UNDERSTANDING LOGICAL CONNECTIVES:
A COMPARATIVE STUDY OF LANGUAGE INFLUENCE.

Rowan Patricia Garnier

Thesis submitted for the degree of Doctor of Philosophy

University of London Institute of Education

May, 1992

ABSTRACT

Operators called 'logical connectives' convey in a precise way the logical relationships between truth functional propositions and hence determine what can be inferred from them. Mathematical reasoning therefore relies heavily on their use. Whilst the operators are free of ambiguity, this is not so for the linguistic items (called 'linguistic connectives') by which they are codified. In English, at least, there is a widely reported mismatch between the logical concepts and the 'meanings' of the linguistic connectives with which they are frequently identified.

This study compares the provision for expressing logical concepts in Japanese, Arabic and English and seeks to ascertain to what extent the problems reported for English are generalisable to the other two languages. It also aims to establish whether the concepts underlying certain logical connectives are 'more readily available' or 'better established' in the speakers of one or other of these languages and, if so, whether this can be attributed to differing provision in the lexicon.

Two experiments were carried out using as subjects adults who were native speakers of either English, Japanese or Arabic. One was designed to determine to what extent the appropriate linguistic connectives in each of the three languages convey the associated logical concepts. The second compared performance on five concept identification tasks where the concepts tested were conjunction, inclusive and exclusive disjunction, the conditional and biconditional.

The results indicated no significant differences between language groups in the understanding of the linguistic expressions of logical connectives. However, the Japanese language group consistently outperformed the other two groups in all five concept identification tasks and also offered descriptions of these concepts which were more succinct and less variable. Possible explanations for the superior performance of the Japanese group are suggested and some implications for the teaching and learning of mathematics proposed.

ACKNOWLEDGMENTS

For a project which has taken as long as this one, there is a risk approaching certainty of failing to mention all those individuals to whom acknowledgment is due. Should this occur, I offer my apologies to those concerned and plead that a regrettable oversight rather than ingratitude be viewed as the cause.

I am grateful to the Faculty and Curriculum Development Committee at Richmond College for the award of finances which helped considerably with the payment of registration fees. For the granting of a semester's study leave which enabled completion of this thesis, I am indebted to the administration and trustees of Richmond College. I also acknowledge with thanks the receipt of a grant from the University of London Central Research Fund towards the cost of honoraria for data collection.

I have heard constant complaints from colleagues and friends of PhD supervisors who were unavailable, unhelpful or otherwise remiss. I therefore consider myself very fortunate to have been supervised by Professor Celia Hoyles who, to my mind, was everything a good supervisor should be. I am very grateful to her for her help and encouragement at every stage.

There are many from whose assistance this project has benefitted. I am indebted to Hania Dvorak and Stefan Sperl for their help in matters relating to the Arabic language and to Malcolm Simpson for similar assistance with Japanese. I am also grateful to Harvey Goldstein for his advice and help with the statistical data analysis. Without Fumiya Sawa, Elham Hassanain and Nahed AlShawa who acted as experimenters, there would have been no data to analyse and I therefore appreciate their many hours of hard work. The library staff at Richmond College dealt efficiently and cheerfully with my interminable requests for inter-library loans and I am grateful to them for their patience. Many friends and colleagues have shown an interest in this project and have been supportive in various ways. To attempt to list them would inevitably result in omissions but I would like to express my thanks to Nels Johnson, Amita Sen Gupta, Martin Winter and particularly to John Taylor whose ability to ferret out a flawed argument I have exploited mercilessly.

Last, but by no means least, my thanks go to my husband whose support has been no less constant with this than with any other of my more self-absorbing projects.

CONTENTS

Page

Page

List of Tables

Page

List of figures

Page

Chapter 1

INTRODUCTION

The context for this research project is the very broad question of whether mathematics might be easier to do and to learn in some languages rather than others. Human languages exhibit an extraordinary diversity; they differ in their phonology, in the categories for which they provide labels and in the grammatical rules which legislate how lexical units are combined. Could it be that the native speakers of certain languages have an advantage when it comes to learning mathematics because their language possesses structural features which in some way facilitate the formation of mathematical concepts?

Mathematical reasoning relies heavily on the use of 'logical connectives'. Essential to any process of mathematical inference is knowing what can be deduced from a set of statements whose truth is established or assumed. Logical connectives convey in a precise and unambiguous way the relationship between propositions and hence determine what can be inferred from them. However, whilst the operators themselves are free of ambiguity, this is not the case for the linguistic terms (often called 'linguistic connectives') which are used to codify the logical operations. For instance, there is ample evidence that, in English at least, a conditional statement such as 'If you clean the car, I'll give you £5' is not understood as conditional in the logical sense, i.e. false only when the addressee cleans the car and doesn't receive £5. In chapter 3 logical connectives are described together with the English language items which are commonly taken to express the logical relations which they define. Also described are the many ways in which these linguistic connectives fail to match up with the associated logical concepts.

Research published in English has dealt almost exclusively with the problems of expressing logical operations in that language. However, languages vary a great deal in the richness of their linguistic connective vocabulary and also in the grammatical structures which are used to convey logical concepts. It is by no means obvious that the problems reported for English language connectives should be universal features of all languages and yet this often seems to be the implicit (but unsupported) assumption.

The purpose of this project is to compare the provision for expressing logical concepts in Japanese and Arabic and to ascertain to what extent the problems reported for English speakers are generalisable to these two language groups. More specifically, do these languages provide connectives which convey the associated logical concepts more precisely than English does? Or is there a mismatch between what are normally assumed to be linguistic expressions of logical forms similar to that reported for English? Also, are certain logical concepts 'more readily available' or 'better established' in the speakers of one or other of these languages and, if so, can this be attributed to differing provision in the lexicon?

There are two reasons why these particular languages were chosen. The first is that Japanese, English and Arabic belong to different 'language families'. Each is believed to stem from a different linguistic origin and hence we might expect significant structural differences between them. The second reason is pragmatic - the availability of a local source of native speakers to use as experimental subjects.

This empirical and exploratory study aims to contribute to two major areas of debate. The first is the validity of the so-called 'Sapir-Whorf hypothesis', a theory which proposes that our thought processes are influenced by the language we speak. The second concerns issues surrounding the relationship between logic and deductive reasoning. It is all too apparent that even intelligent, educated adults make errors of reasoning which any model of the underlying thinking processes must attempt to explain. Whilst the source of these errors is not entirely clear, the comprehension process is clearly a prime candidate when reasoning is from verbal or written material. If a language conveys logical relationships unambiguously, then a potential source of fallacious reasoning is eliminated and one might expect the incidence of certain kinds of error to be minimised. Whilst this study does not attempt to compare deductive reasoning across language groups, it does seek to identify factors which might differentially influence reasoning performance.

Chapter 2

LANGUAGE AND MATHEMATICS

The nature and extent of the relationship between language and thought has been, and remains, the subject of considerable controversy. At one extreme we have the view expressed in the so-called 'Whorfian hypothesis' (alternatively known as the 'Sapir-Whorf hypothesis'), formulated in the 1920s and 30s by the anthropologist Edward Sapir and his student Benjamin Lee Whorf. This proposes that the language we speak imposes upon us a particular way of viewing the world and hence the way in which we think about the world.

> 'We cut nature up, organize it into concepts, and ascribe significances as we do, largely because we are parties to an agreement to organize it in this way - an agreement that holds throughout our speech community and is codified in the patterns of our language. ... We are thus introduced to a new principle of relativity, which holds that all observers are not led by the same physical evidence to the same picture of the universe, unless their linguistic backgrounds are similar, or can be in some way calibrated.' (Whorf, 1956, pp 212 - 214)

Although there is considerable difference of opinion as to the validity of Whorf's theory, it is not now generally accepted in its 'strong' form - that language *determines* behaviour and thought. The view that a language imposes such rigorous constraints on the cognitive functioning of its user has implications which are both depressing and untenable. True communication between different cultural groups would be, at best, severely limited and the problems faced by the second language learner would be sufficiently daunting to discourage the attempt. Yet we are told that some ideas, easily expressible in one language, are difficult to translate into another and that bilinguals think differently depending upon the language they are using (see, for instance: Wierzbicka, 1985; Ervin-Tripp, 1964). It is evident that languages differ greatly in the aspects of the physical world which they label and also in the rules which they provide for stringing together basic semantic units. The notion that dissimilar languages influence the thought processes of their respective users in different ways is therefore intuitively appealing.

Words are used to label concepts - mental representations which group together items on the basis of shared similarities. In effect, the 'meaning' of a word amounts to the concept underlying it. One current view is that knowledge is represented in memory in units called 'schemata' (see, for instance: Rumelhart, 1980; Cohen and Murphy, 1984). Corresponding to a concept is a schema which incorporates its essential features and their interrelations. Attached to schemata are variables, each associated with a feature of the concept which is not constant across exemplars. For example, associated with the concept 'dog' would be such variables as colour, size, length of hair, etc. Constraints on the variables define their normal range.

However, the constraints are not binding and they allow for encountering unusual values and also for correlating the value of a particular variable with those of others to which it is related. Schemata are not simply definitions. They are self-regulating and, when activated, are capable of assessing how well they fit the data currently being processed. It is proposed (Rumelhart, 1980) that schemata represent knowledge at every level of abstraction and that they play a central role in all reasoning and thinking processes. As such they are 'the building blocks of cognition'.

With this model of the mental representation of concepts, language comprehension consists of a process whereby words evoke the appropriate schemata. Failure to comprehend a word may be because no appropriate schema has been formed or because an inappropriate one has been activated. The words used in day-to-day interchanges in a speech community determine the schemata which members must share if they are to understand each other. A language must therefore oblige its speakers to form certain schemata if they are to be able to communicate effectively. To the extent that different languages label different concepts, it seems reasonable to infer that the schemata common to one linguistic group may not be identical to those shared by another, although we would expect there to be a considerable degree of overlap.

The foregoing account should not be taken to imply that schemata can be formed only for concepts which are labelled nor that language is necessary for concept formation. However, linguistic labels provide a means whereby concepts can be accessed from memory, manipulated at will and communicated to others. Although animals may be able to form certain low order concepts, they are unable to isolate them from the examples which gave rise to them in the first place (Skemp, 1987, p 15). Also, the possession of language greatly increases the range of concepts which can potentially be acquired because it provides for access to those which have been abstracted by other individuals. Without language, a concept must be formed by encountering exemplars and distinguishing them from non-exemplars. However, certain abstract concepts with no perceptible exemplars, such as 'infinity', could not be acquired other than through language.

The existence of a word in a given language implies sensitivity on the part of its users to the defining attributes of its referent and the existence of an underlying schema. Bloom (1984) proposes that a cognitive schema for which there is a label facilitates the external representation of that schema and hence communication of it. He suggests that thoughts which are represented by schemata with no labels must be translated into those that do, not only for the purposes of communication, but also as a means of internal representation. His conclusion is that labelled schemata play an influential role in both the external and internal domains and also in the development of thought, leading each generation of speakers towards the specific schemata for which their culture has developed labels.

A 'weak' form of Whorf's hypothesis (usually referred to as 'the theory of linguistic relativity') is generally taken to propose that it is *easier* to think and to talk about certain things in some languages rather than others. There is a 'cost of computation' associated with reasoning about a topic and this is partly determined by language. Hunt and Banaji (1988) explain this by proposing that thinking consists of manipulating mental representations, which presumably correspond to schemata in the model of knowledge representation described above. As such, thinking is a problem of symbolic computation which is carried out in short-term memory drawing upon ready-established concepts held in long-term memory. These concepts can be viewed as the 'pre-fabricated thoughts' provided by language. The mechanism is efficient because the labels for concepts can be utilised in the processing carried out by short-term memory whilst the large data structures which they represent can be stored in long-term memory. Therefore an idea which can be expressed in a single word rather than a lengthy description places fewer demands on expensive space in short-term memory. Although there is a concomitant additional burden on long-term memory, space there is virtually limitless and the net result is that 'a language user thinks most efficiently about those topics for which his or her lexicon has provided an efficient code'.

Historically, it seems that languages evolve in such a way as to remove the computational burden from short term memory to long term memory (Hunt and Agnoli, 1990). This offers an explanation for Zipf's (1935) observation that, the more frequently a word is used by a language group, the shorter it tends to be. In English there are any number of examples of technological innovations which, as they become integrated into the culture, are tagged with shorter labels than those they originally bore. 'Personal computer' becomes 'P.C.', 'motor carriage' contracts to 'car' and 'telephone' to 'phone'. Since a language user chooses words on the basis of their meaning rather than their length, it seems unlikely that the length of a word is the *cause* of its frequency of usage. It must be, therefore, that words are truncated because of an increase in the need to access the concepts they represent. Zipf terms this the 'Law of Abbreviation'. Brown and Lenneberg (1954) suggest that, in addition to the inverse correlation between word length and frequency of usage, there is also a direct correlation between frequency of utterance and the frequency of making the perceptual judgements necessary to select the word. They also propose a further correlation - that between the frequency of such perceptual discriminations and the 'accessibility' of the underlying concept. To summarise, shorter words tend to label more commonly encountered (and therefore more familiar) concepts so that '... more nameable concepts are nearer the top of the cognitive "deck"'.

Whilst the linguistic relativity hypothesis is compelling, experimental evidence in support of it is little more than flimsy. This is not surprising since we do not have direct access to the thoughts of the members of the linguistic groups we might wish to compare. All we can do is to look for differences in non-linguistic behaviour which might be attributable, in whole or in part, to language variation. This is clearly problematic. Whilst languages differ markedly and in ways that are amenable to analysis, the non-linguistic behaviour of their speakers can be attributed to the influence of a large number of inextricably inter-related variables.

Rosch (1974) describes five factors which she considers should apply to any domain used to test the effect of language on thought.

1) The languages under investigation must exhibit differences in their lexicon with respect to the domain.

2) There must be some objective way of measuring the features of the domain which are reflected in lexical differences.

3) The cultures whose languages are being investigated must have much the same experience of the domain so that differences in non-linguistic behaviour cannot be attributable simply to dissimilar encounters with it.

4) Measures of non-linguistic behaviour associated with the domain must be independent of language and not deduced from it.

5) Differences in non-linguistic behaviour should be explainable in terms of an interaction between linguistic and cultural variables.

Domains which satisfy all these criteria are difficult to identify. One which comes very close to doing so however, is the colour spectrum. Languages differ in the way they divide up this continuum; for instance, some have a single word for the colour range which, in English, is coded either 'green' or 'blue'. Brown and Lenneberg (1954) carried out a series of colour recognition tasks in which Ss were exposed to four colours and, after these had been removed, were asked to pick them out on a chart consisting of 120 different colours. Their results led them to propose that highly codable colours were more likely to be remembered and recognised. A measure of codability was arrived at by considering the length of the word used to describe the colour, the amount of hesitation with which Ss responded to the colour with its name and the degree of agreement amongst Ss as to what was the appropriate name for the colour. Thus a colour to which Ss responded quickly with a short name agreed upon by all was said to be highly codable.

Brown and Lenneberg's conclusions have been criticised by, for instance, Rosch (1974). She points to the results of Berlin and Kay (1969) which suggest that certain colours have more perceptual salience than others and, as a direct result, are more codable. Thus memory for colours may derive from perceptual characteristics which are universal rather than culturally specific linguistic factors. This highlights the difficulty in separating the effects of linguistic categories from the effects of those factors which led to the formation of those categories in the first place. A phenomenon with perceptual salience will itself always be especially amenable to reference.

A more recent attempt to identify Whorfian differences in colour perception was carried out by Kay and Kempton

(1984). Their Ss were English speakers and speakers of Tarahumara, a language which does not distinguish green and blue. Ss were asked to indicate which of three coloured chips in the blue/green range was most different from the remaining two. It was found that English speakers' perceptual judgements were distorted when chips were on the blue/green boundary. This, it was proposed, was due to the fact that, given a difficult discrimination task, Ss had resorted to judgements based on lexical rather than perceptual characteristics. On identifying the central of the three chips, the one 'most different' was the one with a different name. There was no such trap available for the Tarahumara speakers who therefore had a greater tendency to make the correct perceptual judgements. Kay and Kempton concluded that '... there do appear to be incursions of linguistic categorisation into apparently nonlinguistic processes of thinking, even incursions that result in judgements that differ from those made on a purely perceptual basis'.

Of the many attempts to confirm Whorf's hypothesis, most have been, at best, inconclusive. Bloom (1981) noted the absence of structures in the Chinese language which mark a counterfactual statement. In Indo-European languages, the counterfactual is signalled by using verb tenses which invite the listener to suspend reality and consider for a moment what might have been. In English, the subjunctive of the verb 'to be' is used as in 'If I were rich, I would own a yacht'. Chinese speaking subjects showed a marked resistance to move into the counterfactual realm stating that it was an 'un-Chinese' (or even a 'Western') way of thinking. Bloom proposed that, because of this limitation of the Chinese language, a counterfactual interpretation would require considerably more cognitive effort for Chinese speakers than it would for their English counterparts. His experimental evidence seemed to support this contention.

Bloom's results provoked a flurry of criticism. Au (1984) found fault with his experimental procedure, claiming that the stories used to test counterfactual understanding were not idiomatic and were therefore difficult for the Chinese speakers to understand regardless of their counterfactual content. She repeated Bloom's experiments with more idiomatic stories and concluded that Chinese speakers have no particular problem in interpreting counterfactuals. Liu (1985) carried out experiments which also seemed to support this claim.

In his response to Au's criticisms, Bloom (1984) points out the short-comings of attempts to verify Whorf's hypothesis using such studies as colour-naming where the possibility of substituting perceptual images for linguistic labels might preclude the emergence of Whorfian effects. He proposes that it is in the abstract realm that language is most likely to influence thought processes. Bloom is not alone in subscribing to this view. Lemon's (1981) research led him to suggest that, where concepts derive their meaning in verbal contexts rather than through direct sensory experience, language may affect categorisation. Cole and Scribner (1974, p 59) hold a similar opinion which they summarise thus:

> 'It may very well be that the "filtering effect" of language is greatest in respect to domains of phenomena that are definable, not in terms of physical properties, but in terms of attributes that are culturally specified. ... Or consider the area of ideology or theoretical work in general, where concepts largely acquire their meanings through their being embodied in explanatory verbal networks. It is here that language may play the greatest role in shaping the person's view of reality, in influencing his memory and thinking processes, and in contributing to his understanding or misunderstanding of other cultures.'

If, as seems likely, the influence of language on cognition is greatest where there is the necessity to acquire and manipulate abstract concepts, then mathematical thinking must surely be susceptible to language effects. Mathematical objects are abstractions. To deduce a result about triangles, we must work with neither an isosceles nor an equilateral triangle but with a prototypical triangle which is an abstract representation embodying the features shared by all triangles. Whilst visual images may help, they will not suffice. Any image of a triangle is the image of a particular triangle with properties which other triangles do not share. In any case, many (perhaps most) mathematical concepts do not lend themselves to any obvious form of imagery. Whilst we may learn something about triangles by manipulating mental 'pictures' it is difficult to accept that such a process would be productive in learning something about 'groups', for instance.

Classification, generalisation and abstraction are essential features of mathematical activity. Whilst mathematical facts may be discovered by experimenting with concrete materials, it is the generalisation of these facts which is the ultimate goal. Such generalisation involves abstraction and the formation of concepts which are independent of any concrete situation. Also important to mathematical thinking is an appreciation of how mathematical objects are classified into categories whose members share common properties. Furthermore, these categories are inter-related. A square has certain properties by virtue of the particular relationship of its sides and angles. However, a square is a rhombus and therefore has properties which are shared by all rhombuses. Rhombuses are parallelograms which also have their own characteristics, and so on. This view of the nature of mathematical abstraction is summarised by Dienes:

> '... it is a process of class formation. Abstract ideas are formed by classifying objects into classes through some common property which, it is discovered, is possessed by these objects. Generalisation is regarded as the extension of an

> already formed class and, therefore, it is more of a logical operation whereas abstraction is regarded as a constructive operation.' (quoted in Philp, 1973).

The current view is that language and mathematical learning and thought processes are inter-related, albeit in a complex way. This raises the obvious question - do different languages affect these processes in different ways? Before 1974, there was little interest in the implications of cross-linguistic factors for mathematics education. However, there was a growing awareness of the problems faced by children who were, for various reasons, forced to receive instruction in a language very dissimilar from their mother tongue. In 1974, researchers in linguistics and mathematics education gathered to discuss which difficulties faced by the learner of mathematics might be attributable to linguistic factors and to identify pedagogical approaches to overcome these difficulties. The introduction to the report of this symposium summarises the problem thus: 'Difficulties in the learning of mathematics thus depend on the language of learning, because different languages "support" mathematical concept formation, precision and systematisation in different ways.' (Nairobi, 1974.) The Nairobi symposium heralded a growth in research designed to discover where it is that a particular language may fail to support the processes which are essential to successful mathematical activity.

The most fundamental way in which a language may fail to make adequate provision for 'doing' mathematics is by not having a vocabulary which meets the subject's requirements, i.e. by having an insufficiently well-developed register of mathematical terms. (Indeed this has been reported as one of the reasons why many African countries have favoured English or French as the medium of instruction rather than native languages which lack the vocabulary necessary for technological subjects (Macnamara, 1967).) This can be solved by adding words to label the necessary concepts although the introduction of vocabulary for certain abstract concepts could prove problematic. For instance, certain African languages lack connectives and quantifiers such as 'all', 'some', 'only' and 'if' (CASME, 1975). The introduction of terms to convey these relational concepts is clearly more problematic than the introduction of nouns, even if the latter are abstract. However, languages are never static and the 'modernisation' of those which have had to accommodate a proliferation of technical terms has been engineered in an interesting variety of ways (see, for instance, Gallagher, 1969). The introduction of new vocabulary is a constant feature of any language and so there is no reason to believe that mathematical registers cannot be expanded using appropriate strategies.

One way in which it has been suggested that a language structure may create difficulties for thinking mathematically is in its lack of provision for classification hierarchies. Philp (1973) argues that the way in which objects are classified depends, at least in part, on the language used. From the results of research by Kelly (unpublished) in New Guinea, Philp proposes that: 'These data ... support the earlier findings that it is as if the accessibility of inclusive words in a language in some way affects and restricts the inclusiveness of classifications which the child is able to make.' He draws attention to the number of languages which incorporate linguistic classifiers thereby encouraging their speakers to classify other than according to Western logical categories. The suggestion is that the logic of mathematics is the logic of the Indo-European languages and it must not be assumed that this is necessarily the logic inherent in other languages.

Watson (1988) also subscribes to this view:

> 'The words and operations of mathematics, as a field, are in the history of mathematics, a discipline which developed in Indo-European cultures, using and developing the language games of Indo-European cultures in specific ways. Mathematics is not just a set of concepts that anyone can learn as easily as anyone else. It is a specific Indo-European product. Learning mathematics will be easier for children whose language is Indo-European.'

Haugen (1977) opposes the view that mathematics as we know it reflects the characteristics of any specific language or language group.

> 'The development of mathematics may be seen as an attempt to overcome the weaknesses of natural languages for the purpose of exact and elegant statement. The terminology of science has been successful to the extent that it has been able to surmount the limitations imposed by natural language and produce an inter- and supralinguistic language.' (Haugen, 1977)

He goes on to say that accounts of scientific theories are expressed in the language of mathematics and are therefore translatable into any language regardless of its grammar, syntax or phonology. However, accounts which utilise natural language are necessarily only approximations because the ideas involved cannot be expressed precisely in any language. This seems to miss the point, however. Whilst the concepts which underlie mathematical symbols are, in the main, precise and unambiguous, those concepts may well be more alien to some cultural groups rather than others.

In the day-to-day use of language, sentences derive their meanings, not only from the words of which they are constituted, but also from the context in which they are uttered. Natural language can therefore afford to be somewhat imprecise and even occasionally ambiguous. When 'talking mathematics', however, we are required to be precise, clear and unambiguous. Although we have a number of tightly defined terms whose use is confined to mathematical parlance, by and large it is through natural language that mathematical concepts must be conveyed and manipulated. The problem is exacerbated by the fact that mathematics takes common words and endows them with meanings very different from those attached to them in normal usage. For example, 'ring', 'function' and 'root' label concepts in mathematics which have little, if anything, in common with their more widely recognised referents. In so many cases, it is as though words used to label mathematical objects have been selected and assigned at random from the natural language lexicon. This cannot but create difficulties for the learner of mathematics who must assign a label, already laden with connotative and denotative meaning, to a new concept.

Whilst current interest in the interaction between language factors and mathematics education covers a wide spectrum, it is cross-linguistic issues which concern us here. For those researchers who have interested themselves in this field, the primary reason has been a concern for the many who must learn mathematics in a language which is not their mother tongue. A potential problem is a lack of proficiency in the second language but in this case the remedy has nothing to do with mathematics education. What are more insidious and intractable are the difficulties which result from the distance between the student's native language and culture and the language and culture of the teacher and curriculum designer. Western curricula and teaching methods are imposed in many countries of the world where they take little or no account of cultural norms and cognitive schemata implicit in the indigenous language. The problem is particularly acute for various African countries (Nairobi, 1974; Morris, 1978; Berry, 1985), but is by no means confined to Africa nor to developing countries. For instance, it has also been described for Aboriginal children in Australia (Crawford, 1988; Watson, 1988). Children learning mathematics in their native language may also suffer from the effects of a curriculum which is not in sympathy with their cultural and linguistic norms (see, for instance, Bernstein, 1971).

It is clear that mathematics curricula must be designed in such a way as to take maximum advantage of whatever cognitive schemata are established and should not assume that there is one universal way in which concepts are formed and manipulated. If methods of cognitive functioning vary across cultures then the implication is that curricula must vary also. Language is just one aspect of the child's culture but it is nevertheless an important one and a better understanding of its influence on cognitive processes must contribute to our understanding of how mathematics can be learned and taught more effectively.

Chapter 3

LOGICAL AND LINGUISTIC CONNECTIVES

In symbolic logic, a set of operators called 'logical connectives' allow the formation of compound truth-functional propositions from simple ones. These operators are binary in the sense that each is used to conjoin two (simple or compound) propositions. The truth value of the resulting proposition is completely determined by two factors - the truth values of the two component propositions and the particular connective used to conjoin them. In addition to the binary connectives, negation acts as a unary operator reversing the truth value of a proposition to which it is applied. If the proposition p is true, then its negation $\bar{p}$ is false and vice versa.

For two propositions p and q, there are four possible ways in which their truth values may be combined. Both may be true (the case denoted by TT), both may be false (FF) or one may be true and the other false (TF and FT). Any proposition formed by conjoining p and q will have a set of four truth values, one corresponding to each of the four combinations of truth values for p and q. This allows the possibility of $2^4 = 16$ different truth value sets. Taking the set of truth values of the compound proposition $p * q$ as the definition of the logical connective $*$, we can conclude that a maximum of 16 different logical connectives ($*_1, \ldots, *_{16}$) can be identified. These are defined in table 3.1 below.

Table 3.1: **Definition of logical connective $*_i$ by truth value set of $p *_i q$ (i = 1, ..., 16).**

p	q	$*_1$	$*_2$	$*_3$	$*_4$	$*_5$	$*_6$	$*_7$	$*_8$
T	T	T	T	T	T	F	T	T	F
T	F	T	T	T	F	T	T	F	T
F	T	T	T	F	T	T	F	T	T
F	F	T	F	T	T	T	F	F	F

p	q	$*_9$	$*_{10}$	$*_{11}$	$*_{12}$	$*_{13}$	$*_{14}$	$*_{15}$	$*_{16}$
T	T	T	F	F	T	F	F	F	F
T	F	F	T	F	F	T	F	F	F
F	T	F	F	T	F	F	T	F	F
F	F	T	T	T	F	F	F	T	F

For practical purposes, this set of 16 distinct logical connectives can be reduced (following Neisser and Weene, 1962). We can eliminate as trivial a connective which results in a proposition which is always true or always false no matter what the truth values of its components. Hence we may discount $*_1$ and $*_{16}$. Also $p *_3 q$ has exactly the same truth value set as $q *_4 p$. Since either of the component propositions may be labelled p or q arbitrarily, one of these connectives is redundant. For the same reason, so is one connective in each of the pairs $*_6$ and $*_7$, $*_{10}$ and $*_{11}$, $*_{13}$ and $*_{14}$.

The ten connectives which remain may be organised into five pairs where the truth values defining one member of the pair are the reverse of those defining the other. This is equivalent to saying that the compound proposition formed by using one member of the pair is simply the negation of that formed by using the other. Table 3.2 below shows, for each of these ten logical connectives, the defining truth value set, the usual symbolic representation of the compound proposition $p * q$ and the term used to refer to that connective.

Table 3.2: **Definitions of logical connectives.**

p	T	T	F	F	Symbolic representation	
q	T	F	T	F	of $p * q$	
$*_2$	T	T	T	F	$p \vee q$	Inclusive disjunction
$*_{15}$	F	F	F	T	$\overline{p \vee q}$	Joint denial
$*_4$	T	F	T	T	$p \rightarrow q$	Conditional
$*_{13}$	F	T	F	F	$\overline{p \rightarrow q}$	Exclusion
$*_{12}$	T	F	F	F	$p \wedge q$	Conjunction
$*_5$	F	T	T	T	$\overline{p \wedge q}$	Alternative denial
$*_8$	F	T	T	F	$p \veebar q$	Exclusive disjunction
$*_9$	T	F	F	T	$p \leftrightarrow q$	Biconditional
$*_6$	T	T	F	F	p	Affirmation of p
$*_{11}$	F	F	T	T	$\bar{p}$	Denial of p

From the table above it can be seen that only the unary operation of negation together with the connectives denoted symbolically by $\vee$, $\wedge$, $\rightarrow$, $\veebar$ and $\leftrightarrow$ are necessary to define all the relevant compound propositions. However, there is a great deal of redundancy even in this set of connectives. One which is clearly redundant is the biconditional since the truth value set defining this connective is the same as that for $\overline{p \veebar q}$. In fact, all of the propositions above could be expressed in terms of negation and, for instance, conjunction. For example, we could replace the proposition denoted by $p \veebar q$ by $(\overline{p \wedge q}) \wedge (\overline{\bar{p} \wedge \bar{q}})$ which has exactly the same truth value set. In a similar way, a binary proposition containing any of the ten connectives defined above can be replaced by an equivalent one which uses only negation and conjunction. (By 'equivalent' we mean one having the identical set of truth values.) We can condense the set even further, dispense with negation and use only the single connective defined by the truth value set for alternative denial. This connective is sometimes known as the 'Scheffer stroke function' and is denoted by $|$. Conjunction, for example, would then be expressed as $(p \mid q) \mid (p \mid q)$.

The examples above demonstrate that any significant reduction in the connectives symbolised results in an increase in the symbolic complexity of compound propositions. So, despite the inherent redundancy in the set, negation together with the five connectives described above are conventionally used to symbolise compound propositions.

In describing logical connectives, we have so far restricted ourselves to their symbolic representation. However, logic is used to establish the validity of certain types of argument. The rules which it provides allow us to assess whether the conclusion drawn from stated premises is consistent with those premises or whether there is some faulty step in the deductive process which claims to support the validity of the conclusion. It is clearly necessary for such arguments to be expressible, not only in symbols, but also in natural language. Furthermore, logical problems frequently involve an assessment of the truth or falsity of a compound proposition when the truth values of its components are known. In order to carry out such an evaluation, it is necessary to identify the particular connectives used. If the argument is expressed in language, this information must be deduced from the linguistic content of the proposition.

The translation of a compound proposition from symbolic to linguistic form is not a particular problem. There are English language items which, it is generally agreed, convey the sense of each of the five logical connectives. A proposition can be negated by inserting 'not' in the appropriate position or by using 'It is not the case that ...' as a prefix. Inclusive disjunction can be expressed using '... or ... or both', conjunction by '... and ...', the conditional by 'if ... then ...', exclusive disjunction by '... or ... but not both' and the biconditional by '... if and only if ...'. Thus given the two propositions: *p*: Tom is American and *q*: Max is a student, the following are the linguistic expressions for the propositions which can be formed by using negation or by combining p and q using each of the logical connectives $\vee$, $\wedge$, $\rightarrow$, $\veebar$, and $\leftrightarrow$.

$\bar{p}$: Tom is not American.

$p \vee q$: Tom is American or Max is a student or both.

$p \wedge q$: Tom is American and Max is a student.

$p \rightarrow q$: If Tom is American then Max is a student.

$p \veebar q$: Tom is American or Max is a student but not both.

$p \leftrightarrow q$: Tom is American if and only if Max is a student.

In this way we may map logical connectives onto linguistic connectives. However, because this mapping is not one-to-one, a difficulty arises when we attempt to map linguistic connectives to their logical equivalents. Specific difficulties in identifying the logical connective in a compound proposition arise from the ambiguity of the linguistic connectives which are commonly taken to be equivalent to the logical forms. A fundamental problem is that logical connectives are only appropriate between statements to which truth values can be assigned and where the meaning of the resulting statement is not affected by context. Language connectives have a much wider use. A more detailed examination of the problems associated with the individual connectives will serve to highlight the mismatch between the linguistic forms and the formal logical operators with which they are frequently identified.

Conjunction

Although 'and' can be used to link two statements to form a compound statement whose truth or falsity can be deduced according to the rules of truth functional logic, there are other statements containing 'and' which cannot be so interpreted. It can be argued that a sentence like 'John and Mary are students' is simply an abbreviation of the conjunction of the two propositions 'John is a student' and 'Mary is a student'. However, no such argument can be sustained for 'John and Mary are colleagues'. Copi (1986) maintains that such propositions must be regarded as simple (i.e. making only a single statement) rather than as compound propositions with two components.

Logical conjunction is symmetric in the sense that it is commutative; $p \wedge q$ is equivalent to $q \wedge p$. However 'She fell down and broke her neck' conjures up a very different scenario from 'She broke her neck and fell down'. Here again the interpretation of 'and' is not identical to logical conjunction. Staal (1966) distinguishes two meanings for 'and'. One is identified with logical conjunction and the other is asymmetric and means 'and then' or 'and consequently'. Dik (1968) defends the view that the second interpretation of 'and' is arrived at on the basis of non-linguistic knowledge about the two events and is not part of the semantic content of 'and'. Lakoff (1971) argues that, with the asymmetric 'and', the truth of the first member of the pair of propositions must be presupposed for the second to have any meaning. Denying the first assertion renders the compound proposition nonsensical and does not therefore invite an analysis based on logical rules.

Despite the fact that 'and' does not always have an interpretation which can be identified with logical conjunction, the divergence between the logical and linguistic forms is less than is the case for other connectives. It is for this reason that Strawson (1952) proposes that all connectives be defined in terms of negation and conjunction.

Disjunction

The linguistic expression of disjunction is usually given by 'or' or 'either ... or'. As we have seen, in symbolic logic two types of disjunction are distinguished - the exclusive and inclusive forms. The only difference between these is in the truth value of the compound statement when both components are true. For inclusive disjunction the TT case is true whereas for exclusive disjunction it is false. The linguistic connective 'or' is ambiguous in indicating which of these alternative meanings is to be conveyed. Where it is important that the disjunction be interpreted inclusively, this sense can be made explicit by adding 'or both'. Where it is to be given the exclusive interpretation, the phrase 'but not both' can be affixed. However, 'or' is frequently used without either of these qualifiers and, when this is the case, it is not clear which interpretation is intended, especially if no contextual cues are available.

In logic, 'or' is conventionally interpreted inclusively whereas, when 'or' is used in natural language, it is claimed that it is usually taken to represent exclusive disjunction (Lakoff, 1971; Sternberg, 1979; Newstead and Griggs, 1983). On the other hand, Pelletier (1977) is of the opinion that in English 'or' always represents inclusive disjunction and that the exclusive interpretation is the result of an individual's judgement that it is impossible or unlikely for both disjuncts to be true simultaneously. Gazdar (1979, p 78) is even more emphatic that the inclusive interpretation is basic and claims 'that there is no clear evidence to the effect that exclusive disjunction has ever been lexicalised in any language'.

There are 'or' statements which are clearly to be given an inclusive reading, for instance 'Applicants for the job must have a degree or three years relevant experience'. There are others which strongly suggest that they are to be interpreted exclusively such as 'Chelsea or Liverpool will win the FA cup'. However, it is debatable which of the two interpretations is basic and therefore how an ambiguous 'or' statement will be interpreted. Damarin (1977) found that the inclusive interpretation of 'or' was favoured amongst pre-service elementary teachers when interpreting statements about mathematical items. However, there was also a marked tendency for her Ss to interpret 'or' statements in this context as if they were conjunctions.

An additional ambiguity of the word 'or' in natural language is that it is sometimes used where the logical sense of 'and' is intended as in 'I eat meat or fish'.

Conditional and Biconditional

A number of interpretations have been found to exist for a statement expressed in the form 'if *p* then *q*' of which one is equivalent to the logical interpretation. Taplin (1971), in an investigation of adults' interpretations of conditional sentences expressed in 'if ... then' form, found that fewer than 50% evaluated the conditional in any truth functional manner. Of those who did, however, the most common truth functional interpretation was that which actually corresponds to the biconditional. Fillenbaum (1976) found that conditional promises and threats were particularly prone to a biconditional interpretation. Indeed the force of a conditional threat or promise is dependent upon the addressee assuming that a false antecedent guarantees a false consequent. Furthermore, the biconditional is clearly the intended interpretation in many 'if ... then' statements. On hearing 'If you don't take an umbrella, then you'll get wet', the addressee will assume that if he does take an umbrella (and uses it appropriately) he won't get wet. On the other hand 'If we go in August, then the weather will be hot' does not seem to suggest that the weather is hot only in August and therefore invites an interpretation equivalent to the logical conditional.

There is no difficulty in interpreting a statement of the form 'if *p* then *q*' when the antecedent, *p*, is true. The statement would normally be regarded as true when *q* is true and false when *q* is false. It is the case of a false antecedent which is problematic. Wason and Johnson-Laird (1972) argue that the antecedent in an 'if ... then' statement is regarded as an indication of presupposition. The listener must assume the truth of the antecedent in order to interpret the sentence and will regard it as null and void if the antecedent turns out to be false. They therefore propose a 'defective' truth table for a conditional expressed in natural language, that is, one with an entry of 'irrelevant' for the two cases where the antecedent is not true.

Paris (1975, pp. 88 - 89) proposes that the defective truth table for conditional statements is the result of an interpretation which assumes a causal relationship between antecedent and consequent and which therefore dictates a symmetry between the two components. Although the FT case is true logically, it is not considered true causally. The failure of the logical conditional to account for comprehension of 'if ... then' statements is attributed to the fact that it does not take account of the semantic relationship between the two component propositions. As Comrie (1986) points out - there is a fundamental difference between conditionals in logic and in natural language. In logic there is no necessity for the antecedent and consequent to be related, causally or otherwise. A conditional proposition such as 'If Rome wasn't built in a day then London is north of the equator' is perfectly acceptable in logic but would evoke some surprise were it stated in normal conversation.

It has been shown (Johnson-Laird and Tagart, 1969) that understanding of a conditional depends, to some extent, on the way in which it is expressed. Their Ss were more likely to treat the equivalent disjunctive form 'Either there isn't *p* or there is *q* (or both)' as a logical conditional. However this form was found to take longer to process and tended to produce other diverse and unstable interpretations. On the other hand 'if *p* then *q*', though faster and more stable in its interpretation, was less frequently interpreted as the logical conditional.

The biconditional, expressed linguistically using 'if and only if', has attracted less attention from researchers than have conjunctive, disjunctive and conditional language connectives. Lemmon (1965, p. 28) suggests that it is 'of rare occurrence in ordinary speech'. Since, as we have already seen, there is a tendency for conditional statements to interpreted as the logical biconditional, we might expect this tendency to be even more pronounced in interpreting statements containing 'if and only if'. If this is the case, a reasonable match might be expected between the linguistic and logical forms of this particular connective even though the language form is rarely used.

Gazdar (1979, chapter 4) claims that, of the 16 possible logical connectives, only two, conjunction and inclusive disjunction, are lexically encoded and that this is probably a universal feature of all languages. He suggests that there is some feature of the human mind which demands a convenient expression for these logical forms and that conjunction in particular is an unlearned innate human concept. He precludes 'if ... then' from having truth functional status because context affects its truth conditions.

There is a fundamental difference between reasoning carried out according to the rules of symbolic logic and the sort of reasoning that takes place in day-to-day natural language exchanges. In the former nothing may be assumed other than the information contained within the propositions stated. In general conversation however, this is not the case. If A says to B 'If you don't take an umbrella, then you'll get wet', B will reason with a variety of inferences and assumptions not conveyed by the statement alone. For instance, he well almost certainly infer that it is raining (or that it is about to rain). Furthermore, A can reasonably assume that B will infer that it is raining (or about to rain) so that this information need not be stated explicitly. He will also assume that B knows that, in normal circumstances, he can avoid getting wet in the rain by using an umbrella appropriately. It may be that A really meant his statement to be a true conditional (rather than a biconditional) because, say, A knew that B's route would take him past some children throwing buckets of water at passers-by. If B subsequently got drenched because he had failed to use the umbrella at the appropriate time, he could not strictly accuse A of uttering a falsehood. However he could justifiably claim that he had been misled and intentionally so.

There are various conventions which are understood by language users to apply in language exchanges and it is because of these that speakers and sometimes writers can afford to be imprecise and potentially ambiguous.

The potential ambiguity in 'On Monday I shall be attending a conference in Paris or visiting a friend in Birmingham' is resolved by the shared knowledge that the two constituent propositions are unlikely to be true simultaneously and that 'or' should therefore be interpreted exclusively.

Grice (1975) subsumes what he believes to be the conventions of normal language utterances under what he terms the 'cooperative principle'. Speakers try, as far as they can, to be truthful, relevant and informative, and listeners assume that speakers will be so. This allows much to be inferred that is not made explicit. He proposes 'conversational implicatures' - information not stated but understood by a listener by virtue of his assumption that the speaker is being co-operative. The discrepancy between linguistic and logical statements is viewed as a consequence of the fact that a natural language statement conveys more than its analytic meaning.

Geis and Zwicky (1971) explain the mismatch between linguistic and logical connectives by suggesting that certain types of compound propositions expressed in natural language have 'invited inferences'. For instance, 'if *p* then *q*' invites the inference 'if not *p* then not *q*' and hence a tendency to 'perfect conditionals to biconditionals'. They suggest that invited inferences are not part of the meaning of 'if' because they can be cancelled without apparent contradiction. For instance, 'If I study hard then I'll pass my exams and if I don't study hard I may or may not pass my exams' is a perfectly acceptable statement. Another invited inference proposed is 'but not both' in the expression '*p* or *q*'. This can also be countermanded by adding 'or both' without producing a contradiction.

Braine and Rumain (1988) suggest that invited inferences are not part of the lexical entry for particles such as 'if ... then' and 'or' but are invoked whenever considered relevant. In a later paper also co-authored by Braine (Braine and O'Brien, 1991), the 'pernicious ambiguities' in such expressions as 'the meaning of "if"' are pointed out. Such expressions could refer to understanding of the particle in some particular context or to the meaning encoded in semantic memory (that is, the lexical entry). It is pointed out that construal in ordinary comprehension takes account of factors over and above the content of the lexical entry. These include the context, the plausibility of possible construals, knowledge of the speaker's motives and intentions and other general knowledge which may be relevant to the discourse. Reasoning proceeds from the integration of all such factors.

The above makes clear the necessity to distinguish logical and linguistic connectives and yet this distinction is frequently neglected. Dik (1968, p 259) makes the point:

> 'Even in the context of the description of natural language, the co-ordinators are often classed as "logical connectives" and treated as if the "meanings" attached to such connectives in logic are directly relevant to the semantic description of natural language. ... It is no disparagement of logic (nor of natural language) when I stress the undeniable fundamental differences between the two. Again and again, students of natural language and of logical systems have rightly noted that the two are objects of fundamentally different natures, both in their aims and in their internal properties.'

In mathematics, the only acceptable form of reasoning is that which conforms to the laws of symbolic logic. Skemp (1987, pp 170 - 171) emphasises the importance of logical understanding for successful mathematical activity. This he considers is 'evidenced by the ability to demonstrate that what has been stated follows *of logical necessity*, by a chain of inferences, from (i) the given premises, together with (ii) suitably chosen items from what is accepted as established mathematical knowledge (axioms and theorems). This involves analysis, and the construction of chains of logical reasoning to produce what we call demonstrations or proofs.'

Correct inferences depend upon knowing what can be deduced from the logical relationships between statements about mathematical objects. For instance, it is important to be able to distinguish conditional and biconditional propositions; given a mathematical statement of the form 'if *p* then *q*', the invited inference 'if not *p* then not *q*' will lead to errors of deduction. Furthermore mathematical objects are abstractions and therefore the contextual cues available in everyday language exchanges are not necessarily available to aid interpretation. Evans' (1982) assertion that '"If it is a dog then it is an animal" obviously does not entail the converse.' is true enough but it presupposes that the addressee is already familiar with the relationship between dogs and animals. In this case the statement is not informative and is therefore redundant. In mathematics, the interpreter of the statement 'If it is a rhombus then it is a parallelogram' may well be in the position of knowing nothing about the relationship between the two quadrilaterals other than that which the statement conveys. In particular, he may have no idea as to whether or not the statement may be taken to entail its converse, especially when his experience with natural language uses of 'if ... then' allows that it may or may not.

Politzer (1986) suggests that the lack of congruence between logical and linguistic connectives may go some way towards explaining poor performance in mathematics.

> '... one of the sources of the low mathematical achievement of students who perform normally on nonscientific subject matters consists of an

> interpretation of the scientific language based on the pragmatic laws of language rather than on the logic conventions. Such a propensity to use pragmatic conventions would affect inferential activities but have little effect on algorithmic activities (such as calculating fractions, solving equations, etc.) hence the distinction that is well-known to Mathematics teachers between being good at Geometry (that requires the most use of hypothetico-deductive thought) and good at Algebra (that requires the most use of algorithms).'

Zepp (1987b) set out to discover whether students who had studied mathematics for several years might use a different set of logical principles for reasoning mathematically than that used for everyday reasoning. For instance, would these students avoid the error of assuming that the converse of a conditional statement held and would they be less likely to assign 'or' its exclusive interpretation in mathematical contexts? He concluded that mathematical content in reasoning problems added to the difficulty of formal logic. Students who performed adequately when a task was free of mathematical content resorted to confused guessing when the same task was presented in a mathematical context. He concluded that many students apply a method of reasoning in concrete situations which, although not the logic of mathematics, is applied fairly consistently. When asked to apply reasoning with mathematical objects, the students did not adopt the correct logic nor did they apply their own 'everyday' logic. The mathematical content simply confused them and they adopted inconsistent strategies.

Given that mathematical reasoning depends critically upon understanding logical relationships, it is not surprising that linguistic connectives are considered important devices for learning and doing mathematics. Imperfect as they are, it is they that must bear the burden of expressing logical relationships. Using children and adults as Ss, Johansson (1977) compared the level of mastery of the words 'and' and 'or' with performance in logical tests of conjunction and disjunction. He found a correlation between the level of understanding of the words and the quality of performance on the logical tests. Dawe (1982) investigated the ability of bilingual children to reason deductively in mathematics. He found that the single most important factor differentiating English children's performance on a test of mathematical reasoning was a knowledge of linguistic connectives and that this factor also separated high from low achievers amongst bilinguals. His conclusion was that development of the ability to use linguistic connectives in their logical sense was important for success in mathematics.

The foregoing account discusses the mismatch between logical and linguistic connectives expressed in English. One might reasonably ask to what extent this is generalisable to other languages where linguistic connectives may well be expressed very differently. 'Different languages, including European languages, vary enormously in the richness of their logical connectives vocabulary.' (Wilson, 1984). In the same paper, Wilson also points out the rather surprising fact that there are Commonwealth languages which lack such connectives (see also CASME, 1975). It is reported (Cohen, 1977, p. 94) that, in Chinese, there are no words equivalent to 'and' and 'or'. The conjunction of two statements *p* and *q* is expressed as 'there is *p*, there is *q*' and disjunction as 'if not *p* then *q*'. Moore (1982) reports that the Navajo language has no standard word for 'if', a fact which he associated with evidence that hypothetical thinking is not readily accepted by some Navajo speaking students. He also felt that this might go some way in explaining the poor performance of Navajo students on Piagetian conservation tasks.

Comrie (1986) attempted a cross-linguistically valid characterisation of natural language conditionals. He hypothesised that, if a language had any conditional construction at all, then it would have one which is equivalent to the logical conditional and not only one which is to be construed as the biconditional. Marking of the antecedent clause seems to be fairly universal although there are exceptions, for instance Mandarin, where most conditional statements are ambiguous and rely on context to supply the correct interpretation. Marking of the consequent is less common and often optional as with the English 'then'. It is also usual for the antecedent to precede the consequent and in some languages this is obligatory. Although there appear to be some universal features of conditionals, it has been found that their expression can differ substantially from the English form 'if ... then' (Traugott et al, 1986, p. 5).

The ambiguity surrounding the word 'or' in English might be resolved, one would suppose, were the inclusive and exclusive forms to have distinct linguistic expressions. Attempts to identify languages which distinguish these forms have not been wholly successful. Copi (1986, p. 272) claims that Latin makes this distinction as does Quine (1974, p. 12). However, Dik (1968, p. 274) denies that the two forms in question parallel the logical distinction. The Kpelle in Liberia were also thought to have a separate linguistic expression for the two forms (Gay and Cole, 1967). Zepp (1989), having examined many world languages, has failed to find one which has an unambiguous word for exclusive as distinct from inclusive disjunction. He points out that this seems to indicate a fundamental difference between the foundations of mathematics and that of language development.

Evidence suggests that there is considerable variation in the provision made by different languages for encoding logical connectives. If we accept the weaker form of Whorf's hypothesis we might speculate as to the likely effect of such differing linguistic provision on the formation of the associated concepts. For instance, were a language to distinguish inclusive and exclusive disjunction more effectively and habitually than English, might it not be the case that the speaker's attention has been focussed in such a way that an appropriate cognitive schema is more readily available to him? If a language forces its speakers to distinguish conditional from biconditional statements, would they be less likely to make the error of reasoning from the converse? Are some or all of the logical concepts 'nearer the top of the cognitive deck' for some language users rather than others and, if this is so, can this be attributed to language factors? Or, are the problems resulting from the mismatch between logical and linguistic connectives universal features of all languages?

Chapter 4

RELEVANT RESEARCH FINDINGS

There are several areas of research pertinent to this study. Amongst these are investigations which have attempted to elucidate how linguistic connectives are understood and how this understanding (or lack of it) influences success in deductive reasoning. Of the various forms of reasoning, it is propositional reasoning which depends critically upon the logical relationships which the connectives define. Therefore we shall not review the extensive research on reasoning with categorical syllogisms where successful inference depends upon understanding quantifiers such as 'some' and 'all'.

The 'classical' view of human concepts is that they can be defined according to logical relationships between criterial features (for example, see Bourne, 1974). For instance, a particular concept might be defined as all those items possessing attribute A together with either B or C, i.e. the logical conjunction of A with the disjunction of B and C. Furthermore, the formation of concepts was thought to occur through a process of discrimination learning by which the criterial attributes and their rule of combination were identified and associated with positive examples of the concept. This view spawned a body of research which investigated the relative difficulty of learning different concepts using non-verbal stimulus materials. This type of investigation virtually ceased in the late 1970s when the classical view was replaced by theories based on schemata (see chapter 2) and prototypes (Rosch and Mervis, 1975). However, these 'concept identification' studies are of interest because they have shown that, given certain types of stimulus elements, some logical concepts are easier to identify than others.

Also of interest are attempts to establish whether the results obtained in the areas outlined above generalise to speakers of languages other than English, particularly non-European languages.

According to Piaget (1957), truth functional aspects of cognition undergo a gradual development reaching maturity when an individual is approximately 11 - 14 years of age, a stage which he terms 'formal operations'. This is the stage at which the capacity for the full range of logical deductions is attained. Our prime interest is in the relation between logic and language factors in those who have reached Piaget's stage of formal operations. We shall therefore do no more than make passing reference to some of the many studies which have examined developmental factors in the interpretation of linguistic connectives and in deductive reasoning. Should this project reveal differences in the performance of adults in different language groups, the burden of future research will be to elucidate the underlying developmental patterns.

Concept identification studies

A number of studies have focussed on concept learning behaviour. In these, typically the S is presented with a sequence of stimulus materials which vary along a number of dimensions, e.g. colour, form, size. The experimenter has in mind a subset defined by, say, one value (termed an 'attribute') of each of two dimensions and a rule for combining these values. Thus for any particular concept, not all dimensions are relevant.

In general, the rule for combining attributes is such that the presence or absence of either or both defines whether or not a particular stimulus is an example of the concept. For instance, if the relevant attributes are A and B, a conjunctive concept has as its only exemplars those stimuli which exhibit both attributes and inclusive disjunction is exemplified by those stimuli which display either or both of the attributes A and B. Stimuli which are examples of exclusive disjunction possess attribute A or attribute B but not both. The conditional concept 'if A then B' is exemplified by all stimuli except those which exhibit A but not B and the biconditional by those with both attributes or neither.

After presentation of each stimulus the S is required to identify it as an example (positive instance) or non-example (negative instance) of the concept and is then told whether or not [1]he is correct. Trials are repeated until the S makes a sequence of correct responses indicating that he has identified the essential features distinguishing examples of the concept from non-examples. The length of this sequence varies although 16 consecutive correct responses is often the criterion. Sometimes previously presented stimuli are allowed to remain in view sorted into their correct response categories, otherwise they are removed. Clearly the former paradigm simplifies the S's task somewhat by alleviating the memory burden associated with the necessity to remember the information obtained from previously presented stimuli. In order to be able to distinguish examples of the concept from non-examples, the S must identify the relevant stimulus dimensions and decide which are the criterial attributes. He must also identify the rule

[1]Throughout this thesis, 'he' has been used to refer to an arbitrary individual whose gender is irrelevant. This is for no reason other than to avoid repetition of the cumbersome construction 'he or she'.

governing how these attributes are combined. In some investigations the S is told the relevant attributes and must determine the rule of combination. In others the rule is known and the relevant attributes must be identified. The former is often termed 'rule learning' and the latter 'attribute identification'. 'Complete learning' is used to describe concept identification tasks where neither the attributes nor the rule of combination is known.

A consistent finding of research into conceptual rule learning is that Ss find some rules easier than others. Neisser and Weene (1962) report that their adult Ss found conjunction, inclusive disjunction and the conditional easier than exclusive disjunction and the biconditional. Within the first group of concepts, the conditional was more difficult than the other two. A practice effect was noted - Ss' performance improved when the same concept was presented a second time.

Neisser and Weene explain their findings by proposing that the conceptual rules tested form a hierarchy of successive complexity if the operations of negation, conjunction and disjunction are regarded as primitives. Level I concepts are those which are defined by the presence or absence of just one dimensional attribute. Concepts classified at Level II are those which can be expressed as combinations of two Level I concepts using the primitive connectives. Hence conjunction, inclusive disjunction and the conditional (interpreted as $\bar{p} \vee q$) are Level II concepts. Level III concepts are combinations (using primitive connectives) of those at Level II and therefore include exclusive disjunction $((\bar{p} \wedge q) \vee (p \wedge \bar{q}))$ and the biconditional $(p \wedge q) \vee (\bar{p} \wedge \bar{q}))$. In an attempt to explain why higher level concepts might be more elusive, it was suggested that a contributory factor might be the difficulty of formulating them verbally. These concepts could therefore be less familiar so that Ss might find them more difficult to keep in mind.

Whilst Neisser and Weene's findings lend some support to their hypothesis, there remains the question as to why it is that conjunction and inclusive disjunction should be proposed as the primitive connectives rather than some other subset of the ten distinct connectives tested. The authors themselves point out that 'the hierarchy is merely a tautology until it is related to empirical findings' but do not explain on what basis those particular connectives were selected in the first place. Even if it is accepted that conjunction and inclusive disjunction are the primitive connectives, there is a problem with the allocation of certain concepts to Level II. For instance, the conditional is expressed as $\bar{p} \vee q$ and assigned to Level II on this basis. However $(p \wedge q) \vee (\bar{p} \wedge q) \vee (\bar{p} \wedge \bar{q})$ is an alternative expression for this connective in terms of conjunction and inclusive disjunction. Were the conditional to be interpreted in this form, it would then be assigned to Level IV.

Haygood and Bourne (1965) conducted a series of concept identification experiments designed to discover whether the differential ability to learn conceptual rules might be eliminated if Ss were not required to identify relevant attributes as well as the rule of combination. Only Neisser and Weene's (1962) Level II concepts were used so that exclusive disjunction and the biconditional were not tested. Of the three concepts conjunction, inclusive disjunction and the conditional, the latter was again found to be more difficult than the other two. Two factors are offered by way of explaining the greater difficulty of identifying the conditional concept. With two attributes on each of three dimensions, more stimulus elements fall into the truth table category FF (i.e. lack both relevant attributes) than any other truth table category. Hence the conditional has as positive instances a larger proportion of the stimulus population than either of the other two concepts. Furthermore, unlike conjunction and inclusive disjunction, the conditional requires the TT cases (both attributes present) to be placed in the same response category as the FF cases (both attributes absent). The set of positive instances of the conditional concept is therefore large and non-homogeneous. An additional factor suggested as contributing to the relative difficulty in identifying the conditional concept is the fact that the non-commutativity of the rule requires Ss to distinguish FT and TF cases and assign them to different categories.

In a subsequent experiment, Haygood and Bourne (1965) tested Ss on four concept identification tasks where the concepts tested were the conditional, biconditional, conjunction and inclusive disjunction. Biconditional and conditional concepts were more difficult than conjunction and disjunction. Again the authors propose the assignment of TT and FF stimuli to the same class as contributing to the greater difficulty of the conditional and biconditional concepts.

The presence/absence of two relevant attributes partitions the stimulus population into four classes. Haygood and Bourne (1965) suggest that, when relevant attributes are known, Ss identify conceptual rules by acquiring information as to the correct assignment (example/non-example) of each of the four classes. The implication is that Ss employ some kind of informal version of the truth table as a mediating device in the identification process. It is proposed that differences in the ease of identifying different concepts may be due to differing amounts of experience with the concepts or to unexpected assignments of certain truth table categories. However, with training, these differences tend to disappear so that Ss develop an appropriate strategy for identifying the concepts with equal facility.

In a later paper, Bourne (1970) describes an experiment to investigate the relative difficulty of identifying four conceptual rules. The order of difficulty of these rules, from least to most difficult, was conjunction, inclusive disjunction, the conditional and biconditional. Examination of the truth table categories which Ss assign to the incorrect response category reveals that these assignment errors vary from concept to concept. In the process of identifying inclusive disjunction, errors are most frequent in the mixed truth table categories TF and FT, i.e. with stimuli possessing one or other of the two relevant attributes but not both. For the conditional the TF category was the one suffering from the most incorrect assignments, whereas for the biconditional, stimuli with neither attribute (FF) were the most problematic.

With conjunction, errors were much the same for all the truth table categories. The assignment errors for disjunction, the conditional and biconditional are to a large extent those which would result from the identification of each concept as conjunction. Bourne therefore concludes that some of the differences in rule difficulty may be accounted for by a predilection for conjunctive concepts which must be suppressed when identifying other concepts. However, this seems a rather unsatisfactory explanation given that the biconditional differs from conjunction in the assignment of only one truth table category (FF) and yet it was the most difficult of the four concepts tested.

There are substantial inter-rule transfer effects when Ss are presented with a sequence of problems requiring the identification of different concepts (Haygood and Bourne, 1965; Bourne, 1970). Again, some form of mediating truth table is offered by way of explanation. Once learned, the truth table strategy can be applied to any of the 16 possible truth functional concepts and no concept is then more difficult to identify than any other. All require information concerning the correct allocation of each of the four truth table categories and Ss should therefore be able to identify any of these concepts after at most one assignment error in each of the four categories.

Hiew (1977) investigated whether appropriate training would facilitate acquisition of the truth table strategy and thereby enhance performance on concept identification tasks. The training consisted of 12 concept identification problems covering the four concepts conjunction, inclusive disjunction, the conditional and biconditional. For one group of Ss, the three problems testing each concept were presented in random order with the concept changing after each problem. For another group, the three problems testing the same concept were presented consecutively, the concept changing after the third, sixth and ninth problems. After the training sequence, all Ss were tested on three tests of generalisation - a different truth functional concept, a truth functional concept involving three (rather than two) attributes and a truth functional concept involving two attributes where the stimulus elements were drawings of faces rather than geometric shapes. The null hypothesis was that Ss for whom the training stimuli were presented in random order would be more likely to learn the superordinate truth table strategy and would therefore perform better on the three generalisation problems. This was indeed the case. Further support for the mediating truth table is given by the results of Dodd et al (1979). Truth table pretraining was found to improve performance in identifying concepts, presumably because it provided Ss with an appropriate strategy for all subsequent tasks.

Yet another explanation for differential rule difficulty is offered by Denny (1969). Where previously presented stimuli are not available to the S, the information obtained from them must be held in short-term memory. It is suggested that there is an interaction between the processing necessary for the more complex conceptual rules and memory for previous stimulus instances. Previous stimuli were more readily forgotten if the rule was more complex. Where the short-term memory burden was reduced by allowing previously presented stimuli to remain in view, sorted into their appropriate response category, differential rule difficulty was greatly reduced. Denny concluded that the more complex conceptual rules are more sensitive to the effects of short term memory burden. However, no explanation is offered as to what are the characteristics of conceptual rules which make some more complex than others. It does not seem that the greater ease in identifying certain concepts can be attributed simply to greater familiarity with them. Were this to be so, one might expect that the relative rule difficulty would be less pronounced for young children. However, it has been shown that disjunctive concepts are more difficult to identify than conjunctive ones even for children as young as 5 or 6 years old (King, 1966; Snow and Rabinovitch, 1969).

Whilst the superordinate truth table may explain the performance of Ss with some pretraining on concept identification problems, it does not explain why naive Ss so consistently find some concepts easier than others nor why the same concepts always cause the most difficulty. Bourne (1974) proposes a model based on inference operations which correctly predicts the observed order of increasing difficulty: conjunction, inclusive disjunction, the conditional and biconditional. The model assumes a predisposition towards conjunctive categories, reported by Bruner et al (1956) to be a feature of cognitive activity in Western culture. Faced with a concept identification task in which the relevant attributes are known, the naive S initially responds on the basis that the rule is conjunction. He therefore assigns TT stimuli to the positive response category and the remaining classes (TF, FT and FF) to the negative response category. During the course of identifying the concept, the S receives informative feedback which will correct errors of allocation if the concept is not conjunction. The inference model assumes that, should these lead the S to change the response category of FF stimuli, then he will similarly change the response categories for TF and FT instances. He will also attempt to maintain different response categories for the FF and TT categories; a change in one of these will cause a change in the other. Conjunction is therefore easy because it is consistent with the S's pre-experimental bias. Inclusive disjunction is more difficult because it requires the TF and FT categories to be assigned differently from FF stimulus elements. It is proposed that an increment in difficulty occurs when any of the inference operations are violated and that the magnitude of this increment is proportional to the number of different stimulus elements whose response category runs counter to a given inference. This is dependent on the proportion of stimulus elements in each truth table category and hence on the number of attributes on each relevant dimension.

There is some evidence that conceptual rule learning performance is not independent of stimulus materials (Bourne, 1979). Certain pairs of attributes tend to integrate. Their combination presents itself as a unit and is therefore more conducive to a conjunctive interpretation. It is suggested that, on the dimension of number (of geometric forms), 'oneness' is likely to integrate with any other dimension so that, for example, the attributes 'one' and 'circle' are likely to be viewed conjunctively. On the other hand, some combinations of attributes are more difficult to integrate and tend to be viewed separately thereby encouraging a disjunctive interpretation. Colour and form seem to fall into this category.

In a later paper (Ketchum and Bourne, 1980), it is proposed that 'the local integrality' hypothesis described above is too simple to explain rule bias. It is suggested that certain attributes such as colour are perceptually more salient. The S therefore tends to respond positively to an instance which is positive with regard to colour regardless of the presence/absence of the other relevant attribute. The tendency to respond to a highly salient attribute as positive results in a propensity for a disjunctive sorting. If no perceptually salient feature is relevant, Ss will tend to sort stimulus elements conjunctively.

It is clear from the foregoing account that, in concept identification tasks, some concepts are easier to identify than others. A consistent finding is that conjunction and inclusive disjunction are identified more successfully than exclusive disjunction, the conditional and biconditional. The reason for this differential rule difficulty is far from clear however, and none of the explanations offered seems entirely satisfactory.

Interpretation of linguistic connectives

A number of studies have attempted to elucidate how linguistic connectives are interpreted in a particular context and to compare this interpretation with the corresponding logical form. In general the experimental tasks involved are such that they elicit some or all of the truth table which a S associates with a particular linguistic connective. Usually this is achieved by providing the truth values of two simple propositions *p* and *q* and requesting Ss to assess the truth value of compound propositions such as 'if *p* then *q*'. Alternatively, the truth value of the compound statement is given and the S's task is to provide possible truth values for the simple components. Errors are taken to have occurred when the truth values elicited do not correspond to those which define the corresponding logical connective.

Bart (1974) investigated the understanding of linguistic connectives in adolescents aged 14 - 19 years. Test items were compound propositions consisting of two simple components conjoined with the usual linguistic form of logical conjunction, disjunction and the conditional. Ss were asked to assume the truth of just the first component and to evaluate the compound proposition as always true, always false or neither. Propositions containing negation were found to be more difficult than those which did not. Also conjunction was slightly easier than disjunction and the conditional was more difficult than either.

Damarin (1977a) required her Ss (pre-service elementary teachers) to evaluate the truth of compound statements given the truth values of their simple components. Two sets A and B were depicted with their elements (geometrical shapes). The first simple proposition proposed that a particular shape was a member of the set A and the second that another was a member of the set B. Truth or falsity of the simple components could therefore be judged by examination of the pictured sets. The major finding reported was that Ss overwhelmingly tended to treat propositions involving conditional and biconditional connectives as though they were conjunctions, declaring them true only when both simple components were true. Although less pronounced, there was also a tendency to treat disjunctive statements in this way.

These results were replicated with another sample of pre-service teachers (Damarin, 1977b) when Ss were given compound propositions and asked to evaluate for which combinations of truth values of the simple components the proposition was true. The compound propositions consisted of two components. The first stated whether a number M was odd or even and the second made a similar statement about a number N. It was also reported that Ss were more likely to assign an inclusive interpretation to the word 'or', a finding which seems to conflict with the widely held view that, in natural language, disjunction is exclusive (see chapter 3).

In a study to investigate adult Ss' interpretation of conditional sentences, Evans (1972) found a tendency for Ss to view these as true when both components were true and false when the antecedent was true and the consequent false. The statement tended to be viewed as irrelevant when the antecedent was false. (This finding has also been reported by Wason (1966).) Evans' results also led him to postulate a 'matching bias' - a S is less likely to respond with 'irrelevant' when the truth table category matches the items named in the conditional statement. He systematically negated the components of the conditional proposition so that the truth of the simple components *p* and *q* constituted a different truth table category for each of the conditional propositions. For example, given the conditional $\bar{p} \rightarrow q$, the truth of p and q renders the first component false and the second true so that the conditional is true. For $p \rightarrow \bar{q}$ however, these truth values give a false conditional statement. Evans found that the cases where the antecedent was false were less likely to be judged as irrelevant when these were the cases where p and q were true.

A similar investigation of the interpretation of disjunctive sentences (Evans and Newstead, 1980) showed that there were more errors of interpretation when just one of the components of the disjunctive proposition was negated. Also the majority of Ss interpreted disjunction inclusively, assessing TT cases as consistent with the disjunctive statement. No evidence of 'matching bias' was found, however.

Johnson-Laird and Tagart (1969), using university students as Ss, showed that the way in which a conditional statement is expressed affects its interpretation. Ss were presented with sentences of the form: 'if *p* then *q*', 'not *p* if not *q*', '(not *p*) or *q*', 'never *p* without *q*'. These correspond to the logically equivalent forms: $p \rightarrow q$, $\bar{q} \rightarrow \bar{p}$, $\bar{p} \vee q$ and $\overline{p \wedge \bar{q}}$ respectively. The classification of TT cases as confirming, TF cases as disconfirming and FT and FF cases as irrelevant was elicited by the first and last forms ($p \rightarrow q$ and $\overline{p \wedge \bar{q}}$). 'Not p if not q' was similarly treated as irrelevant when the antecedent (now $\bar{q}$) was false. In interpreting the disjunctive form $\bar{p} \vee q$, Ss did not classify stimuli as irrelevant but tended to favour the correct conditional interpretation.

Airasian et al (1975) investigated the understanding of logical statements in adolescents presumed to have reached the stage of formal operations. The results led to the proposal of a hierarchy, the levels of which are conjunction, exclusive disjunction, inclusive disjunction, the conditional and biconditional. Understanding of each is a pre-requisite for the understanding of all subsequent connectives in the hierarchy. However, these findings are difficult to interpret since the linguistic form of the test statements used is not given.

A number of studies have focussed on how understanding of logical connectives develops in children. For 4 - 6 year olds, conjunction and exclusive disjunction are easier than inclusive disjunction and a negated component of the compound proposition substantially increases the difficulty of comprehension (Suppes and Feldman, 1971).

Paris (1973) carried out a comprehensive investigation of the understanding of verbally expressed compound propositions utilising 'and', 'or', 'if ... then', '... if and only if then ...' in children and adults. Ss were shown a slide depicting, for instance, a boy on a bicycle and a dog lying down. They were then asked to assess the truth or falsity of a statement such as 'The boy is on a bicycle or the dog is lying down. 'And' was well understood throughout the age range. 'Or' tended to be interpreted as inclusive disjunction but there was a greater tendency to treat it exclusively amongst the older Ss. For younger children, errors in assessing the truth of a statement containing 'or' tended to occur when the TF or FT instance was depicted although this tendency was diminished when the statement contained 'either ... or' rather than just 'or'. Conditional statements were difficult at all age levels and errors in interpretation occurred largely for the FT and FF cases. However, errors in the FF cases decreased amongst older Ss but FT errors persisted and were very high even amongst college students, who therefore tended to treat conditionals as biconditionals. (This tendency is also reported by Taplin (1971).) Biconditionals evoked errors mainly in the FF cases but decreasingly so for older Ss. The order of increasing difficulty in comprehending the connectives is reported as: conjunction, biconditional, inclusive disjunction, conditional.

Sternberg's (1979) results are somewhat at odds with those of Paris, described above. Interpretation of the biconditional was found to be the most susceptible to error and Sternberg's Ss (7 - 13 year olds and college students) favoured an exclusive interpretation of 'or'. However, his experimental method was rather different. Ss were given a compound proposition and required to assess the truth or falsity of all four truth table categories. Also reported is a preference by younger children for the inclusive interpretation of 'or', which diminished with increasing age in favour of the exclusive interpretation. Younger children also showed a preponderance for a biconditional reading for 'if ... then' which was replaced by a greater tendency for the conditional reading amongst older children. However the biconditional interpretation was not uncommon even amongst college students. The transition from a biconditional to a conditional reading seemed to occur earlier for a conditional statement expressed using 'only if'.

Nitta and Nagano (1966) found that, for Japanese children aged 7 to 15, conjunction was easier than inclusive disjunction. Conjunction was well understood at all age levels whereas performance on disjunctive items improved steadily with age. The generality of their findings was investigated by Neimark and Slotnick (1970) who repeated some of their tests with American children in the same age range. Whilst the general results were similar, there was some indication that Japanese children perform better than American children of the same age. Beilin and Lust (1975) suggest that this developmental difference could be attributable to the greater ambiguity in the expression of logical connectives in English.

The research findings on how linguistic connectives are interpreted lack consistency and are therefore difficult to integrate. It is clear that 'and' is nearly always assigned an interpretation equivalent to logical conjunction and that 'if ... then' is frequently not given a conditional reading. Negation in either component of a compound proposition seems to increase the difficulty of comprehension. There is widespread disagreement as to how 'or' is interpreted especially when contextual cues do not indicate the appropriate reading. A recent view is that adults tend to favour an inclusive interpretation with a sizeable minority preferring an exclusive reading (Johnson-Laird and Byrne, 1991, p. 44). However, the experimental evidence does not seem to support such a view with any consistency.

It is interesting to note that the relative difficulty of interpreting linguistic connectives mirrors to a large extent the difficulty of identifying the underlying logical concepts in non-verbal concept identification tasks. As Lenneberg (1962) observes: '... in most instances of experimental concept formation, there is a correlation between ease of naming the concept and ease of attaining it.' A possible explanation of the relative difficulty of identifying logical concepts is that, for those with ambiguous labels, the underlying schemata are inadequately formed or infrequently invoked. In either case, the underlying concept is less familiar and therefore more difficult to access.

An additional factor is the role of verbalisation, proposed as a significant factor in concept identification tasks. Archer's (1964) view is that in such tasks '... there may be a gradual development of attending to stimuli, selection of information, formulation of hypotheses, testing of these hypotheses, identification of relevant and irrelevant information, elimination of redundant relevant information and a gradual but final "firming up" of a verbal statement of the concept'. However, the evidence suggests that such a verbal statement will be more difficult to formulate for certain logical concepts.

Propositional reasoning

Tests of propositional reasoning focus on the conclusions which can be drawn from compound propositions conjoined using linguistic connectives such as 'and', 'or', 'if ... then', and 'if and only if'. In the most commonly used form of such a test (often termed the 'two-premise deductive argument') the S is presented with a compound proposition consisting of two simple components. This constitutes the 'major premise'. A minor premise is also given. This takes the form of one of the component simple propositions (or its negation) of the major premise. The other component (or its negation) of the major premise constitutes the conclusion. The S is asked to assess the truth of the conclusion assuming the truth of the premises. An example is : 'John is rich or he is clever' (major premise); 'John is not clever' (minor premise); 'John is rich' (conclusion). For this example the conclusion is true. In some cases the truth of the conclusion is undecidable. This would be so in the example above if the minor premise were altered to 'John is clever' and 'or' was interpreted inclusively. A slight modification of this type of test is where the S is not given a conclusion for which to test the validity but must attempt to supply a valid conclusion himself. This type of task is similar to those used to investigate understanding of linguistic connectives (described in the last section). However it requires a scanning of the whole truth table rather than just the individual rows.

Newstead et al (1984) investigated disjunctive reasoning in adults with argument forms as described above where the major premise was varied to cover such contexts as promises, threats, choices, qualifications and abstract material. Two inference forms are valid for disjunctive arguments. From the falsity of one of the disjuncts of the major premise, the affirmation of the other disjunct can be inferred. This form of inference is known as the 'denial inference'. When the minor premise is the affirmation of one of the disjuncts, the negation of the other can be inferred only if the disjunction is taken to be exclusive (known as the 'affirmation inference'). If 'or' is read as inclusive disjunction, nothing can be inferred. Over 90% of Ss made the denial inference with context having little effect on this tendency. Ss are also reported to show a strong inclination to draw the affirmation inference appropriate for exclusive disjunction for all contexts other than qualification (e.g. major premise: 'The successful applicant must have a degree or experience in computing'). When offered as further evidence that natural language disjunction is exclusive this is not really convincing since a number of the major premises used to cover other contexts are such that it is highly unlikely that both disjuncts could be true simultaneously. For instance, 'My wife will be either watching television or preparing a meal' and 'It [a poem] was written either by Ian Jennings or by Peter Lambert' do seem to invite an exclusive interpretation.

Roberge (1978) investigated reasoning with major premises in which the connectives were 'or ... but not both', 'or ... or both' and also conditional premises expressed using 'if ... then' and 'only if'. In each case the minor premise was the denial of the second component of the major premise. Reasoning with premises with abstract content (e.g. 'There is a J only if there is a W') was compared with those with concrete content ('If Joan is athletic then she is rich'). Also investigated was the effect of polarity (negation in either or both components) in the major premise. The following were the more relevant findings of this research.

(1) When the first component of the major premise contained negation, performance was superior when the content of the argument was concrete rather than

abstract. When both components of the major premise were negative, performance was better with abstract content.

(2) For both forms of disjunction, Ss found arguments whose minor premise was the denial of an affirmative second component of the major premise easier than when the major premise contained a negated second component. However, there was no such effect for either form of conditional argument.

(3) Conditional arguments where the first component was affirmative were easier than those where the first component was negative. There was no such effect for disjunctive arguments.

(4) Disjunctive arguments in which at least one component was negated were more difficult than the corresponding conditional arguments. When both components of the major premise contained negation, exclusive disjunction was easier than inclusive disjunction.

(5) Conditional arguments expressed using 'only if' were easier than those using 'if ... then'.

(6) Performance was virtually error-free with arguments involving exclusive disjunction and for conditionals using 'only if' when neither component of the major premise was negated. For arguments involving exclusive disjunction, the introduction of one negated component in the major premise reduced performance considerably, especially if the content of the argument was abstract.

The research projects described above are among the few that have investigated other than conditional reasoning. The pre-occupation with this particular form has been justified by its role in evaluating the validity of conclusions drawn from theoretical assertions and hence its importance in any of the scientific (and, presumably, mathematical) disciplines (Ward et al, 1990). There are two valid inferences which can be drawn when the major premise has the form 'if p then q'. Where the minor premise is p, the conclusion q can be inferred, a form of valid argument known as 'modus ponens'. When the minor premise has the form $\bar{q}$, the valid conclusion is $\bar{p}$ ('modus tollens'). Two invalid inferences are: given $\bar{p}$ as the minor premise, drawing the conclusion $\bar{q}$ (known as 'denying the antecedent') and, given the minor premise q, inferring p ('affirming the consequent'). However, each of the latter inferences is valid if the major premise is interpreted as a biconditional statement. Taplin's (1971) results suggest that, for adults, modus ponens is easier than modus tollens and both these forms are easier than denying the antecedent and affirming the consequent. Only 18% of his Ss consistently (and correctly) denied the validity of denying the antecedent whilst the corresponding percentage for affirming the consequent was 29. Nearly half of all Ss erroneously considered these inferences valid. However, expanding the major premise to countermand the invited inference that the major premise was biconditional was found to reduce these errors (Rumain et al, 1983; Byrnes and Overton, 1988). A similar hierarchy of difficulty for the four forms of argument was also found by Jansson (1975) for pre-service elementary teachers.

A number of factors have been found to influence success in a variety of tasks which have been used to test conditional reasoning. Negation in the major premise was found to increase the difficulty in assessing the validity of the conclusion in grade levels 4 to 10 (Roberge, 1969). Pollard and Evans (1980) also found that the polarity of the major premise affected the performance of adults. Ss were found to have a greater tendency to affirm negative conclusions leading these researchers to postulate a 'conclusion bias' in favour of negative conclusions. However performance on modus ponens arguments seems relatively immune to this effect.

The subject matter of the premises and conclusion also seems to affect conditional reasoning performance (Staudenmayer, 1975). Ss are more consistent in their inferences when the argument is concerned with abstract rather than concrete material although in this case they are no less likely to interpret the major premise as a biconditional. Ss reasoning with abstract material also show more consistency than those reasoning with material which is concrete but where the relation between antecedent and consequent in the major premise is anomalous (e.g. 'If she waters the tropical plant then the light will go on').

A rather different form of conditional reasoning test is the so-called 'selection task'. In its original form (Wason, 1966; Wason, 1968) Ss are shown four cards and told that each has a letter on one side and a number on the other. The cards are presented with one side visible showing a vowel, a consonant, an even number and an odd number. Given the rule 'if there is a vowel on one side of the card, then there is an even number on the other side', Ss are required to select all (and only) those cards which must be turned over to discover whether the rule is true or false. The essential feature of the task is, given p, $\bar{p}$, q, $\bar{q}$ and the conditional rule $p \rightarrow q$, which combination of truth values of p and q could disprove the rule? The answer is p and $\bar{q}$ and this determines the correct combination of cards which must be inspected. In the example above, the card with a vowel and the card with an odd number should be turned over. The task as described above is notoriously difficult with fewer than 10% of adult Ss selecting the correct pair of cards (Manktelow and Evans, 1979). Ss frequently select the card corresponding to p and many incorrectly select q as well. However very few indeed choose $\bar{q}$.

In the twenty-five years since the task was first used, a considerable volume of research has been generated attempting to explain why the selection task causes such problems for intelligent Ss. It was found that performance was greatly improved if the task used realistic material rather than abstract symbols (Wason and Shapiro, 1971; Johnson-Laird et al, 1972; Van Duyne, 1974). However, later studies failed to replicate the facilitating effect of realistic material (Manktelow and Evans, 1979; Griggs and Cox, 1982). This led to the suggestion that it is not simply realistic material which facilitates performance but rather that the material must combine realistic content with a scenario familiar to the S and for which he has already learned the appropriate testing strategies (Pollard and Evans, 1987). In this case, the S is not displaying improved logical reasoning but his own relevant experience is cued by the scenario evoked by the experimental materials and he is thus led to the correct response. This explains the earlier finding (Cox and Griggs, 1982) that the facilitation observed with certain forms of realistic material did not transfer when Ss were subsequently presented with the original form of the selection task. George (1991) also reports a form of 'scenario effect' - Ss were more likely to choose $\bar{q}$ when such a choice was pertinent to a specified goal.

Evans (1972) explains the common selection of p and q by way of a 'matching bias' - Ss tend to choose the cards named in the rule. Evans and Lynch (1973), using symbolic material, investigated the effect of polarity in the conditional rule and found that Ss tend to choose the cards corresponding to p and q even when the rule was $\bar{p} \rightarrow q$ or $p \rightarrow \bar{q}$. (In the latter case, p and q is the correct selection.) Manktelow and Evans (1977) confirmed the matching bias effect for realistic material.

Whilst a comprehensive theory explaining performance on the various forms of the selection task has yet to be developed, it is currently thought that a number of factors contribute to the difficulty of the task. George (1991) suggests that these include linguistic factors (such as interpretation of the conditional connective), knowledge factors about the topic of the conditional statement and utility factors related to the S's current goal.

From the foregoing account, it is clear that human beings' reasoning is frequently illogical in the sense that it fails to conform to the laws of formal logic. Conclusions are drawn which are not sanctioned by logic and valid inferences are overlooked. Nevertheless, it is also clear that a certain degree of logical competence is possessed even by young children and that adults, when reasoning illogically, are often able to recognise that their proposed inference is flawed.

A view once fashionable was that formal logic described the cognitive processes involved in reasoning (Boole, 1854; Kant, 1855) or was at least isomorphic to these processes (Piaget, 1957). This now seems unacceptable for a variety of reasons. One is that no linear correlation appears to exist between the number of logical steps involved in a process of inference and the difficulty of making that inference (Osherson, 1975). The inevitable conclusion is that a logical step does not necessarily correspond to a mental step. Another problem is that a number of logically valid and easily inferred conclusions would be viewed as bizarre in normal discourse (Johnson-Laird and Byrne, 1991). For example, from the premises p and q the conclusions $p \vee q$, $(p \vee q) \wedge p$, $p \wedge p$ are valid although there is little evidence for them in the logical repertoire of ordinary individuals. Whilst logic is not now taken to be representative of thinking processes, it is held to be an essential tool for evaluating the output from those processes, however they are carried out. Formal logic can therefore be regarded as normative rather than descriptive.

In any model of deductive reasoning it seems necessary to postulate at least two components. One is responsible for the comprehension of the premises and for encoding their salient features in a form suitable for manipulation in memory. The second is a processing element responsible for carrying out the operations on the encoded premises and producing a valid inference if this is possible. Braine (1978) refers to the former as the 'performance component' and the latter as the 'logical component'. He further proposes that the performance component contains two main programs. One is concerned with comprehension of the premises and determines what information is to be encoded. The other determines the routines and strategies necessary for constructing a line of reasoning.

The current theories which attempt to explain deductive reasoning have concerned themselves exclusively with the logical component. They describe this by proposing one of three cognitive structures. These are general purpose inference rules (Braine, 1978; Rips, 1983), pragmatic reasoning schemata (Cheng and Holyoak, 1985) and mental models (Johnson-Laird, 1982; Johnson-Laird and Byrne, 1991). Inference rules define the deductive steps which can be applied in an argument and thereby determine what can be concluded from propositions which have been established. They are substitution instances of inference rule schemata - formulae which specify the general form of inference rules. An example of such a schema would be:

$$\frac{p \text{ or } q, \quad \text{not } p}{q}$$

where the conclusion below the line is a valid inference from the premises written above it. Pragmatic reasoning schemata differ only in that the rules invoked are dependent upon such factors as the content and context of the problem. They propose that ordinary life experiences induce abstract knowledge structures. Hence an individual reasons with rules which are sensitive to context and which are invoked on the basis of a pragmatic interpretation of the situation thereby activating the relevant schema. The 'mental models' notion is rather different. It assumes that, in a reasoning situation, a model is set up whose structure is identical to that of the state of affairs with which the reasoning is concerned. The essence of the theory is that the models contain maximal implicit information and minimal explicit information. For instance, a disjunctive statement such as 'There is a circle or there is a triangle' invokes two models, one in which there is a triangle and one in which there is a circle. The information 'there is no circle' eliminates the second model leaving only the first from which the inference 'there is a triangle' is made . No inference rules are necessary. Conclusions are drawn and tested for their validity by attempting to find alternative models which do not support them.

Given that human reasoning is so resistant to introspection, there is understandable disagreement as to which of the above more accurately models the reasoning process. However, from the point of view of this study, what is important about all these models of reasoning is that they imply a 'translation' process from a problem represented in language to an abstracted mental representation. No model incorporates any provision for faulty reasoning so presumably, if this encoding process is carried out successfully, error-free deduction will occur subject to such constraints as the availability of space in working memory. It therefore seems that it is the encoding stage which is implicated as the major source of the many and varied errors which are known to occur in human reasoning.

Of the theories of deduction outlined above none has yet addressed the process of comprehension. Indeed Evans (1989, p. 67) considers that '... such theories are seriously incomplete in that they fail to describe the necessary encoding and decoding stages that must precede and follow reasoning.' Whilst errors in the translation process may be due to a variety of factors (for example, motivation, inattention), for problems expressed in language, language comprehension must be one of the major causes of encoding errors. There are a number of ways in which language factors have been implicated. It has been suggested (Politzer, 1986) that one explanation of poor performance on reasoning tasks may be that mental logical structures are constituted but do not co-ordinate well with the laws of language use. The two systems are in some way contradictory. Braine (1978) explains some of the common errors of syllogistic reasoning as due to '...

the intrusion into formal reasoning of habits characteristic of practical reasoning and ordinary language comprehension.' In a comprehensive investigation of reasoning with conjunctive, disjunctive, conditional and biconditional connectives, Sternberg (1979) compared interpretation with reasoning performance. By comparing the overall errors made at the encoding stage with those made in deductive reasoning tasks, he concluded that most of the reasoning errors could be accounted for by encoding errors. Taplin et al (1974) drew a similar conclusion for conditional connectives.

Where linguistic connectives are concerned, part (perhaps all) of their 'meaning' is the specific inferences which they permit. For an inference rule model of reasoning, a failure to encode such a connective correctly will invoke an inappropriate inference rule. Mental models utilise an accurate representation of the reasoning problem and again must depend critically on the correct interpretation of the relationship between premise components. Whilst there are a wide range of factors which could lead to a breakdown in the process of representation, mis-coding of linguistic connectives will inevitably lead to erroneous reasoning. Given the evidence that they are not generally understood in their logical form, there seems no doubt that at least some reasoning errors must be directly attributable to poor comprehension of these language forms.

Cross-linguistic research

Very few researchers have interested themselves in cross-linguistic comparisons of concept identification or deductive reasoning tasks. Yet if, as seems likely, mis-comprehension of linguistic connectives is a source of reasoning errors, one would not be surprised to find some differences in performance amongst the speakers of languages which use different linguistic forms to convey logical concepts. Cole et al (1968) reported that the Kpelle language distinguished between inclusive and exclusive disjunction. (The Kpelle are a tribal people inhabiting North Central Liberia.) A concept identification experiment was therefore carried out where the performance of Kpelle Ss was compared with that of Americans. It was found that the Kpelle identified conjunctive and inclusive disjunctive concepts with equal facility whereas the Americans found conjunction easier than inclusive disjunction. It was therefore proposed that identification of logical rules might be related to the ease with which they could be expressed linguistically. However, it was suggested (Ciborowski and Cole, 1972) that the stimulus materials might have ~~been~~ influenced the experimental results. Conjunctive and disjunctive concepts were conveyed using four patches of coloured cloth presented in two pairs (e.g. red/green and red/yellow). Combinations of two colours, one from each cloth pair, defined the conceptual rule which the S had to identify. However, whilst the stimulus materials might affect performance, it is not clear why Kpelle and American Ss performed differently given that the same experimental paradigm was used for each group.

Newstead and Griggs (1983) do not consider the criticism of Cole et al's experimental method to be a serious one, especially in view of the results obtained by Ciborowski and Cole (1972). In an attempt to clarify whether the differential difficulty of conjunctive and disjunctive concepts was a culture-specific phenomenon, the performance of Kpelle and American Ss was compared using more standard concept identification stimulus materials. Ss were required to identify the relevant attributes and the rule for combining them (either conjunction or inclusive disjunction). Similar results to those of Cole et al were obtained when previously presented stimulus instances were not available. However, when the memory burden was alleviated by allowing previous stimuli to remain in view sorted into their respective response categories, there was no evidence of a difference in performance between the two cultural groups. Although far from conclusive, these studies are amongst the few which provide some support for the Sapir-Whorf hypothesis. However, it has been pointed out (Newstead and Griggs, 1983) that the Kpelle Ss were bilingual in English and Kpelle, a factor which might well have influenced the experimental results.

The ability of the Kpelle to perform deductive reasoning tasks has also been investigated (Cole and Scribner, 1974, pp. 160 - 169). When asked to evaluate the conclusion of a traditional syllogism, uneducated Kpelle appear not to accept the logical nature of the task. They respond on the basis of personally known facts or general knowledge rather than by processing only the information supplied. The underlying reasoning may well be logical but it proceeds from premises other than (or in addition to) those provided by the experimenter. For educated Kpelle however, responses were very much as those reported for Western Ss. Education appears to instil an appreciation of the logical relations implicit in the task and hence diminishes the reliance on factual content.

Zepp (1982) investigated bilingual students' (aged 11 to 15) understanding of the linguistic connectives 'and', 'or' and 'if ... then ...' in English and Sesotho, the latter being the Ss' first language.

These connectives are expressed rather differently in the two languages. The experimental materials used consisted of cards with sides labelled A and B each of which was coloured red, green or black. Ss were shown one side of the card and were required to use a given statement to deduce whatever they could about the other side. For example, when shown a red side A and told 'side A is not red or side B is black' the correct conclusion is that side B is black. Given the same statement and shown a black side A, no conclusion can be drawn about the colour of side B. Students at lower grade levels were found to perform better when tested in Sesotho whilst the older students tested in English outperformed those tested in their native language. One explanation offered is that the younger students have a poor command of English which improves during their secondary education (carried out primarily in English). However, as Zepp points out, it is interesting to note that the logical skills acquired do not seem to transfer readily to the first language. This is further evidenced by the fact that some of the older students tested in Sesotho nevertheless wrote their answers in English. Also interesting is that the older students all agreed (along with their mathematics teacher) that logic was more difficult in Sesotho and that they preferred to learn mathematics in English.

An investigation of possible Whorfian effects on the interpretations of logical sentences was attempted by Zepp et al (1987) using much the same experimental method as that described above. The performance of first language English speakers was compared with that of first language Chinese speakers with each group tested in their first language. Another group of bilingual students whose first language was Chinese was tested in English. Although there was some evidence that the performance of the bilingual group suffered from lack of adequate fluency in the second language, no other significant differences were found. Logical errors were not specific to either Chinese or English. Zepp therefore postulates that, in all languages and cultures, the conditional sentences which children hear are taken (and meant) as biconditionals. Thus the child comes to learn that the converse and inverse of a true conditional statement are also true. This does not however explain failure to reason from the contrapositive. It is suggested that this might be due to the temporal aspects of conditional statements where the antecedent usually precedes the consequent in time. Reasoning from the contrapositive therefore requires a transformation involving an alteration in the time sequence.

Further evidence for the absence of Whorfian differences in logical reasoning was obtained by Zepp (1987a) in a subsequent study using the same experimental paradigm. Chinese-English bilinguals were tested in Chinese (their first language), given instruction in logic in English and then re-tested in Chinese. Performance in the second test showed considerable improvement. This contradicts the results of Zepp's earlier (1982) research and suggests that logical principles learned in a second language transfer readily to situations where the first language is utilised. Zepp (1986) also reports that an illiterate deaf-mute S performed in a manner comparable with that of verbal Ss on a concept identification task involving conjunctive and disjunctive concepts. This seems to refute Cole et al's (1968) conjecture that ease of concept formation is related to the ease of expression of that concept.

There is little that can be concluded from the very few studies reported in English which have attempted a cross-linguistic comparison of the concepts underlying logical connectives. Results are difficult to interpret and, to some extent, mutually contradictory. However, evidence in favour of Whorfian differences, although slim, seems sufficient to merit further investigation.

Chapter 5

LINGUISTIC CONNECTIVES IN JAPANESE AND ARABIC

The following account outlines how linguistic connectives are expressed in Japanese and Arabic. A brief description of the essential features of each language (particularly those which differ from English) is included so as to establish a context for the discussion which follows. The connectives described are those which potentially convey the sense of logical conjunction, disjunction, the conditional and biconditional. Also included are other uses of these lexical and grammatical items so that potential sources of ambiguity can be appreciated.

General sources of the information included in this chapter are: Kuno, 1973; Bloch, 1946; Alfonso, 1980; Naganuma and Mori, 1962; Hakuta et al, 1982; Tritton, 1943; Beeston, 1970; Smart, 1986; Wright, 1971; Haywood and Nahmad, 1965.

Japanese

Japanese is a 'SOV language' - the usual word order in a sentence with a transitive verb is subject, object, verb. There is some flexibility in the order of subject and object but it is obligatory for the verb to occupy the sentence-final position. As is typical of such languages, Japanese has no prepositions and the relations which these convey are expressed using postpositions - particles which follow the word or clause which they qualify. Postpositions are also used to signal the grammatical functions of nouns, for instance in marking subject and object. Co-ordinating particles are also postpositional.

Verb conjugation is not affected by the gender or plurality of its subject and there are only two tenses, present (or non-past) and past. Adjectives behave very similarly to verbs and are inflected to distinguish (amongst other things) present, past, affirmative and negative. Nouns are not generally marked for plurality (although it is possible to do so) nor for gender and Japanese lacks definite and indefinite articles.

The Japanese writing system consists of a mixture of imported Chinese ideographic characters called 'kanji' together with phonetic characters called 'kana' used for their sound alone. There are two kana syllabaries, hiragana and katakana, each consisting of 48 characters. Imported foreign words for which there is no ideographic character in use are usually written phonetically using katakana. Prior to the Japanese language reforms after World War II, most Japanese newspapers stocked over 5000 kanji in type and highly educated Japanese would know as many as 20,000 ideographs (Seward, 1983, p. 4). Currently, just under 2000 kanji are taught to schoolchildren along with the two kana syllabaries and publications aimed at the general populace are recommended to restrict themselves to 2111 specific kanji.

The Japanese language is well known for its linguistic provision for distinguishing levels of politeness and respect. There are four levels of sentence style which are selected on the basis of intimacy of speaker and addressee. For each level of style, two honorific forms exist, selected on the basis of the speaker's respect for either the subject or object of the sentence (Kuno, 1973).

Japanese has been described as a language 'whose insightful description requires more use of notions related to the affect of speakers (such as their point of view, empathy, and camera angle) than does English.' (Hakuta and Bloom, 1986.) For instance, in Japanese the speaker must signal whether an affirmative statement stems from his own first hand knowledge or from hearsay, even when he is certain that what he has heard (and is now reporting) is factually correct.

Context is said to play a more important role in understanding Japanese than it does for English discourse. It is considered impolite to be too overt and explicit. Indirectness is valued and 'reading between the lines' is therefore an important skill (Yotsukura, 1977).

Conjunctions

There are several different morphemes in Japanese which correspond to the English word 'and'. Unlike English, Japanese distinguishes between conjunction of nouns, adjectives, verb phrases and sentences.

Conjunction of nouns is achieved by using 'to' or 'ya' between conjuncts. The former is appropriate when the list of conjoined nouns is complete, i.e. there are none that the speaker has omitted to mention. If used in a sentence such as 'I visited Paris and Rome', 'to' would indicate that I visited only those cities. On the other hand, 'ya' is used where the list is incomplete and includes other items not mentioned. If 'ya' were used rather than 'to' in the example above, the implication would be that I visited other places as well. Nouns (and adjectives) may also be conjoined by juxtaposition as in 'I visited Paris, Rome, London, New York'. However, this can sound childish if over-used. 'To' and 'ya' can also be used to link noun phrases.

'Ni' can replace 'to' but is very much less common. It tends to be used for listing and generally requires more than two conjuncts to be enumerated.

'To' has a number of other meanings. It is used as a signal to indicate that the preceding words constitute directly reported speech, as in 'The boy asked his father "What are you doing?" '. It is also utilised to signal indirectly reported speech. 'To' can convey the sense of a conditional in that it connects an antecedent with its natural, inevitable or immediate consequent (see below). The sense of the English 'or' is conveyed by 'to' in comparisons or choices, e.g. 'Which do you prefer, coffee or tea?' or 'Who is taller, Paul or Jim?'.

A conjunctive linking of nouns and noun phrases can also be achieved using '... mo ... mo' where 'mo' follows each of the conjuncts (including the last). This gives the sense of 'and also' as in 'We have offices in London and also in Paris'. When the verb in the sentence is negated '... mo ... mo' means 'neither ... nor', so that in the last example, merely negating the verb produces the equivalent of 'We have offices in neither London nor Paris'. This is in contrast to 'neither ... nor' sentences in English where the verb retains its affirmative form. It is also interesting to note that, in this particular Japanese construction, negating the sentence $p \wedge q$ is understood as $\bar{p} \wedge \bar{q}$ rather than $\overline{p \wedge q}$ (or $\bar{p} \vee \bar{q}$), the logical interpretation of the negated form.

Verbs and adjectives in Japanese constitute very similar grammatical forms and adjectives are inflected for nine of the ten categories for which verbs are inflected. Conjunction of adjectives and verbs is achieved by inflecting all except the last of the conjuncts with the ending '-te'. The verb occupies the final position in a Japanese sentence so that conjunction of sentences can also be achieved by inflecting the final verb in the first sentence into its '-te' form. However, this tends to be used for 'and then' and is not appropriate when the two conjuncts describe actions or states which occur simultaneously. In this case the '-i' inflexion is used. Although the '-te' ending does not imply any causal relationship between conjuncts, it can be used in the same way that 'and' is, to link cause and effect, e.g. 'I shouted at him and he ran away'. In fact the '-te' inflection is the most frequently used form of the Japanese verb. As well as signalling conjunction, it is used to indicate that an action is (or was) taking place over a period of time. (It is akin to, but not identical to, the English gerund.) In this sense it resembles the progressive form, e.g. 'It is raining'. However Japanese is more insistent upon the progressive form where English would allow a sentence in its present tense. For example, the following are optional forms of the same sentence in English: 'I teach mathematics in London' or 'I am teaching mathematics in London'. The progressive '-te' form of the verb in this sentence would be obligatory in Japanese.

Juxtaposition may be used to combine sentences but this form of conjunction is less common in Japanese than in English. Two sentences of the form 'A is B' are conjoined using 'de' as the conjunctive connective, e.g. 'This is yours and that is mine'. 'De' also has a number of other uses. It is the particle denoting the means or instrument attached to some action and in this sense translates as 'with' or 'by means of' as in 'He writes with a pencil'. It also denotes the place where an action is performed as in 'He learned it at college'. It marks amounts of time or money and indicates measures of what is necessary to complete some action. In addition, it marks a reason for some fact and is equivalent to 'for' in 'She is famous for her cooking'.

'Shi' can be used to link symmetrical clauses about a common topic and is often used to enumerate reasons, e.g. 'Its raining and the wind is blowing; let's go home.' In linking two sentences or clauses where what is described in the first precedes what is described in the second, 'soshite' can be used. This word therefore translates as 'and then' as in 'She sat down and (then) opened the book'.

Were the natural language items for 'and' to be used in logic, presumably the expression of conjunction would be different depending upon the nature of the conjuncts. For this reason perhaps, logical conjunction is expressed using 'katsu' between conjuncts, a form not used in everyday language.

Disjunctions

The particle 'ka' between two nouns or phrases signals disjunction. It has a number of other functions however. When affixed to the end of an affirmative sentence, 'ka' converts it to a question. For example, affixed to the Japanese translation of 'Jack speaks German', 'ka' would give the equivalent of 'Does Jack speak German?'. It is also used in the same way to mark requests for information, such as 'Where are you going?'. The use of 'ka' at the end of a sentence can also be used to convey doubt or disbelief. These two uses are distinguished by means of intonation. A similar intonational distinction can be achieved in English. Consider, for example, 'You passed your examinations?' which, depending upon the intonation, can convey a straightforward question or an expression of surprise or disbelief. Interrogative words such as 'who', 'what' and 'where' are converted to 'someone', 'something' and 'somewhere' respectively if they are followed by 'ka'.

Disjunction can also be expressed using 'nari' after each disjunct. This form is considered rather formal and tends to carry the implication that there may be a better alternative as in 'Ask John or Mary (or someone better)'. 'Aruiwa' (classed as a connective adverb) can also be used between disjunctive clauses particularly where there is a degree of uncertainty in the speaker's mind as to which of the disjuncts is the case. It is also used in a sentence offering a choice of examples, e.g. '... pets such as dogs or cats or rabbits ...'. 'Soretomo' is used between disjunctive questions such as 'Did you buy a car or did you buy a bicycle?'. Its use is optional and it can be omitted. However 'ka' cannot be substituted. In comparisons and choices 'to' is used for 'or' (see above). 'Matawa' is also used to convey disjunction between nouns and phrases and is the form used for logical disjunction.

Conditionals

There are several forms of conditional sentences in Japanese and the variation of meaning conveyed by them is subtle and complex. An important difference between Japanese and English conditionals is that, in the former, it is necessary to make clear the relationship between the antecedent and consequent. On the other hand, Japanese conditionals do not generally distinguish whether the antecedent is hypothetical ('If such-and-such were the case ...'), is to be realised in the future ('When such-and-such is the case ...'), has been realised in the past ('When such-and-such was the case ...') or is currently so ('Such-and-such being the case ...'). There is therefore potential ambiguity about the state of the antecedent, that is, whether the state of affairs which it describes is hypothetical or actual. This distinction can be made clear in English through the choice of 'when' or 'if'.

Conditionals roughly equivalent to the English form 'if ... then ...' or 'provided that ...' can be expressed using the conditional ending - '-ba' for verbs and '-kereba' for adjectives - in the antecedent clause. These endings usually signal a condition which must be satisfied for the performance or completion of something else. However, this is the least common of Japanese conditional forms.

The particle 'to' has already been described in its role as a conjunctive connective. 'To' can also be used between antecedent and consequent in a conditional statement where the consequent always follows the antecedent. It therefore conveys the sense of 'when' or 'whenever' as in 'When winter comes the swallows fly south'. 'To' is also used to connect antecedent and consequent when the latter is the natural or inevitable result of the former, e.g. 'If you jump then you'll hurt yourself'. When 'to' is used as a conditional connective, the consequent must be a statement of fact and not a command or request as in 'If you go shopping, will you buy me some bread?'.

When the antecedent represents an action completed before that described in the consequent, the conditional is normally expressed using the ending '-tara' on the antecedent verb. This would be the appropriate form for 'If you eat that, then you'll be sick'. 'To' can often be used as an alternative to the '-tara' verbal form but its use would emphasise that the consequent is the inevitable result of the antecedent. The '-tara' form on the other hand, would stress the temporal sequencing of the antecedent and consequent.

In counterfactual conditionals, 'to', '-ba', '-kereba' and '-tara' are used, subject to the conditions described above, and with the verb in the consequent clause in the past tense.

A conditional statement may also be expressed using 'nara' between antecedent and consequent where there is no causal sense to be conveyed. Where the idea of a condition needs to be emphasised 'nara' carries '-ba' as a suffix although 'naraba' tends to be viewed as more formal than 'nara'. 'Nara' expresses a presumption about the truth or actualisation of the antecedent but cannot be used when the antecedent describes an event which is certain to occur. The consequent can be any kind of statement other than one of completed fact. The sense of 'naraba' is perhaps best conveyed by 'If it is the case that ... then...'. Where the '-tara' ending cannot be used because the antecedent does not precede the consequent in time, 'nara' or 'naraba' can often be substituted, as in 'If I go shopping, then I'll buy some bread'. 'Naraba' is the connective used to express the logical form of the conditional statement equivalent to $p \rightarrow q$.

The potential ambiguity in Japanese conditionals regarding whether or not the antecedent is hypothetical or an actuality (realised in the past, present or future) can be resolved by prefixing the sentence with 'moshi'. This functions as a signal alerting the addressee that a supposition or hypothesis follows. The interpretation of the conditional is then 'if ... then ...' rather than 'when ...'. However, 'moshi' is not used when it is obvious or certain that the antecedent has occurred or will occur.

Biconditionals

There is no literal translation of 'if and only if' in Japanese and there is no way of expressing a biconditional statement except as the conjunction of two conditionals.

Arabic

Arabic belongs to the Semitic group of languages, all of which employ a cursive script written and read from right to left. The family includes Modern Hebrew and Amharic. A characteristic of these languages is that most words are built on a root consisting of three consonants. Variations in meaning are achieved by vowels added to the root and by prefixes, suffixes and infixes. Arabic is a VSO language - the usual word order for a transitive sentence is verb, subject, object.

Nouns in Arabic are masculine or feminine and are inflected to distinguish singular, dual and plural (more than two). There are three noun cases roughly corresponding to nominative, accusative and genitive. Adjectives follow the nouns they qualify and agree with them in gender, case and number.

Arabic has a highly developed verb system, the structure of which is very different from that of English. Verbs have only two tenses - the perfect (referring to completed actions) and the imperfect (referring to actions in the future or not yet completed). Verbs distinguish the gender and the number of their subject; there are singular, dual and plural forms. The imperfect form of the verb has three so-called 'moods', the indicative, subjunctive and jussive, which are used after certain co-ordinating particles. There are two further moods - the imperative (for commands) and the energetic (rarely used). All tenses and moods may be active or passive.

The term 'Arabic' is used to cover a number of forms of the language. Local dialects vary considerably to the extent that, for instance, those used in Morocco and

Iraq are to a large extent mutually unintelligible. However, the vernaculars are spoken forms and the written form of the language, generally termed 'literary Arabic' is standard throughout the Arabic-speaking world.

In its written and spoken forms, literary Arabic is the language of the mass media, academic lectures and public addresses. Because it is so widely used on radio and television, in its spoken form it is becoming increasingly understood by even illiterate Arabic speakers. Colloquial forms, although well-suited to the practicalities of day-to-day exchanges, are generally inadequate for intellectual exchanges about abstract topics. Conversations amongst the educated are therefore carried out in a mixture of the literary and colloquial forms.

Literary Arabic (or modern standard Arabic, as it is sometimes known) is closely related to classical Arabic. (In fact, some authors do not distinguish these two forms.) This is the language of the Qur'an, considered the greatest linguistic achievement in the Arabic language and, as such, the ultimate authority on questions of grammar and style. Modern Arabic has a wider vocabulary and is somewhat simpler in its grammar and syntax but, to a large extent, the Arabic grammar taught in schools today differs little from that developed in the seventh century.

Conjunctions

Conjunctions are used very much more frequently in Arabic than in English and sentences frequently begin with the equivalent of 'and'. The basic conjunctive co-ordinator is 'wa' (prefixed to the word following) which may be used to link sentences or nouns. It is not generally used to link adjectives governing the same noun, although it can be if these form the predicate of a nominal sentence, e.g. 'Your car is bright and clean'. Otherwise adjectives are linked by juxtaposition, as in English. 'Wa' may also be used to introduce what are sometimes termed 'circumstantial clauses'. These present an attendant circumstance to the main topic of the sentence. Examples are 'They walked to the town while the rain was falling' or 'He walked towards me with his hat in his hand'. In these examples 'wa' would replace 'while' and 'with' respectively.

An alternative to 'wa' is 'fa' used between clauses rather than single words and also prefixed to the word following. However this carries the additional implication of a temporal or causal sequence as in 'He got up and went out' so that it is equivalent to 'and then'.

'Wa' is the particle used in formal logic to convey conjunction.

Disjunctions

The particle 'aw' is used in very much the same way as the English 'or', to connect the disjuncts in statements or questions. This is also the form used in logic to convey disjunction. 'Either ... or' translates directly as 'imaa ... aw'. With the verb following it in its subjunctive form 'aw' means 'unless' or 'until'.

In questions such as 'Is that John or Jim?', 'am' may be used as the disjunctive connective instead of 'aw'. It has been suggested (Wright, 1971, Volume II, p. 308) that there is a difference of meaning between 'aw' and 'am' when used in questions. If 'aw' is used, this suggests ignorance as to whether either or neither of the disjuncts is the case, whereas 'am' implies that it is known that one of the disjuncts is true and that information is being sought as to which.

Conditionals

In conditional sentences the verb in both antecedent and consequent is in the perfect or jussive. However, there is no particular temporal significance in the verbs of conditional sentences and context is often the only guide to the timing of the two components.

In a counterfactual conditional, or one where the condition expressed in the antecedent is unlikely to be fulfilled, 'law' introduces the antecedent and 'la' is prefixed to the first word of the consequent. Usually no distinction for tense is made in the antecedent so that the sense of, for instance, 'If he had arrived ...' or 'If he were to arrive ...' must be construed from the context.

For other conditional statements 'in' or 'idhaa' introduces the antecedent. The consequent must be introduced by 'fa' (prefixed to the first word) if it is anything other than a straightforward positive statement. In the logical form of the conditional 'idhaa' precedes the antecedent and 'fa' is usually prefixed to 'inna' so that the form used is 'idhaa ... fainna'. 'Inna' is sometimes included in natural language conditionals especially where the consequent is a nominal sentence. In this sense it translates loosely as 'indeed'. This particle is also commonly used at the beginning of sentences of the form 'A is B' (with the subject in the accusative). Its use is more a matter of style since it adds nothing to the meaning of the sentence.

Biconditionals

The logical biconditional is expressed using '... idhaa wa idhaa faqaT ...' where 'faqaT' translates as 'only'. The literal translation is therefore 'if and if only'. This form is used only in logic and mathematics and not at all in everyday language. It seems almost certain that it is a construction which has been 'borrowed' and translated directly from the English.

Of these two languages, Arabic seems more like English in the way that it expresses logical forms. The natural language constructions which these two languages use to convey the respective logical concepts seem to have meanings which are very similar in their everyday usage. Japanese, on the other hand, appears to have adopted specific expressions for logical connectives which are not commonly used natural language constructions. One would therefore suppose that they are more likely to be construed in the logical sense. This is particularly the case for 'katsu' (the expression for logical conjunction) which is of such rare usage that it does not appear in most Japanese/English dictionaries.

Chapter 6

EXPERIMENTAL METHOD

Cross-linguistic studies of the understanding of logical concepts are few in number and inconclusive in their findings (see chapter 4). However, they do suggest differential facility with logical concepts across certain linguistic groups and the possibility that such differences could be language-related. Hence the diversity in the expression of linguistic connectives between Japanese, English and Arabic seems to warrant the investigation of whether there are associated differences in certain aspects of the logical performance of their respective speakers. This empirical study should therefore be viewed as exploratory, previous research suggesting no specific hypothesis to be tested. The nature of any significant differences revealed will suggest hypotheses to be examined in the course of further research.

The experiments described below have two primary aims:

(1) to investigate whether there is variation between the three language groups with regard to the understanding of linguistic items which are used to convey logical concepts;

(2) to investigate whether there is variation between the three language groups in performance on logical concept identification tasks.

Should the groups show variation with regard to both these factors, then a further aim is to establish whether the nature of these differences suggests possible cause-effect relations of the type proposed in Whorf's hypothesis. An additional consideration in designing the experimental tasks was that the overall results could be compared with those of other studies which have reported on concept identification and the interpretation of linguistic connectives in English.

Two tests were administered to adult Ss, each of whom was a native speaker of either English, Japanese or Arabic. Every S completed both tests so that individual performances could be compared if appropriate. The first was designed to investigate performance on a concept identification task where the concepts tested were logical conjunction, disjunction (exclusive and inclusive forms), the conditional and biconditional. Also of interest was the description of a concept used by those Ss who were successful in identifying it. The aim of the second test was to enable a comparison of the provision in the three languages for expressing these logical forms. Of interest was the question of whether or not the appropriate connectives elicit the logical interpretation. It was hoped that the results of the first test would indicate whether any differing provision in the lexicon (revealed in the second test) was associated with differing performance in related concept identification tasks.

The tests were administered in the order indicated because it was felt that this would minimise any possibility of transfer effects. No explicit reference to any of the logical forms was made during the administration of the first test whereas the second was a written test containing the linguistic items normally taken as representing the logical forms.

Ss were tested individually and both tests were carried out entirely in the S's mother tongue. This included all written and verbal instructions. In the case of the Japanese Ss, the experimenter was a Japanese who was a graduate in Psychology. He had shown interest in this research project at an early stage and had, as a small part of his undergraduate course, assisted with some of the background research on Japanese language and culture. He also administered the tests in the pilot study. Although his first language was Japanese, his English was virtually fluent by the time the main study was carried out.

The experimenter for the Arabic speakers was a Jordanian graduate in Mathematical Sciences. She had participated in the pilot study as a subject, became interested in the project and was keen to follow its subsequent development. Her first language was Arabic but her English was sufficiently fluent for her to be considered bilingual.

Each of the two experimenters was responsible for the translation of the second test from English into their mother tongue. These were checked using back-translation carried out by two native English speakers, one a graduate in Arabic language and the other a competent speaker and reader of Japanese. Apart from administering both tests, the experimenters were also responsible for translating S's protocols in the first test into English. Each was paid £150 for their assistance.

Subjects

All Ss were native speakers of either English, Japanese or Arabic. (By far the majority of the English language group were from the US and hence 'English' in this study refers to 'American English'.) All members of the Japanese and Arabic groups could speak some English although their proficiency was very varied. Some were fluent, although they did not consider English to be their first language. Others had considerable difficulty in making themselves understood in English.

All of the English and Arabic speakers were students at Richmond College, an international college in London with a curriculum typical of small American liberal arts colleges. (Both experimenters were graduates of Richmond College.) Unfortunately, at the time when the data were collected, the contingent of Japanese students at the college was insufficient to provide the minimum of 50 Ss. Therefore, approximately 30 of the Japanese Ss were Richmond undergraduates and the rest were students at other London colleges. Whilst it would have been preferable to confine the Japanese sample to the same student population, there seemed to be no reason to suppose that the inclusion of students from other colleges would have any influence on experimental results.

Subjects were selected on no particular basis other than a willingness to participate. Approaches to individuals in the English and Arabic language groups were made more or less at random. Virtually all the Japanese students at Richmond College were asked if they would be willing to participate. Of all the potential Ss approached, very few declined the invitation to be tested although there was no reward for doing so.

Data relating to the age and gender distribution of Ss in each of the language categories are given in table 6.1 below.

Table 6.1: **Age and gender distribution of subjects**

	All subjects				Males		Females	
Language group	N	min. age	max. age	mean age	N	mean age	N	mean age
English	50	17	43	22	17	21	33	23
Japanese	51	18	27	22	16	23	35	22
Arabic	52	17	24	20	33	20	19	19

(Japanese Ss originally numbered 53 but two were eliminated because they were upset by the first task and were clearly not in a frame of mind to continue. However, no other S showed any sign of distress and all those remaining seemed interested and involved in the tasks.)

Students are admitted to Richmond College from a very wide variety of national secondary school systems in diverse geographical areas. However, all are required to have completed twelve years of full-time education and to have achieved a minimum grade of C+ (grade-point average 2.5) in the American high school grading system or its equivalent. If admitted from the British school system, students must have attained a minimum of five GCSEs with a grade of A, B or C in what the college prospectus terms 'acceptable academic subjects' including mathematics or a science. In fact, roughly 40% of the Freshman intake fail the diagnostic mathematics test administered during their first week at the college. For those who fail this test, their standard of mathematics would certainly be below that which would earn them an A, B or C at GCSE. These students are therefore required to take a one-semester course designed to raise their mathematical skills to this level.

The college is fee-paying and students are therefore mainly, although by no means exclusively, from financially secure middle class families. Whilst most students study for the BA degree in Business Administration, a number of other majors are offered. A small minority of students (mainly Japanese) leave the college after achieving the AA (Associate of Arts) degree. This has no subject specialisation and can be achieved in two years whereas the BA degree normally takes four years. The distribution of major areas of study for participating Ss is given in table 6.2 below. However, it must be borne in mind that, at Richmond College (as in all American liberal arts colleges), students are required to study a wide range of subjects, in addition to those which are relevant to their field of specialisation, throughout their four-year course.

Table 6.2: **Distribution of major areas of study of participating Ss.**

	Language Group		
	English	Japanese	Arabic
Psychology	17	2	0
Computer Science	5	0	7
Economics	2	3	0
Business	9	11	34
Communications	1	0	0
Mathematics	1	1	0
Engineering	2	1	3
Lens Media	1	0	0
Accounting	1	0	1
English Literature	3	4	2
History	3	4	0
French	0	1	0
Education	1	0	0
Politics	2	2	0
Social Sciences	1	1	1
Fine Art	1	10	1
Philosophy	0	1	0
Science	0	1	3
AA degree	0	4	0
Undecided	0	5	0
Total	50	51	52

Along with questions about their biographical details, Ss were asked whether they had ever studied any formal logic. (All such information was elicited after completion of both tests.) Ten English speakers, 2 Japanese and 39 of the Arabic language group claimed that they were familiar with elementary logic (truth tables and validation of arguments).

An attempt was made to ascertain the level of mathematics attainment for each S, but this proved difficult given the wide variety of school systems in which their secondary education had been completed. It was clear, however, that levels of achievement in mathematics and attitudes to the subject varied very greatly within each language group.

Experiment 1

Test materials

Stimulus materials were adapted from those used by Neisser and Weene (1962) but using geometrical shapes rather than letters of the alphabet, since the latter would clearly not be equally familiar to all three language groups. The basic stimulus element was a string consisting of four shapes each of which could be a circle, triangle, star or square. Six A3 sheets were prepared. On one was printed all possible $4^4 = 256$ such strings. Each of the remaining five sheets contained the subset of these strings corresponding to one of the five concepts under investigation (conjunction, exclusive and inclusive disjunction, the conditional and biconditional). On these five sheets, the strings were printed in columns, each string occupying a 3 cm by 1 cm rectangular space.

For each of the five concepts, the relevant sheet contained precisely those strings which exemplified the concept defined in terms of the presence or absence of two of the four shapes. For example, if the concept was 'triangle and square', the sheet would contain all (and only) those strings which included at least one triangle and at least one square. The two salient shapes were chosen at random for each concept and were printed on the bottom right hand corner of the relevant sheet. The order of strings on each sheet was random.

The five concepts tested and the subset of strings representing each one are described below.

Concept	Strings
'Triangle and star': ($p \wedge q$)	Those containing at least one triangle and at least one star (64 strings).
'Square or circle or both': ($p \vee q$)	Those containing at least one square or at least one circle or both (192 strings).
'Circle or star but not both': ($p \veebar q$)	Those containing at least one circle but no star and those containing at least one star and no circle (128 strings).
'If triangle then square': ($p \rightarrow q$)	Those containing at least one triangle and at least one square and all those containing no triangles at all (192 strings).
'If and only if star then square': ($p \leftrightarrow q$)	Those containing at least one star and at least one square and those containing neither stars nor squares (128 strings).

For each of the five concepts, a sample of 16 strings was selected from the complete set, four corresponding to each of the truth table categories. For instance, if the relevant shapes were 'circle' and 'square', the 16 strings would consist of four containing both shapes (truth table category TT), four with at least one circle but no square (TF), four with at least one square but no circle (FT), and four with neither shape (FF). Within each truth table category, the strings were selected at random. Each of the 16 strings selected for each concept were printed on a 7cm by 1.8 cm rectangular card and the cards shuffled into random order.

Test procedure

The S was shown the sheet containing the full set of 256 strings. It was explained how each string was composed of some or all of the four shapes circle, triangle, star and square. Ss were told that, as far as the test was concerned, the order of shapes within a string was irrelevant as was the order in which the strings were printed on the sheet. It was emphasised that the only important characteristic of any string was whether a particular shape was present or absent. How many times a shape occurred within the string was of no relevance.

Several strings were pointed out and described in terms of the shapes present and absent. This was continued until E was confident that S understood how the strings were constructed and how the sheet which he was currently viewing contained all possible such strings.

The S was then told that he would be shown a series of sheets on each of which was printed a subset of the complete set of strings. It was explained that the strings on a sheet were all those which had a common characteristic relating to the presence or absence of two particular shapes. In each case, S would be told which were the two relevant shapes and these were printed on the sheet to remind him. His task was to examine each sheet in turn and to determine the common property of the strings which it contained. The S was told that the time taken to identify the concept would be recorded but that he may take as long as was necessary to discover the common characteristic. It was explained that, when S had announced that the concept had been identified, the sheet would be removed and that he would be shown 16 strings, one at a time, and asked whether or not each was an example of the concept. After this had been completed, he would be asked to describe the concept, i.e. the common property shared by all the strings.

Any questions concerning procedure were answered whilst care was taken not to pre-empt any of the concepts to be tested. When E was confident that S understood what was required of him, the first sheet was presented.

For each concept, the time was recorded from presentation of the sheet until S announced that he had identified the concept. Responses to the 16 test strings were recorded as was the S's verbatim description of the common characteristic shared by all strings.

For any given concept, the 16 test strings were presented in the same (random) order and no feedback was given to responses to these strings. The order of presentation of the five concepts was random and varied for Ss within a given language group. Across language groups, the n^{th} S (n = 1, 2, ..., 50) in each of the three groups tackled the concepts in the same order.

It is to be noted that this experimental procedure is not that of 'standard' rule learning tasks (as described in chapter 4). In the latter, the S is required to respond to individual, sequentially presented stimulus elements by stating whether or not each is an example of the concept. Through informative feedback, the S is to infer the relevant features which distinguish examples from non-examples. This particular paradigm was used because it was believed to simulate the means by which human concepts are formed. The purpose of this experiment was different, however. It was designed to discover how familiar the concepts were and how easily they could be abstracted from a set of exemplars, rather than how they were learned. In the traditional concept identification task, Ss must cope with the memory burden imposed by the lack of accessibility of previous stimuli if these are not allowed to remain in view. Furthermore, the S must adopt one of a number of possible strategies in order to integrate the information gleaned from each stimulus element (see Bruner et al, 1956). Both of these factors constitute a source of inter-subject variability which was considered largely irrelevant to purpose of the current investigation.

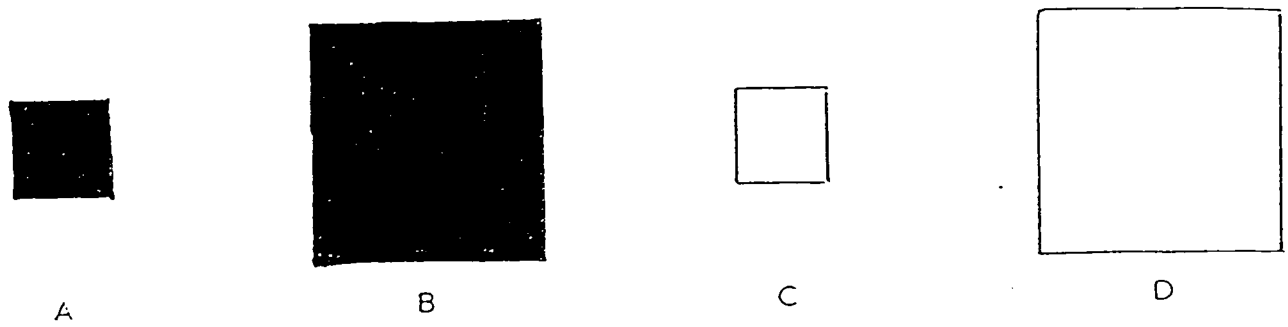
Experiment 2

After all five concept identification tasks had been completed, S was asked to commence the second test.

Test materials

Each S was given a test sheet at the top of which were shown four squares. One was small and black, one large and black, one small and white and one large and white. The squares were labelled A, B, C and D respectively and were displayed as shown in figure 6.1. (The order of the squares was reversed for the Arabic speaking Ss to correspond with the direction of their script.)

Figure 6.1. **Figures used in test 2.**



The sheet also contained 16 numbered statements and Ss were asked to respond to each by stating which of the four squares were consistent with the statement and which were not. Each S recorded his responses on the sheet by placing a tick or cross as appropriate in each of four boxes, one corresponding to each square. Test sheets for each language group are included in appendices 1a - 1c.

Each of the 16 statements was a compound proposition constituted from two simple propositions, one concerning the colour of the square (black/white) and the other concerning its size (small/large). Each compound proposition was formed by conjoining the two simple components using one of the linguistic connectives expressing the logical operations of conjunction, disjunction, the conditional and biconditional. An example of such a compound proposition in English might be 'If the square is large, then it is white'. The appropriate linguistic connectives in Japanese and Arabic were taken from the section of a high school textbook (in the appropriate language) dealing with elementary truth functional logic. The connectives used are those described as labelling the logical operations in chapter 5. As stated in that chapter, there is no way of expressing the biconditional in Japanese except by the conjunction of two conditional propositions. This was therefore the form used.

The statements were ordered in 4 blocks of 4. Within each block, each connective appeared once and each colour/size combination once. Each block contained two statements where size was mentioned before colour (e.g. 'The square is large and it is white') and two where colour was mentioned before size (e.g. 'The square is black or it is small').

Instructions for filling in the test sheet were included. This test was not timed.

Test procedure

The S was given the test sheet and asked to follow the instructions contained therein. Any questions requesting clarification of the test procedure were dealt with.

The approximate total time to complete both tests varied from 30 to 60 minutes.

Pilot study

A pilot study was carried out, the results of which suggested a number of modifications which were incorporated in the main study. It also revealed some minor translation errors which were corrected in the Japanese and Arabic versions of the second test.

In the pilot run, each of the five test sheets used in the first test (one corresponding to each concept) contained all 256 strings arranged in two groups. Those on the left-hand side of the sheet were examples of the concept whilst those on the right-hand side were the remaining non-examples. Ss were therefore shown positive and negative instances simultaneously. Some tended to concentrate on the negative instances (especially where these were fewer in number) and hence to identify and describe the negation of the intended concept. Whilst the relative ease of identification of a concept and its negation would form the basis of an interesting investigation, such a comparison was not intended in this project. It was therefore decided to use only positive instances so that Ss would be obliged to describe the concept itself rather than what the concept was not.

The sixteen test strings in the pilot study were selected at random for each concept. This meant that the four truth table categories were not equally represented and, in particular, for one concept, none of the strings selected exemplified the case where both relevant shapes were absent. It was not therefore possible to ascertain whether a S would have classified these correctly. Including four cards in each truth table category (as in the main study) enables a check for consistency in classifying each category.

The geometric figures used in the second test in the pilot study were four pairs each consisting of a triangle and a circle, with the triangle printed to the left of the circle. The four pairs were: black triangle, white circle (figure A); white triangle, black circle (B); white triangle, white circle (C); black triangle, black circle (D). Ss were provided with eight numbered statements and asked to respond by stating which of the four pairs of shapes were consistent with the statement and which were not. A chance remark from one S revealed that she had difficulty in interpreting a conditional statement due to the relative positions of the two shapes. She could not make sense of the statement 'If the circle is black then the triangle is white' when the triangle was to the left of the circle. It was therefore decided to use single geometric shapes varying on two dimensions (colour and size) rather than two shapes varying on a single dimension (colour).

The eight statements used in the second test were extended to 16 in the main study. These together exhausted all possible combinations of connective/colour/size. This was to enable a fuller investigation of some surprising inconsistencies in responses to statements containing the same linguistic connective which were noted in the pilot study.

In the pilot study, different Ss were used for each of the two tests so that there was no possibility of transfer effects. It was decided that it would be more useful to require the same Ss to complete both tests so that their performance could be compared if necessary. It was felt that transfer effects were unlikely to be significant if the tests were administered in the order indicated.

Chapter 7

RESULTS OF EXPERIMENT 1

The purpose of experiment 1 was to compare performance on five concept identification tasks where the concepts tested were conjunction, inclusive and exclusive disjunction, the conditional and biconditional. The concepts were defined in terms of the presence/absence of two relevant figures (known to the S) in a set of four-figure strings. All positive instances were presented simultaneously for each concept. Whilst we are primarily interested in a comparison of aspects of performance between language groups, we shall also consider the relative performance for concepts within language groups.

In assessing a S's performance on a particular concept identification task, two factors are of interest: (a) responses to the 16 test strings, and (b) the verbal description of the concept. Whilst we might expect a high correlation between these two performance factors, we have already noted that linguistic connectives are used ambiguously. Therefore we cannot assume that successful classification of the test strings is automatically associated with a description which makes clear the distinction between examples and non-examples of the concept.

A S who accurately classifies the test strings as examples or non-examples may be deemed to have identified the concept and to have some adequate mental representation against which the strings can be tested. For the purposes of this test, a S was taken to have identified a concept if the following criterion was satisfied: errors in classifying the strings numbered no more than two and, when two errors were made, these referred to strings in different truth table categories.

This condition may seem somewhat arbitrary and therefore requires some justification. Since there were four test strings in each of the four truth table categories (TT, TF, FT, FF), a S who makes two or fewer errors must have correctly allocated all strings in at least two categories. The condition that two errors must be in different truth table categories ensures that at least three of the four strings in each category have been correctly assigned. With two errors in one category, it is not clear whether or not that category was judged as exemplifying the concept, and hence whether or not the correct concept was being identified. The probability of achieving this criterion on the test strings by chance is sufficiently small ($113(0.5)^{16} = 0.00172$) to be discounted.

For each concept, S's verbal description was classified according to whether or not it provided a correct unambiguous definition of the concept, i.e. whether or not it adequately distinguished examples of the concept from non-examples. (We consider verbal descriptions in more detail in the next section of this chapter.) For a particular concept, we shall denote the set of Ss who satisfy the test string criterion by C and the set of those who gave an accurate description of the concept by D. For each concept the set of all Ss may then be partitioned into four disjoint sets according to which of the two conditions are satisfied:

(i) $C \cap D$: the set of those who achieved criterion on the test strings and who gave a correct unambiguous description of the concept;

(ii) $\bar{C} \cap D$: the set of those whose performance on the test strings did not satisfy the criterion but who gave an accurate description of the concept;

(iii) $C \cap \bar{D}$: the set containing all who achieved criterion on the test strings but who did not describe the concept accurately;

(iv) $\bar{C} \cap \bar{D}$: the set of those whose performance on the test strings did not satisfy the criterion and whose verbal description did not define the distinction between examples and non-examples of the concept.

A summary of the results for each of the concepts biconditional (BIC), conditional (COND), conjunction (CONJ), inclusive disjunction (INCOR) and exclusive disjunction (XOR) is given in table 7.1 below. This gives the numbers in each of the four disjoint subsets, classified by language group and gender.

For each concept, the numbers of males and of females in the set C was calculated, collapsed across language groups. The numbers of males and females in the set D was obtained similarly. Using a log-linear model, the likelihood ratio test was applied to these data to establish, for each concept, whether there were any differences in the proportions of males and females in the sets C and D, i.e. whether males were any more or less likely than females to achieve success on the test string criterion or whether there were gender differences in the ability to describe the concept accurately. The results are given in tables 7.2 and 7.3 below.

Table 7.1: **Results of experiment 1 classified by language group and gender**

		$C \cap D$		$\bar{C} \cap D$		$C \cap \bar{D}$		$\bar{C} \cap \bar{D}$	
		M	F	M	F	M	F	M	F
BIC	English	8	17	2	9	7	0	0	7
	Japanese	8	21	4	9	3	5	1	0
	Arabic	9	10	2	3	5	1	17	5
COND	English	8	17	2	5	3	0	4	11
	Japanese	8	19	3	7	0	3	5	6
	Arabic	9	8	2	2	0	1	22	8
CONJ	English	16	29	1	2	0	2	0	0
	Japanese	16	34	0	0	0	1	0	0
	Arabic	28	17	2	0	1	0	2	2
INCOR	English	16	24	0	1	0	4	1	4
	Japanese	13	35	1	0	1	0	1	0
	Arabic	23	16	4	1	2	0	4	2
XOR	English	10	16	2	2	1	3	4	12
	Japanese	9	26	3	2	2	3	2	4
	Arabic	12	12	3	1	5	0	13	6

(M: number of males, F: number of females,
C: the set of Ss who attained the test sting criterion,
D: the set of Ss who gave a correct description of the concept.)

Table 7.2: **Results of likelihood ratio tests for gender differences in attaining criterion in identifying test strings.**

Concept	χ^2	d.f.	P
BIC	8.64	3	< 5%
COND	3.67	3	> 5%
CONJ	0.03	3	> 5%
INCOR	6.37	3	> 5%
XOR	2.14	3	> 5%

Table 7.3: **Results of likelihood ratio tests for gender differences in correctly describing the concept.**

Concept	χ^2	d.f.	P
BIC	9.05	3	$< 5\%$
COND	2.32	3	$> 5\%$
CONJ	2.50	3	$> 5\%$
INCOR	8.37	3	$< 5\%$
XOR	3.99	3	$> 5\%$

Gender differences, significant at the 5% level, in the proportion of Ss in set C (those who achieved criterion on the test strings) were found for the biconditional concept. Therefore, for this concept, two-tailed tests of the difference in proportions of males and females in the set C were carried out for each language group. A significant difference in the proportion of males (.8824) and the proportion of females (.5152) was found for the English speakers ($z = 2.56$, $P < 1\%$). For the other two language groups there were no significant gender differences for proportions in the set C for the biconditional concept.

For Ss who correctly described the concept (set D), significant gender differences were found for the biconditional and for inclusive disjunction (see table 7.3). For the biconditional, the sole source of this difference was found to be the Arabic group where the proportion of males was .3333 and the proportion of females .6842 ($z = 2.44$, $P < 1\%$). For inclusive disjunction, a significant difference was found between the proportion of males (.8750) and the proportion of females (1.000) in set D for the Japanese language group ($z = 2.13$, $P < 5\%$). There were no other significant gender differences.

The gender differences found are inconsistent with respect to concept and language group. Furthermore, the directions of the differences are inconsistent with respect to gender. For the biconditional and the English language group, males outperformed females in achieving the test string criterion, whereas for the same concept and the Arabic speakers, females outperformed males in correctly describing the concept.

The differences found do not suggest that, for any language group, differential gender performance is a consistent feature of this type of concept identification task. Furthermore, no gender differences have been reported for any of the concept identification tasks referred to in chapter 4 and no significant differences in the performance of males and females have been found in experiments investigating the understanding of logical connectives (Paris, 1973; Sternberg, 1979). It seems probable that the observed gender differences are an experimental artifact and therefore males and females are treated as a single sample in the discussion which follows.

The results of experiment 1 for each concept/language group collapsed across gender are given in table 7.4.

Table 7.4: **Results of experiment 1 by language group**

		$C \cap D$	$\bar{C} \cap D$	$C \cap \bar{D}$	$\bar{C} \cap \bar{D}$
BIC	English	25	11	7	7
	Japanese	29	13	8	1
	Arabic	19	5	6	22
COND	English	25	7	3	15
	Japanese	27	10	3	11
	Arabic	17	4	1	30
CONJ	English	45	3	2	0
	Japanese	50	0	1	0
	Arabic	45	2	1	4
INCOR	English	40	1	4	5
	Japanese	48	1	1	1
	Arabic	39	5	2	6
XOR	English	26	4	4	16
	Japanese	35	5	5	6
	Arabic	24	4	5	19

(Table entries are numbers of Ss)

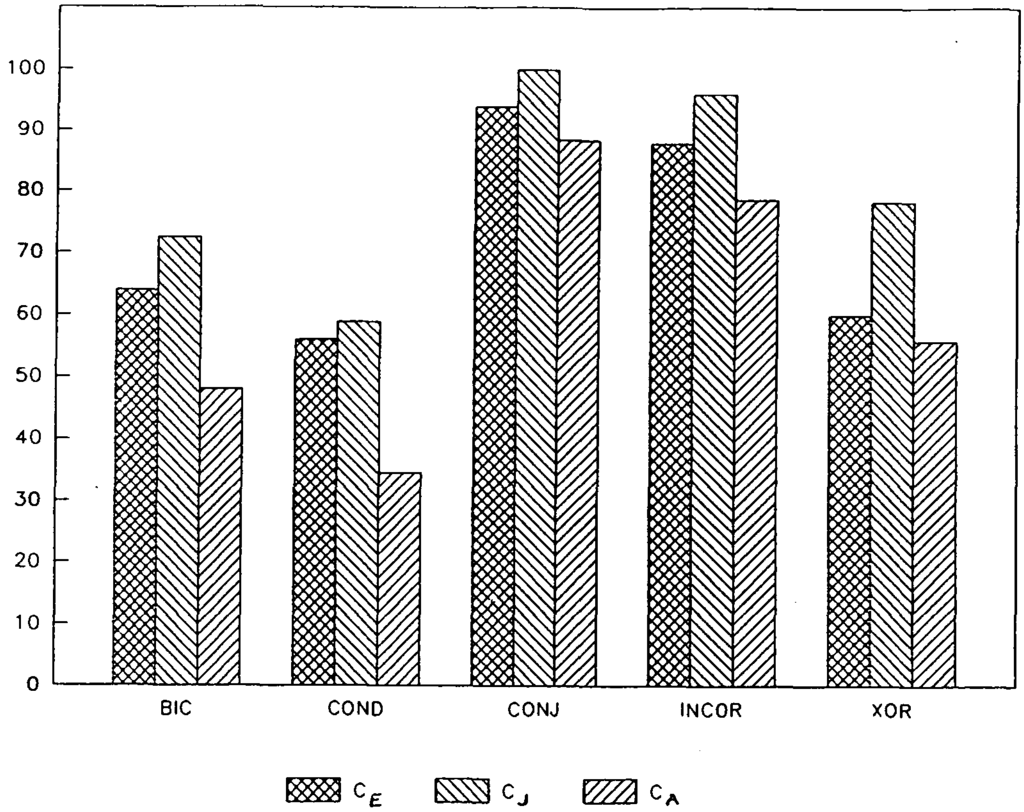
(C: the set of Ss who attained the test sting criterion,
D: the set of Ss who gave a correct description of the concept.)

COMPARISON OF RESULTS BETWEEN LANGUAGE GROUPS

Identification of concepts

Since we know that linguistic expressions of logical concepts do not always convey those concepts unambiguously (see chapter 4), we cannot use a S's description of a concept as an indicator that the concept has been correctly identified. To be confident that a concept has been abstracted, we need to be assured that a S has some mental representation which enables him to distinguish examples from non-examples. We shall therefore take accurate classification of the test strings (i.e. membership of the set C) as the criterion for successful concept identification, although the nature of the descriptions used by these Ss will also be of interest. The percentage of Ss in this set for each language group is shown in figure 7.1. For each concept, we use C_E, C_J and C_A to denote the sets of English, Japanese and Arabic speakers respectively who achieved criterion on the test strings. For every concept, the percentage of successful Ss in the Japanese language group is highest and the percentage of successful Arabic speakers least.

Figure 7.1: **Percentage of language group in set C by concept**



(C_E: set of English speakers who attained the test string criterion;
C_J: set of Japanese speakers who attained the test string criterion;
C_A: set of Arabic speakers who attained the test string criterion.)

Chi-square tests for independence of language group and success in identifying test strings were carried out for each concept. The results, given in table 7.5., are all significant at 5% and therefore indicate that, for all five concepts, success in attaining the test string criterion is not independent of language group.

Table 7.5: **Results of chi-square tests for independence of language group and success in achieving the test string criterion.**

Concept	χ^2	d.f.	P
BIC	6.71	2	< 5%
COND	7.23	2	< 5%
CONJ	6.19	2	< 5%
INCOR	7.04	2	< 5%
XOR	6.51	2	< 5%

To establish the source of these language group differences, significance tests for differences in the numbers of Ss in the set C were carried out, by concept, for each pair of language groups. The results are given in table 7.6.

Table 7.6: **Results of chi-square tests of independence of language group and success in achieving the test string criterion for each pair of language groups.**

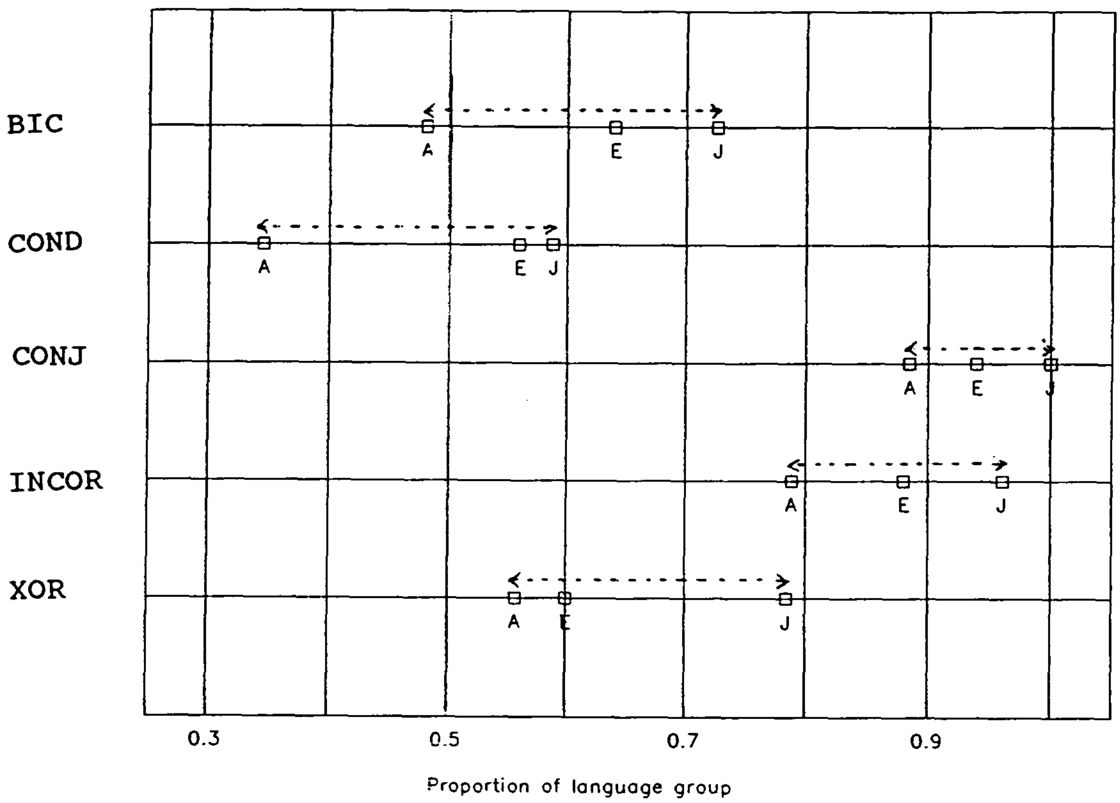
	English/ Japanese	English/ Arabic	Japanese/ Arabic
BIC	0.85	2.62	6.44*
COND	0.08	4.71	6.06*
CONJ	3.15	0.97	6.25*
INCOR	2.26	1.53	6.93*
XOR	4.03	0.19	5.99*

(* indicates significant at the 5% level; d.f. = 2.)

The difference between the proportion of Japanese Ss and the proportion of Arabic speaking Ss who attained criterion on the test strings was significant at the 5% level for all five concepts. These differences are summarised in figure 7.2.

It might be expected that a S who could distinguish examples of a concept from non-examples would also give an accurate description of that concept and vice versa. The results indicate that this is likely but by no means certain. Not all those Ss who identified the concept offered a description which matched their criteria for assigning the test strings and not all those who gave a correct description assigned the test strings in accordance with that description. Further, for those Ss who attained the test string criterion, the likelihood that an accurate description was given varied from concept to concept. Figure 7.3 shows, for each concept and language group, the percentage of those Ss identifying the concept (according to the test string criterion) who also described the concept correctly.

Figure 7.2: **Proportion of each language group in set C by concept**



<------> indicates significant difference at the 5% level

(E: English language group;
J: Japanese language group;
A: Arabic language group.)

Figure 7.3: **Ss in set $C \cap D$ as a percentage of those in set C.**

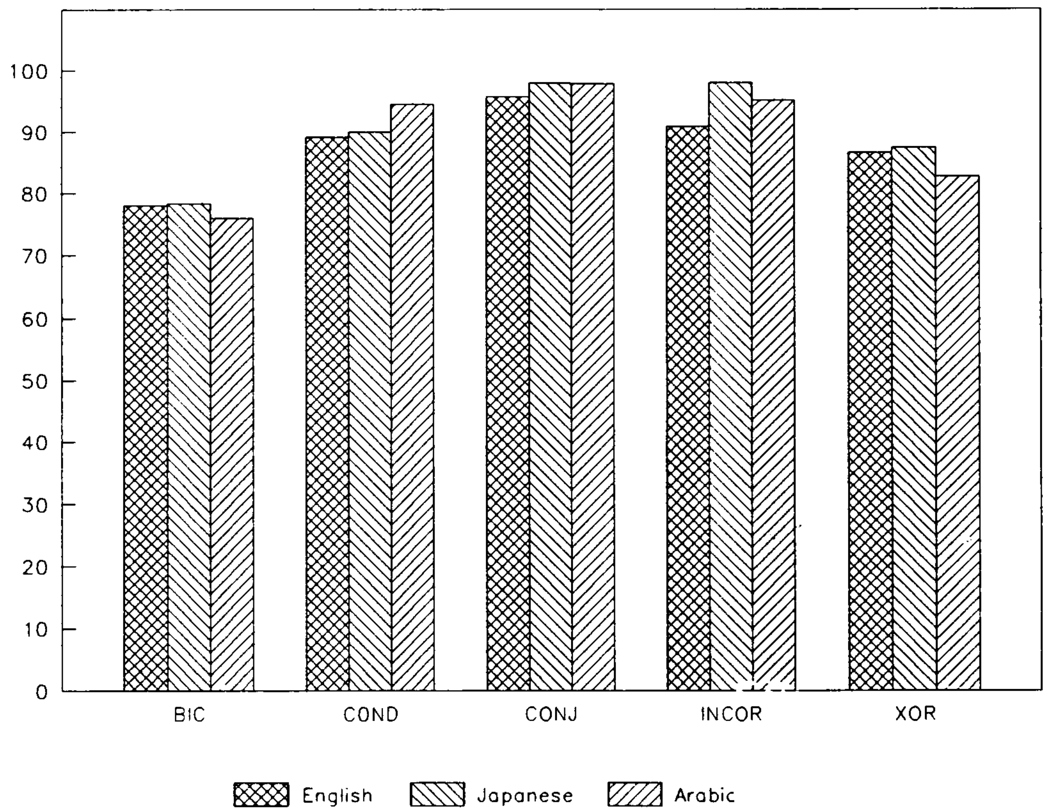


Table 7.7 shows, for each concept, the results of chi-square tests for independence of language group and ability to describe the concept for those Ss who successfully identified it. None are significant at the 5% level. We can conclude that *if* a S had an effective model of the concept which enabled him to distinguish examples from non-examples then, no matter what his language group, he was equally likely to be able to give an adequate description of the concept.

Table 7.7: **Results of chi-square tests for independence of language group and success in describing the concept for Ss in set C.**

Concept	χ^2	d.f.	P
BIC	.05	2	> 5%
COND	.39	2	> 5%
CONJ	.57	2	> 5%
INCOR	2.34	2	> 5%
XOR	.33	2	> 5%

Condensing results across language groups, the proportions of all subjects in set C who also described the concept correctly are as follows: BIC: 0.78; COND: 0.91; CONJ: 0.97; INCOR: 0.95; XOR: 0.86. Hence for these Ss, the order of difficulty of describing the concept was (from least to most difficult): conjunction, inclusive disjunction, the conditional, exclusive disjunction and the biconditional.

We can perform a similar analysis for the proportion of those Ss who described the concept correctly who also attained criterion on the test strings. These proportions are shown in figure 7.4 and the results of the analysis are given in table 7.8.

Figure 7.4: **Ss in set $C \cap D$ as a percentage of those in set D**

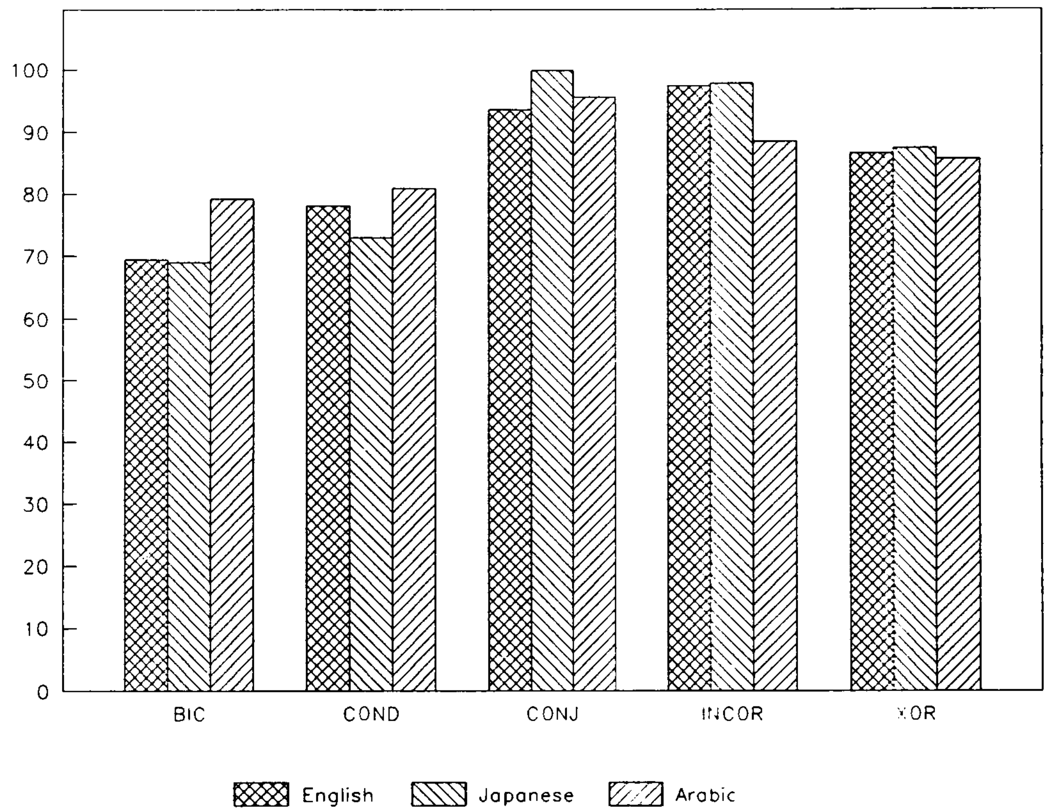


Table 7.8: **Results of chi-square tests for independence of language group and success in identifying test strings for Ss in set D.**

Concept	χ^2	d.f.	P
BIC	0.89	2	> 5%
COND	0.54	2	> 5%
CONJ	3.01	2	> 5%
INCOR	5.00	2	> 5%
XOR	0.05	2	> 5%

Here again, there are no significant differences between language groups. Given that a S described the concept correctly, he was equally likely to attain the test string criterion regardless of his language group. Condensing results across language groups, the proportions of those who gave an accurate description who also identified the test strings are: BIC: 0.72; COND: 0.77; CONJ: 0.97; INCOR: 0.95; XOR: 0.87.

The results of these two tests provide some post-hoc justification for distinguishing performance on the test strings from concept description. Whilst there is clearly some association between these two factors, it is clear that Ss do not necessarily identify examples of the concept by their stated criterion for doing so.

Verbal Descriptions of Concepts

We now look at the types of verbal description used by Ss to define the concepts. Our purpose in doing so is to investigate how logical concepts are coded in natural language. For instance, are there differences between language groups in the use of the linguistic items normally taken to convey the logical forms?

For each concept there are several alternatives which define it in such a way that examples are distinguished from non-examples. In nearly all cases, the description can be translated into its underlying logical form and determining whether it describes the concept accurately consists of comparing its truth table with that of the appropriate connective. If the truth values of the two expressions are identical for each pair of truth values of the simple components, then the expressions are logically equivalent and define the same logical concept.

Translation of Ss' protocols into truth functional form is not always straightforward, particularly where the expression used is complex. Almost invariably, descriptions consist of a string of binary compound propositions (i.e. compound propositions with two simple components) linked by the lexical form of a logical connective, usually (but not always) conjunction or disjunction. The binary propositions are such that, of their two components, one is a statement about one of the two relevant shapes and the second is a statement about the other. We shall refer to these binary propositions as 'elementary propositions'. Examples of elementary propositions are: 'There is a circle and no triangle' or 'If there's a square then there's a star'. A description of, say, the biconditional concept might take the form 'If there's a square then there's a triangle and if there's a triangle then there's a square'. This consists of the conjunction of the two elementary propositions 'If there's a square then there's a triangle' and 'If there's a triangle then there's a square'. If we denote 'there's a square' by p and 'there's a triangle' by q, we can 'translate' this description into the truth functional form $(p \rightarrow q) \wedge (q \rightarrow p)$.

An elementary proposition is either: (a) a binary proposition consisting of two simple components, or (b) the negation of a binary proposition consisting of two simple components. Therefore, all elementary propositions are composed of (or are composed of the negation of) a proposition p (or its negation) conjoined to another proposition q (or its negation) by one of the logical connectives conjunction, inclusive or exclusive disjunction, the conditional or biconditional. Hence the following all symbolise elementary propositions: $p \wedge \bar{q}$, $\bar{p} \rightarrow q$, $\overline{p \vee \bar{q}}$, $\overline{p \veebar q}$. The most succinct description of each concept would, of course, consist of one elementary proposition with p and q joined by the appropriate connective. One description of a concept will be viewed as more complex than another if it contains more elementary propositions. Here we are considering complexity as a function of the length of the description and we shall not therefore attempt to distinguish levels of complexity amongst the elementary propositions themselves.

A method of arriving at a description which can be applied to all five concepts is to classify strings which appear on the sheet into disjoint categories according to the presence/absence of each of the two relevant shapes and then to describe each of the disjoint categories. For example, the strings exemplifying exclusive disjunction may be viewed as falling into the two categories 'circles with no stars' and 'stars with no circles' and described as 'There are circles with no stars or stars with no circles'. This adequately summarises the common property of all strings on the sheet and translates to the truth functional form $(p \wedge \bar{q}) \vee (\bar{p} \wedge q)$. For a description arrived at in this way, each of the disjoint categories corresponds to the conjunction of a statement p or its negation with another statement q or its negation. The concept can then be described by forming the inclusive (or exclusive) disjunction of these conjunctive categories. The application of this 'algorithm' will result in what is termed the 'disjunctive normal form' of the concept. Disjunctive normal forms for each of the concepts are:

biconditional : $(p \wedge q) \vee (\bar{p} \wedge \bar{q})$

conditional : $(p \wedge q) \vee (\bar{p} \wedge q) \vee (\bar{p} \wedge \bar{q})$

conjunction : $(p \wedge q)$

inclusive disjunction : $(p \wedge q) \vee (\bar{p} \wedge q) \vee (p \wedge \bar{q})$

exclusive disjunction : $(\bar{p} \wedge q) \vee (p \wedge \bar{q})$

Since the number of disjoint categories which exemplify each of the concepts is not the same, nor is the number of elementary propositions which comprise the disjunctive normal form. This form requires three elementary propositions for the conditional and inclusive disjunction, two for the biconditional and exclusive disjunction and just one for conjunction.

Although every care was taken in translating Ss' verbal descriptions into propositional form, a certain amount of 'interpretive licence' is unavoidable given the ambiguous way in which linguistic connectives are used in natural language. A description which really did not distinguish those strings exemplifying the concept from those which did not was classified as incorrect. However, there were a few descriptions which communicated the concept adequately although they had a literal logical interpretation which was clearly not what the S intended. For instance, 'and' was sometimes used between compound propositions where inclusive or exclusive disjunction was clearly the intended connective. An example is in describing the biconditional. Some Ss, who had successfully identified the concept according to the test string criterion, defined this as: 'There are stars with squares and no stars with no squares'. The two elementary propositions $p \wedge q$ and $\bar{p} \wedge \bar{q}$ are used here to define the two disjoint sets of strings which appeared on the sheet - those including both shapes and those including neither. If 'and' is translated as conjunction, this would result in the compound proposition $(p \wedge q) \wedge (\bar{p} \wedge \bar{q})$ which would have truth values of 'false' for all truth values of its simple propositions and therefore no strings would be examples of the concept. Since the sheet was clearly not blank, the conclusion is that the correct form $(p \wedge q) \vee (\bar{p} \wedge \bar{q})$ was intended.

There is also the problem with the translation of 'or' as inclusive or exclusive disjunction. Here again, which interpretation was intended was usually obvious from the responses to the test strings. In some cases either form would be correct as in the disjunction of $(p \wedge q)$ and $(\bar{p} \wedge \bar{q})$ to define the biconditional (referred to above) or in any disjunctive normal form. In these cases, and in those where the distinction is not important, 'or' has arbitrarily been taken to indicate inclusive disjunction.

In only a very few cases was it not possible to translate a S's description into truth functional propositional form. This was usually because the S attempted to define the concept in terms of some irrelevant feature of the strings, for example, the frequency of occurrence of a particular shape. Ambiguities in interpretation requiring resolution, other than those referred to above, are described under the results for individual concepts. These follow below.

Results for individual concepts

Biconditional

For Ss whose performance on the test strings indicated that they had identified the biconditional concept, the most commonly used correct description for all language groups was equivalent to the disjunction of $(p \wedge q)$ and one of the logically equivalent forms $(\bar{p} \wedge \bar{q})$ or $(\overline{p \vee q})$. It is sometimes difficult to distinguish which of these is intended. 'No squares and no stars' is clearly to be translated as $\bar{p} \wedge \bar{q}$, whereas 'neither squares nor stars' seems suggestive of $\overline{p \vee q}$. It is less clear whether 'both star and square together or not at all' should be translated as the disjunctive normal form $(p \wedge q) \vee (\bar{p} \wedge \bar{q})$ or as $(p \wedge q) \vee (\overline{p \vee q})$. For this reason, and because the complexity of each expression in terms of number of elementary propositions is the same, responses in both categories are grouped together for the purposes of analysis.

Of the correct descriptions used by Ss in the set C, all but one consisted of at least two elementary propositions. The exception was a Japanese S who gave a description suggestive of the negation of the exclusive disjunction: $\overline{p \veebar q}$ ('The square and the star do not occur separately'). No S in any language group gave a description which had any hint of the lexical 'if and only if ...'.

A breakdown of the correct descriptions used is given in table 7.9 below. The numbers of Ss using each form is divided into those in the set $C \cap D$ (attainment of criterion on test strings and correct description) and those in the set $\bar{C} \cap D$ (failure on test string criterion although correct description).

Amongst the Japanese Ss in $C \cap D$, only two different descriptions of the biconditional were used, whereas six different descriptions were offered by the English speakers in this set and five by the Arabic group. Hence, of all Ss ~~him~~ who identified the biconditional concept (according to the test string criterion) correct descriptions offered by the Japanese group were less varied than those of either the English or Arabic language groups and none consisted of more than two elementary propositions.

Table 7.9: **Breakdown of correct biconditional descriptions**

	English		Japanese		Arabic	
	$C \cap D$	$\bar{C} \cap D$	$C \cap D$	$\bar{C} \cap D$	$C \cap D$	$\bar{C} \cap D$
$(p \wedge q) \vee (\bar{p} \wedge \bar{q})$ or $(p \wedge q) \vee \overline{(p \vee q)}$	18	8	28	13	15	5
$(p \rightarrow q) \wedge (q \rightarrow p)$	3	0	0	0	1	0
$(p \wedge q) \vee [\overline{p \wedge q} \rightarrow (\bar{p} \wedge \bar{q})]$	1	1	0	0	1	0
$(p \vee q) \rightarrow (p \wedge q)$	1	1	0	0	0	0
$(p \wedge q) \vee (\bar{p} \wedge \bar{q}) \vee \overline{(p \vee q)}$	1	0	0	0	0	0
$(p \rightarrow q) \wedge [\bar{p} \rightarrow (\bar{p} \wedge \bar{q})]$	1	1	0	0	0	0
$\overline{p \veebar q}$	0	0	1	0	0	0
$[(p \wedge q) \vee \overline{(p \vee q)}] \wedge \overline{(p \wedge \bar{q})}$	0	0	0	0	1	0
$(p \wedge q) \vee \overline{(p \veebar q)}$	0	0	0	0	1	0
TOTAL	25	11	29	13	19	5

Recall that Ss who failed to attain criterion on the test strings did so because they made three or more classification errors or because they made two errors in the same truth table category. In the case of Ss offering a correct description of the biconditional who failed the test string criterion, nearly all made two or more errors in the FF category, i.e. they misclassified strings containing neither of the two relevant shapes. Of the 11 English Ss in the set $\bar{C} \cap D$, 8 made errors only in the FF category and a further 2 combined two or more errors in that category with single errors in other categories. Similar results apply to the Japanese Ss. Of the 13 in the set $\bar{C} \cap D$, 9 had errors only in the FF class. Of the 5 Arabic Ss in the set $\bar{C} \cap D$, 3 mis-classified all 4 strings where both shapes were absent. Hence, for those Ss who failed the test string criterion even though they gave a correct description of the concept, errors of assignment of strings where both shapes were absent were the most common.

The mis-classification of strings containing neither of the two relevant shapes represents a tendency to respond as if the concept were conjunction. This tendency was also exhibited by some who failed on both the test string and description criteria, i.e. those in the set $\bar{C} \cap \bar{D}$. Of the seven English speakers in this set, 4 described the concept as the conjunction of *p* and *q* and responded to the strings in a manner consistent with this description. (The remaining three Ss gave definitions which could not be translated into propositional form.) Only one Japanese failed both criteria, giving a description which could not be translated into truth functional form but who mis-classified two FF strings. By far the greatest 'failure' rate was in the group of Arabic Ss where 22 failed both criteria. Of these, six described and classified as if the concept were conjunction and two more described conjunction but made classification errors inconsistent with that description. Of the remaining 18, six descriptions could not be assigned a propositional form. The remaining eight Ss gave as many different descriptions and made errors in various and sundry truth table categories.

Amongst those who correctly classified the test strings (according to the criterion used) but described the concept incorrectly (i.e. the set $C \cap \bar{D}$), no pattern was evident. The seven English speakers in this category gave seven different incorrect descriptions. Of the eight Japanese however, six described the concept as the conjunction of *p* and *q* although they were clearly not assigning the strings according to the truth table for conjunction. The six Arabic speakers in the set $C \cap \bar{D}$, like their English counterparts, varied in the descriptions they gave.

Conditional

Of all the five concepts, the conditional gave rise to the greatest variety of descriptions, most of which included redundant elementary propositions. Of the correct descriptions, only two consisting of one elementary proposition were offered. These were the lexical equivalents of $p \rightarrow q$ ('If there was a triangle then there was a square.') and the logically equivalent form $\overline{p \wedge \bar{q}}$ (normally stated as 'There was never a triangle without a square.'). Most other accurate verbal descriptions included one of these two amongst the elementary propositions combined with other forms. Redundancy is therefore a feature of nearly all descriptions involving two or more elementary propositions.

For Ss who attained the test string criterion, as with the biconditional concept, Japanese correct descriptions were very much less varied with 17 of the 27 (63.0%) using one of the two descriptions equivalent to $p \rightarrow q$ or the disjunctive normal form $(p \wedge q) \vee (\bar{p} \wedge q) \vee (\bar{p} \wedge \bar{q})$. The two most popular correct descriptions used by the English speakers in set C were $p \rightarrow q$ and $(p \wedge q) \vee (\bar{p} \wedge q) \vee (\overline{p \wedge \bar{q}})$ but these accounted for only 9 of the 25 (36.0%) Ss in this category. The two descriptions $(p \wedge q) \vee (\bar{p} \wedge q) \vee (\bar{p} \wedge \bar{q})$ and the lengthy $(\overline{p \wedge \bar{q}}) \vee (p \wedge q) \vee (\bar{p} \wedge q) \vee (\overline{p \vee q})$ accounted for 7 of the 17 (41.2%) correct descriptions offered by Arabic speakers who attained the test string criterion. For Ss in set C, only seven different correct descriptions were used by the Japanese language group compared with 10 for the Arabic and 11 for the English speakers.

A breakdown of correct descriptions is given in table 7.10.

Table 7.10: **Breakdown of correct conditional descriptions**

	English		Japanese		Arabic	
	$C \cap D$	$\bar{C} \cap D$	$C \cap D$	$\bar{C} \cap D$	$C \cap D$	$\bar{C} \cap D$
$p \rightarrow q$	5	0	8	4	1	1
$\overline{p \wedge \bar{q}}$	3	0	6	1	0	0
$(\overline{p \wedge q}) \rightarrow \bar{p}$	1	0	0	0	0	0
$q \vee (\bar{q} \rightarrow \bar{p})$	1	0	0	0	0	0
$(\overline{p \wedge \bar{q}}) \vee (\bar{p} \wedge q)$	3	1	0	0	1	0
$(\bar{p} \wedge \bar{q}) \vee (p \rightarrow q)$	1	0	0	0	0	0
$(p \rightarrow q) \vee (\overline{q \rightarrow p})$	0	2	0	0	0	0
$(\overline{p \wedge \bar{q}}) \vee (p \rightarrow q)$	0	1	0	0	1	1
$(p \rightarrow q) \vee (\bar{p} \wedge q)$	0	0	1	1	0	0
$(\bar{p} \wedge q) \vee (\overline{p \wedge \bar{q}})$	0	0	0	0	1	0
$(\bar{p} \wedge \bar{q}) \vee (\overline{p \wedge \bar{q}})$	0	0	0	0	1	0
$(p \wedge q) \vee q \vee (\overline{p \wedge \bar{q}})$	0	0	0	0	1	0
$(p \wedge q) \vee (\bar{p} \wedge q) \vee (\overline{p \wedge \bar{q}})$	4	0	0	2	2	2
$(p \wedge q) \vee (\bar{p} \wedge q) \vee (\bar{p} \wedge \bar{q})$	1	2	9	2	4	0
$(\bar{p} \wedge q) \vee (p \rightarrow q) \vee (\bar{p} \wedge \bar{q})$	2	0	1	0	2	0
$(\overline{p \wedge \bar{q}}) \vee (p \rightarrow q) \vee (\bar{p} \wedge q)$	0	1	0	0	0	0
$(\bar{p} \wedge \bar{q}) \vee (\bar{p} \wedge q) \vee (\overline{p \wedge \bar{q}})$	0	0	1	0	0	0
$(p \wedge q) \vee q \vee [\bar{q} \rightarrow (\bar{p} \wedge \bar{q})]$	1	0	0	0	0	0
$(p \wedge q) \vee (\bar{p} \wedge q) \vee (\overline{p \wedge \bar{q}}) \vee (\bar{p} \wedge \bar{q})$	3	0	1	0	0	0
$(p \wedge q) \vee (\overline{p \wedge \bar{q}}) \vee (\bar{p} \wedge q) \vee (\overline{p \vee q})$	0	0	0	0	3	0
TOTAL	25	7	27	10	17	4

Of the correct descriptions offered by Ss who attained the test string criterion, the proportions which involved only one elementary proposition were 32.0% for English speakers, 51.9% for the Japanese and 5.9% for the Arabic speakers. The Japanese Ss in $C \cap D$ also showed a greater tendency to use the disjunctive normal form $(p \wedge q) \vee (\bar{p} \wedge q) \vee (\bar{p} \wedge \bar{q})$ with 33.3% falling into this category as opposed to 4.0% for the English speakers and 23.5% for the Arabic.

Of Ss who described the conditional concept correctly but who failed the test string criterion, errors in the FF truth table category predominated. Of the 7 English Ss in this class, 6 mis-classified 2 or more strings where both shapes were absent. Of the 10 Japanese in $\bar{C} \cap D$, 8 made two or more FF errors whilst all four Arabic Ss in this class made two or more FF errors.

There were few Ss who achieved the test string criterion but failed to offer an accurate description of the conditional and no particular pattern was evident amongst them. Incorrect descriptions offered by those who failed the test string criterion were very varied in the case of the 15 English and 11 Japanese Ss who fell into this category. Of the 30 Arabic Ss in $\bar{C} \cap \bar{D}$, 7 gave the description $(p \wedge q) \vee (\bar{p} \wedge q) \vee (p \wedge \bar{q})$, the disjunctive normal form for inclusive disjunction. However, none of these gave responses to the test strings consistent with this description.

Conjunction

Table 7.11 below shows the breakdown of correct descriptions of conjunction.

Table 7.11: **Breakdown of correct conjunctive descriptions**

	English		Japanese		Arabic	
	$C \cap D$	$\bar{C} \cap D$	$C \cap D$	$\bar{C} \cap D$	$C \cap D$	$\bar{C} \cap D$
$p \wedge q$	41	3	47	0	44	1
'at least one p and one q'	4	0	3	0	0	0
$(p \wedge q) \vee \overline{(\bar{p} \wedge q)} \vee \overline{(p \wedge \bar{q})}$	0	0	0	0	1	1
TOTAL	45	3	50	0	45	2

The description '(There was) a triangle and a star,' or something very close to it, was offered as the description of the conjunctive concept by nearly all Ss. There were only two other descriptions of conjunction. One of these was 'at least one triangle and one star' which appears to take account of the number of times a shape occurs, an irrelevant property of the strings. Two Arabic Ss used $(p \wedge q) \vee \overline{(\bar{p} \wedge q)} \vee \overline{(p \wedge \bar{q})}$, one of whom failed the test string criterion.

Of the very few Ss who described the concept successfully but failed the test string criterion, errors in the FF category were responsible for both the Arabic failures and one of the three English Ss. The other two English speakers in this group failed because of errors in the mixed (FT, TF) truth table categories.

Disjunction

The ambiguity of the natural language 'or' creates problems in the translation of certain descriptions into truth functional form. For example, a S who describes exclusive disjunction as 'square or circle' cannot be said to have defined the concept unambiguously since we cannot be sure that he does not have inclusive disjunction in mind. The only way to resolve this difficulty is, where a description is simply '*p* or *q*', to examine performance on the test strings and thereby infer the form of disjunction that 'or' is intended to convey. Identification of three or more of the TT strings as examples of the concept is taken to indicate that inclusive disjunction was intended whilst identification of the majority of TT strings as non-examples is assumed to indicate that the disjunction is exclusive. There is, of course, a problem: should a S who fails the test string criterion because of exactly two errors in the TT category be assigned to the set D or $\bar{D}$? Fortunately, this arose only once. One Arabic S described exclusive disjunction using 'either ... or' and made two TT allocation errors (in addition to two TF errors). Since the proportion of Arabic Ss in the set $\bar{D}$ was higher than that for either of the other language groups, it was decided to allocate this S to the set D so as not to inflate any differences which might be of significance.

Where either disjunctive concept is described as 'p or q', the description has been coded as $p \vee q$ or $p \veebar q$ according to performance on the test strings. In descriptions where either interpretation is possible (for example, 'square or circle or both the square and the circle' for inclusive disjunction), a 'default coding' of $p \vee q$ has been used. (This choice of coding is somewhat arbitrary and, given the disagreement over whether natural language use of 'or' corresponds to inclusive or exclusive disjunction, $p \veebar q$ could equally well have been used.)

Inclusive Disjunction

The most succinct description of the inclusive disjunction is 'There is (either) a square or a circle'. (Ss' descriptions did not give any indication that 'either' was used to distinguish exclusive from inclusive disjunction and therefore responses using 'either ... or' are combined with those using 'or' alone.) Of correct descriptions for Japanese Ss who attained criterion on the test strings, 33 (68.7%) fell into this category as opposed to 16 (40.0%) English and 5 (12.8%) Arabic offerings. The difference between each pair of these proportions is significant at the 1% level (one-tailed test). Hence, of the Ss who identified the concept, the Japanese were more likely than either of the other two groups to describe it using the most succinct form with only one elementary proposition.

For those who identified the concept and used 'or' or 'either ... or' to describe it, these clearly have an inclusive sense. Others felt it necessary to add 'or both' suggesting that, for them, 'or' is ambiguous and needs a 'clarifier' or even that 'or' is understood exclusively. These descriptions (which have been 'translated' as $(p \vee q) \vee (p \wedge q)$) accounted for 37.5% of the descriptions of English Ss in $C \cap D$, 16.7% of $(C \cap D)_J$ and 20.5% of $(C \cap D)_A$.

Whilst these two categories of responses accounted for the majority of accurate Japanese and English descriptions, this was not so for the Arabic speakers. For them, the disjunctive normal form $(p \wedge q) \vee (\bar{p} \wedge q) \vee (p \wedge \bar{q})$ accounted for 33.3% of the $C \cap D$ descriptions, whereas for the English and Japanese groups, the corresponding percentages are only 12.5 and 4.2 respectively. A further 7 (17.9%) of $(C \cap D)_A$ described the concept as 'the square and the circle appeared together or separately' which is also suggestive of the disjunctive normal form, although it has been 'translated' as $(p \wedge q) \vee (p \veebar q)$.

A complete breakdown of correct descriptions of inclusive disjunction is given in table 7.12.

Again the Arabic descriptions were more varied - 9 different descriptions of inclusive disjunction were used as opposed to 7 by English Ss and 5 by the Japanese.

Of those Ss who correctly described the concept, only one English and one Japanese speaker failed the test string criterion. Of the Arabic Ss who gave a correct description of the concept, the five who failed the test string criterion used the disjunctive normal form or the related $(p \wedge q) \vee (p \veebar q)$. Of these, two failed to attain criterion because of errors in the FF category and three because of errors in the mixed truth table categories TF and FT.

Table 7.12: **Breakdown of correct descriptions of inclusive disjunction**

	English		Japanese		Arabic	
	$C \cap D$	$\bar{C} \cap D$	$C \cap D$	$\bar{C} \cap D$	$C \cap D$	$\bar{C} \cap D$
$p \vee q$	16	1	33	1	5	0
$(p \vee q) \vee (p \wedge q)$	15	0	8	0	8	0
$(p \wedge q) \vee (\bar{p} \wedge q) \vee (p \wedge \bar{q})$	5	0	2	0	13	3
$(p \wedge q) \vee (p \veebar q)$	0	0	4	0	7	2
$(\overline{\overline{p \vee q}})$	0	0	1	0	1	0
$(p \vee q) \vee (p \wedge q) \vee (\bar{p} \wedge q) \vee (p \wedge \bar{q})$	0	0	0	0	1	0
$(p \vee q) \vee (p \wedge q) \vee (\overline{\bar{p} \wedge \bar{q}})$	0	0	0	0	2	0
$(p \wedge q) \vee (p \wedge \bar{q}) \vee (\bar{p} \wedge q) \vee (\bar{p} \wedge \bar{q})$	1	0	0	0	1	0
$(p \vee q) \vee (p \veebar q)$	0	0	0	0	1	0
$(p \vee q) \vee (\overline{\bar{p} \vee \bar{q}})$	1	0	0	0	0	0
$(p \wedge q) \vee (p \veebar q) \vee (\overline{\bar{p} \wedge \bar{q}})$	1	0	0	0	0	0
$(p \vee q) \vee [(\overline{p \wedge q}) \rightarrow (p \veebar q)]$	1	0	0	0	0	0
TOTAL	40	1	48	1	39	5

Only one Japanese S failed both the test string and description criteria. Of the 6 Arabic and 5 English Ss who did so, a variety of descriptions was offered and no particular pattern is evident.

Exclusive Disjunction

Of the Ss who described this concept using an single 'or' statement, there was a more pronounced tendency to add 'either' than with the descriptions of the inclusive disjunction. However Ss who added the clarifier 'but not both' ($\overline{p \wedge q}$) also had a tendency to use 'either' so again there is no strong evidence that 'either' is being used to distinguish exclusive from inclusive disjunction.

The most popular category of response for all groups was equivalent to $(p \vee q) \vee (\overline{p \wedge q})$, accounting for 53.8% of $(C \cap D)_E$, 45.7% of $(C \cap D)_J$ and 33.3% of $(C \cap D)_A$ descriptions. However the Arabic speakers again favoured the disjunctive normal form $(p \wedge \bar{q}) \vee (\bar{p} \wedge q)$ (16.7% of $(C \cap D)_A$ descriptions) or a form close to it - $(p \wedge \bar{q}) \vee (\bar{p} \wedge q) \vee (\overline{p \wedge q})$ (a further 29.2%). These two forms accounted for 45.8% of all correct descriptions offered by the Arabic Ss in set C, whereas no Japanese and only one English S used either of them.

Five of the 35 Japanese Ss in set C who successfully described the concept and one each of the English and Arabic Ss used 'only one of circle and star' or 'circle and star appeared separately', a form suggestive of $p \veebar q$. These are listed separately in the breakdown of correct descriptions in table 7.13.

Table 7.13: **Breakdown of correct descriptions of exclusive disjunction**

	English		Japanese		Arabic	
	$C \cap D$	$\bar{C} \cap D$	$C \cap D$	$\bar{C} \cap D$	$C \cap D$	$\bar{C} \cap D$
$p \veebar q$ (or/either ... or)	8	0	14	0	4	1
$p \veebar q$ (only one of .../ ... separately)	1	0	5	0	1	0
$(p \vee q) \wedge (\overline{p \wedge q})$	14	4	16	1	8	1
$(p \wedge \bar{q}) \vee (\bar{p} \wedge q)$	1	0	0	0	4	2
$(p \wedge \bar{q}) \vee (\bar{p} \wedge q) \vee (\overline{p \wedge q})$	0	0	0	0	7	0
$(p \vee q) \vee (\overline{p \wedge q}) \vee (\overline{\bar{p} \wedge \bar{q}})$	1	0	0	0	0	0
$(\overline{p \wedge q}) \vee (p \veebar q)$	1	0	0	0	0	0
TOTAL	26	4	35	1	24	4

Of the four English Ss who, although they described exclusive disjunction, failed the test string criterion, two failed because of mis-classification of the FF strings. One mis-classified all four TT cases and was therefore classifying strings according to the truth table for inclusive disjunction even though the description included $(\overline{p \wedge q})$. The only Japanese in $\bar{C} \cap D$ made three FF errors. Of the four Arabic speakers in $\bar{C} \cap D$, only one failed because of errors in the FF category.

Of those who achieved criterion on the strings but who gave a description which was not that of exclusive disjunction, two of the three English and three of the five Japanese described the concept as 'not both the circle and the square', i.e. $\overline{p \wedge q}$. Although this is not equivalent to exclusive disjunction, these Ss all classified the FF strings as non-examples and therefore responded consistently with the correct exclusive disjunction interpretation.

Of those who failed both the description and test string criteria (16 English, 6 Japanese and 19 Arabic speakers), 31% English, 67% Japanese and 58% Arabic speakers described the concept as inclusive disjunction and classified the test strings accordingly. Presumably they had failed to recognise that strings where both shapes were present were not represented on the sheet. The remaining English and Japanese Ss in this class all identified the concept as the negation of the conjunction $\overline{p \wedge q}$ and nearly all classified the test strings consistently with this identification. The remaining Arabic speakers used varied descriptions including $\overline{p \wedge q}$.

In view of the ambiguity of the word 'or' and the question of whether 'or' in natural language normally conveys inclusive or exclusive disjunction, it is useful to compare individual Ss' descriptions of inclusive and exclusive disjunction. Ss in the set $C \cap D$ for both these concepts (20 English, 34 Japanese and

18 Arabic speakers), were grouped according to whether they described either of the concepts simply using 'or' or 'either ... or' without making clear the inclusive or exclusive sense, or whether they used some other description for either form which made clear which sense was intended. Table 7.14 shows the distribution of Ss in $C \cap D$ for both disjunctive concepts grouped into four classes according to the description they used for each form. The four classes are: those who used 'or' or 'either ... or' for both concepts; those who used 'or' or 'either ... or' for inclusive disjunction but not exclusive disjunction; those who used 'or' or 'either ... or' for exclusive disjunction but not inclusive disjunction; those who used other than 'or' or 'either ... or' for both forms of disjunction.

Table 7.14: **Distribution of correct descriptions of inclusive and exclusive disjunction for those Ss who attained criterion on test strings for both concepts.**

ENGLISH		Description of exclusive disjunction		
		'or'/'either ... or'	other	
Description of inclusive disjunction	'or'/'either ... or'	0	5	5
	other	6	9	15
		6	14	20

JAPANESE		Description of exclusive disjunction		
		'or'/'either ... or'	other	
Description of inclusive disjunction	'or'/'either ... or'	12	11	21
	other	2	9	11
		14	20	34

ARABIC		Description of exclusive disjunction		
		'or'/'either ... or'	other	
Description of inclusive disjunction	'or'/'either ... or'	1	2	3
	other	2	13	15
		3	15	18

For these Ss, the pattern of Japanese descriptions is rather different from those of the Arabic and English language groups. For the latter groups, roughly equal numbers within the group used 'or' or 'either ... or' to describe exclusive disjunction as used these to describe inclusive disjunction. No English speaker and only one Arabic speaker used 'or' or 'either ... or' for both concepts. Of the 34 Japanese Ss in $C \cap D$ for both disjunctive concepts, 11 (32%) used 'or' or 'either ... or' for inclusive disjunction but not exclusive disjunction, whilst only 2 (6%) used one of these descriptions for exclusive disjunction alone. However, 12 (35%) used 'or' or 'either ... or' for both concepts. Seven of these used 'or' for inclusive disjunction adding 'either' to describe exclusive disjunction. Only one S used 'or' for exclusive disjunction and 'either ... or' for inclusive disjunction. The remaining 4 Ss used the same description ('or' or 'either ... or') for both inclusive and exclusive disjunctive concepts.

These results suggest that the Japanese Ss had a greater tendency to use 'or' or 'either ... or' for inclusive disjunction rather than exclusive disjunction unlike the other two language groups who were equally likely to use one of these descriptions for either form. Furthermore, the Japanese speakers were more likely to use 'or' or 'either ... or' to describe both forms but when they did so, they tended to use 'or' for inclusive disjunction and to add 'either' to distinguish exclusive disjunction.

Complexity of Verbal Descriptions

In order to compare the complexity of the verbal descriptions, the number of elementary propositions in each was counted for each $C \cap D$ description and the mean calculated for each language group and each concept. The results are given in table 7.15. (A small number of descriptions included mention of just one of the two shapes rather than the usual pair, e.g. '... and if there's a triangle then there's a triangle and a square'. This part of the description was counted as 1.5 elementary propositions.)

Table 7.15: **Numbers of elementary propositions in correct verbal descriptions of concepts used by Ss attaining the test string criterion.**

		English	Japanese	Arabic
BIC	n = 1	0	1	0
	n = 2	22	28	17
	n = 2.5	1	0	2
	n = 3	2	0	0
	Mean	2.10	1.97	2.05
	s^2	0.0800	0.0333	0.0235
COND	n = 1	8	14	1
	n = 1.5	2	0	0
	n = 2	4	1	4
	n = 2.5	0	0	1
	n = 3	7	11	8
	n = 4	4	1	3
	Mean	2.24	1.96	2.79
	s^2	1.2024	1.0727	0.6194
CONJ	n = 1	45	50	44
	n = 2	0	0	0
	n = 3	0	0	1
	Mean	1.0	1.0	1.04
	s^2	0.0	0.0	0.0869
INCOR	n = 1	16	34	6
	n = 2	16	12	16
	n = 3	7	2	15
	n = 4	1	0	2
	Mean	1.82	1.33	2.33
	s^2	0.6444	0.3056	0.6325
XOR	n = 1	9	19	5
	n = 2	16	16	12
	n = 3	1	0	7
	Mean	1.69	1.46	2.08
	s^2	0.2899	0.2482	0.4931

(n = number of elementary propositions)

For every one of the five concepts, the mean number of elementary propositions used by Ss in C ∩ D was least for the Japanese language group, equalled only by the English speakers for conjunction. For all concepts except the biconditional, the mean number of elementary propositions was greatest for the Arabic speakers.

The results of one-way analysis of variance applied to the data in table 7.15 are given in table 7.16 below.

Table 7.16: **Results of one-way ANOVA for equality of mean number of elementary propositions for Ss in C ∩ D across language groups (by concept).**

Concept	Source	SS	d.f.	MS	F	P
	Between groups	0.251	2	.126		
BIC	Within groups	3.413	70	.049	2.58	> 5%
	Total	3.664	72			
	Between groups	7.237	2	3.619		
COND	Within groups	69.552	66	1.054	3.43	< 5%
	Total	76.790	68			
	Between groups	0.060	2	.030		
CONJ	Within groups	3.911	137	.028	1.06	> 5%
	Total	3.971	139			
	Between groups	21.569	2	10.784		
INCOR	Within groups	65.110	124	.525	20.54	< 1%
	Total	86.677	126			
	Between groups	5.590	2	2.895		
XOR	Within groups	28.058	82	.342	8.17	< 1%
	Total	33.647	84			

Significant differences between language groups in the mean number of elementary propositions were found for descriptions of the conditional and both forms of disjunction. To establish the source of these significant differences, Tukey's Honestly Significant Difference (HSD) was calculated for each of these three concepts. These are shown below (table 7.17).

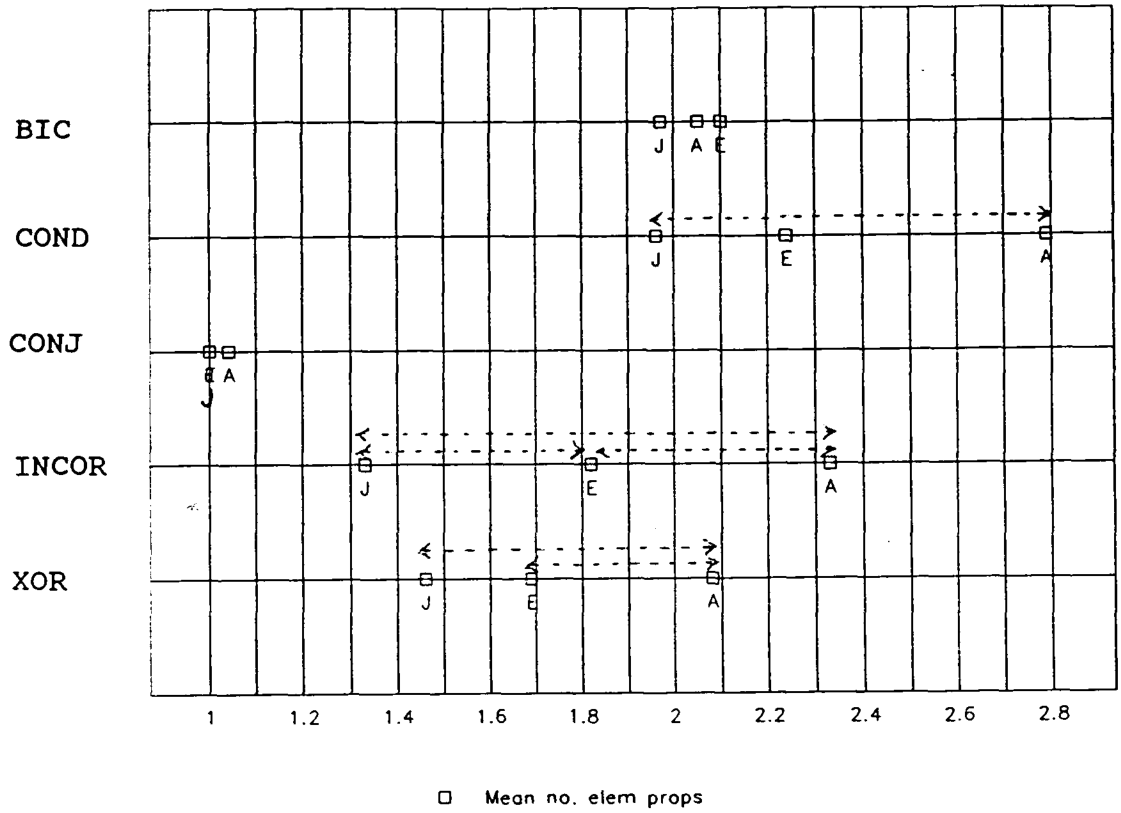
Table 7.17: **Tukey's HSD for mean number of elementary propositions.**

	Tukey's HSD $\alpha = 5\%$
COND	.764
INCOR	.371
XOR	.373

(α = familywise error rate.)

From these we can conclude that, for the conditional, the difference between the means for the Japanese and Arabic language groups is significant. For inclusive disjunction the difference between all three pairs of means is significant. For exclusive disjunction, the difference between the means of the Japanese and Arabic language groups is significant and so is that between the Arabic and English groups. These significant differences are summarised in figure 7.5.

Figure 7.5: **Mean number of elementary propositions in descriptions of concepts**

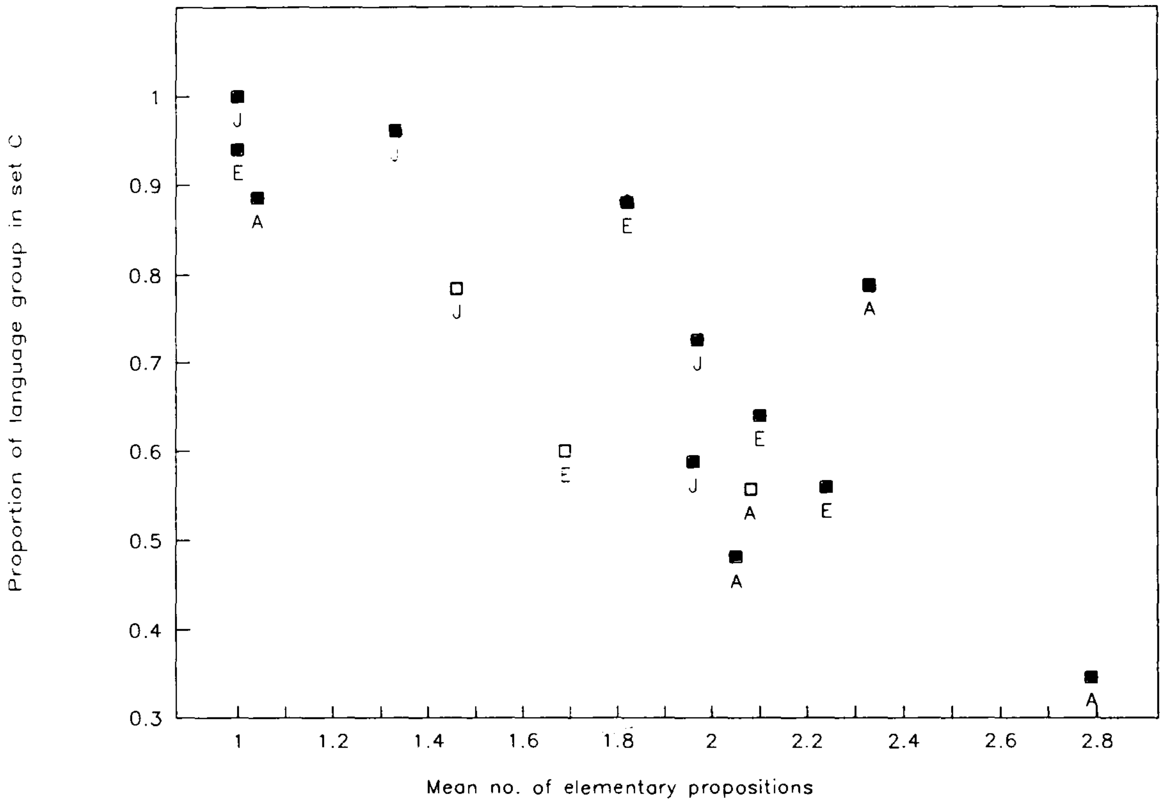


<-------> indicates significant difference (familywise error rate = 5%)

Corresponding differences are also to be found in the number of Ss who described a concept using the simplest possible form (in terms of elementary propositions). For Ss who attained criterion on the test strings, there was a significant effect of language group on the proportions whose descriptions consisted of a single elementary proposition for the conditional (χ^2 = 9.95, d.f. = 2, P < 1%), inclusive disjunction (χ^2 = 27.69, d.f. = 2, P < .1%) and exclusive disjunction (χ^2 = 6.99, d.f. = 2, P < 5%). Differences in these proportions existed for the Japanese and Arabic groups for all three of these concepts. (COND: χ^2 = 9.81, d.f. = 2, P < 1%; INCOR: χ^2 = 27.36, d.f. = 2, P < 0.1%; XOR: χ^2= 6.60, d.f. = 2, P < 5%). For inclusive disjunction, the differences between these proportions were also significant for the English and Japanese language groups (χ^2= 7.30, d.f. = 2, P < 5%) and also for the English and Arabic groups (χ^2 = 7.47, d.f. = 2, P < 5%). These results indicate that, for the conditional and both forms of disjunction, Japanese descriptions tended to be shorter than those offered by the Arabic group and the Japanese descriptions were more likely to be composed of just one elementary proposition.

The differences between the mean number of elementary propositions in descriptions of a concept mirror to a large extent the results for the proportion of Ss in each language group who correctly identified the concept (see figure 7.2). A scatter diagram of mean number of elementary propositions and proportion of Ss attaining criterion on the test strings for each language group/concept combination is given in figure 7.6.

Figure 7.6: **Scatter diagram of mean number of elementary propositions in correct descriptions of concept (for Ss in set C) against proportion of language group identifying the concept.**



: BIC   : COND   : CONJ

: INCOR   : XOR

Across concepts, there appears to be an inverse correlation between success rate in identifying it and the mean length of the description used by successful Ss. In other words, there is a strong tendency for easier concepts to be associated with shorter descriptions.

Times to task completion

We now look at times to task completion for each concept with a view to establishing whether any inter-language group differences exist. (The time to task completion for a concept is the time from presentation of the sheet containing all exemplars to when S announced that he had identified the concept.) Table 7.18 gives the mean and standard deviation of the times to task completion for Ss in the sets C and $\bar{C}$ for each concept/language group combination.

Table 7.18: **Mean and standard deviation of times to task completion for Ss in the sets C and $\bar{C}$.**

		C_E	$\bar{C}_E$	C_J	$\bar{C}_J$	C_A	$\bar{C}_A$
	$\bar{X}$	147.9	97.6	133.9	101.1	103.3	89.2
BIC	s	163.1	90.4	98.6	106.1	62.7	63.6
	n	32	18	37	14	25	27
	$\bar{X}$	309.2	228.7	219.1	232.2	116.2	135.1
COND	s	181.6	196.7	183.1	199.5	83.7	88.4
	n	28	22	30	21	18	34
	$\bar{X}$	58.6	71.0	78.2	-	65.2	72.5
CONJ	s	36.4	39.2	56.5	-	44.3	49.1
	n	47	3	51	0	46	6
	$\bar{X}$	154.1	177.3	110.8	192.0	107.0	104.5
INCOR	s	121.4	200.1	97.3	83.0	79.0	73.3
	n	44	6	49	2	41	11
	$\bar{X}$	57.7	48.2	81.8	82.0	73.8	71.1
XOR	s	50.9	39.2	64.6	80.6	40.5	40.5
	n	30	20	40	11	29	23

(Times are in seconds)

For each concept a two-way analysis of variance was carried out on times to task completion to investigate:

(a) whether these differed across language groups,

(b) whether they differed for the sets C and $\bar{C}$,

(c) whether there were any interaction effects between language groups and the sets C and $\bar{C}$.

The results are given in table 7.19.

Table 7.19: **Summary of two way ANOVA for times to task completion.**

	Source	SS	d.f.	MS	F	P
	Language group	34301	2	17150	1.46	> 5%
	C/$\bar{C}$	34604	1	34604	2.95	> 5%
BIC	Interaction	8046	2	4023	0.34	> 5%
	Error	1722193	147	11716		
	Total	1799144	152			
	Language group	558569	2	279284	10.24	< 1%
	C/$\bar{C}$	10114	1	10114	0.37	> 5%
COND	Interaction	76023	2	38012	1.39	> 5%
	Error	4008553	147	27270		
	Total	4653259	152			
	Language group	9288	2	4644	2.05	> 5%
	C/$\bar{C}$	669	1	669	0.30	> 5%
CONJ	Interaction	49	1	49	0.02	> 5%
	Error	334513	148	2260		
	Total	344518	152			
	Language group	74957	2	37478	3.28	< 5%
	C/$\bar{C}$	4166	1	4166	0.36	> 5%
INCOR	Interaction	11419	2	5709	0.50	> 5%
	Error	1681076	147	11436		
	Total	1771618	152			
	Language group	20420	2	10210	3.47	< 5%
	C/$\bar{C}$	638	1	638	0.22	> 5%
XOR	Interaction	525	2	262	0.09	> 5%
	Error	432011	147	2939		
	Total	453595	152			

There was no effect of success/failure in identifying the concept on times to task completion. In other words, Ss who failed to identify the concept spent no more or less time on the task than those who were successful. Neither was there any interaction between this factor (success/failure) and language group. However, there was an effect of language group for inclusive and exclusive disjunction (significant at 5%) and also for the conditional (significant at 1%).

For the three concepts which showed a language group effect, Tukey's HSDs were calculated to determine significant differences in the mean time to task completion between pairs of language groups. These are given in table 7.20 together with overall means for time to task completion for each language group/concept combination.

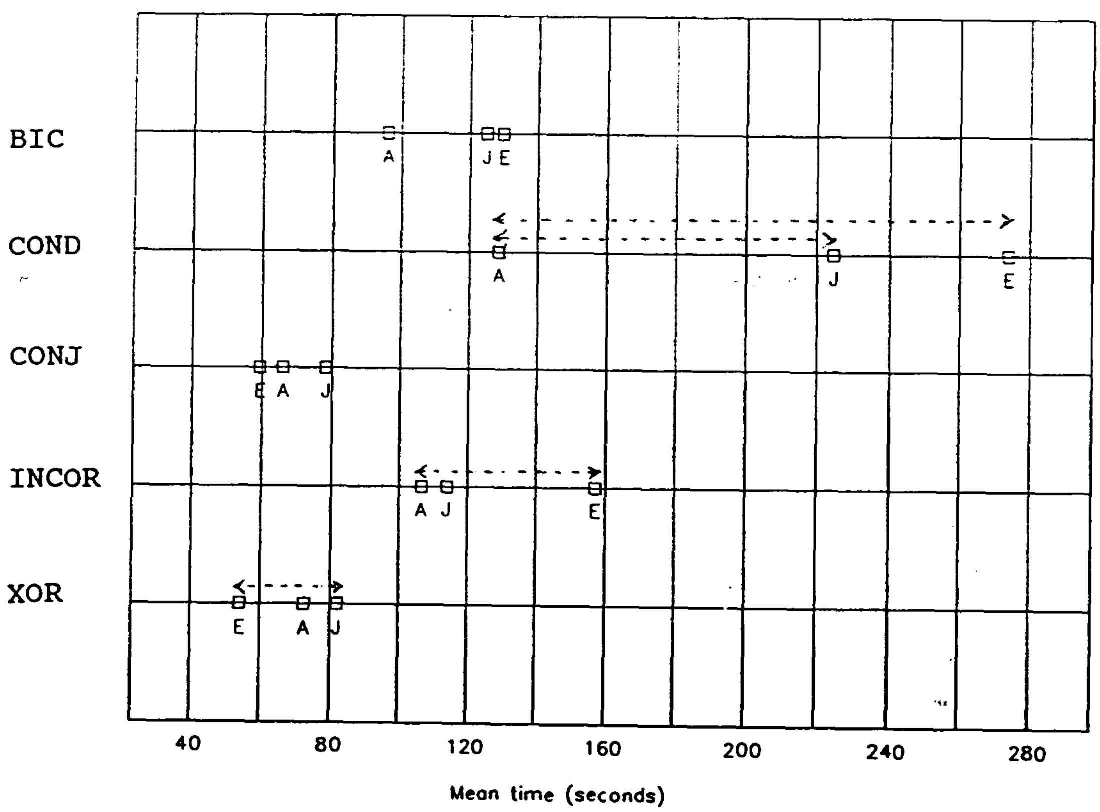
Table 7.20: **Mean times to task completion**

	Language group			
Concept	English	Japanese	Arabic	HSD $\alpha = 5\%$
BIC	129.76	124.90	95.98	-
COND	273.80	224.51	128.56	77.75
CONJ	59.30	78.20	66.06	-
INCOR	156.88	113.96	106.42	50.04
XOR	53.92	81.86	72.58	25.29

(α = familywise error rate.)

For the conditional, the mean time for the Arabic speakers was significantly different from that of the Japanese and that of the English speakers. The mean time for the Arabic group was significantly different from that of the English speakers for inclusive disjunction. For exclusive disjunction, the mean times for the English and Japanese groups were significantly different. These differences are summarised in figure 7.7.

Figure 7.7: **Mean times to task completion**



<------> indicates significant difference (familywise error rate = 5%)

The number of males and females in the Japanese and English language groups was much the same - approximately twice as many females as males. However these proportions were reversed for the Arabic group which contained almost twice as many males as females. To establish whether this factor contributed to the significant differences between mean times for the Arabic speakers and the other two groups, two-sample t tests were carried out for each language group on the times to task completion summed across the five concepts. The results, given in table 7.21, show that there were no gender differences in time spent on the concept identification tasks for any language group.

Table 7.21: **Results of two-sample t tests for total times to complete all five tasks.**

Language group		Males	Females	t	P
English	n	17	33		
	$\bar{X}$	758.4	630.0	1.30	> 5%
	s	329.0	320.0		
Japanese	n	16	35		
	$\bar{X}$	600.4	634.0	0.37	> 5%
	s	280.4	300.5		
Arabic	n	33	19		
	$\bar{X}$	454.1	496.5	0.60	> 5%
	s	234.6	256.5		

COMPARISON OF RESULTS WITHIN LANGUAGE GROUPS

Identification of concepts

We now look at a comparison of performance on each of the five concept identification tasks within the three language groups. Figure 7.8 shows, for each language group, the proportion of Ss who satisfied the test string criterion. It suggests that the order of difficulty in identifying the concepts was much the same for all three groups. The least difficult was conjunction followed by inclusive disjunction and the most difficult for all groups was the conditional.

A multiple comparison test (Grizzle et al, 1969) for equality of the proportions of Ss correctly identifying each concept (i.e. satisfying the test string criterion) was carried out for each language group. The results are given in table 7.22.

Figure 7.8: **Percentage of language group in set C for each concept.**

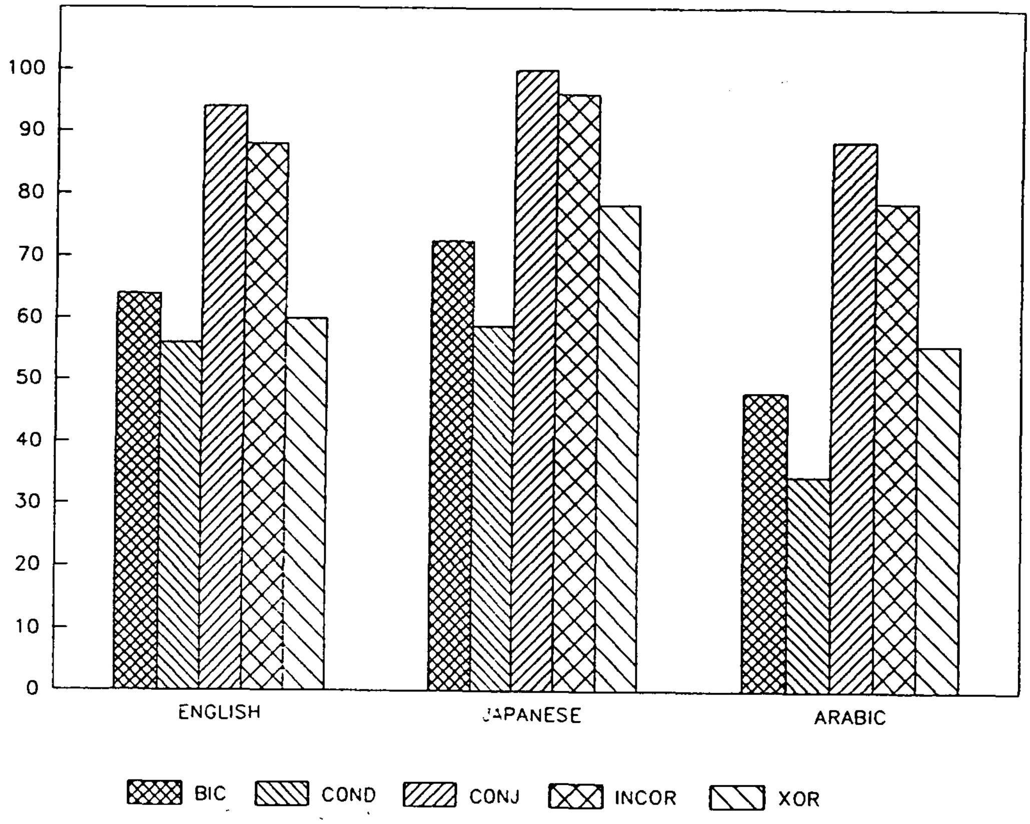


Table 7.22: **Results of tests of equality of proportions of Ss in set C by language group.**

Language group	χ^2	d.f.	P
English	48.8	4	< 1%
Japanese	55.8	4	< 1%
Arabic	42.8	4	< 1%

For every language group, the concepts fell into two homogeneous subgroups, the first containing conjunction and inclusive disjunction and the second containing exclusive disjunction, the conditional and biconditional. There were no significant differences

in proportions for concepts within a subgroup but for those in different subgroups, the difference in proportions was significant. Hence for every language group, the proportion identifying conjunction was significantly differently from each of the proportions identifying exclusive disjunction, the conditional and biconditional and similarly for the proportions identifying inclusive disjunction. In other words, for all groups, conjunction and inclusive disjunction were significantly easier than each of exclusive disjunction, the conditional and biconditional.

Times to task completion

Times to task completion were compared within each language group to establish whether there were any significant differences. An analysis of variance (repeated measures) was carried out for each language group. The results are given in table 7.23. All are significant at the 1% level.

Table 7.23: **Results of ANOVA (repeated measures) for times to task completion within language groups.**

	Source	SS	d.f.	MS	F	P
	Between Ss	1080998	49			
	Within Ss					
ENG	Concept	1603750	4	400938	27.39	< 1%
	Error	2869162	196	14639		
	Total	5553911	249			
	Between Ss	886133	50			
	Within Ss					
JAP	Concept	717853	4	179463	15.1	< 1%
	Error	2377504	200	11888		
	Total	3981490	254			
	Between Ss	617624	51			
	Within Ss					
ARAB	Concept	134793	4	33698	13.9	< 1%
	Error	493178	204	2418		
	Total	1245595	259			

To ascertain the source of the significant concept effects, the Newman-Keuls procedure was used for multiple comparisons of mean times to task completion within language groups. The results are shown in table 7.24.

Table 7.24: **Results of Newman-Keuls tests for significant differences in mean times to task completion within language groups.**

(Figures in parentheses are mean times to task completion for the concept for all Ss in the language group.)

ENGLISH

		XOR (53.9)	CONJ (59.3)	BIC (129.8)	INCOR (156.9)	COND (273.8)	r	W_r
XOR	(53.9)	-	5.38	75.8*	103.0*	219.9*	5	67.1
CONJ	(59.3)	-	-	70.5*	97.6*	214.5*	4	63.1
BIC	(129.8)	-	-	-	27.1	144.0*	3	57.5
INCOR	(156.9)	-	-	-	-	116.9*	2	47.9
COND	(273.8)	-	-	-	-	-		

JAPANESE

		XOR (78.2)	CONJ (81.9)	INCOR (114.0)	BIC (124.9)	COND (224.5)	r	W_r
XOR	(78.2)	-	3.7	35.8	46.7	146.3*	5	59.8
CONJ	(81.9)	-	-	32.1	43.0	142.6*	4	56.3
INCOR	(114.0)	-	-	-	10.9	110.5*	3	51.3
BIC	(124.9)	-	-	-	-	99.6*	2	42.7
COND	(224.5)	-	-	-	-	-		

ARABIC

		CONJ (66.1)	XOR (72.5)	BIC (96.0)	INCOR (106.4)	COND (128.6)	r	W_r
CONJ	(66.1)	-	6.4	29.9*	40.3*	62.5*	5	26.7
XOR	(72.5)	-	-	23.5*	33.9*	56.1*	4	25.2
BIC	(96.0)	-	-	-	10.4	32.6*	3	22.9
INCOR	(106.4)	-	-	-	-	22.1*	2	19.1
COND	(128.6)	-	-	-	-	-		

(* indicates significant difference, α = 5%.)

The results for the English and Arabic speakers are similar - three homogeneous subgroups of concepts. The first contains conjunction and exclusive disjunction, members of the second are the biconditional and inclusive disjunction and the third contains only the conditional. Significant differences exist between the mean times for concepts in different subgroups. For the Japanese language group, there are two homogeneous subgroups. One contains conjunction, both forms of disjunction and the biconditional. The second contains the conditional which is therefore significantly different from all the other concepts. However there are no significant differences among concepts in the first subgroup. Therefore, for all language groups, the conditional task took significantly longer than any other. For the English and Arabic groups, Ss spent less time on conjunction and exclusive disjunction than any other task. However, for the Japanese Ss, the time spent on these two tasks was no different from that spent on inclusive disjunction and the biconditional.

SUMMARY OF RESULTS OF EXPERIMENT 1

An indicator of the difficulty of a concept identification task for a given language group is the proportion of Ss who successfully identified it. Using this measure to rank the concepts, we find that the order of difficulty within each language group is much the same. Conjunction and inclusive disjunction were the easiest concepts with the conditional as the most difficult for all groups. The only difference between language groups was in the relative difficulty of exclusive disjunction and the biconditional. For the English speakers, the success rate for the biconditional was slightly greater than that for exclusive disjunction, whereas for the Japanese and Arabic groups, the success rate for exclusive disjunction was slightly greater than that for the biconditional. However, within language groups, the differences in the proportions of Ss who successfully identified each of these concepts is too small to be of any significance and, for all groups, these two concepts were more difficult than conjunction and inclusive disjunction, and significantly so for the English and Arabic speakers. Hence, if we compare the relative difficulty of concepts, we find no noteworthy differences across the three language groups.

If we compare performance on individual concepts, we find consistent and striking inter-language group differences. For all five tasks, the success rate at identifying the concept was highest for the Japanese group and least for the Arabic speakers. Furthermore, the difference between the proportion of successful Japanese and successful Arabic speakers was significant for every concept.

Striking language group differences were also found in the nature and length of the descriptions used by Ss to describe the concepts. For every concept, the length (measured by number of elementary propositions) of the Japanese descriptions were, on the average, shorter than those of either the Arabic or English speakers. Japanese descriptions were also less varied than those of either of the other two groups for every concept.

The remarkable consistency of these findings across concepts suggests that the superior performance of the Japanese is not an experimental artefact. For some reason, the Japanese as a group were better able to identify the concepts and to describe them economically, although the relative difficulty of the concepts themselves was no different for the Japanese than for either of the other groups.

Chapter 8

RESULTS OF EXPERIMENT 2

The object of experiment 2 was to investigate how the linguistic items normally taken to codify logical connectives were interpreted in each of the three language groups investigated. The test used to obtain this information consisted of 16 compound propositions each of which was composed of two simple components conjoined with the lexical form of either conjunction, disjunction, the conditional or biconditional. (The linguistic connectives used for each of the language groups are described in chapter 5). Of the 16 compound propositions, there were four containing each of the four linguistic forms. This enabled a check for consistency of interpretation.

In each compound proposition, one component made a statement about the size of a square (small/large) and the other was a statement about its colour (black/white). Four squares were printed on the test sheet, one representing each size/colour combination: large/white, large/black, small/white, small/black. Ss were required to respond to each compound proposition by stating which of the four squares were consistent with that proposition and which were not.

For a given compound proposition, the four squares represent each of the four truth table categories TT, TF, FT and FF. For instance, consider the proposition 'If the square is large then it is white'. In this case the large white square constitutes the case where both simple components are true, that is the TT case. Both components are false for the small black square so that this represents the FF truth table category. The small white and large black squares constitute the FT and TF cases respectively.

From his or her responses to each of the 16 compound propositions, a S's truth table for that proposition could be inferred. For instance, suppose a S responds to the proposition 'If the square is large then it is white' by indicating that the large white and small black squares are consistent with it whereas the large black and small white squares are not. This corresponds to an interpretation in which the compound proposition is judged as true only when its simple components are both true (the TT case) or both false (FF). This corresponds to the underlying truth table for a biconditional and we therefore deduce that this is the S's interpretation of the logical relationship between the two simple propositions. In a similar way, truth tables, and hence underlying logical connectives, were inferred for each S for all 16 of the test propositions.

Since each truth table category (represented by the four squares given) could be marked consistent or inconsistent, there are 2^4 = 16 possible responses to each test proposition. However, nearly all responses offered by the Ss tested corresponded to the truth tables for one of the following: $p \wedge q$, $p \vee q$, $p \rightarrow q$, $p \leftrightarrow q$, $p \veebar q$, *p* (affirmation of *p*) or *q* (affirmation of *q*). Responses corresponding to truth functional propositions other than these were rare and varied and have therefore been grouped together in the analysis which follows. Table 8.1 shows, for each of the 16 test propositions, the number of responses in each language group corresponding to the truth table representations of conjunction, inclusive and exclusive disjunction, the conditional, biconditional and the affirmation of *p* (the first component of the compound proposition) and of *q* (the second component of the compound proposition). Shaded cells are those where the inferred logical interpretation is that which the linguistic connective is supposed to convey.

Table 8.1. **Summary of results of experiment 2**

Prop. no.	Response								
	$p \wedge q$	$p \vee q$	$p \to q$	$p \leftrightarrow q$	$p \veebar q$	p	q	other	
1.	49	1							Eng
	51								Jap
	51							1	Arab
2.	43		3	1		1		2	Eng
	35		4	9		1		2	Jap
	38		10	3				1	Arab
3.	7	24			5	8	4	2	Eng
	10	26			1	9	4	1	Jap
	5	35				7	4	1	Arab
4.	41		1	4		2	1	1	Eng
	35			14			2		Jap
	40	2	4	2		1	1	2	Arab
5.	6	29		1	6	5	1	2	Eng
	6	32				4	9		Jap
	7	33			1	6	3	2	Arab
6.	41		2	3		3		1	Eng
	38		6	7					Jap
	37		6	4			2	3	Arab
7.	41		1	3		2	1	2	Eng
	34			13		3		1	Jap
	37	1	4	2		3	1	4	Arab
8.	47			2				1	Eng
	50						1		Jap
	47	1	1	1				2	Arab
9.	41		2	2		2	1	2	Eng
	33			13		1	3	1	Jap
	38	2	4	4		1		3	Arab
10.	39		3	2		4		2	Eng
	36		6	8		1			Jap
	38		6	2		1	1	4	Arab
11.	6	30			5	5	2	2	Eng
	4	36				4	6	1	Jap
	7	33				10	1	1	Arab
12.	47	1	1					1	Eng
	49						1	1	Jap
	45	1	1	2		2		1	Arab
13.	39		1	5		1	2	2	Eng
	34		1	13	1	2			Jap
	39	1	3	4		1		4	Arab
14.	6	29			5	6	1	3	Eng
	5	37				4	5		Jap
	5	31				7	8	1	Arab
15.	49				1			1	Eng
	51								Jap
	48		2	1		1			Arab
16.	40		4	1		2		3	Eng
	35		6	8		1	1		Jap
	38		7	4			1	2	Arab

The results indicate that propositions in which the components were joined by the equivalent of 'and' (1, 8, 12 and 15) were interpreted as logical conjunctions by nearly all Ss regardless of language group. Statements containing the linguistic form of disjunction (3, 5, 11 and 14) tended to be interpreted as inclusive disjunction by the majority of Ss. Exclusive interpretations were negligible for the Japanese and Arabic groups. However, each disjunctive statement elicited an exclusive interpretation from approximately 10% of the English speakers. Of those responses which did not correspond to either of the logical forms of disjunction, nearly all indicated that the statement had been construed either as a conjunction or as the affirmation of one of its two components (usually the first) with the other disregarded.

If we view the logical interpretation as the 'correct' one, by far the greatest error rate occurred in interpreting the linguistic conditionals and biconditionals. Of the conditional statements (2, 6, 10 and 16), statement 2 elicited a conditional interpretation from nearly 20% of the Arabic speakers. Otherwise a conditional response to any of these statements did not exceed 12% of any language group. Linguistic biconditionals (statements 4, 7, 9, and 13) were even less successful in eliciting an interpretation consistent with a logical biconditional for the English and Arabic groups. However, each one of the biconditional statements was given the correct logical reading by approximately 25% of the Japanese group. Of those who failed to interpret conditionals or biconditionals as logic dictates, the majority of responses to these statements revealed an underlying truth table corresponding to logical conjunction, i.e. only the TT case was considered to be consistent with the test proposition.

In table 8.1, the rows corresponding to statements containing the same linguistic connective are not identical. We can therefore deduce that Ss were not always consistent in their interpretation of a given linguistic form. For instance, a S might interpret two of the four propositions expressed using 'if ... then' as conditionals whereas the other two might elicit (from the same subject) the truth table corresponding to conjunction. To establish whether the interpretations occurred with any degree of consistency, each of the four responses to the propositions containing the same connective were compared for each S. If an individual's responses to three or more of the four statements containing the same linguistic form indicated the same logical interpretation, then that S was deemed to have interpreted that connective consistently. Furthermore, S's logical interpretation of the connective was inferred to be that corresponding to the truth table of the three or more consistent responses. If no three responses corresponded to the same truth functional form, then the S's interpretation of that linguistic connective was classified as inconsistent.

Table 8.2 shows, for each group of four propositions containing the same linguistic connective, the frequency distribution of consistent logical interpretations. For each group of four propositions, table entries are numbers of Ss in each language group who gave three or more responses corresponding to the logical form indicated.

Table 8.2: **Frequency distribution of truth functional interpretations of linguistic connectives.**

Ling. conn.	Logical interpretation								
	$p \wedge q$	$p \vee q$	$p \rightarrow q$	$p \leftrightarrow q$	$p \veebar q$	p	q	incon-sistent	
'and'	48							2	Eng
	51								Jap
	48		1	1				2	Arab
'or'	6	27			5	3		9	Eng
	5	33				2	1	10	Jap
	4	31				3	1	13	Arab
'if ... then'	40		2	1		1		6	Eng
	37		6	8					Jap
	36		8	3				5	Arab
'if and only if'	41		1	2		1	1	4	Eng
	33			12				6	Jap
	38		4	2				8	Arab

Results for individual connectives are considered below.

Results for individual connectives

Conjunction

Four propositions were of the form 'p and q', e.g. 'The square is black and it is large'. For Ss in all three language groups, virtually all of the responses to these propositions indicated an interpretation consistent with the truth table for conjunction. Of the English speakers, 48 (96%) responded in this manner to three or more of the propositions containing 'and'. Corresponding figures for Japanese and Arabic speakers are 51 (100%) and 48 (92.3%). Only 2 Ss (Arabic speakers) gave consistent responses indicative of an interpretation other than logical conjunction and, for the few whose responses were inconsistent, no particular pattern of errors was evident.

Disjunction

Of Ss who interpreted three or more of the four propositions containing 'or' consistently, the most popular response in all language groups corresponded to inclusive disjunction. Ss who interpreted the majority of 'or' propositions in this way numbered 27 (54%) English, 33 (64.7%) Japanese and 31 (59.6%) Arabic speakers respectively. Differences between these proportions are not significant ($\chi^2 = 1.20$, d.f. = 2, $P > 5\%$). Whilst none of the Japanese and Arabic speakers showed any tendency to interpret 'or' as exclusive disjunction, 5 English speakers gave a majority of responses which indicated an exclusive reading. This suggests that 'or' in English may be more likely to be read exclusively than disjunctions in Japanese and Arabic.

In all three language groups, there were a small number of Ss who interpreted 'or' propositions as conjunctions. The numbers of Ss in the English, Japanese and Arabic groups who exhibited this tendency were 6, 5 and 4 respectively.

Of the four linguistic connectives, 'or' had a greater tendency to elicit inconsistent interpretations. Responses which fell into this category generally included inclusive disjunction along with the affirmation of *p* or of *q*.

Conditional

No matter what their language group, there was a pronounced tendency for Ss to respond to conditional propositions as if they were conjunctions. Ss whose responses to the conditional propositions indicated a conjunctive reading numbered 40 (80%) English, 37 (72.5%) Japanese and 36 (69.2%) Arabic speakers. There is no significant difference between these proportions ($\chi^2 = 1.60$, d.f. = 2, $P > 5\%$).

Ss who interpreted these propositions as conditionals were few in number - 2 English, 6 Japanese and 8 Arabic speakers. The only other interpretation worthy of note was equivalent to the biconditional. The tendency to give a biconditional reading to the conditional statements was greatest amongst the Japanese of whom eight interpreted the majority of conditional propositions as biconditionals. Only 3 of the Arabic language group and one English speaker's responses indicated a consistent biconditional interpretation.

Biconditional

As with the conditional propositions, there was a pronounced tendency, independent of language group, to interpret biconditional propositions as conjunctions of the two simple components. Of the English speakers, 41 (82%) responded to the majority of biconditional propositions as if they were conjunctions. Corresponding figures for the Japanese and Arabic speakers are 33 (64.7%) and 38 (73.1%) respectively.

Of the three language groups, a higher proportion of the Japanese speakers responded in a manner consistent with a biconditional interpretation. Twelve Japanese interpreted the majority of biconditional propositions as such whereas only 2 each of the English and Arabic speakers did so. However, the apparently superior performance of the Japanese group is probably attributable to the fact that the biconditional propositions which they were interpreting were expressed as the conjunction of two conditionals (see chapter 5).

Interpreting the biconditional as a conditional was rare - only 4 Arabic speakers and 1 English speaker did so.

SUMMARY OF RESULTS OF EXPERIMENT 2

A striking result of this test is the overwhelming tendency for Ss, irrespective of their language group, to respond to all propositions other than disjunction as if they were conjunctions, i.e. true if both their simple components were true and false otherwise. Thirty-two (64%) of English speakers, 28 (54.9%) of Japanese and 30 (57.7%) of the Arabic language group gave truth table interpretations corresponding to logical conjunction for all linguistic connectives other than 'or'. The tendency to respond in this way is independent of language group (χ^2 = .91, d.f. = 2, P > 5%).

Although 'or' statements were more likely to be interpreted 'logically' than either conditionals or biconditionals, of the four linguistic connectives tested, disjunctions had a greater tendency to elicit responses which were inconsistent across the four items testing interpretation of the same linguistic form.

The results of experiment 2 revealed no major differences between language groups in the truth table interpretations of the linguistic connectives which are generally taken to convey the logical forms. Minor differences of interest are:

(a) 10% of the English speakers interpreted 'or' statements as exclusive disjunction whereas no S in either of the other two language groups did so;

(b) Japanese and Arabic speakers were slightly more likely than English speakers to interpret statements containing 'if ... then' as logical conditionals, and Japanese speakers were more likely than either English or Arabic speakers to interpret these as biconditionals;

(c) for the biconditional, the truth table corresponding to the logical form was elicited from more Japanese speakers than either of the other two groups but this is probably attributable to the fact that, for reasons given, it was expressed in Japanese as the conjunction of two conditionals.

Chapter 9

DISCUSSION OF RESULTS

This study was designed to investigate whether the widely reported mismatch between logical connectives and the associated English language items is generalisable to speakers of other languages. It also seeks to establish whether differences in the interpretation of these lexical items might be linked with differential familiarity with the associated logical concepts. Success at identifying logical concepts from non-verbal materials was tested in experiment 1, whilst experiment 2 sought to obtain information regarding the comprehension of linguistic connectives when no contextual clues are available to aid interpretation. Since the aim was to explain dissimilar concept identification performance in terms of differential linguistic provision for expressing the underlying logical concepts, we consider first the results of experiment 2.

UNDERSTANDING OF LINGUISTIC CONNECTIVES

Comparison across language groups

Experiment 2 revealed only minor differences between language groups in the comprehension of the linguistic items normally taken to convey the logical forms of conjunction, disjunction, the conditional and biconditional. There were no language group differences at all in the comprehension of 'and' statements which were understood as logical conjunctions by at least 92% of the Ss in each of the three groups.

Linguistic disjunctions were less successful in eliciting an interpretation which corresponded to either of the two logical forms. Between 60% and 65% of each language group interpreted three or more of the 'or' statements as either inclusive or exclusive disjunction. For all language groups the inclusive interpretation was overwhelmingly favoured. However, there was a greater tendency for the English speakers to read 'or' exclusively than was the case for either of the other two groups. Of those who consistently interpreted the disjunctions in accordance with the truth table for either logical form, 16% of English speakers construed them exclusively. However, no S in either the Japanese or Arabic language groups gave 'or' an exclusive reading. This suggests that it is possible that an ambiguous disjunction in English, although likely to be read inclusively, may be more susceptible to an exclusive interpretation.

Linguistic conditionals were poorly understood by all language groups, if we consider the logical interpretation to be the correct one. The overwhelming tendency for all three groups was to interpret conditionals as if they were conjunctions, i.e. true only when their two components were true and false otherwise. A conditional interpretation was marginally more likely to be offered by the Arabic speakers than by the Japanese who in turn were slightly more likely than the English speakers to interpret conditionals 'logically'. There does, however, seem to be a greater tendency for a Japanese conditional to be interpreted as a biconditional. Of the consistent interpretations of the conditional, 19% of those offered by the Japanese were biconditional. Corresponding figures for the English and Arabic speakers were 2% and 7%.

Biconditionals were as poor at eliciting the logical interpretation as conditionals and were also understood as conjunctions by the majority of Ss. Whilst the Japanese biconditionals were more often construed as such, the reason is almost certainly attributable to their expression as the conjunction of two conditionals. Of interest is whether biconditionals expressed in this way would be more successful at eliciting the logical interpretation for the English and Arabic speakers.

Of the four connectives tested, only linguistic conjunction and disjunction seem to codify adequately the underlying logical concepts. The generality of this finding across language groups lends some support to Gazdar's (1979, chapter 4) proposal that lexical encoding of logical connectives in all languages is confined to conjunction and disjunction. It does indeed seem to be the case that these two logical concepts have some particular salience which necessitates their encoding and that this may be independent of culture.

Discussion of results for individual connectives

The results of experiment 2 suggest that the linguistic codings of logical connectives in English, Japanese and Arabic have interpretations which are very similar in each language. However, the connectives themselves are not equally successful at conveying the associated logical concept. We now discuss the general results for individual connectives in the light of other research findings which have attempted to elucidate how linguistic connectives are understood in English.

Experiments, such as experiment 2, which have been designed to elicit all or part of the truth table underlying a particular linguistic connective generally use one of two types of task. In one, the S is given a truth table category and asked to evaluate the truth of a compound proposition. For instance, he may be asked to assess the truth or falsity of 'M is odd or N is even' given 'M = 3' and 'N = 5', the TF case. Alternatively the task may take the form used in this study - the S is given a compound statement which he is to assume to be true and asked to assess one or more of the truth table categories. Using the example above, the S would be given 'M is odd or N is even' and asked to evaluate 'M = 3, N = 5'. Either design can allow identification of the truth table underlying a connective but only if all four categories (TT, TF, FT, FF) are tested and the results analysed for individual Ss.

Understanding of linguistic conjunction

Regarding the interpretation of 'and' statements, there is little to say other than to point to the fact that these seem to be universally construed as logical conjunctions when context or conversational conventions do not suggest another reading.

Understanding of linguistic disjunction

The overwhelming tendency of Ss in this study to interpret the disjunction of two statements inclusively rather than exclusively conflicts with the oft-reported (although contentious) view that natural language disjunction is exclusive. Sternberg (1979) used a similar experimental design where Ss were given a compound proposition and asked to assess each truth table category as true, false or maybe true/maybe false. He found that Ss were far more likely to interpret 'or' as exclusive, rather than inclusive, disjunction. Given a true proposition of the form '*p* or *q*', most Ss responded 'false' to the TT case whereas a inclusive interpretation would lead to the response 'maybe true/maybe false'. Newstead et al (1984) also report a preponderance of exclusive interpretations for a variety of 'or' statements. However, Braine and Rumain (1981) found that their adult Ss favoured an inclusive interpretation in a task very similar indeed to that used by Sternberg. Damarin, using a task very similar to that used in experiment 2, found that no S consistently interpreted 'or' exclusively and, in the current study only 5 Ss (all English speakers) of the total of 153 felt that only the mixed truth table categories were consistent with the disjunctive statement.

There seems no way of explaining the considerable discrepancy between various results other than through some feature of the experimental procedure or materials. Newstead and Griggs (1983) suggest that what is apparently an inclusive interpretation of 'or' statements may simply result from a strategy of responding in certain experimental situations. Given a statement of the form '*p* or *q*' to evaluate from one of the truth table categories, a S first looks to see whether *p* is true. If it is, he looks no further and responds 'true'. If *p* is false, he then seeks to ascertain whether *q* is true. If so he responds 'true', otherwise 'false'. However, this explanation of the apparent inclusive interpretation of 'or' is far from satisfactory. Presumably this strategy could be applied to either of the two types of task used to elicit underlying truth tables. Therefore, why should it be that neither task consistently produces one or other of the two possible interpretations? Furthermore, if natural language disjunction is exclusive (a view which Newstead and Griggs seem to favour), why should so many Ss adopt a strategy which is at odds with their presumed understanding of the connective 'or'?

A possible explanation of the different interpretations of 'or' is suggested by the results of a study carried out by Braine and Rumain (1981). Wooden blocks were used varying along three dimensions - colour, shape and size. For adult Ss given the command 'Give me all the green things or give me all the round things', 91%

responded as though they interpreted 'or' exclusively, i.e. they offered the experimenter blocks having one of the two properties (usually the first-mentioned) regardless of the other. However, when given the command 'Give me all those things that are either green or round', 27% of the same Ss responded with blocks exhibiting either or both of the two properties. The difference between these two commands is that the former seems to suggest the disjunction of two actions - 'giving the green things' and 'giving the round things'. On the other hand, the latter seems to indicate alternative properties of the things to be given - they must be green or round. It is possible that where a statement is the disjunction of two alternative courses of action, these are viewed as mutually exclusive choices even when both could be performed, either simultaneously or consecutively. With alternative attributes of a single item, perhaps a S is more likely to accept the disjunction as allowing both to be true simultaneously, provided that this is possible. This would explain the results of experiment 2 where the disjunction took the form 'The square is white or it is large'. These are alternative properties of the square and not alternative actions. Similarly Damarin's (1977b) disjunctions, which were of the form 'M is odd or N is even', could be viewed as describing properties of a pair of numbers.

Some support for this explanation of the different interpretations of 'or' statements is provided by the results of Newstead et al (1984) who investigated the effect of context on the interpretation of realistic disjunctive statements. Threats, promises and choices, all of which consisted of alternative actions, elicited exclusive interpretations from at least 85% of adult Ss. For disjunctive qualifications, on the other hand, just over one-half of these Ss favoured the inclusive interpretation. All such statements defined alternative properties qualifying an individual for some role, e.g. 'The successful applicant must have a degree or experience in computing', 'The man I marry will have to be either rich or handsome'. A higher proportion of inclusive interpretations (23%) was also obtained for what Newstead et al classified as 'concrete' disjunctives. These were also suggestive of alternative properties rather than alternative actions, e.g. 'Either the pants are dark brown or they belong to David', 'My son will either turn out to be rich or he will be intelligent'.

It remains to account for the surprising difference in the results of Sternberg (1979) and those of Braine and Rumain (1981). The disjunctions to which Ss had to respond were almost identical. In both cases they referred to the contents of a box - fruits or shapes in the former study, toy animals in the latter. Sternberg's disjunctions were of the form 'There is a circle in the box or there is a square in the box' and Braine and Rumain's 'Either there's a horse or there's a dog in the box'. Why the former should elicit exclusive interpretations from an overwhelming majority of Ss whereas for the latter inclusive disjunction was favoured is puzzling until one examines the experimental procedure more closely.

For Sternberg's Ss the experimental task was presented as a game. The experimenter placed unseen objects into a box and, given the disjunctive statement, Ss were asked to assess the truth table categories, presented as the possible contents of the box. It seems feasible that the disjunctive statement could be construed as stating the results of alternative courses of action on the part of the experimenter. Hence the preponderance of exclusive interpretations. On the other hand, in the Braine and Rumain study, four boxes containing animals were on view with their contents visible. Each box represented one of the four truth table categories relative to the disjunctive statement given which had to be evaluated by reference to one box at a time. The statement 'Either there's a horse or there's a dog in the box' therefore seems to refer to alternative properties of the box and, by the theory proposed, would be more likely to invite the inclusive interpretation.

That individuals are aware of the potential ambiguity of statements containing 'or' is indicated by the results of experiment 1. For those Ss who identified either of the disjunctive concepts and who described them correctly, there was a strong tendency in all language groups to add the clarifier 'or both' or 'but not both' to the 'or' statement. Furthermore, in the English and Arabic groups, Ss who used a simple 'or' statement to describe either or both of the disjunctions were equally divided between the two forms. The pattern for the Japanese was rather different however, with an apparent favouring of 'or' alone to describe inclusive, rather than exclusive, disjunction (see table 7.14).

It seems reasonable to assume that contextual cues affect the interpretation of disjunctions in everyday language exchanges. For instance, pragmatic criteria determine the construal of certain disjunctions as exclusive simply because it is clear that both disjuncts cannot be true simultaneously. However, the theory outlined above seems promising as an explanation for the different interpretations of 'or' statements where context does not necessarily suggest which is the appropriate interpretation. Clearly further research is necessary to establish the extent of its validity and also whether it extends to speakers of languages other than English. The results for disjunctions in experiment 2 suggest that English speakers may, in any case, be slightly more likely than the speakers of certain other languages to interpret an ambiguous disjunction exclusively.

Understanding of linguistic conditionals and biconditionals

Since the results of experiment 2 indicated very similar interpretations for conditionals and biconditionals, and because the latter connective has received little attention from researchers, we discuss these two connectives jointly.

The tendency, observed in experiment 2, to interpret biconditionals and conditionals as conjunctions has not been widely reported. However, there are few studies which have investigated the understanding of both of these linguistic connectives in adults and, for some of these, the nature of the experimental task does not allow identification of the complete underlying truth table.

In Damarin's (1977b) study, Ss (pre-service elementary teachers) were presented with a task very similar to that used in experiment 2. They were given a compound proposition and asked to identify which of the four truth table categories could be true when the proposition was true. The results, like those of experiment 2, revealed a strong predilection for a conjunctive interpretation for biconditionals and conditionals. However Bart (1974) used test items in which Ss were asked to assess whether a compound proposition was always true, always false or neither given the truth of its first component. Whilst this allows for the detection of certain errors of interpretation, it tests only for the evaluation of two truth table categories - TT and TF. A correct conditional or biconditional reading would therefore be indistinguishable from a conjunctive interpretation. In Paris' (1973) study, Ss were asked to assess the truth of a compound proposition given one truth table category. However, the 32 item test contained several propositions involving the same connective and thus allowed overall truth category errors to be identified, although individual truth tables were not isolated. Errors for the conditional were high for the FT and FF cases and errors for the biconditional were high for the FF case. These are precisely the errors which would result from construing the conditional and biconditional as conjunction.

Byrnes and Overton (1988) used a similar type of task from which Ss' truth tables could be inferred from an 'if ... then ...' statement. Ss were given a rule such as 'If it has rained, then the grass is wet' together with the four associated truth table categories and asked to say whether each proved the rule true, proved it false or proved nothing. The responses of 69% of the college student Ss were consistent with a conditional reading and a further 28% appeared to interpret the conditional as a biconditional. However, the apparent success of this task in eliciting a conditional interpretation may well be due to the nature of the response categories. The correct response to the TT, FT and FF truth table categories is 'proves nothing' - these are consistent with the rule but do not prove it to be true. However, Ss were scored correct if they responded 'proves false' to the TF case and either 'proves nothing' or 'proves true' to the remainder. A response scored as correct could therefore arise from the 'defective truth table' proposed by Wason (1966). A S might well respond 'proves nothing' to the FT and FF categories, not because he recognises that both render the conditional true, but because he considers the case where the antecedent is false to be irrelevant to the truth value of the conditional rule. Such an option was not available in experiment 2 where each truth table category had to be rated as consistent or inconsistent with the rule. Were the Ss to respond 'inconsistent' to the truth categories judged to be irrelevant, the inferred truth table would be indistinguishable from that of conjunction.

In propositional reasoning tasks however, there is no indication that conditional premises are construed as conjunctions. In their investigation of adults reasoning from simple abstract conditional sentences, Taplin and Staudenmayer (1973) tested Ss on all the eight possible forms of the 'two-premise deductive argument' (see chapter 4). One group of Ss was required to assess the conclusion as always true, sometimes true or never true and the other to assess it as always false, sometimes false or never false. Roughly 30% of each group responded in a manner consistent with a conditional interpretation whilst approximately 14% were inferred to have interpreted the major premise as a biconditional. A further 40% were inconsistent in their interpretations. This raises the obvious question - why is the conditional premise in the two premise deductive argument apparently so much more successful than the task used in this study in eliciting an interpretation consistent with the logical conditional? From the point of view of truth table evaluation, Taplin and Staudenmayer's task is more difficult. In order to respond 'sometimes true' or 'sometimes false' a S must examine the whole truth table to establish for how many of the four categories the major and minor premises are true. The task used in experiment 2 requires only the individual truth table categories to evaluated for truth or falsity of the conditional statement. Given that Ss are less successful in the latter task than the former, one can only conclude (along with Osherson (1975), Wason and Johnson-Laird (1972) and others) that, for individuals untutored in formal logic, propositional deductive reasoning does not normally utilise truth tables.

The discrepancy between the results for conditionals in experiment 2 and those which are typical of propositional reasoning with conditionals are particularly puzzling when it is recognised that the task in experiment 2 could be treated as a two-term deductive argument and tackled using the same strategy. For example, given the statement 'If the square is large, then it is black', the S could treat this as the major premise. Upon examining, say, the large white square, he could note that it was large and add 'The square is large' as the minor premise. Responding 'consistent' or 'inconsistent' then amounts to assessing the validity of the conclusion 'The square is white'. The argument now consists of premises $p \rightarrow q$ and p and conclusion $\bar{q}$. Were Ss to approach the task in this way, we would expect the results for conditionals in experiment 2 to correspond with performance on deductive reasoning tasks, i.e. to indicate a considerably greater proportion of both conditional and biconditional interpretations (see for example: Taplin, 1971; Taplin and Staudenmayer, 1973; Taplin, Staudenmayer and Taddonio, 1974). The fact that they do not suggests that deductive reasoning tasks may not be tackled in the same way as truth category evaluation.

Given the evidence that adults do, in certain contexts, reason successfully with conditional statements, the results of experiment 2 cannot be taken to indicate that Ss find conditionals (and biconditionals) indistinguishable from conjunctions. As Braine and O'Brien (1991) point out, a conjunctive response pattern cannot be taken to reflect the lexical entry for 'if ... then' because arguments such as denying the antecedent and modus tollens would have contradictory premises. We are therefore forced to conclude that the experimental task induces some kind of response bias. One possibility is the effect of a 'matching bias'

(Evans, 1972) - Ss tend to attach particular salience to the items named in the rule and consider the mismatching cases irrelevant. The effect of a matching bias with biconditional propositions has not been investigated but presumably the conjunctive interpretation of these propositions could be accounted for similarly. However, what could not be explained is why the disjunctive propositions in experiment 2 were relatively immune to its effects. The absence of 'matching bias' for disjunctions has also been reported by Evans and Newstead (1980).

A more promising explanation of the conjunctive interpretations, at least for conditionals, is to be found in the inference schemata model of reasoning proposed by Braine and O'Brien (1991). It is suggested that part of the lexical entry for 'if ... then' is a conditional proof schema described as follows: 'To derive or evaluate *If p then* ... first suppose *p*; for any proposition *q* that follows from the supposition of *p* taken together with other information assumed, one may assert *If p then q*.' When confronted with a truth table category in which *p* is true, a S applying this schema would be led to respond 'true' when *q* is also true and 'false' when *q* is false. When the antecedent is false, the schema cannot be applied - a reasoner is not willing to make a supposition and then treat it as false. In this case, the conditional cannot be evaluated and the result is the defective truth table referred to above. It is proposed that the conjunctive pattern is the result of a pragmatic judgement that an irrelevant truth table category is more appropriately considered false than true if that is the only choice. If we assume that, in experiment 2, the S proceeds by examining each truth table category and then uses it to assess the validity of the conditional statement, this would explain the 'inconsistent' response to the FT and FF cases. 'Inconsistent' is considered a more appropriate response than 'consistent' for these irrelevant categories.

A study by Evans and Newstead (1977) provides some support for this notion. Adult Ss were given a conditional rule and asked to classify each truth table category as rendering the rule true, false or irrelevant. For a conditional rule where both components are affirmative (as used in experiment 2), combining Ss' responses in the categories 'false' and 'irrelevant' gives results similar to those obtained in experiment 2 for the response category 'inconsistent'. Of interest is to what extent Ss in the study described here would have chosen the 'irrelevant' category had it been available to them. The theory predicts that the response 'inconsistent' would be preferred for the TF category and 'irrelevant' for the two cases where the antecedent is false.

Although Braine and O'Brien (1991) do not deal with biconditional reasoning, their explanation of the conjunctive interpretation of the conditional can be extended to the biconditional if this is viewed as the conjunction of two conditionals. The biconditional could then be evaluated by applying the conditional proof schema to each conditional component. In this case, the biconditional would be considered irrelevant when the antecedent of either conditional was false, i.e. for all truth table categories other than TT. If this is so, we would expect Ss with the option of responding 'irrelevant' to do so for the categories TF, FT and FF. Hence the biconditional response pattern would differ from that of the conditional only in the TF case.

We can now go some way towards explaining the apparent discrepancy between the results of the different experimental tasks from which conditional truth tables have been inferred. For tasks like that used in experiment 2 which involve the assessment of individual truth table categories, the apparent truth table underlying a conditional statement depends upon the response categories available. Where there are only two, the underlying truth table for a conditional may appear to be that of conjunction as in this study and those of Damarin (1977a, 1977b) and Paris (1973). However, when the additional response category 'irrelevant' is allowed, Ss appear to associate a conditional with the defective truth table described above (Johnson-Laird and Tagart, 1969; Evans and Newstead, 1977).

Why is it that in propositional reasoning tasks, Ss are more likely to interpret an 'if ... then' statement as a logical conditional? Braine and O'Brien (1991) propose that reasoning from a conditional premise together with a simple component or its negation utilises the modus ponens schema: 'Given *If p then q* and *p*, one can infer *q*.' Deductions such as modus tollens are difficult because the inference schema must be used to produce a contradiction: if *p* is true then *q* is true, but *q* is false, therefore *p* is false'. Hence errors in reasoning may result from a failure to execute the necessary steps in a more complex deduction or from succumbing to the invited inference 'if not *p* then not *q*' (Geis and Zwicky, 1971). In any case, the truth table which is inferred from a particular pattern of responses to conditional arguments may result from a variety of causes other than a defective lexical entry for 'if ... then'. Therefore, it is not surprising that it does not always correspond with that inferred from the evaluation of truth table categories. These types of error may also explain why as many as 40% of Taplin and Staudenmayer's (1973) Ss gave inconsistent responses.

PERFORMANCE ON CONCEPT IDENTIFICATION TASKS

For the concept identification tasks used in experiment 1, two performance factors are of interest. These are:

(a) success at abstracting the concept from the set of all exemplars as evidenced by the ability to distinguish examples from non-examples;

(b) the ability to describe the concept in such a way that examples are distinguished from non-examples.

For each of the concepts tested (conjunction, inclusive disjunction, exclusive disjunction, the conditional and biconditional), there were Ss whose descriptions of the concept clearly did not match their criterion for responding to the test strings and this tendency was independent of language group. Therefore we can conclude that giving an accurate description of the concept does not necessarily imply successful concept identification in terms of the ability to distinguish examples from non-examples.

For those Ss who had a mental representation of a concept which enabled them discriminate between examples and non-examples, the success rate in describing the concept was not constant across concepts. For conjunction, inclusive disjunction and the conditional, more than 90% of those who achieved the test string criterion offered an accurate description. However the biconditional was more difficult to describe with only 78% of those who could distinguish examples from non-examples succeeding in offering a correct description of their criterion for doing so.

It was also found that an accurate description was not necessarily associated with the successful distinction of examples from non-examples. This too was independent of language group. Again the biconditional was the most problematic with only 72% of those who correctly described the concept achieving criterion on the test strings.

These results suggest that the mental representation of a concept and the words used to describe it do not necessarily match. An individual may well be able to distinguish examples of a concept from non-examples with a high degree of consistency but fail to offer a description which adequately describes his criterion for doing so. This does suggest that concept identification studies which have used an accurate description as indicating successful concept identification (Bourne, 1966, p 4) were using an inappropriate criterion.

Relative difficulty of concepts

The order of difficulty of the five concepts (ranked by the proportion of Ss who achieved the tests string criterion) was much the same for all language groups. Conjunction was easiest, followed by inclusive disjunction and the success rate for each of these concepts was significantly higher than that for either exclusive disjunction, the conditional or the biconditional. The conditional was the most difficult concept for all three groups.

These results are in broad agreement with those reported for standard concept formation tasks utilising sequentially presented stimuli and where concepts are ranked for difficulty using mean (or median) trials to criterion. The only difference is that, for the few studies which have investigated the identification of biconditional concepts, these are usually reported as being more difficult than conditionals (Bourne, 1970; Bourne, 1974; Neisser and Weene, 1962).

Neisser and Weene's (1962) assignment of concepts to levels of difficulty predicts the ranking obtained for all concepts except the conditional. The hierarchy is based on the expression of each concept in terms of the 'primitive' connectives conjunction and disjunction together with the unary operation of negation. The conditional, expressed as $\bar{p} \vee q$, was assigned to level II along with conjunction and disjunction. The results of experiment 1 suggest that a more appropriate expression would be the disjunctive normal form $(p \wedge q) \vee (\bar{p} \wedge q) \vee (\bar{p} \wedge \bar{q})$ which would correctly predict that the conditional would be more difficult than the level III concepts - the biconditional $((p \wedge q) \vee (\bar{p} \wedge \bar{q}))$ and exclusive disjunction $((\bar{p} \wedge q) \vee (p \wedge \bar{q}))$. Also, to explain the greater difficulty of conjunction over disjunction, we would have to propose that the former is in some sense 'more primitive' than the latter.

Bourne's (1974) inference model for conceptual rule learning (outlined in chapter 4) also predicts greater difficulty of the biconditional over the conditional for standard rule learning tasks where (as in experiment 1) stimulus elements are equally divided between truth table categories. Also, contrary to the findings of this study, it predicts that exclusive disjunction should be more difficult than the conditional. However, the model describes concept learning when the S is given informative feedback to responses to sequentially presented stimuli and may fit less well to the design of experiment 1. Bourne's model also predicts, for each concept, the truth table categories whose assignment will cause particular problems. These are: for the biconditional: TF, FT, FF; the conditional: TF, FF; inclusive disjunction: TF, FT; exclusive disjunction: TT, TF, FT. These show a fair degree of agreement with the errors of assignment of the test strings for Ss who used a correct description of the concept. However, FF errors far outnumbered those in the other categories for the conditional and biconditional and, although less common, were not absent for either form of disjunction.

Times to task completion

The rankings of times to task completion were also much the same within each language group, with the conditional having a significantly greater mean than each of the other four concepts. For all groups, the mean times for conjunction and exclusive disjunction were each less than the mean times for the biconditional and inclusive disjunction, and these differences were significant for the English and Arabic groups. In an experiment where a S is required to persevere until he achieves success at a task, mean times to task completion can be used as an indicator of the difficulty of the task. However, in experiment 1 this was not the case. The ranking of concepts by mean times to complete the task corresponds roughly to the number of test strings exemplifying the concept. It seems likely that inclusive disjunction took longer than the more difficult concepts exclusive disjunction and the biconditional because there were more strings to scan. Presumably the conditional took significantly longer than any other concept for the same reason and also because it was inherently more difficult.

Relative performance of language groups

Whilst the relative difficulty of concepts appears to be much the same for all three language groups, it is when we compare the results across language groups that we observe consistent and striking differences in performance. For every concept, the proportion of Japanese Ss who successfully identified it was greatest and the proportion of Arabic speakers the least. Furthermore, the differences in the proportions of successful Japanese speakers and successful Arabic speakers were significant for each of the five concepts.

Descriptions of concepts

If we examine the descriptions of each concept used by those Ss who were successful in identifying it, we again find some surprising language group differences. For all concepts, the mean length of the description (measured by the mean number of elementary propositions) was least for the Japanese speakers, although equalled by the English speakers for conjunction. The mean length of descriptions used by Arabic speakers was greatest for all concepts except the biconditional. Furthermore, the differences between the means for the Arabic and Japanese groups were significant for all concepts except the biconditional (where the disjunctive normal form was overwhelmingly favoured by all language groups) and conjunction (where nearly all Ss used the description corresponding to '*p* and *q*'). Also, for these three concepts (the conditional and both forms of disjunction), Japanese Ss who identified the concept were more likely than Arabic speakers to describe it using the most succinct form with just a single elementary proposition.

Within language groups too, the order of difficulty of the concepts (measured by the number of Ss in set C) is very much the same as their ordering by the mean length of correct descriptions used by Ss in set C. The only noteworthy exception is the biconditional for the Arabic speakers which was exceeded in difficulty only by the conditional but had a mean description length which was exceeded by both forms of disjunction in addition to the conditional. This does suggest some association between the proportion of successful Ss and the length of the description of the concept used by those Ss. For each concept, the highest success rate was associated with the least mean description length. In fact, examination of figure 7.5 reveals a high degree of inverse linear correlation between these two variables for all concepts other than conjunction (which exhibits a 'ceiling effect') and the biconditional. This lends support to Lenneberg's (1962) observation that 'in most instances of experimental concept formation, there is a correlation between ease of naming a concept and ease of attaining it'.

In addition to offering shorter descriptions, the Japanese as a group used fewer different descriptions than either of the other two groups for every concept. The nature of the more 'popular' descriptions was also rather different for this group particularly when compared with the Arabic speakers. For the biconditional the disjunctive normal form was favoured by all language groups. However, amongst the Japanese speakers only one S who attained the test string criterion offered a description other than that corresponding to the disjunctive normal form.

Japanese descriptions of the conditional also tended to refer to the disjunctive normal form when they did not correspond to one of the single elementary propositions $p \rightarrow q$ or $\overline{p \wedge \bar{q}}$. English and Arabic speakers used a wide variety of descriptions of the conditional but did not show a particular preference for the disjunctive normal form. For both varieties of disjunction however, the Arabic speakers showed a stronger tendency to use this form than either of the other two groups, each of which favoured a description of the form '*p* or *q*' with or without 'or both' or 'but not both' to make clear the sense of 'or'.

The results seem to indicate that a succinct coding of each concept was in some way more readily available to the Japanese speakers, particularly when compared with the Arabic group which they outperformed in the identification of every concept. Furthermore, when a succinct expression for the concept was not used, the Japanese tended to resort to the disjunctive normal form. This seems a sensible strategy since it singles out each of the disjoint categories of strings and defines them in terms of the presence/absence of the salient shapes. Also, the only logical connectives used in the disjunctive normal form are conjunction and disjunction which the results of experiment 2 suggest are well coded in all three languages.

For a concept coded in its disjunctive normal form, when classifying test strings as examples/ non-examples it is presumably necessary to test each string against each of the conjunctive categories. Since the disjunctive normal form for the conditional and inclusive disjunction involves more elementary propositions than any other of the concepts, it might well be more difficult to remember and therefore more likely to lead to errors than a more succinct coding. However, the results of experiments 1 and 2 seem to suggest that Ss may find difficulty in coding a conditional concept with a single elementary proposition and therefore resort to a lengthier, more error-prone, coding.

Although the disjunctive normal form for inclusive disjunction also involves three elementary propositions, Japanese and English speakers were able to find a more succinct coding corresponding to one of $p \vee q$ or $(p \vee q) \vee (p \wedge q)$. For some reason the Arabic speakers tended to code this concept using the lengthier disjunctive normal form which had a high error rate when it came to distinguishing examples from non-examples. Of the 16 Arabic speakers who offered the disjunctive normal form as a description of inclusive disjunction, 3 (approximately 19%) failed the test string criterion.

GENERAL DISCUSSION

It will be recalled that the purpose of this project was to ascertain whether there were differences amongst language groups in the construal of the linguistic expressions of logical connectives. The aim of experiment 2 was to identify any such differences. The purpose of experiment 1 was to establish whether differential understanding of linguistic connectives was associated with differing performance on concept identification tasks. Given that experiment 2 revealed no major differences in the understanding of those linguistic connectives, we are now faced with finding an alternative explanation for the consistent differences revealed by experiment 1.

A potential explanation for the relatively superior performance of the Japanese group over the Arabic speakers would be that the latter group did not spend sufficient time on the tasks. However, there is no evidence to suggest that this was the case. As figure 7.6 shows, the only concept for which the mean time to task completion for Arabic speakers was significantly less than that for the Japanese was for the conditional and, in this case, it was the English language group which had the greatest mean time. Otherwise the differences in the mean times for Japanese and Arabic groups were not significant. In any case, for conjunction and exclusive disjunction, it was the English group rather than the Arabic speakers who had the shortest mean time. Only for conjunction did the Japanese have the greatest mean and this was not significantly different from that of either of the other two groups.

Another possible explanation for differential performance in the concept identification tasks is prior training in logic. However, only two Japanese claimed that they were familiar with elementary logic and the subject does not seem to be included in the Japanese national curriculum in Mathematics (Howson, 1991). On the other hand, 39 of the Arabic speakers claimed that they had studied some logic. Were this information to be verifiable, we would be obliged to entertain the notion that prior training in logic is detrimental to performance on concept identification tasks. In the event, it was not possible to ascertain the extent of any familiarity with formal logic. Furthermore, the results of experiment 2 gave no indication that the Arabic speakers were any more familiar with the logical concepts tested than either of the other groups. We can only conclude that any prior knowledge of logic which might be expected to have enhanced performance in either of the experimental tasks had either not been acquired or had been forgotten.

The higher success rate of the Japanese Ss could stem from superior performance at either or both of two stages. The possibilities are:

(a) as a group the Japanese were better at identifying the common characteristics of the strings exemplifying each concept;

(b) for the Japanese group, the coding in memory of the common characteristics of the strings was more likely to lead to successful classification of the 16 test strings.

Of these two alternatives, the second seems less likely, if we assume that the verbal description offered by a S in some way reflects his internal representation of the concept. Of those descriptions which should have been associated with successful discrimination of exemplars from non-exemplars, those offered by the Japanese were no more effective for any concept (see table 7.6).

We are faced with explaining why it should be that the Japanese Ss were consistently more successful at abstracting the logical concepts from the strings and why their coding of these concepts was more succinct and less variable. There is no particular reason to suppose that the Japanese sample of Ss was any less representative of the underlying population than either of the other two samples and there is little in the research literature which might suggest the reasons for their consistently superior performance on the concept identification tasks.

In international studies of mathematics attainment, a consistent finding is that Japanese children are superior to certain others of the same age (Husen, 1967; Robitaille and Garden, 1989). Unfortunately these studies did not consider nations where the indigenous language is Arabic. However, in no study has a sample of US students outperformed a comparable sample of Japanese students and the difference is reported as existing at all grade levels (Mayer et al, 1991). In a comparative study of 15 - 17 year olds, the average Japanese score was better than 98% of the Americans (Harnisch et al, 1985). However US and

Japanese students do not differ in scores on tests of basic cognitive ability (Mayer et al, 1991) and their superiority in mathematics does not extend to other subjects (Lynn, 1989). It is suggested that the comparatively low level of mathematics achievement in American children is attributable to the fact that the quantity of classroom instruction which they receive is less than that of the Japanese and that less homework is required of them (Stevenson et al, 1986). The general opinion is that it is exposure to mathematics rather than differential innate ability which leads to the superior performance of the Japanese children.

The English speaking Ss used in this study were nearly all from the US and virtually all the Japanese Ss had been educated in Japan and, as a group, had almost certainly been exposed to more mathematics teaching in their secondary schools than either of the other two groups. Whilst logic is not included as a specified item in the national curriculum in Japan, it is possible that the logical concepts inherent in mathematics are simply more familiar to the Japanese because they have been exposed to them more frequently. On the other hand, if logical concepts are more familiar to the Japanese because of some other aspect of their language or culture, then this may be a contributory factor in their superior performance in mathematics. There is a problem here as to which is cause and which is effect.

Another possible contributory factor to the greater success of the Japanese in the task used in experiment 1 is that they have a greater facility with abstracting meaning from visually presented symbolic material due to the nature of their character-based writing system. Japanese newspapers and magazines generally utilise around 2000 ideographic characters (kanji), 900 of which are learned by children during the first 6 years of their elementary education (Miller, 1967). Kanji vary in their visual complexity which is measured by the number of strokes of which they are constituted. More complicated kanji may contain as many as 26 strokes (Paradis et al, 1985). Visual symbols are much more a feature of the Japanese cultural experience and it may be that this would facilitate performance in a task of the type used. Of interest would be whether the speakers of other ideographic languages (for example, Chinese) would also excel at such a task.

Some support for this explanation of language group differences is offered by the evidence that there may be cognitive differences in the speakers of languages with ideographic, as opposed to phonetic, writing systems at least with regard to the processing of written material. Zepp (1988) reports that an experiment performed in China revealed that damage to the temporal bone could result in destruction of the capacity to write in Caucasians but had no effect at all on the writing capacity of the Chinese. Damage to the parietal bone, on the other hand, did not affect the reading and writing ability of Caucasians whereas Chinese who suffered such an injury completely lost their ability to cope with written language. As Zepp points out '... educators should at least be aware of the possibility that students who speak character-based languages may attack problem solving and other theoretical concepts using strategies quite different from those used by students using phonetic languages.'

The disparity in the performance of the Japanese and Arabic groups in the concept identification tasks does seem to indicate some difference in the cognitive functioning of the two groups. The greater success rate of the Japanese and the way in which they referred to the concepts suggest that either the concepts as presented were more familiar or that, as a group, the Japanese were more effective at identifying the defining characteristics and coding them. If the former is the case, prior training in logic does not seem a likely explanation although greater exposure to mathematics could account for greater familiarity with logical concepts. If the latter explains the superior performance of the Japanese, then this could be attributed to a writing system which facilitates the abstraction of concepts from visually presented material.

The underlying reasons are far from clear but the consistency with which the Japanese outperformed the English and Arabic speakers argues for further research to establish which factors - linguistic, educational or cultural - are implicated.

Chapter 10

CONCLUSIONS

Experimental studies such as those reviewed in chapter 4 have repeatedly demonstrated that the concepts underlying certain logical connectives are not adequately conveyed by the English language items with which they are normally associated. However, it is not entirely clear whether this mismatch between linguistic and logical forms is a general feature of a wide range of languages or whether it is restricted to certain specific language groups.

This study constituted an exploratory investigation to determine to what extent the results reported for English linguistic connectives were generalisable to Japanese and Arabic, two languages unrelated to English and to each other. A comparison of conjunctions, disjunctions, conditionals and biconditionals in these languages (chapter 5) revealed sufficient structural differences to suggest that there might be differences in the comprehension of these linguistic forms.

Understanding of the linguistic coding of logical connectives in Japanese, Arabic and English was investigated in experiment 2 using a task which elicited the truth table associated with the natural language expression of each of the logical concepts.

Ss were native speakers of English, Japanese or Arabic and were tested in their mother tongue. Whilst the primary aim was to compare results across language groups, the task was such that the results could be compared with those of other studies (reviewed in chapter 4) which have used similar tasks to investigate the comprehension of linguistic connectives in English.

Inherent in cross-cultural studies is the problem of obtaining comparable samples. Random assignment of Ss to experimental groups is not possible because the nature of the comparisons to be made determines that the groups be composed of Ss with a specified characteristic - in this case, that they be a native speaker of one of the languages investigated. In this study, an attempt was made to obtain comparable samples by using, as far as possible, students at a liberal arts college (Richmond College) which makes every effort to apply constant admissions standards (in terms of academic achievement) to students from diverse geographical regions. Whilst all the English and Arabic speaking Ss satisfied this criterion, unfortunately it was necessary to supplement the Japanese group with students at other London colleges. However, there is no particular reason to suppose that these had characteristics which rendered the Japanese sample different from one composed solely of students drawn from Richmond College. Furthermore, there were no indications that the Japanese Ss who were not Richmond College students performed any differently from those who were.

An additional difficulty in comparing the results of the three samples is that, because of the requirement that Ss be tested in their native language, three different experimenters were necessary - one for each language group. In anticipation of the possibility that this might result in dissimilar treatment of language groups, the Japanese and Arabic experimenters underwent considerable training prior to the testing of Ss. This included several trial runs testing English speaking Ss (not used in the main study) in English under observation. Both were well acquainted with the purpose of the study and each was sufficiently fluent in English to understand exactly what was required of them. There is therefore no reason to suppose that the observed differences in language group performance can be attributed to the behaviour of the experimenters.

Whilst it would have been preferable to use Ss whose only language was their native one, this constraint would have rendered the study impractical. In the event all the Arabic and Japanese speaking Ss had some familiarity with English, although even the few whose

English was fluent maintained that it was not their first language. It is recognised that, in a study which seeks to attribute observed differences in performance to language factors, contributory effects due to second language interference are possible. However, the likely nature of such effects is far from clear. By using only the S's native language throughout the testing process, it was hoped that any such influence would be minimised.

In addition to performing a task designed to investigate their understanding of linguistic connectives, Ss also performed a series of concept identification tasks where the concepts to be identified were conjunction, inclusive and exclusive disjunction, the conditional and biconditional. This constituted experiment 1. Studies using such tasks (reviewed in chapter 4) have typically used geometrical figures varying on a number of dimensions such as size, colour and shape. A concept is then defined by the presence/absence of one value on each of two dimensions. For example, if the relevant attributes are 'red' and 'large', conjunction would be defined by all figures which are both red and large and inclusive disjunction by figures which have one or both of these two characteristics. Less frequently, concept identification studies have utilised strings of letters or shapes. In this case a concept is defined by the presence/absence of either of two relevant shapes. The latter paradigm was selected for this study because there is some evidence that attributes such as colour, shape and size may not have the same perceptual salience (Ketchum and Bourne, 1980). On the other hand, there seemed no reason to suppose that, of the shapes used in this study (circle, square, star and triangle), any one was more salient perceptually than any other.

The concept identification tasks used fall into the category of 'rule learning'. The relevant shapes were known to the S who therefore had to identify the concept only in terms of how those shapes defined the characteristics shared by its exemplars. Each task differed somewhat from 'standard' rule learning tasks in that the concept was to be abstracted from a complete set of exemplars rather than from a sequence of individual examples and non-examples. This eliminated the memory burden associated with the necessity to consolidate information obtained by previously presented instances.

The results obtained from experiment 2 suggest that, for the languages investigated, if there are any differences in the understanding of linguistic connectives, these are very minor. English speakers were slightly more likely to interpret an 'or' statement as exclusive disjunction than the other two groups and the Japanese had a greater tendency to interpret 'if ... then' statements as biconditionals. The necessity of expressing a biconditional in Japanese as the conjunction of two conditionals is proposed as explaining the greater proportion of readings equivalent to the logical biconditional found for this group.

When we consider the results for individual connectives, we find that 'and' statements are almost universally understood as logical conjunctions and that this finding is independent of language group. 'Or' statements are slightly less successful in eliciting either of the two possible disjunctive interpretations but, when they do, the inclusive interpretation is overwhelmingly favoured. This, too, is independent of language group. It was also found that, for all groups, linguistic conditionals and biconditionals were only rarely interpreted 'logically' and appeared to have an underlying truth table corresponding to that of logical conjunction.

This hierarchy of understanding of linguistic connectives is the same as that reported by a number of studies which have compared the comprehension of English language connectives (for instance: Airasian, 1975; Damarin, 1977a, 1977b; Bart, 1974). However, the tendency to construe linguistic disjunctions inclusively is at odds with the results of a number of studies which have reported a preference for an exclusive reading of 'or' statements in English (Sternberg, 1979; Newstead et al, 1984; Braine and Rumain, 1981). To explain the discrepancy between these results, it is proposed that the interpretation of an 'or' statement may well depend on the nature of the disjuncts. When they are actions, the preferred interpretation is likely to be exclusive and where they are properties, an inclusive interpretation will be favoured.

We propose that the apparent interpretation of conditionals and biconditionals as conjunctions is due to the nature of the response categories used in this study. The results for conditionals can be explained by the 'defective truth table' (Wason, 1966) whereby it is proposed that Ss view the truth table categories where the antecedent is false to be irrelevant to the truth value of a conditional statement. When faced with the alternatives of responding 'consistent' or 'inconsistent' to these categories (as in this study), Ss default to 'inconsistent' - but not because they believe that these render the conditional statement false. They simply consider them irrelevant and consider 'inconsistent' to be the more appropriate of the two alternative responses.

The three language groups performed similarly in their interpretation of the linguistic connectives tested despite the difference in the construction of these items in Arabic, Japanese and English. However, there were striking and consistent differences in performance in the concept identification tasks. The Japanese as a group outperformed the other two groups on every concept and the differences between the proportions of successful Japanese speakers and successful Arabic speakers was significant for all five concepts.

The hierarchy of difficulty of the concepts amongst themselves was similar for all language groups, however. Every group was more successful at identifying conjunction than inclusive disjunction and both of these concepts were easier than exclusive disjunction, the conditional and biconditional. The conditional was the most difficult concept for all groups. These findings are in broad agreement with other studies which have investigated relative difficulty of concepts in similar (although not identical) tasks (for example: Neisser and Weene, 1962; Haygood and Bourne, 1965; Bourne, 1970). However, these have reported the biconditional to be more difficult than the conditional.

Why the Japanese were so much more successful in the concept identification tasks and why their descriptions of these concepts were consistently less varied and shorter remains to be explained. Had differences been found in the comprehension of the linguistic coding of logical connectives, we could have attempted to explain dissimilar familiarity with the logical concepts by the differing provision within each language for conveying those concepts. However, this study revealed no such differences and it therefore constitutes yet another investigation which has failed to find Whorfian effects amongst the speakers of dissimilar languages. However, it cannot be concluded that such differences do not exist. There is no obvious explanation for the consistently superior performance of the Japanese group, particularly over the Arabic speakers, in identifying all five of the concepts tested. Whilst it does not appear to be attributable to Whorfian differences in the understanding of the appropriate linguistic forms, language factors which might explain greater facility with these tasks cannot be ruled out.

The fact that shorter and less variable descriptions were used by the Japanese suggests that there are certainly related language factors, although these may not be causative. Furthermore, although their representation was visual, the nature of the concepts themselves is abstract, a realm where language has been proposed as exerting a greater influence (Bloom, 1984; Lemon, 1981). For some reason, the schemata associated with the logical concepts seem more 'accessible' to the Japanese Ss. This seems to lend some support to Brown and Lenneberg's (1954) proposal that more familiar concepts are tagged with shorter labels so that '... more nameable concepts are nearer the top of the cognitive deck'. However the results of this study suggest no obvious reason why the 'cognitive deck' of the Japanese should be organised differently with regard to logical concepts.

Implications for learning mathematics

A cognitive factor which has been found to be positively correlated with mathematical ability is 'speed of closure' (Pullman, 1979). This is defined as 'the ability to unify an apparently disparate perceptual field into a single percept'. Whilst experiment 1 did not set out to assess this factor, the nature of the tasks used is such that success would seem to be associated with the general skills subsumed under 'speed of closure'. It may well be that the superior performance of the Japanese in experiment I is indicative of a cognitive structure which favours speed of closure. If so, this would go some way towards explaining the greater success in mathematics reported for the Japanese (Husen, 1967; Robitaille and Garden, 1989). However, we are still faced with explaining why the Japanese should be more generously endowed with this cognitive factor.

We could also explain the superior performance of the Japanese on the tasks used in experiment 1 by proposing that, for some reason, they are more familiar with the underlying logical concepts. This could result from greater exposure to mathematics and hence to the logical concepts inherent in the subject. On the other hand, if it stems from some other source, it could explain superior mathematical performance.

It is clear that linguistic connectives used to convey logical concepts frequently fail to do so. Furthermore this problem is not peculiar to the English language, but seems to be a feature of a range of languages with distinct linguistic origins. What then are the implications for the teaching and learning of mathematics?

If, as has been suggested, codability is a measure of how readily available a concept is (Brown and Lenneberg, 1954; Zipf, 1935), then we must conclude that conditional and biconditional concepts are not part of the cognitive 'stock-in-trade' of logically naive individuals. Yet, these connectives play a fundamental role in mathematical reasoning. It therefore seems surprising that the teaching of logical concepts is a neglected area of the mathematics curriculum in a wide variety of national settings (Travers and Westbury, 1989; Howson, 1991). Why this should be so is not clear but it seems likely that there is a tacit assumption that, because linguistic items such as 'if ... then' are so widely used in natural language, they need no further clarification. However, although they may be familiar terms, the concepts which they label are not necessarily the logical concepts so fundamental to mathematical inference.

There are many mathematical concepts which are labelled with linguistic items which have very different meanings when they are used in everyday conversation. Examples are 'factor', 'group', 'differentiate' and 'function'. When teaching these concepts, we take considerable pains to ensure that the terms are defined adequately. Yet there is little evidence to suggest that we do the same for linguistic connectives despite the evidence that they suffer from a similar lack of congruity between their 'mathematical meaning' and their 'everyday meaning'.

The situation may well be exacerbated by the fact that, in mathematics itself, we are often careless about the way in which we use conditionals. Mathematical definitions are almost always expressed as conditionals when the intended interpretation is that of a biconditional. For instance, having equipped students with the definition 'If a triangle has three equal sides, then it is called an equilateral triangle', we will take it that all future references to equilateral triangles will be understood as confined to triangles having three equal sides. We simply assume that the invited inference will be made, that the conditional will be 'perfected to a biconditional' and interpreted as we intended. On the other hand, if the same students were to succumb to the invited inference associated with 'If a function is differentiable then it is continuous', we would undoubtedly accuse them of having made a fundamental reasoning error and exhort them to be careful to avoid it in the future.

It is interesting to note that the direct method of proving a conditional statement in mathematics is parallel to Braine and O'Brien's (1991) inference schema for a natural language conditional: 'To derive or evaluate *If p then* ... first suppose *p;* for any proposition *q* that follows from the supposition of *p* taken together with other information assumed, one may assert *If p then q*'. The difference in proving a mathematical conditional is that the 'other information assumed' must consist of axioms and proved theorems and 'follows from' must be justified using a sequence of steps sanctioned by the laws of logic.

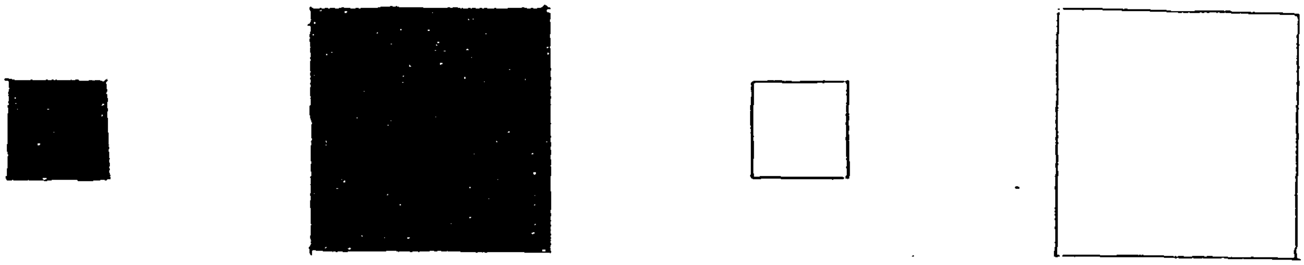
It is not surprising that logical errors are routinely committed when the burden of conveying logical relationships rests on linguistic forms known to invite a variety of interpretations. Drawing valid inferences is fundamental to successful mathematical activity and yet logical concepts as such are rarely taught formally to other than specialist mathematicians and the nature of a deductive proof receives little attention in secondary school texts. Furthermore, as mathematics teachers and writers of textbooks, we ourselves are often guilty of making logical statements which do not mean what we intend them to mean and which do not justify the inferences which we condone.

Whilst this study has added little to the 'Whorfian debate', it does offer some contribution to the discussion of how linguistic connectives are interpreted in a variety of languages and the implications for the teacher and learner of mathematics. It suggests that, in teaching mathematics, we should be aware of the potential ambiguity in the interpretation of logical relationships conveyed through natural language. It also points to the need to include some instruction in formal logic and deductive proof in secondary school curricula.

Appendices

NAME: ____________________

Appendix 1a



A B C D

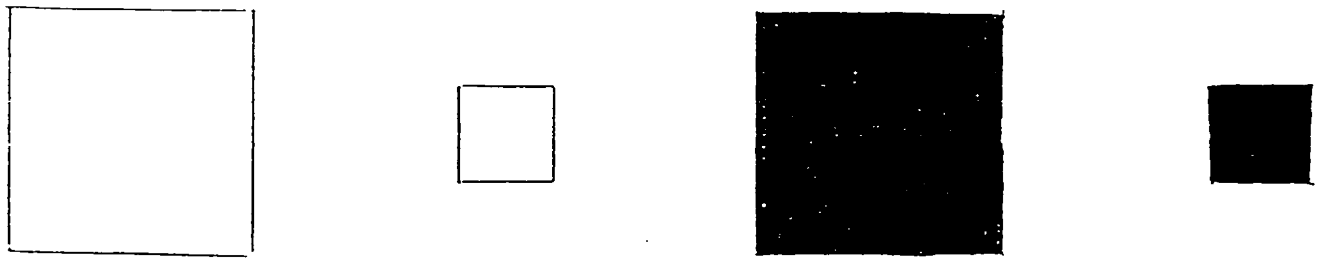
You are given below 16 numbered statements. Above are 4 diagrams labelled A, B, C and D each showing a square. Two of the squares are large and two are small. Two of the squares are black and two are white.

For each statement, look at the diagrams A, B, C and D and decide, for each one, whether it is consistent with that statement or not. Tick (✓) the appropriate boxes for the diagrams which you think are consistent with the statement and place a cross (X) in the appropriate boxes for those which are not consistent with the statement.

	A	B	C	D
1. The square is black and it is large.				
2. If the square is large then it is white.				
3. The square is white or it is small.				
4. The square is small if and only if it is black.				
5. The square is large or it is black.				
6. If the square is black then it is small.				
7. The square is white if and only if it is large.				
8. The square is small and it is white.				
9. The square is large if and only if it is black.				
10. If the square is white then it is small.				
11. The square is black or it is small.				
12. The square is large and it is white.				
13. The square is white if and only if it is small.				
14. The square is large or it is white.				
15. The square is black and it is small.				
16. If the square is large then it is black.				

Appendix 1b

الاســم ________________



د ج ب أ

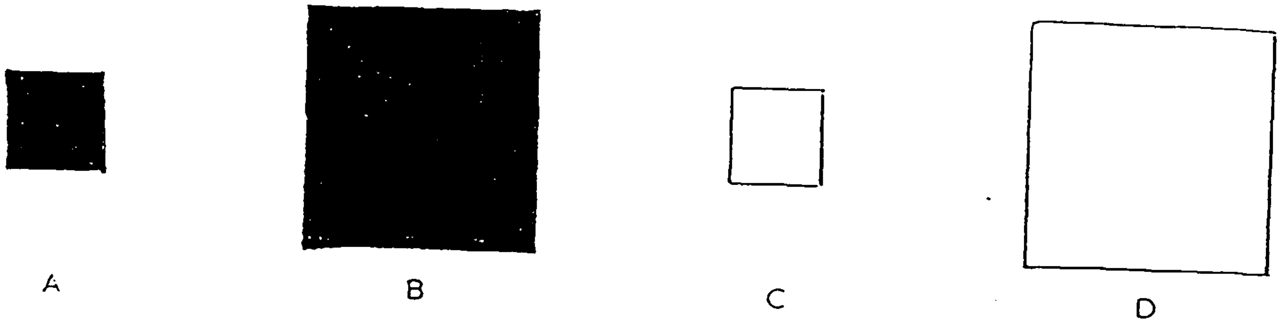
معطى لديك ادناه ١٦ جملة مرقمة، وفي الاعلى اربعة اشكال معنونة أ،ب ،ج ،د وكل منها يمثل مربعا ،اثنان من المربعات كبيران واثنان صغيران اثنان من المربعات لونهما اسود واثنان لونهما ابيض .

لكل جملة انظر للاشكال أ ،ب ،ج ،د وقرر لكل منها ان كان يناسب الجملة اولا . ضع اشارة (✓) في الخانة المناسبة للاشكال التي تطابق الجملة و ضع اشارة (X) في الخانة المناسبة للاشكال التي لا توافق الجملة .

د	ج	ب	أ	
				١- المربع اسود وهو كبير .
				٢- اذا كان المربع كبيرا فانه ابيض .
				٣- المربع ابيض او هو صغير .
				٤- المربع صغير اذا واذا فقط هو اسود
				٥- المربع كبير او هو اسود
				٦- اذا كان المربع اسود فانه صغير
				٧- المربع ابيض اذا واذا فقط هو كبير
				٨- المربع صغير وهو ابيض .
				٩- المربع كبير اذا واذا فقط هو اسود
				١٠- اذا كان المربع ابيض فانه صغير
				١١- المربع اسود او هو صغير
				١٢- المربع كبير وهو ابيض
				١٣- المربع ابيض اذا واذا فقط هو صغير
				١٤- المربع كبير او هو ابيض .
				١٥- المربع اسود وهو صغير
				١٦- اذا كان المربع كبيرا فانه اسود

名前＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

Appendix 1c



下記の１６の命題は上にある４つの四角形A，B，C，Dに関するものです。４つの四角形のうち２つの四角形は大きく残りの２つは小さくまた２つは黒く２つは白いものとします。

各命題を図A，B，C，Dと照らし合せ一致しているかしていないか答て下さい。もし一致していればティック（✓）をそうでなければクロス（✗）を記入して下さい。

		A	B	C	D
1	四角形が黒くかつ大きい。				
2	四角形が大きいならば白い。				
3	四角形が白いまたは小さい。				
4	四角形が黒いならば小さくかつ小さいならば黒い。				
5	四角形が大きいまたは黒い。				
6	四角形が黒いならば小さい。				
7	四角形が大きいならば白くかつ白ければ大きい。				
8	四角形が小さくかつ白い。				
9	四角形が黒いならば大きくかつ大きいならば黒い。				
１０	四角形が白いならば小さい。				
１１	四角形が黒いまたは小さい。				
１２	四角形が大きくかつ白い。				
１３	四角形が小さいならば白くかつ白いならば小さい。				
１４	四角形が大きいまたは白い。				
１５	四角形が黒くかつ小さい。				
１６	四角形が大きいならば黒い。				

References

Airasian, P.W., Bart, W.M. and Greaney, B.J. The analysis of a propositional logic agame by ordering theory. *Child Study Journal*, 1975, **5**, no. 1, 13 - 24.

Alfonso, A. *Japanese language patterns. Volumes I and II.* Tokyo: Sophia University, 1980.

Archer, E.J. On verbalizations and concepts. In A.W.Melton (Ed), *Categories of human learning.* NY: Academic Press, 1964.

Au, T.K. Counterfactuals: In reply to Alfred Bloom. *Cognition*, 1984, **17**, 288 - 302.

Bart, W.M. Adolescents' understanding of logical connectives. *Psychological Reports*, 1974, **35**, 41 - 42.

Beeston, A.F.L. *The Arabic language today.* London: Hutchison, 1970.

Beilin, H. and Lust, B. Connectives: Logical, linguistic and psychological theory. In H. Beilin (Ed), *Studies in the cognitive basis of language development.* NY: Academic Press, 1975.

Berlin, B. and Kay, P. *Basic colour terms; their universality and evolution.* Berkeley and Los Angeles: University of California Press, 1969.

Bernstein, B. *Class codes and control.* London: Routledge and Kegan Paul, 1971.

Berry, J.W. Learning Mathematics in a second language: Some cross-cultural issues. *For the Learning of Mathematics*, 1985, **5**, no 2, 18 - 23.

Bloch, B. Studies in colloquial Japanese II: Syntax. *Language*, 1946, **22**, 200 - 248.

Bloom, A.H. *The linguistic shaping of thought: A study of the impact of language on thinking in China and the West.* Hillsdale, N.J.: Lawrence Erlbaum, 1981.

Bloom, A.H. Caution - the words you use may affect what you say: A response to Au. *Cognition*, 1984, **17**, 275 - 287.

Boole, G. *An investigation of the laws of thought.* London: Macmillan, 1854.

Bourne, L.E., Jr. *Human conceptual behaviour*. Boston: Allyn and Bacon, 1966.

Bourne, L.E., Jr. Knowing and using concepts. *Psychological Review*, 1970, **77**, no. 6, 546 - 556.

Bourne, L.E., Jr. An inference model for conceptual rule learning. In R.L. Solso (Ed), *Theories in cognitive psychology: The Loyola Symposium*. Potomac, Md: Lawrence Erlbaum, 1974.

Bourne, L.E., Jr. Stimulus-rule interaction in concept learning. *American Journal of Psychology*, 1979, **92**, no. 1, 3 - 17.

Braine, M.D.S. On the relation between the natural logic of reasoning and standard logic. *Psychological Review*, 1978, **85**, no. 1, 1 - 21.

Braine, M.D.S. and O'Brien, D.P. A theory of If: A lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, 1991, **98**, no. 2, 182 - 203.

Braine, M.D.S. and Rumain, B. Development of comprehension of 'or': Evidence for a sequence of competencies. *Journal of Experimental Child Psychology*, 1981, **31**, 46 - 70.

Braine, M.D.S. and Rumain, B. Logical reasoning. In P. Mussen (Ed), *Handbook of child psychology. Vol 3: Cognitive development*. NY: Wiley, 1983.

Brown, R.W. and Lenneberg, E.H. A study in language and cognition. *Journal of Abnormal Psychology*, 1954, **69**, 454 - 462.

Bruner, J., Goodnow, J. and Austin, G. *A study of thinking*. NY: Wiley, 1956.

Byrnes, J.P. and Overton, W.F. Reasoning about logical connectives: A developmental analysis. *Journal of Experimental Child Psychology*, 1988, **46**, 194 - 218.

CASME. *Languages and the teaching of science and mathematics with special reference to Africa*. London: The Commonwealth Secretariat, 1975.

Cheng, P.W. and Holyoak, K.J. Pragmatic reasoning schemas. *Cognitive Psychology*, 1985, **17**, 391 - 416.

Ciborowski, T. and Cole, M. A cross-cultural study of conjunctive and disjunctive concept learning. *Child Development*, 1972, **43**, 774 - 789.

Cohen, G. *The psychology of cognition.* London: Academic Press, 1977.

Cohen, B. and Murphy, G.L. Models of concepts. *Cognitive Science*, 1984, **8**, 27 - 58.

Cole, M., Gay, J. and Glick, J. Some experimental studies of Kpelle quantitative behavour. *Psychonomic Monograph*, 1968, 2(10), whole no. 26.

Cole, M., Gay, J., Glick, J. and Sharp, D. *The cultural context of learning and thinking.* NY: Basic Books, 1971.

Cole, M. and Scribner, S. *Culture and thought.* London: Wiley, 1974.

Comrie, B. Conditionals: A typology. In E.C. Traugott et al (Eds), *On conditionals.* Cambridge: Cambridge University Press, 1986.

Copi, I.M. *Introduction to logic* (seventh edition). London: Collier Macmillan, 1986.

Cox, J.R. and Griggs, R.A. The effects of experience on performance in Wason's selection task. *Memory and Cognition*, 1982, **10**, 496 - 502.

Crawford, K.P. *Language and technology in classroom settings for students from non-technological cultures.* Paper presented at ICME-6, Budapest, Hungary, 1988.

Damarin, S.K. The interpretation of statements in standard logical form by pre-service elementary teachers. *Journal for Research in Mathematics Education*, 1977a, **8**, part 2, 123 - 131.

Damarin, S.K. Conjunctive interpretations of logical connectives: Replication of results using a new type of task. *Journal for Research in Mathematics Education*, 1977b, **8**, part 3, 231 - 233.

Dawe, L. Bilingualism and mathematical reasoning in English as a second language. *Educational Studies in Mathematics*, 1983, **14**, 325 - 353.

Denny, N.R. Memory load and concept-rule difficulty. *Journal of Verbal Learning and Verbal Behaviour*, 1969, **8**, 202 - 205.

Dik, S.C. *Coordination: Its implications for the theory of general linguistics.* Amsterdam: North-Holland, 1968.

Dodd, D.H., Kinsman, R.A., Klipp, R.D. and Bourne, L.E., Jr. Effects of logic pretraining on conceptual rule learning. *Journal of Experimental Psychology*, 1971, **88**, no. 1, 119 - 122.

Ervin-Tripp, S. Interaction of language, topic and listener. *American Anthropologist*, 1964, **66**, 86 - 102.

Evans, J.St B.T. Interpretation and matching bias in a reasoning task. *Quarterly Journal of Experimental Psychology*, 1972, **27**, 297 - 306.

Evans, J.St B.T. *The psychology of deductive reasoning*. London: Routledge and Kegan Paul, 1982.

Evans, J.St B.T. *Bias in human reasoning: Causes and consequences*. Hove, Sussex: Lawrence Erlbaum, 1989.

Evans, J.St B.T. and Lynch, J.S. Matching bias in the selection task. *British Journal of Psychology*, 1973, **64**, 391 - 397.

Evans, J.St B.T. and Newstead, S.E. Language and reasoning: A study of temporal factors. *Cognition*, 1977, **8**, 265 - 283.

Evans, J.St B.T. and Newstead, S.E. A study of disjunctive reasoning. *Psychological Research*, 1980, **41**, 373 - 388.

Fillenbaum, S. Inducements: On the phrasing and logic of conditional promises, threats and warnings. *Psychological Research*, 1976, **38**, 231 - 250.

Gallagher, C.F. Language rationalisation and scientific progress. In R.H. Silvert (Ed), *The social reality of scientific myth*. New York: American Universities Field Staff, Inc., 1969.

Gay, J. and Cole, M. *The new mathematics and an old culture*. NY: Holt, Rinehart and Winston, 1967.

Gazdar, G. *Pragmatics. Implicature, presupposition and logical form*. London: Academic Press, 1979.

Geis, M. and Zwicky, A.M. On invited inferences. *Linguistic Inquiry*, 1971, **2**, 561 - 566.

George, C. Facilitation in the Wason selection task with a consequent referring to an unsatisfactory outcome. *British Journal of Psychology*, 1991, **82**, 463 - 472.

Grice, H.P. Logic and conversation. In P. Cole and J.L. Morgan (Eds), *Syntax and semantics. Vol 3: Speech acts.*
NY: Academic Press, 1975.

Griggs, R.A. and Cox, J.R. The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, 1982, **73**, 407 - 420.

Grizzle, J.E., Starmer, F.C. and Koch, G.G. Analysis of categorical data by linear models. *Biometrika*, 1969, **25**, 489 - 504.

Hakuta, K. and Bloom, L. The search for cross-linguistic invariants and variation in language development. In H. Stevenson, H. Azuma and K. Hakuta (Eds), *Child development and education in Japan.* NY: W.H.Freeman, 1986.

Hakuta, K., de Villiers, J. and Tager-Flusberg, H. Sentence co-ordination in Japanese and English. *Journal of Child Language*, 1982, **9**, 193 - 207.

Harnisch, D.L., Walberg, H.J., Tsai, S.L., Sata, T. and Fyans, L.J.Jr. Mathematics productivity in Japan and Illinois. *Evaluation in Education*, 1985, **9**, 277 - 284.

Hatano, G. and Suga, Y. Understanding and use of disjunction in children. *Journal of Experimental Child Psychology*, 1977, **24**, 395 - 405.

Haugen, E. Linguistic relativity: Myths and methods. In W.C. McCormack and S.A. Wurm (Eds), *Language and thought: Anthropological issues.* The Hague: Mouton, 1977.

Haygood, R.C. and Bourne, L.E., Jr. Attribute and rule learning aspects of conceptual behaviour. *Psychological Review*, 1965, **72**, 175 - 195.

Haywood, J.A. and Nahmad, H.M. *A new Arabic grammar of the written language.* London: Lund Humphries, 1965.

Hiew, Chok C. Sequence effects in rule learning and concept generalisation. *American Journal of Psychology*, 1977, **90**, no. 2, 207 - 218.

Howson, G. *National curricula in mathematics.* Southampton: The Mathematical Association, 1991.

Hunt, E. and Agnoli, F. The Whorfian Hypothesis: A cognitive psychology perspective. *Psychological Review*, 1991, **98**, no. 3, 377 - 389.

Hunt, E. and Banaji, M.R. The Whorfian Hypothesis revisited: A cognitive science view of linguistic and cultural effects on thought. In J.W. Berry, S.H. Irvine and E.B. Hunt (Eds), *Indiginous cognition: Functioning in cultural context*. Dordrecht: Martinus Nijhoff, 1988.

Husen, T. *International study of achievement: A comparison of twelve countries*. Stockholm, Almqvist and Wiksell, 1967.

Jansson, L.C. The judgement of simple deductive arguments by pre-service elementary school teachers. *The Alberta Journal of Educational Research*, **XXI**, no. 1, 1 - 10.

Johansson, B.S. Levels of mastery of the coordinators 'and' and 'or' and logical test performance. *British Journal of Psychology*, 1977, **68**, 311 - 320.

Johnson-Laird, P.N. Thinking as a skill. *Quarterly Journal of Experimental Psychology*, 1982, **34A**, 1 - 29.

Johnson-Laird, P.N. and Byrne, R.M.J. *Deduction*. Hove, Sussex: Lawrence Erlbaum, 1991.

Johnson-Laird, P.N., Legrenzi, P. and Legrenzi M.S. Reasoning and a sense of reality. *British Journal of Psychology*, 1972, **63**, no. 3, 395 - 400.

Johnson-Laird, P.N. and Tagart, J. How implication is understood. *American Journal of Psychology*, 1969, **82**, 367 - 373.

Kant, I. *Introduction to logic and essay on the mistaken subtlity of the four figures*. (Translated by T.K. Abbott.) London: Longmans, 1855.

Kay, P. and Kempton, W. What is the Sapir-Whorf hypothesis? *American Anthropologist*, 1986, **86**, 65 - 69.

King, W.L. Learning and utilisation of conjunctive and disjunctive classification rules: A developmental study. *Journal of Experimental Child Psychology*, 1966, **4**, 217 - 231.

Kuno, S. *The structure of the Japanese language*. Cambridge, Mass: MIT Press, 1973.

Lakoff, G. *Women, fire and dangerous things*. Chicago: University of Chicago Press, 1987.

Lakoff, R. If's, and's and but's about conjunction. In C.J. Fillmore and D.T. Langendoen (Eds), *Studies in linguistic semantics*. NY: Holt, Rinehart and Winston, 1971.

Lemmon, E.J. *Beginning logic*. London: Nelson, 1965.

Lemon, N. Language and Learning: Some observations on the linguistic determination of cognitive processes. In B. Lloyd and J. Gay, (Eds), *Universals of human thought*. Cambridge: Cambridge University Press, 1981.

Lenneberg, E.H. The relation of language to the formation of concepts. *Synthese*, 1962, **14**, 103 - 109.

Liu, L.G. Reasoning counterfactually in Chinese: Are there any obstacles? *Cognition*, 1985, **21**, 239 - 270.

Lynn, R. Mathematics teaching in Japan. In B. Greer and G. Mulhern, (Eds), *New directions in mathematics education*. London: Routledge, 1989.

Macnamara, J. The effects of instruction in a weaker language. *Journal of Social Issues*, 1967, **23**, 121 - 135.

Manktelow, K.I. and Evans J.St B.T. Facilitation of reasoning by realism: Effect or non-effect? *British Journal of Psychology*, 1979, **70**, 477 - 488.

Mayer, R.E., Tajika, H. and Stanley, C. Mathematical problem solving in Japan and the United States: A controlled comparison. *Journal of Educational Psychology*, 1991, **83**, no 1, 69 - 72.

Miller, R.A. *The Japanese language*. Tokyo: Tuttle, 1967.

Moore, C.G. *The Navajo culture and the learning of mathematics*. Final report to the National Institute of Education, Washington, DC, 1982.

Morris, R.W. The role of language in learning mathematics. *Prospects*, 1978, **VIII**, no. 1, 73 - 81.

Naganuma, N. and Mori, K. *Practical Japanese*. Tokyo: Japan Publications Trading Co., 1962.

Nairobi. *Interactions between Linguistics and Mathematical Education*. UNESCO, Paris, 1974.

Neimark, E.D. and Slotnick, N.S. Development of the understanding of logical connectives. *Journal of Educational Psychology*, 1970, **61**, no. 6, 451 - 460.

Neisser, U. & Weene, P. Hierarchies in concept attainment. *Journal of Experimental Psychology*, 1962, **64**, no. 6, 640 - 645.

Newstead, S.E. and Griggs, R.A. The language and thought of disjunction. In J.St B.T. Evans (Ed), *Thinking and reasoning*. London: Routledge and Kegan Paul, 1983.

Newstead, S.E., Griggs, R.A. and Chrostowski, J.J. Reasoning with realistic disjunctives. *Quarterly Journal of Experimental Psychology*, 1984, **36A**, 611 - 627.

Nitta, N. and Nagano, S. Basic logical operations and their verbal expressions - child's conception of logical sum and product. *Research Bulletin of the National Institute for Educational Research*, no. 7, Tokyo, 1966.

Osherson, D. Logic and models of logical thinking. In R.J.Falmagne (Ed), *Reasoning: Representation and process in children and adults*. Hillsdale, N.J.: Lawrence Erlbaum, 1975.

Paradis, M., Hagiwara, H. and Hildebrandt, H. *Neurolinguistic aspects of the Japanese writing system*. Orlando, Florida: Academic Press, 1985.

Paris, S.G. Comprehension of language connectives and propositional logical relationships. *Journal of Experimental Child Psychology*, 1973, **16**, 278 - 291.

Paris, S.G. *Propositional logical thinking and comprehension of language connectives*. The Hague: Mouton, 1975.

Pelletier, F.J. Or. *Theoretical Linguistics*, 1977, **4**, 61 -74.

Philp, H. Mathematical education in developing countries - some problems of teaching and learning. In A. Howson (Ed), *Developments in mathematical education*. Cambridge: Cambridge University Press, 1973.

Piaget, J. *Logic and psychology*. NY: Basic Books, 1957.

Politzer, G. Laws of language use and formal logic. *Journal of Psycholinguistic Research*, 1986, **15**, no. 1, 47 - 92.

Politzer, G. Comparison of deductive abilities across language. *Journal of Cross-Cultural Psychology*, 1991, **22**, no. 3, 389 - 402.

Pollard, P. and Evans, J.St B.T. The influence of logic on conditional reasoning performance. *Quarterly Journal of Experimental Psychology*, 1980, **32**, 605 - 624.

Pollard, P. and Evans, J.St B.T. Content and context effects in reasoning. *American Journal of Psychology*, 1987, **100**, no. 1, 41 - 60.

Pullman, H.W. *Cognitive structure and performance in mathematics*. Ohio: Youngstown University Press, 1979.

Quine, W.V. *Methods of logic* (third edition). London: Routledge and Kegan Paul, 1974.

Rips, L.J. Cognitive processes in propositional reasoning. *Psychological Review*, 1983, **90**, no. 1, 38 - 71.

Roberge, J.J. Negation in the major premise as a factor in children's deductive reasoning. *School Science and Mathematics*, November 1969, 715 - 722.

Roberge, J.J. Linguistic and psychometric factors in propositional reasoning. *Quarterly Journal of Experimental Psychology*, 1978, **30**, 705 - 716.

Robitaille, D.F. and Garden, D.F. (Eds) *The IEA study of mathematics II: Contexts and outcomes of school mathematics*. Oxford: Pergamon, 1989.

Rosch, E. Linguistic Relativity. In A. Silverstein (Ed), *Human communication: Theoretical explorations*. Hillsdale, N.J.: Lawrence Erlbaum, 1974.

Rosch, E. and Mervis, C.B. Family resemblances: Studies in the internal structures of categories. *Cognitive Psychology*, 1975, **7**, 563 - 605.

Rumain, B., Connell, J. and Braine, M.D.S. Conversational comprehension processes are responsible for reasoning fallacies in children as well as adults: If is not the biconditional. *Developmental Psychology*, 1983, **19**, no. 4, 471 - 481.

Rumelhart, D.E. Schemata: The building blocks of cognition. In R.J. Spiro, B.C. Bruce, W.F. Brewer (Eds), *Theoretical issues in reading comprehension*. Hillsdale, NJ: Lawrence Erlbaum, 1980.

Seward, J. *Japanese in action.* (2nd edition). NY: Weatherhill, 1983.

Skemp, R.R. *The psychology of learning mathematics.* Hillsdale, NJ: Lawrence Erlbaum, 1987.

Smart, J.R. *Arabic.* London: Hodder and Stoughton, 1986.

Snow, C.E. and Rabinovitch, M.S. Conjunctive and disjunctive thinking in children. *Journal of Experimental Child Psychology*, 1969, **7**, 1 - 9.

Staal, J.F. Analyticity. *Foundations of Language*, 1966, **2**, 67 - 93.

Staudenmayer, H. Understanding conditional reasoning with meaningful propositions. In R.J. Falmagne (Ed), *Reasoning: Representation and process in children and adults*. Hillsdale, N.J.: Lawrence Erlbaum, 1975.

Sternberg, R.J. Developmental patterns in the encoding and combination of logical connectives. *Journal of Experimental Child Psychology*, 1979, **28**, 469 - 498.

Stevenson, H.W., Lee, S-Y. and Stigler, J.W. Mathematics achievement of Chinese, Japanese and American children. *Science*, 14 Feb 1986, **231**, 693 - 699.

Strawson, P.F. *Introduction to logical theory.* London: Methuen, 1952.

Suppes, P. and Feldman, S. Young children's comprehension of logical connectives. *Journal of Experimental Child Psychology*, 1971, **12**, 304 - 317.

Taplin, J.E. Reasoning with conditional sentences. *Journal of Verbal Learning and Verbal Behaviour*, 1971, **10**, 219 - 225.

Taplin, J.E. and Staudenmayer, H. Interpretation of abstract conditional sentences in deductive reasoning. *Journal of Verbal Learning and Verbal Behaviour*, 1973, **12**, 530 - 542.

Taplin, J.E., Staudenmayer, H. and Taddonio, J. Developmental changes in conditional reasoning: Linguistic or logical. *Journal of Experimental Child Psychology*, 1974, **17**, 360 - 373.

Traugott, E.C., ter Meulen, A., Reilly, J.S. and Ferguson, C.A. (Eds). *On conditionals.* Cambridge: Cambridge University Press, 1986.

Travers, K.J. and Westbury, I. *The IEA study of mathematics I: Analysis of mathematics curricula.* Oxford: Pergamon, 1989.

Tritton, A.S. *Arabic.* London: Hodder and Stoughton, 1943.

Van Duyne, P.C. Realism and linguistic complexity in reasoning. *British Journal of Psychology*, 1974, **65**, no. 1, 59 - 67.

Ward, S.L., Byrnes, J.P. and Overton, W.F. Organisation of knowledge and conditional reasoning. *Journal of Educational Psychology*, 1990, **82**, no. 4, 832 - 837.

Wason, P.C. Reasoning. In B.M.Foss (Ed), *New Horizons in Psychology.* London: Penguin, 1966.

Wason, P.C. Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 1968, **20**, 273 - 281.

Wason, P.C. and Johnson-Laird, P.N. *Psychology of reasoning.* London: Batsford, 1972.

Wason, P.C. and Shapiro, D. Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, 1971, **23**, 63 - 71.

Watson, H. Language and mathematics education for Aboriginal Australian children. *Language and Education*, 1988, **2**, no. 4, 255 - 273.

Whorf, B.L. *Language, thought and reality.* London: Chapman and Hall, 1956.

Wierzbicka, A. The double life of a bilingual. In R. Sussex and J. Zubrzycki (Eds), *Polish people and culture in Australia.* Canberra, Australia: Australian National University, 1985.

Wilson, B.J. Communication in science, technology and mathematics education in the Commonwealth. *Science and Public Policy*, August, 1984.

Wright, W. *A grammar of the Arabic language.* (3rd edition). Cambridge: Cambridge University Press, 1971.

Yotsukura, S. Ethnolinguistic introduction to Japanese literature. In W.C. McCormack and S.A. Wurm (Eds), *Language and thought: Anthropological issues.* The Hague: Mouton, 1977.

Zepp, R.A. Bilinguals understanding of logical connectives in English and Sesotho. *Educational Studies in Mathematics*, 1982, **13**, 205 - 221.

Zepp, R.A. Concept formation and verbalization. *Perceptual and Motor Skills*, 1986, **62**, p 370.

Zepp, R.A. The transfer of logical principles from second language to first language. *Psychologia*, 1987(a), **30**, 49 - 52.

Zepp, R.A. Logic in everyday language and mathematics. *Chinese University Education Journal*, 1987(b), **15**, no. 1, 45 - 49.

Zepp, R.A. *New directions in research on language and mathematics*. Paper presented at ICME-6, Budapest, Hungary, 1988.

Zepp, R.A. *Language and mathematics education*. Hong Kong: API Press, 1989.

Zepp, R.A., Monin, J. and Lei, C.L. Common logical errors in English and Chinese. *Educational Studies in Mathematics*, 1987, **18**, 1 - 17.

Zipf, G.K. *The Psycho-biology of language*. Boston: Houghton Mifflin, 1935.