responses was very high on *p q* cases. In the same way, the ¬*p* ¬*q* cases elicited a high rate of true response for both promises and threats with no effect of grades, $F(3, 84) = 1.35$, $p = .26$, $\eta_p^2 = .03$, no effect of type of conditionals, and no interaction, $Fs < 1$. For the two other cases, ¬*p q* cases elicited a high rate of "false" responses in third, sixth, and ninth graders that contrasted with a lower rate in adults, $F(1, 84) = 43.18$, $p < .001$, $\eta_p^2 = .37$. This lower rate of "false" responses in adults was due to 32% and 33% of *indeterminate* responses for promise and threats respectively in this age group. Concerning *p* ¬*q* cases, the results revealed a high rate of "false" responses in each grade for both promise and threat conditionals. Thus, for the *p q*, ¬*p* ¬*q* and *p* ¬*q* cases, the results corroborated our hypotheses. The effect of grades on ¬*p q* cases was unexpected, but the following response pattern analysis will shed light on this phenomenon.

These responses were categorized on the basis of the interpretations already observed in Experiments 1 and 2 plus the *equivalence* pattern corresponding to true, true, false, and false responses to *p q*, ¬*p* ¬*q*, ¬*p q*, and *p* ¬*q* cases respectively. In line with our hypotheses, we observed for each grade a high rate of *equivalence* responses for both promises and threats (Fig. 4). Neither the effect of grades, $F(3, 84) = 1.22$, $p = .31$, $\eta_p^2 = .01$, of conditionals, nor the interaction, $Fs < 1$, was significant. Apart from these equivalence patterns, participants produced some conjunctive patterns the rate of which decreased with age, defective biconditional patterns that exhibited the developmental qua-

Fig. 4. Percent of response patterns categorized as conjunctive, defective biconditional (Def Bicond), defective conditional (Def Cond), equivalence, and others as a function of grades for promises and threats in Experiment 3.

dratic trend already described with NN conditionals, and defective conditional patterns that increased with age.

As in the previous experiments, it is worth noting that, although the promise and threats did not differ in the rate of equivalence responses they elicited (74%, 78%, 72%, and 74% for the four promises and 76%, 78%, 74%, and 74% for the four threats), the effect of pragmatic implicature was not universal. When it did not occur (i.e., when the interpretation differed from the equivalence reading), the standard developmental trend from a conjunctive, to a defective biconditional and finally a defective conditional interpretation reappeared and, in line with the 33% of indeterminate responses on ¬p q cases, adults produced 30% of defective conditional patterns (Fig. 4).

Discussion

Newstead et al. (1997) observed that adults deem promises and threats true for both *p q* and ¬*p* ¬*q* cases and false for ¬*p q* and *p* ¬*q* cases. Our theory therefore predicted that these conditionals elicit a two-model initial representation leading to a predominant equivalence reading in all the age groups. This is exactly what we observed with more than 75% of equivalence readings in children and adolescents with no significant effect of age. This result strongly contrasts with the noticeable developmental changes that were observed with basic and causal conditionals and lends strong support to the hypothesis that promises and threats elicit a complex initial model underpinned by heuristic processes relatively immune to developmental changes. This initial model blocks in turn any enrichment of the representation by analytic processes, leading to a predominant equivalence response in all the age groups.

The probability of conditionals

The task of evaluating the probability of conditional statements has became most popular in the study of adult reasoning and has provided the main arguments supporting the probabilistic as well as the suppositional theory of conditional (Evans et al., 2003; Oberauer & Wilhelm, 2003). According to Evans' suppositional theory (Evans & Over, 2004; Evans et al., 2005; Evans, 2007b), when evaluating an *If p then q* conditional statement, people would use a procedure known as the Ramsey test by which they hypothetically add *p* to their stock of knowledge and would argue on that basis about *q*. A psychological consequence of this procedure is that people's degree of belief on a conditional should equate the conditional probability of *q given p*, $P(q/p)$. Indeed, adding *p* to their stock of knowledge, people disregard the *not p* cases that are mainly judged as irrelevant to the truth of the conditional. Instead, they focus on the *p* possibilities and compare the probability of the *p q* cases with that of the *p* ¬*q* cases. To the extent that *p q* is judged more probable than *p* ¬*q*, the conditional probability of *q given p* is high and a high probability is assigned to the conditional *if p then q*. It is worth to note that evaluating the probability of the conditional as the conditional probability departs from the prescriptions of formal logic, which favors an interpretation of the conditional in terms of material implication according to which conditionals are true for *p q*, ¬*p q*, and ¬*p* ¬*q* and false for *p* ¬*q*. The probability for a conditional to be true would thus equate $P(p\ q) + P(\neg p\ q) + P(\neg p\ \neg q)$.

Evans et al. (2003) tested the prediction issued from the suppositional theory by giving adult participants frequency information about the cases *p q*, *p* ¬*q*, ¬*p q*, and ¬*p* ¬*q* in a pack containing cards that were either yellow or red and have either a circle or a diamond printed on them (say 1 yellow circle, 4 yellow diamonds, 16 red circles, and 16 red diamonds). The participants were then asked to evaluate the probability that a claim like "If the card is yellow then it has a circle printed on it" is true for a card drawn at random from the pack. As it was predicted, most of the participants responded by giving the conditional probability $P(q/p)$ that is equal to $P(p\ q)/[P(p\ q) + P(p\ \neg q)]$. In other words, they evaluate the probability to find a card with a circle among the yellow cards. Unexpectedly, a substantial minority of participants responded by giving the conjunctive probability $P(p\ q)$, which is the proportion of yellow cards with a circle among the entire pack. Evans et al. (2003) assumed that this latter pattern of response is given by shallow processors who cut short the Ramsey test and stop at the *p q* cases, an assumption corroborated by the fact that the participants giving conjunctive responses to the probability task have lower cognitive capacities than the others (Evans, Handley, Neilens, & Over, 2007).

Oberauer and Wilhelm (2003) reported approximately the same results that have been extended to causal conditionals (Over, Hadjchristidis, Evans, Handley, & Sloman, 2007). The evaluation of the probability of conditionals as a conditional probability, along with the defective table observed in truth table tasks, was considered by Evans et al. (2005) as the main evidence contradicting Johnson-Laird and Byrne's mental model theory. Indeed, according to Evans et al. (2003), the mental model theory proposes a material implication account of basic conditionals (see above) and thus would predict the probability $P(p\ q) + P(\neg p\ q) + P(\neg p\ \neg q)$. However, this kind of response is practically never observed in the evaluation of probability task.

There are 8 cards
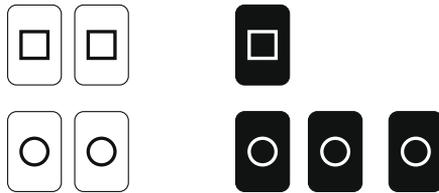
How likely is the following statement to be **true** of a card drawn at random from the pack?

If the card is black then there is a square printed on it

The probability that the statement is true is ☐ out of ☐

Fig. 5. Example of material given to participants in the probability task.

How does our theory account for the way people evaluate the probability of conditional statements and what are its developmental predictions? Our hypothesis is that people evaluate the probability of a given conditional statement from the mental models they have constructed by focusing on those cases that are relevant for the truth or falsity of this conditional (i.e., that make it either true or false). The probability that the conditional is either true or false would be given by the ratio between those cases that make the conditional either true or false and the relevant cases. We have seen that, for adults, the relevant cases are most often the $p\,q$ and $p\,\neg q$ cases whereas the $\neg p$ cases are irrelevant, leaving indeterminate the truth value of the conditional. Thus, the probability of a basic conditional of being true or false should be $P(p\,q)/[P(p\,q) + P(p\,\neg q)]$ and $P(p\,\neg q)/[P(p\,q) + P(p\,\neg q)]$ respectively, as Evans et al. (2003) observed. For sake of simplicity, we will call this response the defective conditional response in reference to the corresponding interpretation in the truth table task. For example, with the pack of cards represented in Fig. 5 and the conditional "If the card is black, then there is a square printed on it", the defective conditional response for true is 1/4 and 3/4 for false. However, what would be these evaluations at the other developmental levels? The different levels on the interpretation of conditionals previously identified permit critical predictions. For children and adolescents who favor a conjunctive interpretation, there is no irrelevant case: $p\,q$ cases make the conditional true whereas the other cases make it false. Thus, the probability that the conditional is true should be $P(p\,q)$ (i.e., 1/8 in the example of Fig. 5) whereas the probability that it is false should be $P(p\,\neg q) + P(\neg p\,q) + P(\neg p\,\neg q)$, that is 7/8 in the example. More interesting is the intermediary level of interpretation we observed above, which involves a defective biconditional reading of conditionals. Within this interpretation, $p\,q$ is the sole case making the conditional true, $p\,\neg q$ and $\neg p\,q$ cases make it false, $\neg p\,\neg q$ being the sole case deemed as irrelevant. As a consequence, at this developmental level, the relevant cases are $p\,q$, $p\,\neg q$, and $\neg p\,q$. Thus, the probability for "true" should equate $P(p\,q)/[P(p\,q) + P(p\,\neg q) + P(\neg p\,q)]$, whereas the probability for "false" should be $[P(p\,\neg q) + P(\neg p\,q)]/[P(p\,q) + P(p\,\neg q) + P(\neg p\,q)]$. In our example, this leads to a probability of 1/6 for true and 5/6 for false. We will consider this response as a defective biconditional response. Thus our theory predicts that evaluating the probability of the conditionals as the conditional probability should be a developmental achievement related to the defective conditional interpretation observed in adults. This level should be preceded by different evaluations related to the incomplete interpretations resulting from immature analytic processes, as we observed in the truth table task.

How could the frequent conjunctive responses observed in adults be explained within this theoretical framework? As Evans et al. (2003), we do not believe that more than 40% of adults really have a conjunctive understanding of basic conditionals, and it can be noted that we practically never observed such a response pattern in the previous experiments, nor Evans et al. (2007) with causal conditionals. However, we do not endorse Evans' explanation of an incomplete Ramsey test, our theory providing a fairly simple account of this phenomenon. Adults have the capacities to go beyond a

one-model conjunctive interpretation when solving the truth table task, but evaluating the probability of a statement whose meaning conveys itself uncertainty is probably experienced as a very difficult, if not awkward, task by many adults. It is possible that, puzzled by the requirements of the task, those participants do not go beyond the heuristic step in solving the problem. Thus, they would only represent the *p q* case and give the conjunctive response observed by Evans et al. (2003) and Oberauer and Wilhelm (2003). Though resulting in the same response, this process probably differs from the conjunctive response given by children and young adolescents in the truth table task. Indeed, it is fairly possible that some of these children and adolescents detect the need to go beyond the heuristic processes, but do not have the capacities to carry out analytic processes. By contrast, many adults have probably the mindware necessary to run analytic processes, but they do not detect the need for further analytic processes in this precise task. This illustrates the fact that poorly elaborated responses such as conjunctive responses can result from an inadequate representation (e.g., children's conjunctive responses in the truth table task), but also from inadequate or incomplete processing (e.g., adults' conjunctive responses in the probability task). Thus, the conjunctive responses triggered by the exclusive reliance on heuristic processes when evaluating the probability of conditionals should be observed at all ages and should increase the rate of conjunctive responses underpinned by the conjunctive interpretations that were already observed in the truth table tasks.

Though the probability task should elicit a high rate of conjunctive responses, we should observe a developmental trend corresponding to the three developmental levels described with the truth table task. A first developmental level should be characterized by a high rate of conjunctive responses that should progressively decrease, with defective biconditional responses as an intermediary level before the predominance of the conditional defective responses observed by Evans et al. (2003) in adults. These predictions were tested in the following experiment.

Experiment 4

Twenty sixth graders (mean age = 11.9 years, SD = 0.2, 12 females), 20 ninth graders (mean age = 15.3 years, SD = 0.3, 13 females), and 23 students from the University of Geneva (mean age = 25.2 years, SD = 1.72, 19 females) performed the probability task. The task was also proposed to third graders, but they usually failed to grasp the notion of probability. Each participant completed a booklet containing nine pages: one for general instructions and one for each of eight problems. The procedure was adapted from Evans et al. (2003) to adolescents. The problems concerned a pack of cards either black or white, with either a circle or a square printed on them. The total number of cards varied from 6 to 9, the number of cards of each type varying from 1 to 3 in such a way that two different interpretations could not give rise to the same responses. For each problem, the entire pack was represented on a page of the booklet followed by a question of the form: "How likely is the following statement to be true of a card drawn at random from the pack? If the card is black then there is a square printed on it" (Fig. 5). Half of the problems were presented with the "true" question and the other half with the "false" question, "true" and "false" problems being counterbalanced across participants. Eight different booklets were created in which the order of presentation of the eight problems was randomly varied as well as the order of presentation of the four types of cards within each problem. In order to avoid complex calculations and to make clear the notion of probability, participants were invited to give their response by filling sentences of the following form with two numbers:

Table 5
Probabilities of true and false conditionals for the different interpretations of basic conditionals.

Interpretation	Probability	
	True	False
Conjunctive	$P(p\,q)$	$P(p\,\neg q) + P(\neg p\,q) + P(\neg p\,\neg q)$
Defective biconditional	$P(p\,q)/[P(p\,q) + P(p\,\neg q) + P(\neg p\,q)]$	$[P(p\,\neg q) + P(\neg p\,q)]/[P(p\,q) + P(p\,\neg q) + P(\neg p\,q)]$
Defective conditional	$P(p\,q)/[P(p\,q) + P(p\,\neg q)]$	$P(p\,\neg q)/[P(p\,q) + P(p\,\neg q)]$
Material implication	$P(p\,q) + P(\neg p\,q) + P(\neg p\,\neg q)$	$P(p\,\neg q)$

Table 6
Percent of type of responses in each age group in evaluating the probability that the conditional is either true or false in Experiment 4.

Type of problem	Age group					
	Sixth grade *n* = 20		ninth grade *n* = 20		Adults *n* = 23	
	True	False	True	False	True	False
Type of response						
Conjunctive	79	81	38	35	39	29
Defective Biconditional	9	8	35	35	0	0
Defective Conditional	0	0	25	24	58	62
Material Implication	0	0	0	0	0	9
Other	13	11	3	6	3	0

"the probability that the statement is true is ... out of ...". Before the experiment, the participants were informed that they had eight problems to solve and received instructions about the fact that each pack contained always four different kinds of cards.

Analysis of responses

The different responses taken into account in the following analyses are given in Table 5. We added to the three types of responses expected (i.e., conjunctive, defective biconditional, and defective conditional) the response that should result from a material implication interpretation. Almost all the responses given by the participants (94%) corresponded to one of these four types with negligible differences between "true" and "false" problems (Table 6). As we predicted, the conjunctive responses predominated in younger participants and then decreased with age, $F(2, 60) = 15.42$, $p < .01$, $\eta_p^2 = .14$, with a leveling between the two older groups. The responses reflecting a defective biconditional interpretation followed the predicted developmental trend, increasing from sixth to ninth grade, $F(1, 38) = 7.87$, $p < .05$, $\eta_p^2 = .13$, before disappearing in adults who favored responses corresponding to the conditional probability. These latter responses that were never observed in sixth graders appeared in ninth grade, their rate significantly increasing toward adulthood, $F(1, 41) = 6.27$, $p < .05$, $\eta_p^2 = .13$.

Those participants who were consistent on at least 6 out of 8 responses on a given type of interpretation were considered as being consistent on this interpretation. The results were particularly clear. In sixth graders, 17 out of 20 children were consistent, 15 giving the conjunctive probability and 2 a defective biconditional probability. In ninth graders, 19 out of 20 participants were consistent, with 7 participants producing conjunctive responses, 7 other producing defective biconditional responses and 5 participants giving the conditional probability. Finally, 20 out of the 23 adult participants were consistent, 7 giving the conjunctive probability and 13 the conditional probability (Fig. 6). Two additional adults gave a meaningful pattern of response consisting in *P* (*p q*) for the probability of a true conditional and *P* (*p ¬q*) for the probability of a false conditional, considering the same true and false cases as in the defective conditional pattern but evaluating the probabilities against the entire set of cards and not only on *p* cards.

Discussion

Our theory easily accounts for the facts revealed by the probability task. Assuming that participants evaluate the probability of a conditional to be true or false from the mental models they construct as the ratio between the favorable and the relevant cases, the three developmental levels identified by our modified mental model theory lead to three successive levels of responses. The conditional probability given by a majority of adults appears as a developmental achievement resulting from the defective conditional interpretation in which the ¬p cases are considered as irrelevant and thus disregarded. However, the nature of the cases considered as irrelevant evolves with age, and so evolves the evaluation of the probability of conditional statements, the defective conditional response being preceded, as we expected, by a conjunctive and then a defective biconditional response. As it
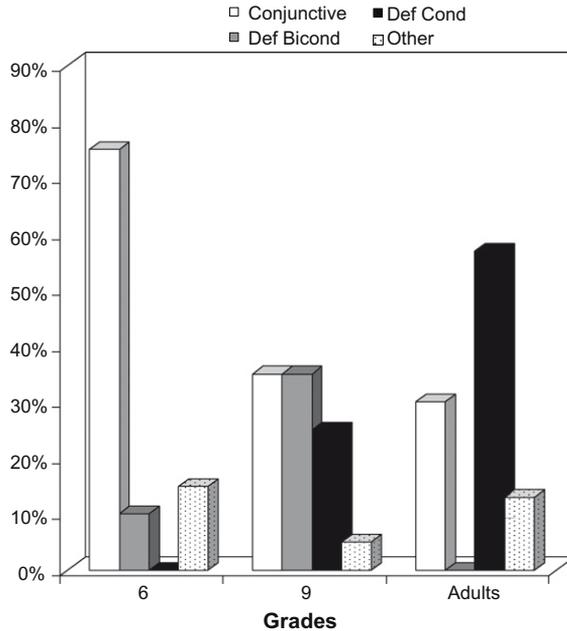
Fig. 6. Percent of response patterns categorized as conjunctive, defective biconditional (Def Bicond), and defective conditional (Def Cond) responses to the probability task in Experiment 4.

could be expected from previous studies (Evans et al., 2003; Oberauer & Wilhelm, 2003), conjunctive responses were very frequent, even in adults. Our interpretation is that the difficulty of the task leads many participants to base their evaluation on the sole initial model provided by heuristic processes. As a consequence, it can be observed that the developmental trend resulting from the intervention of the analytic system is delayed in the probability task, with sixth graders producing almost 80% of conjunctive responses, a rate never observed with the truth table task in the present study or the inference task in previous researches (e.g., Barrouillet et al., 2000).

Though our theoretical account of the evaluation of the probability of conditionals is akin to Evans' conception, the two proposals differ in important aspects. According to Evans, the meaning of the conditional is suppositional in nature. This means that understanding or evaluating a conditional involves the construction of mental models of the antecedent and, on that basis, the assessment of the believability of the consequent through the Ramsey test (Evans et al., 2003). In this account, the irrelevance of ¬p cases is a consequence of the suppositional nature of the conditional: it is because the Ramsey test focuses on p possibilities that ¬p cases are disregarded as irrelevant. Our account is different because the developmental approach makes clear that the irrelevance of ¬p cases is a gradual construction. Importantly, some ¬p cases such as ¬p ¬q can be considered as irrelevant while others (i.e., ¬p ¬q) are still considered as falsifying the conditional. Thus, understanding the irrelevance of ¬p developmentally precedes the suppositional level of interpretation, reversing the causal relation between the suppositional nature of the conditional and the irrelevance of false-antecedent cases postulated by Evans. In our theory, the main fact is the developmental increase in the number of models constructed through fleshing out, ¬p cases being considered as irrelevant as far as the corresponding mental models can be constructed through analytic processes. Thus, in our approach, the suppositional nature of the conditional is not the cause but the consequence of the irrelevance of ¬p cases, which results in turn from the complete fleshing out of the initial representation. We do not deny that most adults use a procedure like the Ramsey test to assess conditionals, but we regard this procedure as resulting from the process of constructing a complete set of mental models rather than as the basis of the meaning of conditionals.

Moreover, the two theories differ in their account of the frequent conjunctive responses in adults. According to Evans' proposal, these responses result from an incomplete Ramsey test in which shallow processors cut short the test. They hypothetically add p to their stock of knowledge and focus on p cases, but would stop at the $p\ q$ cases, disregarding $p\ \neg q$ cases. This results in evaluating the probability of the conditional as equivalent to $P\ (p\ q)$. This theory accounts very well for the conjunctive responses on the "true" problems, but what does this theory predict for false problems? In these problems, shallow processors should engage in the Ramsey test by taking $p\ \neg q$ cases into account and stop the process at this level, disregarding $p\ q$ cases. Thus, they should give $P\ (p\ \neg q)$ as the probability for a conditional to be false in the same way as they give $P\ (p\ q)$ as the probability for a conditional to be true. Indeed, we observed this pattern in two adults, but it does not correspond to the frequent conjunctive response in which the probability for false is $P\ (p\ \neg q) + P\ (\neg p\ q) + P\ (\neg p\ \neg q)$. Thus, by assuming that the Ramsey test is the procedure by which people understand and evaluate conditionals, the suppositional theory fails to account for very frequent responses such as the conjunctive interpretations observed in all the experiments so far conducted.

By contrast, by assuming that the conjunctive responses result from a reasoning restricted to the mental model constructed by the heuristic processes, our theory explains why participants consider the entire pack of the card as relevant instead of the sole p cases: it is because analytic processes are needed to consider $\neg p$ cases as irrelevant, and when these analytic processes do not intervene, $\neg p$ cases become relevant. As a consequence, when evaluating the probability of the conditional to be true, $p\ q$ cases are compared to the entire pack of the cards, and when evaluating the probability of the conditional to be false, all the cases except $p\ q$ are considered as making the conditional false because the corresponding mental models have not been constructed. Thus, our theory accounts for the fact that adults evaluate the probability of a conditional as the conditional probability, it makes developmental predictions that were verified, and it gives a more convincing account of the different responses observed in adults.

General discussion

The present study tested the developmental implications of a theory assuming that two different systems concur to the construction of the mental models by which individuals represent the meaning of conditional sentences, the models produced by each system having different epistemic status. A first tacit, automatic, and largely unconscious system would provide individuals with an initial representation that constitutes the psychological meaning of the statement. Thus, any state of affairs matching this representation is considered as making the conditional true. Because this heuristic system does not strongly evolve with age, the initial model and, as a consequence, those cases considered as making the conditional true should remain quite stable from childhood to adulthood. However, the structure of this initial model can vary from one kind of conditional statement to another because the heuristic system involves pragmatic implicatures that can enrich the $p\ q$ representation with other models. A second system, which is optional in nature and based on analytic processes, can intervene to enrich the initial representation with additional models that are not part of the basic meaning of the conditional sentence but are compatible with it. As a consequence, those states of affairs matching these additional models are not considered as making the conditional true, because they do not come spontaneously to mind, neither false because they are compatible with it. Because this analytic system involves controlled and attention demanding processes, its efficiency increases with age. Thus, those cases considered as leaving indeterminate the truth value of the conditional should evolve with age.

Predictions issued from this theory were tested by investigating and comparing how the evaluation of several types of conditionals (basic, causal, and inducements) develops with age and how these interpretations determine the evaluation of the probability of basic conditionals. Overall, the results confirmed our proposals. The cases considered as making the conditionals true can vary from one type of conditional to another, but they remain remarkably constant across age. By contrast, the number of cases considered as leaving the truth value of the conditionals indeterminate increases with age. Interestingly, and in line with Evans (2006) who assumes that the two systems are interdependent because the heuristic processes supply content continuously to consciousness for analytic processing, the output of the analytic system systematically varies in a predictable way as a function of the structure of knowledge used to construct additional models. The present study illustrates how the interplay between the two systems

leads to a great variety of developmental trajectories, which goes far beyond what has already been reported in the literature concerning the development of conditional reasoning. In some cases, for example with basic or weak causal conditionals, the development appears as a steady evolution from childhood to adulthood. In other cases, this evolution is either interrupted in the early adolescence, for example with basic conditionals involving binary terms or strong causal relations, or totally disappears, children and adults exhibiting the same response patterns. With complex tasks like the evaluation of probability, analytic processes seem more difficult to implement and development is delayed.

Moreover, it is interesting to note that the developmental trajectories we observed cannot be reduced to a universal succession of stages that different contents to be processed would interrupt and freeze at different levels. We have seen, for example, that the predominant interpretation for promises and threats adopted by children and adults does not correspond to any of the developmental steps observed with basic or causal conditionals. These new findings have developmental and theoretical implications that we will discuss in turn.

Mental models and the development of conditional

As far as development is concerned, the truth evaluation task used in this study reveals a complex set of developmental trajectories that goes beyond Markovits and Barrouillet's (2002) account. Of course, these authors assumed that the development differed from one kind of conditional to the other, but they presented these differences as being mainly due to the familiarity of the relation between the antecedent and the consequent. Familiarity was thought to facilitate the retrieval of different classes of objects or events represented in mental models. For example, reasoning from class-based conditionals (e.g., "If an animal is a cow, then it has four legs") would facilitate the retrieval in young children of elements from the complementary class (animals that are not cows and do not have four legs) and the "alternatives" class (animals that are not cows, but that have four legs). However, Markovits and Barrouillet (2002) implicitly assumed that all the models constructed have essentially the same epistemic status. They did not anticipate that elements from the complementary class (i.e., $\neg p \ \neg q$ cases) could differ in status from one kind of conditional to the other, some of them confirming the conditional and making it psychologically true (for promises and threats) whereas others are only compatible with the conditional and leave its truth-value indeterminate (for causal conditionals for example), leading to two different forms of biconditional reading (either equivalence or defective biconditional).

According to our theory, these different epistemic statuses would result from two types of retrieval from memory, either automatic or controlled, that Markovits and Barrouillet (2002) did not distinguish. The present results confirmed this point of view. Consider for example the $\neg p \ \neg q$ cases. The previous versions of our developmental approach (Barrouillet & Lecas, 1998; Barrouillet et al., 2000) assumed that the model representing $\neg p \ \neg q$ possibilities is issued from a demanding and time-consuming fleshing out process. Accordingly, there were important developmental differences in the responses elicited by $\neg p \ \neg q$ cases for causal or basic conditionals, suggesting the intervention of a controlled, explicit, and demanding retrieval process through the analytic system. However, there were no age-related differences in the way children, adolescents and adults processed these cases for promises and threats, suggesting an automatic retrieval process immune to strong developmental changes. In line with our assumptions, promises and threats were deemed true for these cases, suggesting that they are represented in the initial model constructed by the heuristic system.

Assuming that the initial model is produced by heuristic processes immune to strong developmental changes led us to predict that the possibilities matching this initial model would be judged to make the conditional true and that the cases eliciting these "true" responses would remain unchanged from childhood to adulthood, whatever the conditional under study. In line with our previous assumptions (Barrouillet & Lecas, 1998; Markovits & Barrouillet, 2002), the development of the interpretation of conditional would mainly be due to an increased capacity to construct additional models through fleshing out, a time-consuming and demanding process that involves the analytic system. However, it must be acknowledged that we do not have direct evidence that the initial model is provided by heuristic processes while the additional models are constructed by the analytic system. Nonetheless, among the methodologies listed by Evans (2007a) that theorists use to provide evidence for dual-pro-

cessing accounts, response latencies and individual differences can provide some support to our assumptions. As stressed by Evans (2007a), all dual-process theorists agree that heuristic processes are faster than analytic processes. Accordingly, when studying the time course of conditional inferences, Barrouillet et al. (2000) observed that the inferences supported by the initial model *p q* (i.e., the endorsement of Modus Ponens and Affirmation of Consequent) are much faster than the inferences that require a fleshing out process and the construction of additional models, findings that were replicated by Grosset, Barrouillet, and Markovits (2005) with causal conditionals. Concerning individual differences, most dual-process theories assume that analytic processes are loaded on cognitive capacities whereas heuristic processes are assumed to be independent of cognitive ability. Accordingly, Barrouillet and Lecas (1999) observed in a truth table task involving the construction of cases compatible with basic conditionals that working memory capacities are a very good predictor of children's and adolescents' capacities to construct ¬*p* cases, whereas all the children, whatever their working memory capacities, were able to construct *p q* cases. Thus, convergent evidence suggest that, as we assume, the initial model could be produced by System 1 heuristic processes whereas the construction of additional models requires the intervention of analytic processes pertaining to System 2.

Implications for the suppositional and the standard mental model theories

The present results have also implications for the theories of conditionals. More precisely, though our theory is an integration of Evans' and Johnson-Laird and Byrne's approaches of conditional, our results lead to modify both of these accounts.

In their theory of the meaning of conditionals, Johnson-Laird and Byrne (2002) propose principles of modulation by which knowledge, pragmatics, and semantics can modulate the meaning of the conditional by adding information to the models or by preventing the construction of possibilities. This modulation accounts very well for the effects related to binary terms and strong causal relations. In the same way, the principle of pragmatic modulation assumes that the knowledge about the context of a conditional is normally represented in explicit models that can modulate the core interpretation of a conditional. This corresponds to what we observed with promises and threats. Our analyses suggested that the pragmatic implicatures concerning ¬*p* ¬*q* cases are represented in the initial model leading to "true" responses on these cases. The ¬*p* ¬*q* model resulting from these implicatures is therefore explicit, as Johnson-Laird and Byrne (2002) suggested, and we have seen that it modulates the core interpretation of the conditionals by blocking the construction of the ¬*p q* model, resulting in equivalence readings.

However, other aspects of our results do not fit Johnson-Laird and Byrne's (2002) approach. For example, Johnson-laird and Byrne suggest that there are ten interpretations of conditionals of the form *If A then C* depending on the nature of the possibilities represented and, among the interpretation involving two possibilities, they distinguish the biconditional (*a c* and ¬*a* ¬*c*), the strengthened antecedent (*a c* and *a* ¬*c*), and the relevance interpretations (*a c* and ¬*a c*). Our results suggest that there are probably more interpretations, because we observed that two different interpretations can be distinguished when the *a c* and ¬*a* ¬*c* possibilities are represented (i.e., equivalence and defective biconditional). This distinction can not be made by the current mental model theory that does not distinguish between models referring to possibilities that either make the conditional true or leave its truth-value indeterminate. Thus, this theory has also difficulties in accounting for the findings related to the defective truth table and the evaluation of the probability of conditionals by adults, as Evans et al. (2005) noted, and also by children as we observed in the present studies.

Of course, the suppositional approach defended by Evans and Over (2004) who inspired Barrouillet et al. (2008) theoretical proposal has less difficulty in accounting for the part of our results related to the two forms of biconditional interpretation. Indeed, Evans and Over (2004) stressed that the converse (adding *If q then p* to *If p then q*) and the inverse (adding *If not p then not q*) implicatures are equivalent for a truth functional conditional but not for their suppositional conditional. Adding the converse implicature conjoins two conditionals that have a defective table leading to the defective biconditional reading, as we observed for example with basic and causal conditionals. By contrast, adding the inverse implicature leads to the equivalence interpretation we observed with promises and threats. This account raises an interesting question because it is possible that, in many cases,

the defective biconditional patterns we observed reflected a complex initial model containing two representations, namely one representing a relation from p to q and the other a relation from q to p, that can not be differentiated by the truth evaluation task because both elicit the "true" response for the p q cases. However, our results tend to contradict the idea that defective biconditional interpretations mainly result from adding the converse implicature. We observed a developmental quadratic trend for defective biconditional interpretations that were rare in children and adults but predominant in adolescents. It is difficult to conceive a development by which the production of implicatures would not be available in childhood, thus leading to the conjunctive interpretation, then would appear in the earlier adolescence before disappearing in adulthood. Another difficulty would be to explain why it is precisely the converse and not the inverse implicature that would be added. It is far simpler to imagine, as we did, that the defective biconditional reading results from the incomplete fleshing out process of a single p q initial model. Children who do not flesh out this initial model do not produce biconditional readings and adults go beyond this level by producing a complete fleshing out leading to the three-model representation and the defective conditional interpretation.

Another important difference between our theory and Evans' suppositional approach concerns the meaning of conditionals. As we noted above, there is no need in our approach to assume that conditionals are suppositional in nature to account for the fact that $\neg p$ cases are considered as irrelevant by adults. It is sufficient to assume that the initial model of basic conditionals is produced by heuristic processes whereas additional models involve the intervention of an optional analytic system that produces models that are compatible with the conditional statement without pertaining to its core meaning. Thus, the evaluation of the probability of conditionals as the conditional probability, which is the main fact supporting the suppositional approach, can also be accounted for within the mental model theory framework. Our theory successfully predicted an intermediate developmental level in the probability task corresponding to the defective biconditional interpretation in which $\neg p$ cases are not necessarily considered as irrelevant. The suppositional theory could account for this developmental trend by supposing that this response results from adding the converse implicature. However, we have already noted above that this hypothesis is quite convoluted. Thus, although it gives a better account of our results, the suppositional approach of conditional defended by Evans and Over (2004) could have difficulties in accounting for some developmental phenomena. Moreover, as we have seen, conjunctive responses in evaluating the probability of conditionals cannot be accounted for by the Ramsey test postulated by the suppositional theory, whereas our mental model approach provides a simple explanation of these surprising responses.

Relationships with other dual-process theories

It is worth noting that our attempt to account for conditional reasoning and its development using a dual-process theoretical framework is far from being the first, and that several models have already been proposed. For example, Klaczynski et al. (2004) distinguished an *analytic* and an *experiential* (instead of heuristic) system in accounting for the development of conditional inferences from deontic and causal conditionals. Conditional inferences are inferences permitted by a conditional premise of the form *If p then q* and a minor premise either affirming or denying the antecedent p or the consequent q, resulting in four basic logic forms, *Modus Ponens* (MP; *If p then q*; p is true), the *affirmation of consequent* (AC; *If p then q*; q is true), *Modus Tollens* (MT; *If p then q*; q is not true), and the *denial of antecedent* (DA; *If p then q*; p is not true). The standard logic considers as valid the determinate inferences invited by MP (*If p then q*; p is true, therefore q is true) and MT (*If p then q*; q is not true, therefore p is not true), whereas the determinate inferences invited by AC (*If p then q*; q is true, therefore p is true) and DA (*If p then q*; p is not true, therefore q is not true) are considered as fallacies because alternative possible antecedents of the form $\neg p$ could lead to the same consequent q, rendering uncertain the determinate inferences *p is true* and *q is false*.

Klaczynski et al. (2004) suggested that on MP and MT, the predominance of either the analytic or the experiential system should lead to determinate inferences (i.e., concluding q and *not p* respectively), because both contextualized and decontextualized representations facilitate the invited inference on these forms. By contrast, on AC and DA, the experiential system should lead to the acceptance of the invited inferences (i.e., p and *not q* respectively), whereas the analytic system should lead to

indeterminate inferences by permitting the evocation of alternative ¬p q cases that block both p and *not q* conclusions. Because older children have better-developed executive functions and higher inhibition capacities, the authors predicted and observed that the analytic system should be more likely to predominate in older than in younger children, in whom the experiential system should predominate and lead them to endorse determinate inferences on all the basic logic forms. In further studies, Klaczynski and Daniel (2005) moderated these views. Indeed, it appeared that thinking dispositions to rely on analytic processing predicted inferences on problems that do not elicit conflict between the two systems, such as MP and in a lesser extent MT for which both the analytic and the experiential systems lead to determinate inferences.

Though they concern different processes than those addressed in the present studies, namely inference production, these suggestions are not incompatible with our conception. Klaczynski's model assumes that the experiential system leads to endorse and produce invited inferences on the four canonical forms. According to the mental model theory, this pattern of inference results from the construction of a two-model representation containing p q and ¬p ¬q. Thus, Klaczynski's proposal differs from our theory in suggesting that the heuristic experiential system produces also the ¬p ¬q model, whereas we have assumed that the heuristic processes do not deliver a double but a single initial model p q, at least for the causal conditionals as studied by Klaczynski et al. (2004) and Klaczynski and Daniel (2005). This single model supports the determinate inferences on MP and AC by permitting to conclude q from p and p from q, but not on MT and DA, which necessitate the construction of a second model of the form ¬p ¬q. However, it should be noted that the conditionals studied by Klaczynski et al. (2004) and Klaczynski and Daniel (2005) involved highly familiar relations which probably permit the automatic retrieval of ¬p ¬q cases from the complementary class, as assumed by Markovits and Barrouillet (2002). By contrast, alternative antecedents (of the form ¬p q) permitting to reject the determinate inferences on AC and DA are more difficult to evoke (Barrouillet et al., 2001). This would explain why, in Klaczynski et al. (2004) study, inferences on MP and MT were not related to inferences on AC and DA in older children who are probably more able to construct a three-model representation including a ¬p q model.

Using artificial relations that do not permit the retrieval from long-term memory of ¬p ¬q cases, Barrouillet et al. (2000) observed a developmental pattern that fits the predictions of our model. In a first developmental level, third graders frequently produced the determinate inferences on MP and AC, but not the determinate inferences on DA and MT, which necessitates the construction of a second model of the form ¬p ¬q. This first developmental step corresponds to the intervention of the sole heuristic (experiential) system producing the initial model p q that underpins determinate inferences on MP and AC. It is only in older children that was observed the production of determinate inferences on the four logic forms, which necessitates the construction of two models (p q and ¬p ¬q). The production of determinate inferences on MP and MT along with responses of indeterminacy on AC and DA, which require the intervention of the analytic system to produce a 3-model representation, were not observed before the end of adolescence and adulthood. Though our theory differs from Klaczynski et al. (2004), it is not incompatible with the idea that with highly familiar relations, the heuristic system can support determinate inferences on DA and MT through the automatic activation and retrieval of ¬p ¬q cases. By contrast, it remains unclear how Klaczynski et al. approach could account for the findings observed through the truth table and the probability tasks.

Although it does not address developmental questions, another relevant dual-process account of conditional inferences was proposed by Verschueren, Schaeken, and d'Ydewalle (2005). They proposed to consider the probabilistic and the mental model accounts of conditional reasoning from a dual-process perspective. They suggested that the two kinds of information on which reasoning is based according to these two accounts, namely probabilistic information for the probabilistic account and counterexamples for the mental model theory, might fit the heuristic/analytic polarity. Indeed, the reasoning processes based on probability and likelihood can be considered as heuristic, whereas the conscious search for counterexamples and alternative models can be considered as their analytic counterpart. In line with the hypothesis that heuristic are faster than analytic processes, Verschueren et al. observed that fast conditional inferences were better accounted by estimates of the likelihood of the conditional relation, whereas slow reasoning was better accounted by the availability of counterexamples. Subsequently, they showed that the analytic reasoning based on the retrieval of counterexamples can override the conclusions based on likelihood information.

Though this account is in some aspects akin to our approach, there are also important differences. Among the similarities is the idea that the heuristic processes provide probabilistic information. Markovits and Barrouillet (2002) assumed that children and adults have an understanding of conditionals that is inherently relational and that brings to bear a rich pragmatic experience. Thus, it can be imagined that the heuristic processes produce an initial model of the relation between p and q that involves implicit knowledge about the likelihood of this relationship (provided that it is familiar), and that this information can fix the degree of belief on the inferences that can be drawn from this representation. Of course, we also agree that analytic reasoning involves the search for alternative models. However, there are also important differences. We agree that analytic reasoning involves mental models manipulation, but reasoning by models does not necessarily rely on analytic process in the sense that it would involve the systematic search for counterexamples. In line with the basic proposals of the mental model theory, we assume that even heuristic reasoning is based on mental models, more precisely initial models. Moreover, Verschueren et al.'s account seems to endorse what Evans (2007a) called a parallel-competitive model, which assumes that the two systems operate in parallel to deliver a putative response, whereas we favor, following Evans (2006), a default-interventionist model according to which the heuristic system cues default responses that may or may not be modified by the analytic system. This latter option is richer because it suggests more complex relations between the two systems, the heuristic processes influencing the outcome of the analytic system by supplying content to its operations. Finally, Verschueren et al. account, as the probabilistic account of reasoning, is limited to everyday contextualized thinking, whereas one strength that our theory inherits from the mental model theory is to provide an account of reasoning processes that encompasses everyday and abstract reasoning within the same theoretical framework.

Finally, one of the most influential dual-process theory of reasoning and its development is probably the fuzzy-trace theory proposed by Brainerd and Reyna (2001). In a nutshell, this theory assumes that there is three core principles of reasoning. According to the principle of *gist extraction*, when children and adults encode the background facts accompanying reasoning problems, the memories that are stored go beyond the surface form of these premises to include the sense and meanings that the facts instantiate. Reasoners therefore process memories other than the verbatim memories that represent detailed surface content of the information. In several studies were demonstrated reasoning-remembering independence effects indicating that remembering rely on different types of representations processed by different types of operations (Brainerd & Reyna, 1992; Reyna, 1992). Thus, and this is the second principle, the memories that are processed vary in their degree of specificity along *fuzzy-to-verbatim continua*. More importantly, the third principle assumes that there is a *fuzzy-processing preference* and that adult's reasoning is disposed to operate on the simplest and least precise representation that will generate a solution. Thus, there is an opposition between analytical reasoning that involves sequential logical processing of precise surface content of background facts, and intuitive reasoning that involves global heuristic processing of gist memories. Whereas classical approaches of cognitive development and information processing theories usually assume an intuitive-to-analytic developmental shift, the fuzzy-trace theory departs from these approaches by assuming that adult reasoning preferentially operates on fuzzy traces and gistified information rather than on precise verbatim memory, and that this preference becomes more pronounced with development. Thus, this theory rejects the idea that intuitive reasoning atrophies and is progressively replaced by analytical reasoning. Instead, it is assumed that both analytic and intuitive reasoning improve with age, development being characterized by a computation-to-intuition shift.

Though, at the first glance, this theory could appear as diametrically opposed to our conceptions, there are many points of agreement. First, it is one of the main assumptions of the mental model approach that people do not reason on verbatim memories of premises, but on mental models that capture the meaning and structure of the facts the premises describe. Mental models are even explicitly described as capturing the gist of sentences rather than their verbatim details (Mani & Johnson-Laird, 1982). Thus, the mental model theory has no difficulty with the remembering-reasoning independence and the idea that reasoning is based on gist traces. The production of the initial model through heuristic processes is very akin to the gist retrieval described by Brainerd and Reyna (1993). The construction of additional models and the processes of coordinating models and eliminating inconsistencies is not so far from the process of editing the gist store, even if Brainerd and Reyna prefer a comparison with Kintsch's

(1988) construction-integration model. This editing mechanism is assumed to develop with age (Brainerd & Reyna, 1993), and this could account for the relationship observed between working memory capacity development and the construction of models observed by Barrouillet and Lecas (1999). Second, we agree that, even in adults, there is a fuzzy-processing preference because the heuristic processes cue default responses whereas analytic processes are only optional. Thus, as Brainerd and Reyna assume, intuitive reasoning based on heuristic processes does not atrophy with increasing development.

As we have already discussed above, the main difference between the fuzzy-trace and our theory concerns the development of the two systems of reasoning. We have argued that, at least for conditional reasoning and during the period from school years to adulthood that we have investigated, the analytic system strongly develops while the heuristic system remains relatively unchanged. However, it should be noted that the fuzzy-trace theory is more complex than the brief presentation we gave above. Brainerd and Reyna (2001, p. 54) explicitly specify that "certain forms of intuitive reasoning increase with age", implicitly suggesting that other forms would not increase in the same extent, and they acknowledge that some tasks are more analytical than others, giving conditional inference as the first example. Indeed, concerning conditional inference, it seems that the computation-to-intuition developmental shift assumed by fuzzy-trace theory does not occur. Both Klaczynski's and Verschueren's dual-process theories agree that producing indeterminate inferences on the AC and DA forms rely on analytic processes. Actually, there is abundant evidence that the production of these indeterminate inferences increases with age from childhood to adulthood, strongly suggesting that, more often than children, adults rely on analytic processes when reasoning from conditional statements (Barrouillet et al., 2000; Barrouillet et al., 2001; De Neys & Everaerts, 2008; Ennis, 1976; Jenveau-Brennan & Markovits, 1999; Markovits, Fleury, Quinn, & Venet, 1998; O'Brien & Overton, 1982).

Concluding comments

In summary, introducing the distinction suggested by Evans (2006), Evans (2008) between heuristic and analytic systems in our revised mental model theory of conditionals (Barrouillet et al., 2008) proved to be particularly fruitful by permitting a series of developmental hypotheses that were verified in the present study. The interplay between the two systems and their differential sensitivity to development lead to a variety of contrasted developmental trajectories from childhood to adulthood depending on the type and the contents of the conditionals, and to a variety of interpretations that goes beyond what Markovits and Barrouillet (2002) and even Johnson-Laird and Byrne (2002) expected. Further studies investigating the development of deontic, temporal or couterfactual conditionals could reveal an even wider range of developmental courses and interpretations. At the moment, our modified mental model theory of conditionals accounts for the development on the production and evaluation of cases compatible or incompatible with a conditional rule (Barrouillet & Lecas, 1998; Barrouillet & Lecas, 1999), on the production of conditional inferences (Barrouillet et al., 2000), on various effects of content and context in children and adolescent conditional reasoning (Barrouillet & Lecas, 2002; Barrouillet et al., 2001), on truth table tasks (Barrouillet et al., 2008, and the present studies), and on the evaluation of the probability of conditionals. Moreover, this theory is not restricted to developmental phenomena and provides an integrative account of the way adults understand and reason from conditional statements (Barrouillet et al., 2008). Thus, though this theoretical approach needs to be extended to the forms of conditionals we have not yet explored, integrating Evans' heuristic analytic theory in our modified mental model theory of conditionals extends considerably our previous developmental model and probably provides one of the most comprehensive account of conditional reasoning in children, adolescents, and adults.

Acknowledgments

Appendix

BB conditionals used in Experiment 1

"If the pupil is a boy then he wears glasses".
"If the door is open then the light is switched on".
"If the student is a woman then she wears a shirt with long sleeves".
"If the piece is big then it is pierced".

NN conditionals used in Experiment 1

"If the card is yellow then a triangle is printed on it".
"If there is a star on the screen then there is a circle".
"If he wears a red t-shirt then he wears a green trousers".
"If there is a rabbit in the cage then there is a cat".

Strong causal relations used in Experiment 2

"If the button 3 is turned then the blackboard's lights are switched on".
"If the lever 2 is down, then the rabbit's cage is open".
"If the second button of the machine is green then the machine makes sweets".
"If I pour out pink liquid in the vase then stars appear on it".

Weak causal relations used in Experiment 2

"If the touch F5 is pressed then the computer screen becomes black".
"If the boy eats *alkali* pills then his skin tans".
"If the fisherman puts flour in the water then he catches a lot of fishes".
"If the gardener pours out *buntil* in his garden then he gathers a lot of tomatoes".

Promises used in Experiment 3

"If you gather the leafs in the garden then I give you 5 francs".
"If you score a goal then I name you captain".
"If you exercise the dog then I cook you a cake for dinner".
"If you clean your room then you watch the TV".

Threats used in Experiment 3

"If you break the vase then I take your ball".
"If you do not buy the bread then you do not play video games".
"If you do not do your homework then you do not go to the attraction park".
"If you have a bad mark then you do not go to the movie".

References

Artman, L., Cahan, S., & Avni-Babad, D. (2006). Age, schooling and conditional reasoning. *Cognitive Development, 21*(2), 131–145.
Barra, B. G., Bucciarelli, M., & Johnson-Laird, P. N. (1995). Development of syllogistic reasoning. *American Journal of Psychology, 108*(2), 157–193.
Barrouillet, P., Gauffroy, C., & Lecas, J. F. (2008). Mental models and the suppositional account of conditionals. *Psychological Review, 115*(3), 760–771.
Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Memory span development: A time-based resource-sharing model account. *Developmental Psychology, 45*(2), 477–490.
Barrouillet, P., Grosset, N., & Lecas, J. F. (2000). Conditional reasoning by mental models: Chronometric and developmental evidence. *Cognition, 75*, 237–266.
Barrouillet, P., & Lecas, J. F. (1998). How can mental models account for content effects in conditional reasoning? A developmental perspective. *Cognition, 67*, 209–253.
Barrouillet, P., & Lecas, J. F. (1999). Mental models in conditional reasoning and working memory. *Thinking and Reasoning, 5*(4), 289–302.
Barrouillet, P., & Lecas, J. F. (2002). Content and context effects in children's and adults' conditional reasoning. *Quarterly Journal of Experimental Psychology, 55*(3), 839–854.

Barrouillet, P., Markovits, H., & Quinn, S. (2001). Developmental and content effects in reasoning with causal conditionals. *Journal of Experimental Child Psychology, 81*(3), 235–248.

Brainerd, C. J., & Reyna, V. F. (1992). Explaining memory free reasoning. *Psychological Science, 3*(6), 332–339.

Brainerd, C. J., & Reyna, V. F. (1993). Memory independence and memory interference in cognitive development. *Psychological Review, 100*(1), 42–67.

Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neuroscience. In H. W. Reese & R. Kail (Eds.). *Advances in child development and behavior* (Vol. 28, pp. 41–100). San Diego, CA: Academic Press.

Byrnes, J. P., & Overton, W. F. (1988). Reasoning about logical connectives – A developmental analysis. *Journal of Experimental Child Psychology, 46*(2), 194–218.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term-memory span. *Journal of Experimental Child Psychology, 33*(3), 386–404.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17*, 391–416.

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how human reason? *Memory & Cognition, 24*(6), 823–829.

Cummins, D. D. (1995). Naïve theories and causal deduction. *Memory & Cognition, 23*(5), 646–658.

Cummins, D. D. (1996). Evidence of deontic reasoning in 3- and 4-year-old children. *Cognition, 31*, 187–276.

Cummins, D. D., Lubart, T., Alksnis, O., & Rist, R. (1991). Conditional reasoning and causation. *Memory & Cognition, 19*(3), 274–282.

Daniel, D. B., & Klaczynski, P. A. (2006). Developmental and individual differences in conditional reasoning: Effects of logic instructions and alternative antecedents. *Child Development, 77*(2), 339–354.

Davidson, D. (1995). The representativeness heuristic and conjunction fallacy effect in children's decision making. *Merrill-Palmer Quarterly, 41*, 328–346.

De Neys, W., & Everaerts, D. (2008). Developmental trends in everyday conditional reasoning: The retrieval and inhibition interplay. *Journal of Experimental Child Psychology, 100*(4), 252–263.

Ennis, R. H. (1976). An alternative to Piaget's conceptualization of logical competence. *Child Development, 47*, 903–919.

Evans, J. St. B. T. (2006). The heuristic–analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review, 13*(3), 378–395.

Evans, J. St. B. T. (2007a). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning, 13*(4), 321–339.

Evans, J. St. B. T. (2007b). *Hypothetical thinking: Dual processes in reasoning and judgment.* Hove, UK: Psychology Press.

Evans, J. St. B. T. (2008). Dual processing accounts of reasoning, judgment and social cognition. *Annual Review of Psychology, 59*, 255–278.

Evans, J. St. B. T., Handley, S. J., Neilens, H., & Over, D. E. (2007). Conditionals and conditional probability. *Memory & Cognition, 35*(7), 1772–1784.

Evans, J. St. B. T., Handley, S. J., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, memory and cognition, 29*(2), 331–335.

Evans, J. S. T. B., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction.* Hove, UK: Lawrence Erlbaum.

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning.* Hove, UK: Psychology Press.

Evans, J. St. B. T., & Over, D. E. (2004). *If.* Oxford: Oxford University Press.

Evans, J. St. B. T., Over, D. E., & Handley, S. J. (2005). Suppositions, extensionality and conditionals: A critique of the mental model theory of Johnson-Laird and Byrne (2002). *Psychological Review, 112*, 1040–1052.

Fillenbaum, S. (1975). IF – Some uses. *Psychological Research Psychologische Forschung, 37*(3), 247–260.

Fillenbaum, S. (1976). Inducements – Phrasing and logic of conditional promises, threats and warnings. *Psychological Research Psychologische Forschung, 38*(3), 231–250.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177–190.

Goldvarg, E., & Johnson-Laird, P. N. (2001). Naïve causality: A mental model theory of causal meaning and reasoning. *Cognitive Science, 25*(4), 565–610.

Grosset, N., Barrouillet, P., & Markovits, H. (2005). Chronometric evidence for memory retrieval in causal conditional reasoning: The case of the association strength effect. *Memory & Cognition, 33*(4), 734–741.

Harris, P. L., & Nunez, M. (1996). Understanding of permission rules by preschool children. *Child Development, 67*, 1572–1591.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence.* New York: Basic Books.

Jacobs, J. E., & Potenza, M. (1991). The use of judgement heuristics to make social and object decisions: A developmental perspective. *Child Development, 62*(1), 166–178.

Jenveau-Brennan, G., & Markovits, H. (1999). The development of reasoning with causal conditionals. *Developmental Psychology, 35*(4), 904–911.

Johnson-Laird, P. N. (2006). *How we reason?* Oxford: Oxford University Press.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction.* Hillsdale, NJ: Lawrence Erlbaum.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review, 109*, 646–678.

Kintsch, W. (1988). The role of knowledge in discourse comprehension – a construction integration model. *Psychological Review, 95*(2), 163–182.

Klaczynski, P. A. (2001). Analytic and heuristic processing influences on adolescent reasoning and decision making. *Child Development, 72*, 844–861.

Klaczynski, P. A., & Cottrell, J. M. (2004). A dual-process approach to cognitive development: The case of children's understanding of sun cost decisions. *Thinking & Reasoning, 10*(2), 147–174.

Klaczynski, P. A., & Daniel, D. B. (2005). Individual differences in conditional reasoning. *Thinking & Reasoning, 11*(4), 305–325.

Klaczynski, P. A., Schuneman, M. J., & Daniel, D. B. (2004). Theories of conditional reasoning: A developmental examination of competing hypotheses. *Developmental Psychology, 40*(4), 559–571.

Kokis, J. V., MacPherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology, 83*(1), 26–52.

Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills*. San Diego: Academic Press.

Kushsnir, T., & Gopnik, A. (2007). Conditional probability versus spatial contiguity in causal learning: Preschoolers use new contingency evidence to overcome prior spatial assumptions. *Developmental Psychology, 43*, 186–196.

Lecas, J. F., & Barrouillet, P. (1999). Understanding of conditional rules in childhood and adolescence: A mental models approach. *Current Psychology of Cognition, 18*, 363–396.

Light, P., Blaye, A., Gilly, M., & Girotto, V. (1989). Pragmatic schemas and logical reasoning in 6- to 8-year-old children. *Cognitive Development, 4*, 49–64.

Mani, K., & Johnson-Laird, P. N. (1982). The mental representation of spatial descriptions. *Memory & Cognition, 10*(2), 181–187.

Marcus, S. L., & Rips, L. J. (1979). Conditional reasoning. *Journal of Verbal Learning and Verbal Behavior, 18*, 199–223.

Markovits, H. (1984). Awareness of the possible as a mediator of formal thinking in conditional reasoning problems. *British Journal of Psychology, 75*, 367–376.

Markovits, H., & Barrouillet, P. (2002). The development of conditional reasoning: A mental model account. *Developmental Review, 22*(1), 5–36.

Markovits, H., & Dumas, C. (1999). Developmental patterns in the understanding of social and physical transitivity. *Journal of Experimental Child Psychology, 73*(2), 95–114.

Markovits, H., Fleury, M. L., Quinn, S., & Venet, M. (1998). The development of conditional reasoning and the structure of semantic memory. *Child Development, 69*(3), 742–755.

Morsanyi, K., & Handley, S. J. (2008). How smart do you need to be to get it wrong? The role of cognitive capacity in the development of heuristic-based judgment. *Journal of Experimental Child Psychology, 99*, 18–36.

Moshman, D. (1990). The development of metalogical understanding. In W. F. Overton (Ed.), *Reasoning, necessity, and logic: Developmental perspectives* (pp. 205–226). Hillsdale, NJ: Lawrence Erlbaum.

Moshman, D., & Franks, B. A. (1986). Development of the concept of inferential validity. *Child Development, 57*(1), 153–165.

Newstead, H. E., Ellis, M. C., Evans, J. St. B. T., & Dennis, I. (1997). Conditional reasoning with realistic material. *Thinking and Reasoning, 3*(1), 49–76.

O'Brien, D. P., & Overton, W. F. (1982). Conditional reasoning and the competence-performance issue: A developmental analysis of a training task. *Journal of Experimental Child Psychology, 34*, 274–290.

Ohm, E., & Thompson, V. A. (2006). Conditional probability and pragmatic conditionals: Dissociating truth and effectiveness. *Thinking and Reasoning, 12*(3), 257–280.

Oberauer, K., & Wilhelm, O. (2003). The Meaning(s) of Conditionals: Conditional Probabilities, Mental Models, and Personal Utilities. *Journal of Experimental Psychology:learning, memory and cognition, 29*(4), 680–693.

Over, D. E., Hadjchristidis, C., Evans, J. St. B. T., Handley, S. J., & Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology, 54*(1), 62–97.

Overton, W. F. (1990). Competences and procedure: Constraints on the development of logical reasoning. In W. F. Overton (Ed.), *Reasoning, necessity, and logic: Developmental perspectives* (pp. 1–32). Hillsdale, NJ: Lawrence Erlbaum.

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. Human Development, 15(1).

Quinn, S., & Markovits, H. (1998). Conditional reasoning, causality, and the structure of semantic memory: Strength of association as a predictive factor for content effects. *Cognition, 68*(3), 93–101.

Reber, A. S. (1993). *Implicit learning and tacit knowledge*. Oxford: Oxford University Press.

Reyna, V. F. (1992). Reasoning, remembering, and their relationship: Social, cognitive, and developmental issues. In M. L. Howe, C. J. Brainerd, & V. F. Reyna (Eds.), *Development of long-term retention* (pp. 103–127). New York, NY: Springer.

Reyna, V. F., & Ellis, S. C. (1994). Fuzzy-trace theory and framing effects in children risky decision-making. *Psychologica Science, 5*(5), 275–279.

Schroyens, W. (1997). Meta-propositional reasoning about the truth or falsity of propositions. *Psychologica Belgica, 37*(4), 219–247.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Lawrence Erlbaum.

Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2005). A dual-process specification of causal conditional reasoning. *Thinking & Reasoning, 11*(3), 239–278.

Vinter, A., & Perruchet, P. (2000). Implicit learning in children is not related to age: Evidence from drawing behavior. *Child Development, 71*(5), 1223–1240.