This article was downloaded by: [National Chung Cheng University], [In-mao Liu]
On: 01 August 2015, At: 00:02
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London, SW1P 1WG

CrossMark
Click for updates

Journal of Cognition and Development
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hjcd20

A Probabilistic Approach to Conditional Reasoning Development

In-mao Liu[a] & Ting-hsi Chou[b]

[a] National Chung-Cheng University and National Taiwan University, Taiwan

[b] Asia University, Taiwan

Accepted author version posted online: 21 Feb 2014.Published online: 21 Feb 2015.

To cite this article: In-mao Liu & Ting-hsi Chou (2015) A Probabilistic Approach to Conditional Reasoning Development, Journal of Cognition and Development, 16:3, 522-540, DOI: 10.1080/15248372.2013.879874

To link to this article: http://dx.doi.org/10.1080/15248372.2013.879874

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms &

Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

JOURNAL OF COGNITION AND DEVELOPMENT, 16(3):522–540
Copyright © 2015 Taylor & Francis Group, LLC
ISSN: 1524-8372 print/1532-7647 online
DOI: 10.1080/15248372.2013.879874

Routledge
Taylor & Francis Group

A Probabilistic Approach to Conditional Reasoning Development

In-mao Liu
National Chung-Cheng University and National Taiwan University, Taiwan

Ting-hsi Chou
Asia University, Taiwan

How likely is the glass to break, given that it is heated? The present study asks questions such as this with or without the premise *if the glass is heated, it breaks*. A reduced problem (question without premise) measures the statistical dependency (conditional probability) of an event to occur, given that another has occurred. Such statistical dependency represents knowledge-based reasoning (inferring from ''glass heated'' to ''its breaking'') and is a component of the response to the complete problem (question with premise). The complete problems therefore measure not only knowledge-based reasoning in terms of statistical dependencies (inductive component) but assumption-based reasoning (deductive component). Two experiments revealed: a) Knowledge-based reasoning continues to develop and attains adult levels at 7th grade for the problems tested, and b) assumption-based reasoning (deductive component) is reliable only for secondary school students (7th graders).

There seems no dispute that children are capable of deductive reasoning from early childhood (e.g., Hawkins, Pea, Glick, & Scribner, 1984; Leevers & Harris, 1999). Thus, Hawkins et al. (1984) had children aged 4 to 5 years old solve a fantasy type of syllogistic problem in which premises described mythical creatures foreign to practical knowledge. Their results indicate that young children are capable of making deductive inferences required in solving fantasy problems. Using class-based syllogisms similar to those of Hawkins et al., Leevers and Harris (1999) showed that instructions explicitly encouraging children to consider a premise and its implications could benefit logical performance without a prompt to use the imagination.

Without relying on contrived problems (e.g., Hawkins et al., 1984; Leevers & Harris, 1999) to uncover a deductive component in solving reasoning problems, the present study attempts to identify a deductive component by separating an inductive component (or knowledge-based component) from the total reasoning response involved in everyday reasoning problems.

Correspondence should be sent to In-mao Liu, Department of Psychology, National Chung-Cheng University, 160 San-Hsing, Min-Hsiung, Chia-Yi 621, Taiwan. E-mail: psyiml@ccu.edu.tw

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

The main difference between the present study and previous studies is in administering a control problem in addition to a reasoning problem, exemplified as follows:

If *the glass is heated* (*p*), then *it breaks* (*q*).	Major premise
The glass is heated (*p*).	Minor premise
Therefore, *it breaks* (*q*).	Conclusion

This argument is called a conditional reasoning argument, because the major premise is a conditional statement. In this argument, given the major premise (If *p*, then *q*) and the minor premise (*p*), the conclusion (*q*) certainly follows. For convenience of exposition, *p* and *q* are often used in the following text to represent events with which people are familiar as in the present example.

There are four forms of conditional reasoning: *P*-*Q* reasoning and three others. These are four forms of conditional reasoning with the major premise (complete problems, Table 1). There are also four forms of control problems without the major premise: *p*-*q* reasoning and three others (reduced problems, Table 1).

The conclusion of each reasoning problem in Table 1 is probabilistic, because very few people endorse everyday reasoning problems with perfect certainty (e.g., George, 1995; Stevenson & Over, 1995). According to probability theory (e.g., Jeffrey, 1981; Liu, 2003), what is measured from a control problem (e.g., *p*-*q* reasoning problem) represents a component of a *P*-*Q* reasoning response. Inferring from the glass being heated to its breaking (*p*-*q reasoning*) is apparently based on our world knowledge and hence is inductive. Moreover, following probability theory, an upward increase from the control problem to the *P*-*Q* reasoning problem, if it exists, stands for a deductive component (Liu, 2010).

This article is organized as follows: First, the development of an inductive component (knowledge-based reasoning) and a deductive component is hypothesized. Second, two experiments are reported in support of the hypotheses. Third, implications of the two experiments are discussed in the context of previous findings.

TABLE 1
The Reduced and Complete Problems in the Probabilistic Form

Reduced problem (Conditional reasoning without premise)	*Complete problem (Conditional reasoning with premise)*
p-q reasoning: Given *p*, how probable is *q*?	*P-Q* reasoning: If *p*, then *q*. Given *p*, how probable is *q*?
not-q-not-p reasoning: Given *not-q*, how probable is *not-p*?	*not-Q-not-P* reasoning: If *p*, then *q*. Given *not-q*, how probable is *not-p*?
not-p-not-q reasoning: Given *not-p*, how probable is *not-q*?	*not-P-not-Q* reasoning: If *p*, then *q*. Given *not-p*, how probable is *not-q*?
q-p reasoning: Given *q*, how probable is *p*?	*Q-P* reasoning: If *p*, then *q*. Given *q*, how probable is *p*?

Note. The type of complete problem is named after the type of reduced problem.

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

A PROBABILISTIC MODEL OF CONDITIONAL REASONING DEVELOPMENT

In the literature, investigators have particularly been interested in how early children could have an understanding of cause-and-effect relationships. Gelman, Bullock, and Meck (1980) used three-picture causal sequences (e.g., a cup, a hammer, and a shattered cup) and found an early (3-year-olds) understanding of many cause-and-effect relations. More recently, Gopnik and colleagues (e.g., Gopnik, Sobel, Schulz, & Glymour, 2001; Gopnik & Wellman, 2012, for review) showed that children as young as 2 years old could make causal inferences about genuine relationships.

More generally, people continuously learn statistical dependencies (conditional probabilities) from their childhood that a certain event will occur given that a specific prior event has occurred. A great deal of knowledge acquisition can therefore be understood in terms of *attained levels* of the conditional relationships (''if cause then effect,'' ''if category [diamond] then property [hardness],'' ''if member [dog] then class [animal].'').

As with Cummins (1995), Thompson (1995), and Weidenfeld, Oberauer, and Hornig (2005), such conditional relationships are referred to as forward, while the reversed conditional relationships are referred to as backward. Thus, inferring from the glass being heated to its breaking (*p-q* reasoning) is an instance of forward reasoning, while inferring from the glass breaking to its being heated (*q-p* reasoning) is an instance of backward reasoning. Gelman et al. (1980) found that the forward reading of causal sequences was easier for their 3-year-olds than was the reversed reading.

Development of Inductive (Knowledge-Based) Reasoning

To consider statistical dependencies between two events generated by *p* (glass heated) and *q* (glass breaking), the first event may be chosen from the four events, *p* and *q* and their complements *not-p* (glass not heated) and *not-q* (glass not breaking). There are four dependencies (*p-q, q-p, not-p-not-q, and not-q-not-p*) and four others (e.g., *p-not-q*, etc.). However, the *p-q* (glass breaking, given that it is heated) dependency and the *p-not-q* (glass not breaking, given that it is heated) dependency are complementary (*p-q* dependency = 1 – *p-not-q* dependency). This is the reason why we consider only the four types of dependencies (see the left half of Table 1) in the following.

The four types of dependencies. The conditional ''If it is a dog, it is an animal'' suffices to illustrate how the four types of dependencies develop to adult levels. In this example (''if member, then class''), the *dog-animal* dependency is measured by the subjective probability of being an animal, given that it is a dog. After learning that a dog is an animal, children become increasingly familiar with both concepts of dog and animal. At the same time, children become more certain that a dog is an animal and reach adult levels of *dog-animal* dependency.

The second type of dependency is the *not-animal-not-dog* dependency. Representing events as sets, it can be shown that the *not-animal-not-dog* dependency increases to adult levels as the *dog-animal* dependency increases to adult levels.

Third, the *animal-dog* dependency is measured by the probability of a creature's being a dog, given that it is an animal. It is easy to show that this probability decreases with knowledge acquisition, but this result is novel. To prove a decrease in the *animal-dog* dependency with knowledge acquisition, suppose children become knowledgeable about many nondog animals, such as donkeys, goats, etc. Then, the probability of an animal's being a dog becomes small.

Fourth, the *not-dog-not-animal* dependency is measured by the probability of a creature not being an animal, given that it is not a dog. There are two cases to consider. First, suppose children acquire knowledge of many nonanimal living things, then this probability tends to increase. Second, suppose children acquire knowledge of many nondog animals, then the *not-dog-not-animal* dependency tends to decrease. Consequently, there is no clear tendency observable for the *not-dog-not-animal* dependency.

Finally, it is well documented that negative sentences take longer to comprehend and are more prone to errors in verification than are affirmatives even for adults. Reasoning from a negative sentence (it is not an animal) to a negative sentence (it is not a dog) is, for instance, certainly more difficult than just comprehending negative sentences. Therefore, it generally follows that *not-animal-not-dog* reasoning develops to adult levels slower than does *dog-animal* reasoning.

Development of Deductive (Assumption-Based) Reasoning

In the present approach, estimating the *wearing glasses-intelligent* dependency (*p-q* reasoning) under the assumption of the conditional statement (if a person wears glasses, then this person is intelligent) gives rise to the *Wearing Glasses-Intelligent* reasoning (*P-Q* reasoning, see Table 1). Together with three other dependencies, four forms of conditional reasoning would be generated.

The probability of a girl being intelligent given her wearing glasses would be about .50 (inductive component). However, if this probability is estimated under the assumption that ''if a person wears glasses, then she is intelligent,'' then there would be an upward increase in this probability from the original .50. According to probability theory, this upward increase in probability represents a deductive component of the *P-Q* reasoning (Liu, 2010). In the present approach, therefore, it is possible to locate where and how a deductive component arises in conditional reasoning.

Thus, the presence of a deductive component in *P-Q* reasoning requires that reasoners hold the major premise (if wearing glasses, intelligent) and the *p-q* reasoning problem (given wearing glasses, intelligent?) in their working memory to detect their relationship (see Table 1). This prerequisite for working-memory capacity also explains why the deductive component involved in *not-Q-not-P* inferences (containing two negatives) is much more difficult than that involved in *P-Q* inferences. Second, children should be able to solve *p-q* reasoning problems to a degree of stabilization or to a degree of reaching adult levels for solving *P-Q* reasoning problems efficiently. In other words, it is only when children's knowledge-based reasoning (inductive component) has attained adult levels that sufficient cognitive resources could be allocated to perform assumption-based reasoning (deductive component). The phrase ''degree of stabilization'' is used here under the assumption that adult levels generally conform to objective levels of statistical dependency, if they exist.

The third prerequisite for children to be capable of performing deductive reasoning involved in *P-Q* reasoning depends on their ability to detach the major premise (if wearing glasses, then intelligent) from reality. Otherwise, reasoners would be unable to see the hypothetical nature of the major premise and would tend to see the major premise to reflect the *p-q* reasoning (''*wearing glasses*, therefore *intelligent*''; i.e., reality). Piaget (1972) considered children to

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

perceive the independence of its form from reality content as one of the essential characteristics of formal thought. Similarly, Stanovich (1999) and Stanovich and West (1998) refer to such skill as decontextualization skill enabling reasoning processes to operate independently of interfering context such as world knowledge. George (1995) observed higher *P*-*Q* inferences with premises referring to an imaginary person or object.

The same analyses apply to the presence of a deductive component in the *not*-*Q*-*not*-*P* reasoning. With respect to the *Q*-*P* reasoning, there is no upward increase from the *q*-*p* reasoning, because it is impossible to see how ''intelligent, therefore wearing glasses'' could be related to ''if wearing glasses, then intelligent,'' unless the latter is interpreted as ''if intelligent, then wearing glasses.'' This is known as a biconditional response. The same analysis for the *Q*-*P* reasoning applies to the case of the *not*-*P*-*not*-*Q* reasoning.

It should be noted that a deductive component in *P*-*Q* reasoning may not be observable because of the ceiling effect, when *p* is perceived as highly sufficient for *q*. Therefore, it is necessary to include a set of conditionals for which *p* and *q* are arbitrarily related. In this case, a deductive component in *P*-*Q* reasoning could be measured in its intact whole, if it is present. The same argument applies to the case of *not*-*Q*-*not*-*P* reasoning. On the other hand, the set of conditionals for which *p* and *q* are arbitrarily related will not be included in testing predictions about the development of statistical dependencies.

EXPERIMENTAL METHOD

Plan for Experiments

Variables to be manipulated. The effect of perceived sufficiency on *P*-*Q* and *not*-*Q*-*not*-*P* reasoning is well documented (e.g., Byrne, 1989; Cummins, Lubart, Alksnis, & Rist, 1991; Staudenmayer, 1975). With introduction of the reduced problems, however, it is clear that perceived sufficiency affects *P*-*Q* and *not*-*Q*-*not*-*P* reasoning responses by affecting their knowledge-based reasoning (e.g., Liu & Chou, 2012). As a matter of fact, the *p*-*q* dependency is a measure of perceived sufficiency. It is also known that *Q*-*P* and *not*-*P*-*not*-*Q* reasoning responses are affected by perceived necessity (e.g., Bucci, 1978, Experiment 2; Rumain, Connell, & Braine, 1983; Markovits, 1984; Thompson, 1994, 1995). As the *p*-*q* dependency is a measure of perceived sufficiency, the *q*-*p* dependency is a measure of perceived necessity.

The present study, therefore, included two experiments. For conditionals of the form ''if *p* then *q*,'' in Experiment 1, perceived sufficiency of *p* for *q* was manipulated from low to medium to high, while perceived necessity of *p* for *q* was kept low in all the conditions. Perceived sufficiency was manipulated in Experiment 1 because two forms of conditional reasoning (*P*-*Q* and *not*-*Q*-*not*-*P* reasoning) are affected by perceived sufficiency, while two other forms (*Q*-*P* and *not*-*P*-*not*-*Q* reasoning) are unaffected.

In Experiment 2, the antecedent and consequent clauses of each conditional used in Experiment 1 were reversed to obtain conditionals of the form ''if *q* then *p*.'' In Experiment 2, therefore, perceived necessity was manipulated from low to medium to high, while perceived sufficiency was kept low in all the conditions. Perceived necessity was manipulated in Experiment 2 because two forms of conditional reasoning are affected by perceived necessity, while two other forms are unaffected.

Predictions. As was presented earlier, in the course of knowledge acquisition, it may be predicted that the *p-q* dependency increases to adult levels (Prediction 1), the *not-q-not-p* dependency increases to adult levels (Prediction 2), the *q-p* dependency decreases to adult levels (Prediction 3), the *not-p-not-q* dependency has no clear way of developing to adult levels (Prediction 4), and the *not-q-not-p* dependency lags behind the *p-q* dependency in attaining adult levels, although both increase to adult levels (Prediction 5).

With respect to the presence or absence of a deductive component in *P-Q* reasoning, it is conjectured that a deductive component may be observable in *P-Q* reasoning from fifth or/and seventh graders, if the three prerequisites for the appearance of a deductive component could be satisfied for fifth or/and seventh graders.

Participants' age levels. In consideration of the fact that children of a rural area in Chia-Yi were to serve in the experiments, three age levels were originally selected for the present experiments: third, fifth, and seventh graders. Fifth and seventh graders were selected because the former are primary school students, while the latter are junior high school students. We did not select participants of older ages because previous studies involved not only college students (e.g., Liu & Chou, 2012), but also senior high school students (e.g., Liu, Lo, & Wu, 1996). The previous studies from our laboratory showed that observed conditional reasoning responses were generally comparable to those of Western participants. More specifically, a deductive component is consistently observable in both *P-Q* and *not-Q-not-P* reasoning for college students, while it is consistently observable in *P-Q* reasoning but only starts to appear in *not-Q-not-P* reasoning for senior high school students.

After finding from a preliminary experiment that third graders were not suitable to participate in probability rating experiments, only fifth and seventh graders participated in the two experiments. In the following, the preliminary experiment conducted with third and fifth graders is reported first, followed by presentation of the two experiments.

Preliminary Experiment

Because the measurement of both knowledge-based (inductive) and assumption-based (deductive) reasoning assumes that participants are capable of estimating probabilities, it was necessary to conduct a preliminary experiment to ascertain whether third and fifth graders could understand and estimate the probability of some familiar event (i.e., effectively use the probability measure). A class of 32 third graders (aged 8–9 years old) and a class of 34 fifth graders (aged 10–11 years old) at an elementary school in a rural area of Chia-Yi participated in this preliminary experiment. One of the two practice problems to be used in Experiments 1 and 2 was administered to these two groups of children: ''Given that Mary is an *A* Primary School student, how likely is it that she is going to a picnic today?'' They were to answer the problem by indicating their judged probability on an 11-point scale that ranged from 0 to 100, with 0 standing for ''completely improbable'' and 100 for ''completely certain.'' In addition, they were asked to write down reasons why they gave their answer.

If children's reason for giving a probability estimate contained a component of uncertainty or ignorance, the estimate was counted as indicating children's understanding of probability. If children's reason was incompatible with their probability estimate, the estimate was counted as indicating that children did not understand how to estimate probabilities.

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

The results showed that 69% of third graders could estimate the probability of a familiar event, with the mean estimate of .47, while the majority (88%) of fifth graders could, with the mean estimate of .46. For this reason, the two experiments to be reported did not include third graders as the participants.

Experiment 1: Manipulating Perceived Sufficiency

We used three sets of conditionals for generating reduced and complete problems. When tested with senior high school students and adults (e.g., Liu et al., 1996), one set of high-sufficiency conditionals is characterized by high sufficiency (mean ratings from .85 to .95) and low necessity (from .45 to .55), exemplified by ‘‘given that a substance is a diamond, it is very hard.’’ Another set of medium-sufficiency conditionals is characterized by medium sufficiency (mean ratings from .65 to .75) and low necessity (from .45 to .55), exemplified by ‘‘given that a person moves to a new house, this person adds some furniture.’’ The third set of low-sufficiency conditionals is characterized by low sufficiency (mean ratings from .45 to .55) and low necessity (from .45 to .55), exemplified by ‘‘given that a woman has long hair, she is a quiet woman.’’

Participants. The participants were 62 fifth graders (about half girls, half boys), aged 10 to 11 years old, and 42 seventh graders (about half girls, half boys), aged 12 to 13 years old, at an elementary school and a secondary school, respectively, both in a rural area of Chia-Yi. They did not serve in any other part of the present study.

Conditional statements. All the problems used in this experiment were generated from 12 conditionals. They were slightly modified from those used in the Liu et al. (1996) study: four conditionals with high perceived sufficiency and low perceived necessity, four conditionals with medium perceived sufficiency and low perceived necessity, and four conditionals with low perceived sufficiency and low perceived necessity. The modifications were made in consultation with two teachers of third graders and were aimed at attaining the goal of easy comprehension by third graders. The 12 conditionals used as the conditional premises are presented in the Appendix.

Procedure. The participants served in the experiment in large groups. They worked out two practice problems printed on the front page of a booklet before attempting to solve experimental problems. The instructions informed participants that there were two types of problems in the booklet: ‘‘The task is to judge the probability of some event for each type of problem. A rating scale attached to every problem represents probabilities of some event, from ‘completely improbable’ (0%) to ‘completely probable’ (100%). Please make your judgment by ticking on an appropriate place on the rating scale.’’

The first practice problem (Example 1) was in the *reduced* form: ‘‘Given that Mary is an *A* Primary (High) School student, how likely is it that she is going to a picnic today?’’ Participants then made their judgment by ticking on an appropriate place on the 11-point scale. The second problem (Example 2) was in the *complete* form: ‘‘If Mary is an *A* Primary (High) School student, then she is going to a picnic today. Given that Mary is an *A* Primary (High) School student, how likely is it that she is going to a picnic today?’’ They were told that the second type of problem consisted of a conditional sentence and a question sentence.

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

The further instructions were as follows: ''The conditional sentence of Example 2 explains that it is the date for *A* Primary (High) School students to go to a picnic today. Therefore, in the case that Mary is an *A* Primary (High) School student, she goes to a picnic today. Under this circumstance, the probability of Mary's going to a picnic could be different from Example 1.'' Participants also made their judgment by ticking on an appropriate place on the 11-point scale.

They were told, ''There are similar problems in the following. Please leaf through the following pages of the booklet, and start to give your answer to problems.'' They were reminded to write down their answer and rely on their own judgment. They were also reminded not to change their answers by going back to earlier problems. Then, participants rated 48 experimental problems (to be described) at their own pace. Two class teachers, one for fifth graders and one for seventh graders, served as the experimenters.

For about half of the fifth graders (32) and half of the seventh graders (21), the first 24 experimental problems were in the reduced form and the last 24 were in the complete form. The order was reversed for the remaining participants. For each participant, the first set of 24 experimental problems was constructed by randomly selecting two out of each set of four conditionals of different degrees of perceived sufficiency. Because each conditional could be used for constructing four types of arguments, there resulted in 24 experimental problems altogether. The complementary set of six conditionals was used to construct the second set of 24 experimental problems. For half of the participants, one set of 24 problems was in the reduced form, while the other set of 24 problems was in the complete form. Thus, when half of the participants saw 24 problems in the reduced form, the other half saw these same 24 problems in the complete form. Within each set of the reduced or complete forms, there were two randomized orders and two respective reverse orders of presenting the 24 problems.

Design. The design was a 2 (grade level: fifth or seventh) $\times$ 2 (problem type: reduced or complete) $\times$ 4 (forms of conditional argument) $\times$ 3 (perceived sufficiency: high, medium, or low) mixed design. Both grade level and problem type were between-subjects variables, and both conditional argument form and perceived sufficiency were within-subjects variables.

Experiment 2: Manipulating Perceived Necessity

Thompson (1994) and Cummins (1995) introduced the technique of reversing the antecedent and consequent clauses of conditional statements and attempted to study the effect of content independently from the syntactic forms (forward vs. reversed conditional relationships). We used their technique for interchanging degrees of perceived sufficiency with degrees of perceived necessity in Experiment 2.

Participants and problems. The participants were 55 fifth graders (about half girls, half boys), aged 10 to 11 years old, and 45 seventh graders (about half girls, half boys), aged 12 to 13 years old, at the same elementary school and the same secondary school of Experiment 1, respectively. They did not serve in any other part of the present study.

By reversing the antecedent and consequent clauses in the 12 conditional statements of Experiment 1, each set of 4 conditionals of high, medium, or low perceived necessity was generated. Each newly generated conditional was now characterized by low perceived sufficiency.

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

Both complete and reduced problems were then constructed on the basis of these 12 new conditional statements.

Procedure and design. The 55 fifth graders and 45 seventh graders were each randomly divided into two groups of approximately the same numbers of participants for receiving different orders of the reduced and complete problems. With perceived sufficiency replaced by perceived necessity, all the other details of the procedure and design were the same as in Experiment 1.

RESULTS AND DISCUSSION

In Experiments 1 and 2, for fifth and seventh graders, about half the participants received the reduced problems first and the complete problems second, while the order was reversed for the remaining participants. There were some order effects, so we only analyzed the first task participants performed.

Data Structure

So far as knowledge-based reasoning is concerned, Experiment 2 is a replication of Experiment 1. Therefore, the two experiments will be reported together. The mean ratings obtained from fifth and seventh graders in both experiments are shown separately in Tables 2 and 3 as a function of sufficiency (or necessity), type of argument, and type of problem. The mean rating observed from reduced problems stands for knowledge-based reasoning for each type of argument.

Tables 2 and 3 differ from conventional tables, as follows. First, the type of argument is named after the type of reduced problem. Second, independent variables (perceived sufficiency and necessity) are defined not with respect to the major premise, but with respect to the reduced problems. Thus, identical reduced problems are easily identified across the two tables. The assignment of participants to conditions was randomized in each experiment. Therefore, whenever the same conditions were involved in the two experiments, they were treated as replications in the following analyses of the development of knowledge-based reasoning (inductive component).

There is one caveat with respect to the notations for the four forms of conditional reasoning as follows. The major premises differ between Experiments 1 and 2. Thus, *P*-*Q* reasoning of Experiment 1 is comparable to *Q*-*P* reasoning of Experiment 2, which should not be confused with *Q*-*P* reasoning of Experiment 1. This is the reason that prime symbols are used to distinguish the four forms of conditional reasoning in Experiment 2 from the four forms of conditional reasoning in Experiment 1. Therefore, *P*-*Q* reasoning of Experiment 1 is comparable to Q'-P' reasoning of Experiment 2, and the development of deductive components will be reported separately for Experiments 1 and 2.

Development of Knowledge-Based Reasoning (Inductive Component)

There could not be knowledge acquisition across the two age levels in the low-sufficiency/necessity conditions, because *p* and *q* are arbitrarily related in these conditions. It can be seen from Tables 2 and 3 that participants tend to rate the probabilities of one event given

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

TABLE 2
Mean Probability Ratings for the Reduced and Complete Problems in Experiment 1

		Conditional argument							
		P-Q		*not-Q-not-P*		*not-P-not-Q*		*Q-P*	
Perceived sufficiency (necessity)	*Type of problem*	*Fifth graders*	*Seventh graders*	*Fifth graders*	*Seventh graders*	*Fifth graders*	*Seventh graders*	*Fifth graders*	*Seventh graders*
High (Low)	Reduced (*SD*)	.76 (.19)	.88 (.17)	.66 (.25)	.77 (.26)	.53 (.22)	.53 (.16)	.68 (.20)	.60 (.19)
	Complete (*SD*)	.84 (.15)	.89 (.18)	.66 (.24)	.66 (.25)	.55 (.19)	.61 (.22)	.66 (.22)	.60 (.20)
	(Increase)	.08	.01	.00	−.11	.02	.08	−.02	.00
Medium (Low)	Reduced (*SD*)	.68 (.19)	.81 (.13)	.57 (.28)	.65 (.27)	.59 (.29)	.71 (.25)	.71 (.19)	.63 (.20)
	Complete (*SD*)	.80 (.15)	.81 (.17)	.58 (.23)	.55 (.24)	.53 (.23)	.54 (.23)	.72 (.20)	.77 (.18)
	(Increase)	.12	.00	.01	−.10	−.06	−.17	.01	.14
Low (Low)	Reduced (*SD*)	.52 (.18)	.48 (.15)	.57 (.20)	.51 (.13)	.58 (.24)	.49 (.14)	.61 (.22)	.66 (.17)
	Complete (*SD*)	.58 (.22)	.62 (.16)	.53 (.16)	.55 (.19)	.52 (.22)	.52 (.18)	.62 (.23)	.70 (.17)
	(Increase)	.06	.14	−.04	.04	−.06	.03	.01	.04

Note. Perceived sufficiency (high, medium, or low) applies to the left half of the table; perceived necessity (low, low, low) applies to the right half.

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

TABLE 3
Mean Probability Ratings for the Reduced and Complete Problems in Experiment 2

		Conditional argument							
		Q′-P′		*not-P′-not-Q′*		*not-Q′-not-P′*		*P′-Q′*	
Perceived sufficiency (necessity)	*Type of Problem*	*Fifth graders*	*Seventh graders*	*Fifth graders*	*Seventh graders*	*Fifth graders*	*Seventh graders*	*Fifth graders*	*Seventh graders*
High (Low)	Reduced (*SD*)	.65 (.25)	.51 (.14)	.58 (.21)	.58 (.11)	.60 (.28)	.76 (.23)	.80 (.25)	.88 (.14)
	Complete (*SD*)	.62 (.26)	.69 (.17)	.55 (.24)	.55 (.19)	.68 (.26)	.73 (.24)	.76 (.23)	.88 (.16)
	(Increase)	−.03	.18	−.03	−.03	.08	−.03	−.04	.00
Medium (Low)	Reduced (*SD*)	.68 (.21)	.56 (.14)	.64 (.26)	.68 (.17)	.63 (.19)	.60 (.17)	.69 (.21)	.75 (.13)
	Complete (*SD*)	.68 (.25)	.76 (.17)	.62 (29)	.65 (.22)	.62 (.29)	.61 (.16)	.75 (.24)	.81 (.14)
	(Increase)	.00	.20	−.02	−.03	−.01	.01	.06	.06
Low (Low)	Reduced (*SD*)	.64 (.18)	.57 (.14)	.51 (.23)	.50 (.15)	.64 (.20)	.51 (.10)	.56 (.21)	.48 (.17)
	Complete (*SD*)	.61 (.24)	.78 (.15)	.45 (.23)	.53 (.20)	.50 (.24)	.59 (.16)	.59 (.23)	.63 (.18)
	(Increase)	−.03	.21	−.06	.03	−.14	.08	.03	.15

Note. Perceived sufficiency (high, medium, or low) applies to the right half of the table; perceived necessity (low, low, or low) applies to the left half. Prime symbols (P', Q') indicate that the conditional relationships are backward.

another slightly greater than .50, mostly in the range of .50 to .60. An analysis of variance (ANOVA; 2 [experiments] × 2 [ages] × 4 [reduced problem types]) performed on the low-sufficiency/necessity conditions showed that the age effect (.59 for fifth graders and .53 for seventh graders) was significant: $F(1, 98) = 4.52$, $MSE = .067$, $p < .05$, $\eta_p^2 = .044$. Thus, it is noteworthy to find that the conditional probability of two arbitrarily related events tends toward the adult level of .50 (e.g., Liu et al., 1996) across the two age levels. The only other significant effect was obtained for reduced problem type: $F(3, 294) = 10.27$, $MSE = .024$, $p < .01$, $\eta_p^2 = .095$. An inspection of Tables 2 and 3 showed that this significant effect is apparently due to a tendency to overestimate *q-p* dependencies in comparison to *p-q* dependencies.

Four types of dependencies. With respect to *p-q* dependencies, an ANOVA (2 [ages] × 2 [high- and medium-sufficiency conditions] × 2 [experiments]) showed that an increase in this probability (.73 vs. .83) from fifth to seventh graders was significant (Prediction 1), $F(1, 98) = 9.93$, $MSE = .046$, $p < .01$, $\eta_p^2 = .092$. The effect of sufficiency was significant, $F(1, 98) = 20.45$, $MSE = .023$, $p < .01$, $\eta_p^2 = .173$. All the other effects were not significant.

For *not-q-not-p* dependencies, an ANOVA (2 [ages] × 2 [high- and medium-sufficiency conditions] × 2 [experiments]) showed that an increase in this probability (.61 vs. .67) from fifth to seventh graders was significant (Prediction 2), $F(1, 98) = 4.26$, $MSE = .076$, $p < .05$, $\eta_p^2 = .042$. The effect of sufficiency was significant, $F(1, 98) = 7.45$, $MSE = .045$, $p < .01$, $\eta_p^2 = .071$. All the other effects were not significant.

For *q-p* dependencies, an ANOVA (2 [ages] × 2 [low–low-necessity conditions] × 2 [experiments]) showed that a decrease in this probability (.68 vs. .58) from fifth to seventh graders was significant (Prediction 3), $F(1, 98) = 10.31$, $MSE = .053$, $p < .01$, $\eta_p^2 = .095$. All the other effects were not significant.

For *not-p-not-q* dependencies, an ANOVA (2 [ages] × 2 [low–low-necessity conditions] × 2 [experiments]) showed that the difference in these probabilities (.59 vs. .63) between fifth and seventh graders was not significant (Prediction 4). The effect of necessity was significant, $F(1, 98) = 12.74$, $MSE = .041$, $p < .01$, $\eta_p^2 = .115$. All the other effects were not significant.

An inspection of Table 2 showed that this significant effect of necessity is due to *not-p-not-q* responses in the ''medium'' row being particularly high (.59 and .71 for fifth and seventh graders, respectively) in comparison with the other conditions. In the ''medium''-sufficiency condition, the four reduced problems generally involve school or social regulations, such as, ''Given that M4 does not cheat in the exam, how likely is it that M4 will not be punished by the school?'' It is understandable that some children tend to rate the probability of such problems as high. This is an example of rating the *not-p-not-q* dependency as high, in spite of the low *q-p* dependency. This finding will be referred to as a *not-p-not-q* bias, because it appears several times in the sequel.

Contrasting p-q with not-q-not-p dependencies. Two ANOVAs (2 [experiments] × 3 [sufficiency conditions] × 2 [ages]) for *p-q* dependencies and *not-q-not-p* dependencies were performed to see how fifth and seventh graders developed to the adult levels of knowledge-based reasoning. An ANOVA for *p-q* dependencies showed that only the effects of sufficiency and its interaction with age were significant: $F(2, 196) = 104.46$, $MSE = .025$, $p < .01$, $\eta_p^2 = .516$; $F(2, 196) = 8.20$, $MSE = .025$, $p < .01$, $\eta_p^2 = .077$. This significant interaction indicates that *p-q* dependencies continued to develop from fifth to seventh graders. Further analyses of simple

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

effects showed that the effect of sufficiency on *p-q* dependencies was significant for both fifth and seventh graders: $F(2, 62) = 19.40$, $MSE = .024$, $p < .01$, $\eta_p^2 = .385$; $F(2, 40) = 45.65$, $MSE = .021$, $p < .01$, $\eta_p^2 = .695$.

An ANOVA for *not-q-not-p* dependencies also showed that only the effects of sufficiency and its interaction with age were significant: $F(2, 196) = 13.04$, $MSE = .037$, $p < .01$, $\eta_p^2 = .117$; $F(2, 196) = 8.87$, $MSE = .037$, $p < .01$, $\eta_p^2 = .083$. This significant interaction indicates that *not-q-not-p* dependencies also continued to develop from fifth to seventh graders. Further analyses of simple effects showed that the effect of sufficiency on *not-q-not-p* dependencies was significant only for seventh graders: $F(2, 40) = 8.59$, $MSE = .079$, $p < .01$, $\eta_p^2 = .300$ (Prediction 5).

Thus, the effect of sufficiency on *p-q* dependencies is observable from fifth graders as well as from seventh graders (see the first two columns of Table 2 and the last two columns of Table 3), while the effect of sufficiency on *not-q-not-p* dependencies is observable only from seventh graders (see the third and fourth columns of Table 2 and the fifth and sixth columns of Table 3). These findings support Prediction 5 that *not-q-not-p* dependencies develop to adult levels slower than *p-q* dependencies do.

Although Markovits and colleagues (e.g., Markovits, 2000; Markovits, Fleury, Quinn, & Vennet, 1998; Markovits & Thompson, 2008) observed that *Q-P* and *not-P-not-Q* indeterminate responses increase as a function of age, their results are not directly comparable in the present formulation. The main reasons are as follows: First, indeterminate responses are not comparable to probability responses. Second, these investigators did not administer reduced problems separately from complete problems. Therefore, they were unable to test Predictions 1 through 5. Although Prediction 3 is related to *Q-P* responses, it is still difficult to test Prediction 3, because it is not known whether observed *Q-P* responses are confounded with the biconditional interpretation of the major premise. This confounding can only be assessed by administering reduced problems separately from complete problems.

Development of Assumption-Based Reasoning (Deductive Component)

It was noted that there are three prerequisites for the appearance of a deductive component in *P-Q* or/and *not-Q-not-P* reasoning. Only one prerequisite could be assessed to determine whether it is satisfied on the basis of the observed data. It is only when children's knowledge-based reasoning (inductive component) has attained adult levels that sufficient cognitive resources could be allocated to perform assumption-based reasoning (deductive component).

For this purpose, let us consider whether *p-q* and *not-q-not-p* dependencies had developed to the adult pattern (.85 to .95, .65 to .75, and .45 to .55 for the high-, medium-, and low-sufficiency conditions, respectively, for *p-q* dependencies, while about .10 less in the high- and medium-sufficiency conditions for *not-q-not-p* dependencies [e.g., Liu et al., 1996]). It can be seen from Tables 2 and 3 that the three figures for seventh graders are .88, .78, and .48 for *p-q* dependencies, and .77, .63, and .51 for *not-q-not-p* dependencies. It may therefore be concluded that seventh graders had attained the adult levels for knowledge-based reasoning.

P-Q (Q′-P′) and not-Q-not-P (not-P′-not-Q′) responses. An increase in the mean rating from the reduced to complete problems for each type of argument represents the presence of a

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

deductive component. Tables 2 and 3 also present increases in the mean ratings in three rows, each row for one condition of perceived sufficiency or necessity. Negative signs indicate decreases in the mean ratings. To test the significance of the presence of a deductive component in *P-Q* and *not-Q-not-P* reasoning, ANOVAs were performed only in the low-sufficiency condition because of nearly no involvement of knowledge-based reasoning in these sufficiency conditions.

In Experiment 1, an ANOVA (2 [ages] × 2 [types of problem]) was separately performed on *P-Q* and *not-Q-not-P* inferences in the low-sufficiency condition. For *P-Q* inferences, only the effect of problem type was significant: $F(1, 100) = 7.18$, $MSE = .033$, $p < .01$, $\eta_p^2 = .067$. For *not-Q-not-P* inferences, every effect was not significant. With respect to *P-Q* inferences, therefore, a deductive component is seen to appear for both fifth and seventh graders. However, a deductive component in *not-Q-not-P* inferences is absent for both fifth and seventh graders.

In Experiment 2, an ANOVA (2 [ages] × 2 [types of problem] × 3 [low–low–low-necessity conditions]) was separately performed on *Q′-P′* and *not-P′-not-Q′* inferences. For *Q′-P′* inferences, both the effect of problem type and its interaction with age were significant: $F(1, 96) = 8.10$, $MSE = .070$, $p < .01$, $\eta_p^2 = .078$; $F(1, 96) = 12.47$, $MSE = .070$, $p < .01$, $\eta_p^2 = .115$. Because the interaction between age and problem type was significant, further analyses showed that the deductive component (.74–.55) was significant only for seventh graders: $F(1, 43) = 31.49$, $MSE = .014$, $p < .01$, $\eta_p^2 = .423$. All the other effects were not significant.

For *not-P′-not-Q′* inferences, the effect of problem type was not significant. The interaction between age and problem type was also not significant.

To conclude from the results of Experiments 1 and 2, although a deductive component in *P-Q* inferences starts to appear for fifth graders, it is still unstable. For seventh graders, however, it is stable and consistently observable. These results support the prediction from the present model that a deductive component becomes consistently observable only after the inductive component had attained the adult levels of development.

Q-P (P′-Q′) and not-P-not-Q (not-Q′-not-P′) responses. According to the present model, *Q-P* and *not-P-not-Q* responses are directly reflected in the *q-p* and *not-p-not-q* problems because there is no deductive component in these responses. Two ANOVAs were conducted for each experiment to see whether there is a deductive component in *Q-P* and *not-P-not-Q*: 2 (problem types: reduced, complete) × 2 (ages: fifth grade, seventh grade) × 3 (necessity conditions: low, low, low).

With respect to *Q-P* responses in Experiment 1, the effect of problem type was not significant. Its interaction with age was not significant, nor was its interaction with necessity. For *P′-Q′* responses in Experiment 2, the results were identical. Thus, the effect of problem type was not significant. Its interaction with age was not significant, nor was its interaction with necessity.

With respect to *not-P-not-Q* responses in Experiment 1, the effect of problem type was not significant. Its interaction with age was not significant. Its interaction with necessity was, however, significant: $F(2, 200) = 4.38$, $MSE = .044$, $p < .05$, $\eta_p^2 = .042$. This interaction is apparently due to the *not-p-not-q* bias, which was referred to earlier.

For *not-Q′-not-P′* responses in Experiment 2, problem type was not significant. Its interaction with age was not significant. Its interaction with sufficiency was also not significant. The three-way interaction (Problem Type × Age × Sufficiency) was significant: $F(2, 192) = 5.58$, $MSE = .031$, $p < .01$, $\eta_p^2 = .055$. This significant interaction could arise mainly because the effect of sufficiency started to appear only for seventh graders.

In conclusion, *Q*-*P* and *not*-*P*-*not*-*Q* responses do not differ from *q*-*p* and *not*-*p*-*not*-*q* responses, respectively. No exception was observed in the *Q*-*P* case. Although a few exceptions were observed in the *not*-*P*-*not*-*Q* case, this was caused by sample variations in the estimation of *not*-*p*-*not*-*q* responses. A further complication could be caused by the *not*-*p*-*not*-*q* bias.

GENERAL DISCUSSION

In the present study, four forms of conditional reasoning were designated after reduced problem type (see Table 1). Thus, *P*-*Q* argument stands for modus ponens (MP), while P'-Q' argument stands for affirmation of the consequent. However, both have identical reduced problems, and both are affected by perceived sufficiency. This is because sufficiency or necessity is conventionally defined with respect the conditional relationship in the major premise. However, the present study had to be presented in the context of previous studies. This is the reason why the conventional use of sufficiency and necessity was still adopted in presenting the ''Experimental Method.''

The present results are summarized as follows. The first part of this study involves ways in which knowledge is acquired as people learn statistical dependencies from their childhood that some event tends to occur given that some other event has occurred. It is possible to conceptualize four types of dependencies (*p*-*q, q*-*p, not*-*p*-*not*-*q*, and *not*-*q*-*not*-*p*) involving two events (*p, q*) of any forward conditional relationship.

As predicted, it was found that *p*-*q* and *not*-*q*-*not*-*p* dependencies increase to adult levels and that *q*-*p* dependencies decrease to adult levels. It was also found that *not*-*q*-*not*-*p* dependencies develop to adult levels slower than *p*-*q* dependencies do. When two events are arbitrarily related, unexpectedly, we found some slight overestimation of such dependencies by primary school children to tend toward the adult level of .50 for junior high school students. Finally, it was also found that *p*-*q* and *not*-*q*-*not*-*p* dependencies had developed to the adult pattern for junior high school students, but not for primary school children (fifth graders).

The second part of this study involved how the assumption-based reasoning (deductive component) emerges in solving MP problems. As it is possible in the present approach to locate where and how a deductive component arises in conditional reasoning, it was conjectured that it is only when children's knowledge-based reasoning (inductive component) has attained adult levels that sufficient cognitive resources could be allocated to perform assumption-based reasoning (deductive component). It was found that a deductive component is consistently observable in MP inferences for junior high school students (seventh graders), but it only starts to appear for primary school children (fifth graders). This finding is consistent with the conjecture that relevant statistical dependencies should have developed to the adult pattern for the deductive component to be present in MP inferences.

Much confusion in the interpretation of developmental processes could arise simply because MP (or *P*-*Q*) and modus tollens (or *not*-*Q*-*not*-*P*) responses could reflect only the inductive component in the literature (e.g., Hawkins et al., 1984; Leevers & Harris, 1999). Let us consider the logical reasons behind this explanation in some details as follows.

Hawkins et al. (1984) had children of 4 to 5 years of age solve three types of syllogistic problems: a) fantasy problems, in which premises described mythical creatures foreign to practical knowledge; b) incongruent problems, in which premises were in contradiction to practical

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

knowledge; and c) congruent problems, in which premises were compatible with practical knowledge. They concluded from their results that young children are capable of making deductive inferences required in solving fantasy and congruent problems but not in solving incongruent problems.

Hawkins et al. (1984) used fantasy problems such as:

All purple animals sneeze at people.
A banga is a purple animal.

Therefore, a banga sneezes at people.

Because participants are instructed to pretend that everything the stories say is true (i.e., to believe that all purple animals sneeze at people), the children's task is actually as follows:

A banga is a purple animal (that sneezes at people).

Therefore, a banga sneezes at people.

Thus, children are actually reasoning from the second premise with some enrichment to the conclusion. The enrichment refers to the fact that some information (''that sneezes at people'') has been stored temporarily in memory to make *p-q* responses but not assumption-based responses.

Hawkins et al.'s (1984) congruent problem is as follows:

All dogs bark.
It is a dog.

Therefore, it barks.

For this type of problem, children are actually reasoning from ''It is a dog (that barks)'' to ''Therefore, it barks.'' The only difference between congruent problems and fantasy problems is that the information in parentheses is stored in long-term memory in the former, while it is stored temporarily in the latter.

Let us finally consider an example problem, which is classified as Hawkins et al.'s (1984) incongruent problem:

All cats bark.
Hamlet is a cat.

Therefore, Hamlet barks.

For this type of problem, to make deductive inferences correctly, children have to take into account not only the second premise but also the first premise in arriving at the conclusion. Hawkins et al. found that children are incapable of performing assumption-based (or deductive) reasoning correctly for incongruent problems.

Using class-based syllogisms similar to those of Hawkins et al. (1984), Leevers and Harris (1999) showed that instructions explicitly encouraging children to consider a premise and its

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

implications could benefit logical performance without a prompt to use the imagination. More specifically, young children in the Leevers and Harris study are able to reply ''it is black'' given that something is snow (the second premise), because *further instructions* have raised the perceived sufficiency of p (it is snow) for q (it is black). Thus, when given syllogistic problems in which the major premise is incongruent with their empirical knowledge (e.g., ''All snow is black''), young children with further instructions are capable of answering the problems correctly.

In conclusion, the present results support conclusions about developmental progress in conditional reasoning between ages 10 to 11 years old and 12 to 13 years old that are consistent with other research using different methods (e.g., Markovits & Vachon, 1989; Moshman & Franks, 1986). Moshman and Franks (1986) asked students to sort sets of deductive arguments. They found that none of the fourth graders used validity as a basis for distinguishing arguments, while 45% of the seventh graders and 85% of the college students did so. They concluded that the concept of validity typically develops between ages 10 and 12. Markovits and Vachon (1989) studied the abilities of participants at four age levels (10, 13, 15, and 18 years old) to accept if-then premises as a basis for reasoning. They found that the 10-year-olds, and to a lesser extent the 13-year-olds, did have difficulty in accepting contrary-to-fact premises.

Finally, the present study has the following implications for further studies. First, it may be worthwhile to investigate how various environmental surroundings and training programs affect the development of children's perception of statistical dependencies of surrounding events. Second, it is possible to assess the effectiveness of different training programs for fostering children's deductive reasoning by administering both reduced and complete problems.

ACKNOWLEDGMENTS

We are grateful for helpful comments by David Moshman and an anonymous reviewer.

FUNDING

Preparation of this manuscript and the research reported were supported by the ROC National Science Council Grant 96-2413-H-194-010-MY2.

REFERENCES

Bucci, W. (1978). The interpretation of universal affirmative propositions. *Cognition*, *6*, 55–77.

Byrne, R. M. J. (1989). Suppressing valid inferences with conditionals. *Cognition*, *31*, 61–83.

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

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

Gelman, R., Bullock, M., & Meck, E. (1980). Preschoolers' understanding of simple object transformations. *Child Development*, *51*, 691–699.

George, C. (1995). The endorsement of the premises: Assumption-based or belief-based reasoning. *British Journal of Psychology*, *86*, 93–111.

Gopnik, A., Sobel, D. M., Schulz, L. E., & Glymour, C. (2001). Causal learning mechanisms in very young children: Two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. *Developmental Psychology*, *37*, 620–629.

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin*, *138*, 1085–1108.

Hawkins, J., Pea, R. D., Glick, J., & Scribner, S. (1984). 'Merds that laugh don't like mushrooms': Evidence for deductive reasoning by preschoolers. *Developmental Psychology*, *20*, 584–594.

Jeffrey, R. C. (1981). *Formal logic, its scope and limits* (2nd ed.). New York, NY: McGraw-Hill.

Leevers, H., & Harris, P. (1999). Transient and persisting effects of instruction on young children's syllogistic reasoning with incongruent and abstract premises. *Thinking and Reasoning*, *5*, 145–173.

Liu, I.-M. (2003). Conditional reasoning and conditionalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 694–709.

Liu, I.-M. (2010). A successive-conditionalization approach to conditional reasoning. In M. Oaksford (Ed.), *The psychology of conditionals* (pp. 219–232). Oxford, UK: Oxford University Press.

Liu, I.-M., & Chou, T.-H. (2012). A successive-conditionalization approach to disjunctive and syllogistic reasoning. *Psychological Research*, *76*, 676–692.

Liu, I.-M., Lo, K.-C., & Wu, J.-T. (1996). A probabilistic interpretation of 'if-then'. *Quarterly Journal of Experimental Psychology*, *49A*, 828–844.

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

Markovits, H. (2000). A mental model analysis of young children's conditional reasoning with meaningful premises. *Thinking and Reasoning*, *6*, 335–347.

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

Markovits, H., & Thompson, V. (2008). Different developmental patterns of simple deductive and probabilistic inferential reasoning. *Memory & Cognition*, *36*, 1066–1078. http://www.ncbi.nlm.nih.gov/pubmed/18927025

Markovits, H., & Vachon, R. (1989). Reasoning with contrary-to-fact propositions. *Journal of Experimental Child Psychology*, *47*, 398–412.

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

Piaget, J. (1972). *Judgment and reasoning in the child*. Totowa, NJ: Littlefield, Adams. (Original work published 1924)

Rumain, B., Connell, J., & Braine, M. D. S. (1983). Conversational comprehension processes are responsible for reasoning fallacies in children as well as adults. *Developmental Psychology*, *19*, 471–481.

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

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, *127*, 161–188.

Staudenmayer, H. (1975). Understanding conditional reasoning with meaningful propositions. In R. J. Falmagne (Ed.), *Reasoning: Representation and process* (pp. 55–79). New York, NY: Wiley.

Stevenson, R. J., & Over, D. E. (1995). Deduction from uncertain premises. *Quarterly Journal of Experimental Psychology*, *48A*, 613–645.

Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory & Cognition*, *22*, 742–758.

Thompson, V. A. (1995). Conditional reasoning: The necessary and sufficient conditions. *Canadian Journal of Experimental Psychology*, *49*, 1–60.

Weidenfeld, A., Oberauer, K., & Hornig, R. (2005). Causal and noncausal conditionals: An integrated model of interpretation. *Quarterly Journal of Experimental Psychology*, *58A*, 1479–1513.

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

Downloaded by [National Chung Cheng University], [In-mao Liu] at 00:02 01 August 2015

APPENDIX

Conditional Statements (Translated From Chinese)

High Perceived Sufficiency

If a substance is a diamond, then it is very hard (如果這物體是鑽石, 這物體就是堅硬的)

If this liquid is gasoline, then it is combustible (如果這液體是汽油, 它就可以燃燒).

If this is a dog, then it is an animal (如果這是狗, 這就是動物).

If H1 is 5 years old, then H1 is a child (如果小華是五歲, 小華就是小孩子).

Medium Perceived Sufficiency

If M1 moves to a new house, then M1 adds some furniture (如果李四搬家, 李四就買新傢俱)

If M2 comes back home late, then M2 will be scolded by his wife (如果林先生晚回家, 林先生就被太太責罵)

If M3 falls ill, then M3 will take a 1-day leave from the company (如果某乙生病, 某乙就向公司請假一天)

If M4 cheats in the exam, then M4 will be punished by his teacher (如果小偉考試作弊, 小偉就被老師處罰)

Low Perceived Sufficiency

If a woman has long hair, then she is a quiet woman (如果一個女人的頭髮很長, 這個女人就是文靜的)

If a person wears glasses, then this person is intelligent (如果一個人戴眼鏡, 這個人就是聰明的)

If a person puts on white clothes, then this person is a principal (如果一個人穿白衣服, 這個人就是校長)

If L4 puts white sport shoes on, then L3 goes to play ping-pong (如果小蔡穿上運動鞋, 小蔡就去打乒乓球)

Note. H1, M1, M2, M3, M4, L3, or L4 stands for a boy or girl's name in Chinese.