'if p then q' . . . and all that: Logical elements in reasoning and discourse

Counihan, M.E.

Link to publication

'if p then q' . . . and all that: Logical Elements in Reasoning and Discourse

Marian Counihan

Abstract   In this paper we explore differences in use of the so-called 'logical' elements of language such as quantifiers and conditionals, and use this to explain differences in performance in reasoning tasks across subject groups with different educational backgrounds. It is argued that quantified sentences are difficult natural bases for reasoning, and hence more prone to elicit variation in reasoning behaviour, because they are chiefly used with a pre-determined domain in everyday speech. By contrast, it is argued that conditional sentences form natural premises because of the function they serve in everyday speech. Implications of this for the role of logic in modelling human reasoning behaviour are briefly considered.

Keywords   Reasoning · Logic · Interpretation · Quantifiers · Literacy

1 Introduction

Research into the logical reasoning abilities of ordinary people is widely understood to have turned up negative results. Witness, for example, the best-selling popular science book *Inevitable Illusions* (Piattelli-Palmarini 1994), which sets out to describe "how mistakes of reason rule our minds". Such popular perceptions are based partly on conclusions drawn from psychology of reasoning research, namely that subjects' reasoning is heavily influenced by non-logical features of a task. To name a few, experimental subjects allow stereotyping (Sloman 1998), their own beliefs (Oakhill et al. 1989), and surrounding linguistic context (Byrne 1989) to affect their choices in reasoning tasks. It might appear, therefore, that there is limited prospect for modelling

M. Counihan (✉)
Institute for Logic, Language and Computation, University of Amsterdam, Room W-2.18,
Nieuwe Doelenstraat 15, 1012 CP Amsterdam, The Netherlands
e-mail: m.e.counihan@uva.nl

the cognitive achievements of real people with logic. Logic remains the unreachable norm for real people, who, in contrast, are 'ruled by mistakes of reason'.

But what are the assumptions about logic underlying this verdict? In particular, where does the boundary between 'logical' and 'non-logical' lie in language? The current paper is part of a programme which sets out to interrogate the assumptions about logic and its relation to language underlying such research, and on the basis of this to reassess conclusions such as those mentioned above. Here we achieve this by examining differences in performance across subject populations with different levels of literacy. These subject groups had different ways of interacting with language, and we can use this discrepancy to tell us about the factors which shape reasoning in a linguistic context. Once we have an idea of the relevant factors we can re-examine whether such negative conclusions about human reasoning are warranted.

The paper is structured as follows. First, results from a recent study of logical reasoning tasks across different subject groups are presented, focussing on a new finding of variation across premise types. This finding motivates the main purpose of this paper, which is to show how reasoning behaviour is shaped by the typical use of the language employed in reasoning tasks. Specifically, we explain variation across premise sets by differential use of the 'logical' terms (used in reasoning tasks), '*all*' and '*if . . . then*' in everyday spoken speech. This will lead us to reassess how logical-semantic models should be used to understand and evaluate the reasoning behaviour of real people.

2 A Curious Empirical Finding

In this section we report on a particular aspect of the findings of an interview study conducted with subjects with varying educational levels.[1] The motivation for studying reasoning across such groups derives from the fact that almost all empirical work in the psychology of reasoning has been conducted within the university undergraduate population. In terms of interaction with the presented material we might well wonder how representative this group is of the rest of the human population. Which parts of what they do are an artifact of their specific, and perhaps specifically literate, approach to the problem? Which aspects are driven by truly universal human cognitive traits? We just do not know. As such, the validity of their response profile on any one task for a general theory of reasoning is unclear. Above this, there is a more subtle confound in work with undergraduate subjects. Reasoning researchers themselves come from this same pool of subjects, at the high end of the literate scale, and thus there might be many assumptions about linguistic material which are shared by both the experimenter and their subjects, but which are not more widely shared. Towards this end, research with less literate groups works to 'make the familiar strange' and enables us as researchers to see and interrogate our assumptions about language which are operative in the context of reasoning tasks.

The limited research which has been conducted outside the undergraduate subject group has focussed on a narrow range of reasoning tasks. The seminal studies with

---

[1] Details can be found in Counihan (2008).

illiterate subjects ([Luria 1976](); [Scribner 1997]()) covered a range of cognitive tasks but within reasoning focussed on syllogistic-type tasks with a quantified or generalised 'major' premise and a particular statement as the 'minor premise', followed by a question.[2] An example is:

> All bears in the far north are white. Novaya Zemlaya is in the far north. What colour are the bears there?

Broadly speaking, when presented with this kind of material, schooled subjects have little problem answering correctly, but unschooled subjects have great difficulty in answering succinctly and correctly. Yet we know very little about *why* this pattern of responses occurs and whether it is related only to syllogistic-type arguments, or to more general tendencies in reasoning. We may say, the representativeness of this reasoning material has yet to be tested. A broader range of reasoning materials was used in the current study as a means to address this problem.

The study was conducted in South Africa and the participants in the experiment were of varying age and education level, all resident in the small town of Hamburg on the coast of South Africa's Eastern Cape province. Of the 29 subjects, six had no education at all, 13 had less than 10 years of education, and had been out of the education system for more than 10 years, and 10 had completed high school within the last 20 years. I shall maintain these three groupings when discussing my data.

Group 1: No education (6 subjects)
Group 2: 4–10 years of education, left the educational system more than 10 years ago (13 subjects)
Group 3: Graduated from high school within the last 20 years (10 subjects)

All subjects were interviewed individually, in Xhosa.[3] Each interview lasted approximately 30 min and participants were paid for taking part. At the beginning of each interview, the subject was asked about their age, language skills and educational history. Then the tasks were introduced by saying that the experimenter wanted to see how the subject used language and would describe situations and ask them questions on the basis of these. It was emphasised that there were no right or wrong answers and that the primary interest was how they used Xhosa.[4]

The materials used in the interviews were drawn from the range of tasks previously used in psychology of reasoning studies, including syllogistic-style tasks, the suppression effect task, and a thematic version of the Wason selection task. Only the syllogistic and suppression effect tasks will be discussed here, and within these tasks, we will focus on one particular aspect of the findings. This is a comparison between

---

[2] Note that what in this literature is labelled a syllogism is a far broader class of problem than strictly understood by the term. For Luria, a pair of premises, in which the first is "in the nature of a general judgement"—for example, 'Precious metals do not rust' and the second is a particular proposition, such as 'Gold is a precious metal', qualify as a syllogism.

[3] Xhosa is one of South Africa's eleven official languages, and, along with Zulu, one of the most widely spoken. Classified as a Bantu language, it is uniquely recognisable by its range of click-consonants.

[4] The experimenter not being a Xhosa speaker, it was assumed that this would help mitigate against potential authority issues in the interaction.

Table 1 Comparing initial responses across premise forms

Group	% correct with …			
	Quantified premises		Conditional premises	
	Ratio	Percentage	Ratio	Percentage
1	7/23	30	9/18	50
2	25/38	66	28/32	81
3	24/33	73	28/33	85

conclusions drawn from two-premise sets with either an initial quantified premise or an initial conditional premise. As such, we do not actually discuss the 'suppression effect', this being the label for the pattern of responses garnered when a second conditional statement is added to the premise set. Examples of the material reported on here, for respectively the syllogistic task and suppression effect task, are as follows:

> All birds in Cape Town are penguins. My friend saw a bird in Cape Town. What kind of bird did she see?
>
> If Thembi wants to visit her friend, she goes to East London. And she does want to visit her friend. Does she go to East London?

The comparison for simple inferences across premise sets is presented in Table 1. The results tabled show a discrepancy between the groups according to the type of premise they had to reason with. The rate of immediate elicitation of correct answers for the syllogistic task in the unschooled group was 30%—comparable to Luria's and Scribner's findings of 40%[5] and 22.3% respectively. The elicitation rate for the schooled group (73%) was not as high as for instance Luria found (100%), but this fits well with other findings (such as Sloman 1998) which show more variation in the schooled group. As for the conditional elicitation rate of 50%–85%, the only figures with which to compare them would be those from studies with undergraduate students, which have been typically somewhat higher (e.g. 96% in Byrne 1989, for *modus ponens* in the simple argument condition), but note however there is much research showing variation in rates of elicitation depending on the material used. Qualitatively, we can summarize the findings thus: similarly to previous studies on reasoning with unschooled subjects, we found that unschooled subjects commonly refused to reason with given quantified premises or provided their own premises as a basis for further reasoning. This contrasts with the response of schooled subjects to quantified premises, when coupled with a particular premise. Also as previously found, with such premise combinations schooled subjects had less difficulty and more frequently gave the 'logical' answer straightaway. Note that this difference in response, elicited in syllogistic tasks, has been the main experimental finding in this area thus far, and as such has served as the basis for much theorising about the differences between literate and illiterate cognition. But in contrast to previous studies, the current study

---

[5] This is a composite figure, based on 9/15 for syllogisms "associated with experience" and 2/13 for those "not associated with experience" (Luria 1976, p. 116).

also yielded the new finding that conditional reasoning tasks elicited a more uniform response across subject groups—most especially in justifications/elaborations—and a higher rate of inference across all groups than the 'quantified' reasoning task.

Such a finding does indeed suggest that reasoning performance is influenced by specificities of linguistic interaction. It looks as though quantified premises are more difficult to reason with than conditional premises, for all groups, but even more so for unschooled groups. This raises questions, such as: How do the different premise forms shape the responses of schooled subjects and unschooled subjects? Does a subject respond correctly for the same reason in different tasks? And do such findings tell us anything more general about human reasoning? The remainder of this paper addresses questions.

3 Premises in Everyday Language

From any given proposition, a multitude of inferences can be drawn. Yet most of them are not. Take a simple proposition, such as "Today is Tuesday". You would hardly expect anyone to conclude, on hearing that, "Oh, so it's Tuesday or Saturday". More generally, for any p we may infer the weaker $p \lor q$. But this is not something we can expect to see commonly occurring. Why not? One explanation is that $p \lor q$ is less informative than p, and this violates Grice's maxim of quantity: be as informative as possible. As such, we would expect the use of p to conclude $p \lor q$ to be infrequent in reasoning and conversational contexts. Undoubtedly there are contexts to be thought up, in which it would be a natural conclusion to draw—for example, when the disjunction is needed as input for further reasoning—but these are circumscribed.[6]

Now observe that reasoning research doesn't focus on the percentages of people who generate $p \lor q$ on presentation of p, nor on developing theories of why people don't draw such inferences. Though there is some work on reasoning *from* disjunctive premises—see for instance (Van der Henst, Yang and Johnson-Laird, 2002)—reasoning research overwhelmingly focuses on premises with conditional and quantified phrasing. Why should this be so? There are at least two places to look for an answer. One is in the history of the field. But a second reason for the dominance of certain types of premises might be that inferences based on them are more natural, recognisable, common, than others; that they 'make sense' to experimenters and subjects alike. Given the close connections between conditional phrases and inferences, a conditional premise would seem to be just such a candidate.

If we grant reasoning researched is thus biased, then reasoning data needs to be understood through the filter of normal language use. Specifically, in order to better understand the plausible range of interpretations generated by subjects on pre-

---

[6] Imagine a tax form, with Category B defined as the group of people who either earn less than a certain amount or don't work at all. Then if you are working as a PhD student, say, you would conclude that you fit into Category B. Under some description you've concluded the disjunction holds on the basis of one of the disjuncts, and used this to conclude that you belong in Category B. Compare this also with normal use of disjunction, which carries the implication that the speaker does not know which of the disjuncts is true: "When is Peter leaving?" "Monday or Tuesday." Note also that *inferring* a disjunction is a different discourse 'move' from that when a disjunction is offered as a *correction* of a simple proposition.

sentation of the premises in reasoning tasks, we need to look beyond the tasks themselves to see where and how the premise sentences are used in natural discourses: are they commonly occurring, and when they do occur, what work do they do in the discourse? If it can be shown that unschooled subjects are trying to respect very general rules or conventions about language use in the tasks as explained, we lessen the charge of illogicality against them. It will here be argued that in fact they are normal conversants, and that schooled subjects in a sense collude with the experimenter, in a special kind of artful discourse.

Taking interpretational processes into account in reasoning data is by no means a new idea. As far back as Henle (1962) attention has been given to the topic. A broad paradigm of taking semantic and pragmatic factors into account when analysing reasoning behaviour has been advocated by Stenning and van Lambalgen (2001, 2004, 2008). Recently, Politzer (2004) (and see also Politzer and Macchi 2002) has lain out arguments for the necessity of both a *macroanalysis* and a *microanalysis* of the semantic structure of experimental tasks. The latter term is Politzer's label for precisely what we hope to achieve in the current study: based on pragmatic theory, the aim of the microanalysis is to output "the determination of the interpretation of the premises, conclusion or question which the participant is likely to work out." As I do here, Politzer sets this against previous work which "assumed that some kind of literal meaning was communicated" (2004, p. 95).

In the current paper it is argued the response profile across premise sets can be seen as a function of the use of the premise sentence types in everyday spoken discourse. First, a pilot corpus study, which distinguishes categories of usage of *all*, is reported. The significance of the categories, especially with respect to the relation between the quantifier and its domain, is discussed in the light of formal work on quantifiers and domains. It is argued that different types of generalisation are associated with differential relations of *all* to its domain, but that this association is distorted in syllogistic reasoning materials.

Next, the better response to conditional premises despite the aforementioned apparent similarity with universally quantified forms will be explained. Analysis of the semantics of the types of conditionals used in reasoning tasks indicates that how they are used in reasoning tasks is very similar to how they are typically used in spontaneous speech. On the basis of this their inferential properties can be contrasted with the more problematic inferential properties of universal generalisation.

3.1 Reasoning with *all*-sentences

3.1.1 All *in Everyday Language*

To be able to judge whether the quantifier is being used in a natural or recognisable way, we need to know how it is used in spontaneous speech contexts. A pilot study was conducted to this end, with the express intention of explaining the reported data, so the results should be read only as preliminary guides for further investigation. Data

Table 2 Categories of *all* usage in a sample of spoken English

	Category	%
1.	Stock phrases	23
2.	Emphatic usage	21
3.	Anaphoric or deictic use	23
4.	Forward quantification	27
5.	Miscellaneous	6
	Total	100

was drawn from the British National Corpus, found online at http://www.natcorp.ox. ac.uk/.[7]

The results of the study are shown in Table 2. It is worth bearing in mind that the treatment of the categories as disjunct is somewhat artificial, but there were only a few overlaps, and these do not affect the contour of the findings, as will become clear. A notable initial finding is that there were but a few occurrences of sentence initial *all*: three in the sample I looked at, and all in the subcategory of law-giving contexts. Here I report only on the English term *all*, but note that a similar study was made of the Xhosa equivalent, the suffix *-onke*, and the categorisation applied equally well there. In fact, there was an even stronger tendency towards anaphoric/deictic use, this accounting for almost half (47%) of *all* translations, with forward quantification accounting for only 11% of occurrences.

These categories are now discussed in more detail:

1. *Stock phrases* included conventionalised constructions such as 'all around him', 'all but impossible', 'all the same', 'all in all'. In English this appears to be a very common usage, but it remains to be seen whether this category is significant cross-linguistically. A rule of thumb to judge this category is to try replacing *all* with other quantifiers—try 'most' or 'some'—and see if the new phrase is useable. If it is, then it doesn't fit into this category. This heuristic also suggests that stock phrases are a kind of fossilised universal quantification. For this reason we'll exclude them from further analysis.

2. *Emphatic usage*. This is related to the above category in that there doesn't seem to be true quantification going on. Rather, *all* seems to serve to emphasise what's being said, often also in conventionalised formulations. The heuristic I used to assign this category was the following: if *all* is omitted, or replaced with a determiner (as in 'all three countries' becoming 'the three countries'), would the sentence get a different reading? If not, then it belongs here. Examples are:
   − '*Let's all* get the hell out of here'
   − 'in *all three countries*'

   In the majority of cases of emphatic usage, the quantifier is 'floating'—that is, it occupies the position normally occupied by an adverb.

   A discussion of the relations between such sentences is beyond the scope of the current study; for our purposes it is enough to know that they are considered to

---

[7] Details of the corpus data can be found in Counihan (2008).

be logically equivalent to each other in the linguistic literature (Bobaljik 2002; Sportiche 1988).

3. *Anaphoric or deictic usage*. This is the coupling of the quantifier with an anaphoric or deictic term such as 'which', 'this', 'the others', so that the *kind* of entities to be quantified over are given either in the previous discourse, or by the non-linguistic context. Bare *all* counts here too. Sometimes the phrasing is quite conventionalised, but replacement by for instance 'most (of)' is generally unproblematic—suggesting there's live quantification going on. Note that anaphoric and emphatic usage serve a similar function: in many cases the emphatic could be seen as a reinforcing or contrasting with the default reference to a group; the anaphoric serves to do this where the bare anaphoric or deictic element might not clearly distinguish between a group and subgroups of its members. Anaphoric use was also a large category—accounting for 23% of the sample. Examples from the corpus are:

   − 'Of course I know *all that's* no reason I can't have a bike'
   − '*all this* will take time to negotiate'
   − '. . . some of which have a less visible protestant ethos, but *all of which* have a loyalist ethos as well.'

   In this category I have also included anaphoric/deictic reference which also have some descriptive content, such as the following:

   − 'In spite of acknowledging *all these factors*'
   − '. . . parliamentary leaders agreed to take *all such decisions* by a show of hands

4. *'Forward' quantification*. This is the category which most closely resembles how *all* is used in syllogistic arguments. The entities to be quantified over are explicitly mentioned after the *all*, as in 'all ethnic minorities' or 'all or part of the primary school cycle'. A number of subcategories have been splintered off here, based on the clear context for use of such sentences. These are: statements of policy or law; as part of religious doctrine; in writing about scientific research. These subcategories deserve special attention because they are probably the only place where quantification truly 'globally' quantifies, and, perhaps related to this, they are more or less directly derived from textual discourses. But first examples which fall outside these subcategories:

   − 'The bank gave consideration to *all matters relating to the company in question's affairs*'
   − 'It has rendered untenable the simplistic belief that members of *all ethnic minorities* are part of one undifferentiated black mass . . .'

   Then we come to the subcategories:
   (a) Statement of **laws/rules/policies**:

      − '*All penalties* are cumulative'
      − 'A complementary excursion to the Dolomites for *all guests staying 14 nights*'

(b) **Religious doctrine**:
  − 'It is the mystery of the Creation, the God of *all Jews . . .* who transcends *all beings*'
  − 'Christ had died for all, *all men and women*'

(c) **Scientific research**:
  − 'Patterning in *all systems* occurs in small groups of cells, . . .'

  Although there is explicit description of entities to be quantified over in this category, note that in almost all instances there must be further domain restriction in order to pick out the appropriate group quantified over—the latter excerpt is a good example, where the 'systems' quantified over are clearly of a pre-specified sort, or range, given elsewhere, but which are currently under discussion. In fact the only cases for which quantification can truly be said to operate over a universal domain seem to be the religious categories!

(d) *Miscellaneous*
  − Proper names: '*All* Saints', '*All* Angels'

Discounting the appearance of the quantifier in stock phrases and emphatic usage, we are left with two large categories where *all* (or equivalent) is used: anaphoric/deictic, to the linguistic or non-linguistic contexts, and that of 'forward' quantification where what is quantified over is specified in linguistic terms. In the anaphoric/deictic category, *all* seems to be acting as a summarizing element, a kind of fishing net for aforementioned or contextually given referents, about which something *further* can then be said. It's almost like a group pronoun. This is more plausibly a universal usage—witness the frequency of this usage in Xhosa. Here *all* may be truly quantificational, but it operates on an extremely restricted domain: that given by the previous discourse or the context of utterance.

What about the cases labelled 'forward' quantification? These cases don't seem to sum up anything previously mentioned, but be 'forward-looking' in the sense that what they quantify over is introduced after *all* is. It is in these cases that *all* is most likely to be available as a premise (the argument 'Christ died for us all, therefore Christ died for you' sounds vaguely like Catechism class).

Interestingly, contexts in which 'forward' quantification was used were often clearly derived from written discourses—for instance religious and legal contexts. What's interesting is that the quantification in these subcategories is the closest thing to strict quantification. Any exceptions to the universality would at least have to be specified. For example, consider the following, 'A complementary excursion to the Dolomites for all guests staying 14 nights'. We might well expect exceptions: guests who benefit from some other special offer, or stay on reduced rates, might not be entitled to their free trip to the Dolomites—but this would have to be explicitly mentioned (the small print!), as a caveat to the rule 'all guests get a free trip'.

The common condition to these uses of the *all* as quantifier is the pairing of its usage with determination of domain restrictions. In both these large categories there is a division of semantic labour between linguistic and contextual (either from the previous discourse or the non-linguistic environment) input by which the domain of

quantification is determined. In some cases—the legal, religious, and science contexts, the burden lies more on the linguistic side—but it is never contained totally in the sentence of quantification. Although 'anaphoric' and 'forward' usages have been presented as separate categories, they could be better viewed as points on a slide of determining quantification—the one side being quantification items being determined by new linguistic information, the other side by previously or contextually given information. The large category of quantifier usage for which the domain is determined 'purely' anaphorically should thus be understood as an articulated continuation of what extends below the threshold of linguistic explicitness i.e. determination of the domain by context.

The matter of fixing an appropriate domain of quantification is clearly central to any usage of quantifiers (when they are really quantifying, that is). How exactly the type of quantification interacts with its domain of quantification is the topic of the following section.

3.1.2 Domain Sensitivity and Quantification

The pervasiveness of the anaphoric/deictic category highlights a distinction which has already been proposed in formal work on quantifiers. For instance, Westerståhl (1985) argues that in providing a semantic analysis for *all* one needs to distinguish three types of universe—as opposed to just two as is usually proposed in the 'flexible universe' strategy, in which pragmatic processes are assumed to continuously adjust the discourse universe appropriately.[8]

In more detail, Westerståhl's account is as follows. Model-theoretic semantics routinely makes reference to a universe of models, or *discourse universe*, M, in a model $\mathcal{M} = \langle \mathcal{M}, [[.]] \rangle$[9] as well as the denotation of the noun in the model (where NP = determiner + noun), which can be viewed as the *NP universe*. But Westerståhl (1985) argues that semantics should distinguish also a *context set*, a contextually selected sub-universe of M. The role of selecting this sub-universe is usually assigned to pragmatics; as Westerståhl says, "in practice this means identifying context sets with (temporarily chosen) model universes" (p. 46). He offers two types of argument why context sets cannot be identified with discourse universes, the first methodological, and the second via concrete examples. Methodologically the two do not match because discourse universes are large and constant over pieces of discourse while context sets are not, and determiners are 'universe-independent' (pp. 48–51) in the sense that their interpretation is not affected by the discourse universe in which they occur, as long as it's large enough (a property referred to as EXT in many logic textbooks[10]).

---

[8] He goes on to show how this three-way distinction can be implemented in an enriched version Barwise and Cooper's Logic for Generalised Quantifiers.

[9] [[.]] is an interpretation function assigning interpretations to natural language expressions.

[10] Formally expressed by

(EXT)   If $A, B \subseteq M \subseteq M'$ then $D_M AB$ iff $D_{M'} AB$

The vivid second argument offered by Westerståhl is in the form of examples for which there is simply "no way to make sense of [the] sentences if the discourse universe is identified with the context set" (p. 49), viz:

(1) The English love to write letters. *Most* children have *several* pen pals in many countries. (emphasis mine)

Now this sentence only makes sense if we understand *most* to be quantifying over children in England but *several* to be quantifying over children *anywhere* in the world. Were we to identify the context set with the discourse universe, either the first NP-universe would be 'most children in the world' or the second one would be 'several children in England'—clearly neither of which is intended in the context of utterance. In fact, the first NP does not operate on the discourse universe but on a restricted context set, given by the previous sentence, while the second NP operates over a bigger set. Obviously, the discourse universe must include this bigger set of children anywhere. The possibility remains open that in certain sentences the context set coincides with the universe of discourse *M*—for instance in the amended example below the context set remains the same for both NPs and can be taken to coincide with the discourse universe:

(2) Children love to write letters. Most children have several pen pals in many countries.

Specifying a role for the context set as apart from the NP-universe in determining the domain of quantification can explain how anaphoric and deictic usage works. In these cases, the context set gives the sub-universe in which the anaphoric or deictic elements are interpreted—and the NP-universe restriction is minimal. Moreover, the size of the category anaphoric/deictic illustrates the importance of the context set in determining the domain of quantification in everyday use of the quantifier *all*. For demonstratives such as 'that', 'this', 'those', they can only get the appropriate denotation when interpreted in a restricted sub-universe of the discourse universe, i.e. in a context set. Westerståhl proposes just such an explanation for bare or pronominal use of *all*, arguing that "the lack of an argument is a visible *context set indicator*, which signals the implicit occurrence of a context set" (*ibid*, p. 49, his emphasis).

Recent work in linguistics makes this same kind of distinction, albeit with different terminology. Apart from explicitly quantified formulations such as

(3) All the girls jumped in the lake.

there are 'bare' plural noun phrases such as

(4) The girls jumped in the lake.

Yet these are often taken to imply quantified formulations such as

(5) Every girl jumped in the lake.

This phenomenon has lead many theorists to posit a so-called "D operator", an implicit distributivity operator which introduces quantification in the denotation of the plural

'the girls' (Brisson 2003). Note though that the first sentence above is slightly 'weaker' in that, as will be discussed below, its truth can in many situations withstand a few non-swimming girls. In linguists' lingo, the quantification is nonmaximal. The influential paper by Brisson (2003) uses Schwarzschild's (1996) idea "that a D-operator has a domain variable in its restriction *whose value is contextually specified*" (Brisson, p. 130, emphasis added), to propose that *all* is in fact not a determiner-quantifier, "but rather interacts with the quantification introduced by the D operator to rule out the nonmaximality that a D operator normally allows" (*ibid*, p. 141). 'Nonmaximality' here means allowing for exceptions, so in this respect resembles genericity. Brisson's proposal is that *all*, as witnessed in (3), functions as a means to adjust the domain to make it maximal, *where the domain is contextually-selected by a variable within the D operator*. (The variable serves the same function as the context set in Westerståhl's account.) In this function *all* is different from quantifiers such as *every* or *each*, which also cannot combine with plural noun phrases, viz:

(6) *Every the girls jumped in the lake.

Although Brisson's proposal is very suggestive in the light of the current analysis, the further details go beyond our purpose and scope. Westerståhl's relatively simple proposal suffices make the general idea of local domain determination precise, but it should be clear from the above that alternative formulations are available.

To return to Westerståhl's terminology: as we will see, in the context of the reasoning task the context set is considered identical with the discourse universe, from the experimenters' point of view, but in certain conditions the two could—and even should—sensibly be distinguished, and this leads to a divergence in denotation for the determiner NP for the experimenter and the unschooled subject. But under which conditions is the domain of quantification given by context set as different from the universe of discourse, and how does the quantification relate to the context set in such cases? We now address this question.

3.1.3 Law-like and Contingent Generalisations

Two main interpretations for *all* sentences are considered and discussed in relation to domain determination and the role of the context set. The first type of interpretation considers statements of the form '*All x are y*' as expressing some kind of theory-supported or causal generalisation. This does not necessarily mean that the quantification is interpreted strictly. Certainly in English *all* can be used to make a generic statement.[11] Consider the sentences

a. All doctors wear white coats.[12]
b. All bears have four legs.[13]

One can read the sentences as expressing something about the prototypical doctor, or bear, and thus equivalent to the bare generic formulation

---

[11] In Counihan (2008) it is argued that such a reading of *all* statements might be an artifact of linguistic theorising and not one common in naturally occurring language use.

[12] Example from Partee (1985).

[13] Example from Pelletier and Asher (1997).

a.′ Doctors wear white coats.

b.′ Bears have four legs.

This only goes through in the case of a conventional or causal, that is, a law-like, connection. In such cases, the generalisation can withstand counterexamples, or rather, exceptions, so that the odd three-legged bear, having perhaps been caught in a hunter's trap, does not shake our belief that 'All bears have four legs'. This does not work for contingent generalisations. Borrowing from an example in Pelletier and Asher (1997), if by some cruel twist of fate all bears in the world lost a leg, we would not assent to 'Bears have three legs', but would maintain the generic 'Bears have four legs'—even while admitting that 'All the bears have three legs'. In other words, the generic reading of a universally quantified statement is not available when it expresses a mere contingent generalisation. Here genericity is expressed with a bare noun phrase—a point to which we shall return. In general, the degree of robustness to exceptions seems to vary with the degree of theoretical basis for the generalisation.

Notice however that in certain contexts the generic reading of *all* is ruled out and it gets a strict reading while paired with a law-like connection. This is the case for statements expressing scientific or theory-based laws, such as 'All bears are mammals', and Goodman's example 'All butter melts at 150° F', as will be discussed in more detail further on. In these cases one can read the quantification as applying to *kinds* (of bears, or butter) rather than individuals (single bears or pats of butter). Combination with deontic modals in imperative statements also yields an exceptionless reading: 'all doctors must wear white coats', although an exception here does not falsify the rule but violate it.[14] In fact many subjects take a deontic interpretation of some reasoning premises, such as 'All people who own houses pay house tax'.

A second available interpretation for an *all* sentence is that of a contingent generalisation, where the connection described by the predication is accidental. Imagine, if you will (after Goodman 1947), that upon checking my pockets this morning at home, I found only silver coins in my right pocket. Then the universal generalisation 'All the coins in my right pocket are silver' is true. But now suppose that on the way to work I bought a coffee and slipped the change into my right pocket. Does the generalisation still hold? Maybe, maybe not. It depends what I put in my pocket. Certainly we wouldn't say that the additional coins *became* silver on being put into my right pocket. If there was a copper coin among my change then it's no longer a true generalisation. The generalisation may be 'universal', but only in the small and rather gloomy universe comprising the inside of my pocket on a particular winter morning. It's a very circumscribed and contingently constituted domain.

These can be compared to universal generalisations which express physical laws, such as

(1) All butter melts at 150° F

This is a statement of a law-like relationship, and can be distinguished from true contingent generalisations like

(2) All the coins in my pocket are silver

---

[14] See Stenning and van Lambalgen (2001) for an extended discussion of this in the context of the Wason selection task.

by the fact that the first statement can be accepted as true before all cases of it have been determined—these undetermined cases being predicted to conform with the law. In contrast, a statement like (2) "is accepted as a description of contingent facts *after* the determination of all cases, *no prediction of any of its instances being based upon it*" (Goodman 1947, p. 124, second emphasis mine). In other words, there are different criteria of acceptance for the two kinds of statement. For the case of a law, a few positive instances may lead us to accept the statement as true, but in the case of an accidental generalisation we need to have tested all instances before we can accept it as true.

Why did I assent to the sentence 'All the coins in my pocket are silver' this morning? I could do so precisely because the sentence is taken to refer to all coins that were in my pocket *at the time of utterance*—and not the coins which have been or will be in my pocket. This is what makes it a contingent generalisation. It's a description of a part of the world at a particular point in time. We might say, the generalisation operates on a context set which can be, and probably is, much smaller than the discourse universe. The only instances which counted were those known of in the specific situation. Sometimes this can also involve past instances—'All my boyfriends are short'—but essentially nothing can be claimed about future instances which will fall under the generalisation, i.e. future coins in my pocket, or future boyfriends, or even instances which fall beyond the domain in some other sense—like coins in the trousers hanging over a chair in my bedroom. So we see that, in cases where we do accept contingent, un-lawlike generalisations it is because there is a limited range of instances to which they apply, and these have been checked.

There is a further difference in the examples offered by Goodman contrasting law-like with contingent generalisation: the latter is expressed with a definite article. In English, the use of the definite article 'the' after *all* supports a contingent reading because it suggests an identifiable and finite domain. For example, compare

(7)   All women have two children

to

(8) All the women have two children.

In the latter it is clear we are talking about a specific group of women to whom the generalisation applies; in the former this reading is not available without considerable contextual support. But although the bare version does not get a contingent reading, the reverse is not always the case. For example, 'All the bears in the North shed their winter coat' can be read as generic, but, again, in that case it does suggest quantification over *types* of bears rather than individual animals. Note that an accompanying definite article rules out tolerance of counterexamples, so that a single woman in the relevant domain with only one child falsifies the generalisation 'All the women have two children'.[15]

For English materials the use of the definite article would create a lurking confound when testing syllogistic premises, if it is indeed associated with contingent generalisations, because, as we'll see, a contingent generalisation makes for an awkward premise.

---

[15] This is often called the 'maximising' character of *all*, contrasting with the nonmaximality of 'The women have two children (each)' (Brisson 2003). See previous section for more discussion of this feature.

In the current study, we can ignore this issue because in Xhosa, there is no distinct part of speech corresponding to the definite article in English (Tsolwana 1996), so that, for example, 'all women' and 'all the women' both get translated as 'bonke abafazi'. This means that the difference between generic/law-like and contingent generalisation is not expressed by the use of the definite article. This will turn out to be an important point in the later analysis.

The two kinds of generalisations can be seen as opposite poles on a scale of domain-sensitivity. On the law-like end, we have generalisations which are often based on causal relations, and which can therefore be judged on the basis of few instances. They are in this way true universal generalisations, *beyond* any given domain (the lab, say, where the experiment is conducted). There are two things to note about this: (1) even, or rather, especially, here, *ceteris paribus* clauses hold, qualifying extrapolation to new cases; (2) these generalisations are usually not explicitly quantified: for example, 'butter melts at 150° F' expresses the same law as that expressed in (1). Simply *by virtue of being butter* it falls under the generalisation. In fact, law-like regularities are probably not expressed in natural language with universal quantifiers, but with generic statements. Goodman's distinction between the two types of generalisation might only be relevant in the domain of scientific language. Should this be so, it only strengthens the current claim that universal statements are usually used contingently and with a pre-specified domain.

What about *all* statements which are read generically such as described above, but which are not interpreted strictly and are thus not falsified by single 'counterexamples'? These are the generalisations which describe stereotypes, patterns, habits, typicalities, which are more predictable than simple accidents, but which are not as reliable as laws of nature. These can cross domains but the *ceteris paribus* clause will become more difficult to enforce and the presence of instances which serve as exceptions will increase. A key aspect of these exceptions is that they are not real counterexamples because they need not negate the quantification. As illustration here consider the example

(9)   All chairs have four legs

or the previous example

(10)   All doctors wear white coats.

One might agree that these statements hold in some general sense while also granting that in certain contexts—respectively, for instance, an *avant-garde* design exhibition and a surgical operating theatre—exceptions will abound.

On the other end of the scale there are descriptions of states of affairs which express entirely contingent or accidental generalisations, which do not extend beyond a known domain. This is what we might call local universal generalisation. Here quantifying expressions do real work because no inherent property of the entities involved makes them fall under the generalisation: there's nothing causal connecting my pocket and silver coins. As such, should anything change about the domain—such as new coins get added—we can say nothing about the status of the generalisation. It's hyper domain-dependent in a way that laws aren't. Another way to see this is

to compare the quantified formulation of generalisation with conditional formulation of it: 'if something is butter, it melts at 150°F' holds, under *ceteris paribus* clauses, for undetermined entities. But 'if something is a coin in my pocket, it is silver', apart from being a very awkward paraphrase of the quantified statement, we can only be sure to be true for entities in the original domain. (See the next section for more discussion of conditional formulations.) Here especially, we might expect context sets to play a key role in domain determination—i.e. contextually-given, surveyable sub-universes in which instances of the generalisation are known or visible.

If we grant that quantifiers are usually interpreted on a contextually given sub-universe of the discourse universe, then we also have a means to explain the difference between law-like and contingent generalisations. The current proposal is that contingent generalisations are interpreted on a context set which is varying from context to context, while law-like generalisations do not rely on a context set for their interpretation: they can be interpreted on any universe of discourse. The phenomenon that Goodman observed, that one can extrapolate to new untested instances which were not necessarily part of the original domain, stems from the theoretical basis for the generalisation.

The next section shows why, in experimental conditions, the failure to distinguish a role for the context set might lead the experimenter to project an inaccurate and often unwarranted interpretation for the quantified premises onto unschooled subjects' reasoning.

3.2 When *all* Sentences Make Lousy Premises

In the syllogistic task, the subject is first presented with a universal generalisation: "All bears in Novaya Zemlya are white." One can interpret this either as a strict law-like, or as a generic, or as a contingent generalisation. Next, the subject is presented with a possible instance of this generalisation, in the second premise "My friend saw a bear in Novaya Zemlya." Then, depending on which reading of the generalisation is taken, the subject should answer the question: "What colour was the bear?". Herein lies the anomaly:

- **Option 1: Strict law-like**
  All instances of bears are predicted to conform, so answering that the bear is white is minimally informative, given that the subject presumes that their interlocutor knows what she's asserting (i.e. the premises). The answer would be more informative if there is some reason to believe that the bear was a counterexample to the law, or if the question is understood to be about the basis for a law-like connection, including specifying what the *ceteris paribus* clauses involves—i.e. what determines the certainty of prediction. A response which fits this reading is found in the following excerpts (see especially emphasised parts):

  Abdurakhm., age 37, illiterate.
  S: There are different sorts of bears.
  [The syllogism is repeated.]

> S: I don't know; I've seen a black bear, I've never seen any others ... *Each locality has its own animals: if it's white, they will be white; if it's yellow, they will be yellow.*[16]

Ishankul, age 63, illiterate.

> S: If you say that they are white from the cold, they should be white there too. Probably they are even whiter than in Russia.

Recall also our discussion of the strictly unnecessary use of *all* for expression of law-like regularities, something which might contribute to the improbability of the law-like reading of the premise.[17]

Preliminary evidence that subjects in the current study have problems taking a 'law-like' reading of the generalisation are given in the following two excerpts. Firstly, the sheer impossibility of a truly universal reading of the quantification is voiced:

> **Florence (group 2)**:
> E: OK. So one more question. So suppose that all lawyers are alcoholics. And all alcoholics smoke cigarettes. Do all lawyers smoke cigarettes?
> S: No, they can't all smoke.
> E: But suppose we make it true.
> S: *But even though we make it true, there will be others that don't smoke.*
> E: It's impossible to make it true?
> S: Yes.

The exchange indicates a tension in her to accept the generalisation on a universal domain because of its inherent implausibility. She seems to be saying: it can only be true in a limited domain.

The second excerpt contains an incisive query about the sense of a law-like reading—recall the coins which surely do not turn silver on being slipped into my pocket. This query comes from a schooled subject.

> Nontembeko (group 3):
> E: So suppose there's this imaginary country called Markia. And all the women who live in Markia are married. And Fatma is a woman who lives in Markia. Is Fatma married?
> S: If Fatma is a lady then definitely she's married because all the women who live in Markia are married.
> ...
> S: *Is it the law that all the women there are married? Is that possible if it's me, I'm going there, and I'm not married?*
> E: You can go there as a visitor. If you visit, you don't have to be married.

The outcome, as seen in the last turn, is that the experimenter is forced to qualify the generalisation to exclude visiting women!

---

[16] Where I read 'it' as referring to the locality. An intriguing suggestion of natural selection?

[17] Obviously, this only applies to generalisations expressed using *all*.

- **Option 2: Generic**

  As discussed above, generics tolerate counterexamples. So for example, bears in Novaya Zemlya might usually be white, and the odd roaming brown bear wouldn't threaten this generalisation. Any individual bear could thus be an exemplar *or* an exception. It might even be that singling one out suggests exception rather than exemplification. Support for this is given by Clark and Bangerter (2004) review of research which shows that subjects identify referents according to salience against the common ground, where common ground includes for instance 'given' information. Witness, for example, Nofezile from my study (group 2), upon presentation of the problem 'In Markia all women are married, and Fatma is a woman who lives in Markia. Do you think she's married?':

  S: Does she stay alone?
  E: We don't know. All we know is that all the women in Markia are married.
  S: I don't think she's married if she stays there.
  E: Why?
  S: *I will say so because you said all the women in Markia are married and then you say her, living in Markia alone.*

  There are two ways to interpret the subject's initial response, 'Does she stay alone?' Firstly, one can read this as a question about the law-like basis for the generalisation: on what basis can we go beyond the known instances? This suggests a law-like reading is available to the subject, but note that we can get away with positing a deontic law-like reading, something like: 'All women in Markia must marry their live-in boyfriends'. Alternatively we can understand the question as one after further specification of the domain, trying to establish a context set so to speak. The last turn indeed suggests that the subject has introduced her own context set, and decided that the protagonist must fall outside the putative—restricted—domain of application, that is, *cohabiting* women in Markia. And indeed, why would we mention Fatma *unless* she was somehow exceptional?

- **Option 3: Contingent**

  Now either this is an untested instance about which we strictly can say nothing, or it's a known instance in which we're in situation of Option 1. The countless refusals along the lines of 'I don't know, I've never been there' can be understood as one of two possibilities on this analysis: either the subject is saying that she has interpreted the premise as a contingent generalisation, in which case it's a moot point whether any new instance falls under it or not—she would have to know what the context set was as it were—or that she has no grounds for a theoretical basis for the generalisation, which would enable her to predict the colour of the next bear simply on the basis of the given generalisation. An informative answer here would be to give conditions under which you could determine which of the cases you're in—i.e. to specify whether or not and why the bear in question falls under the domain of the generalisation or not. This is what we find: witness Nozuko (group 3) seemingly resorting to a limited-domain interpretation when no theoretical grounds for a law-like reading are forthcoming:

E: Suppose all the women in Nigeria are married. Now there's a woman called Connie and she's not married. Can we say she lives in Nigeria or not?

S: What kind of clothes do they wear in Nigeria?[18]

E: Just suppose the world is a strange one in which all the women in Nigeria are married.

S: We can say she's a Nigerian but she hasn't got married yet.

This last turn—especially the 'yet'—can also be understood as a plea for a temporally-delimited interpretation of the concept "woman" as females of marrying age.

After surveying these three options which are open to a subject on hearing the question 'What colour is the bear?', we see that answering 'white', is not as straightforward as it might seem. In fact, the given range of options suggests that the subtext of the question is a question after the applicability of the generalisation. Especially the generic reading of the generalisation suggests the question be best interpreted thus. On the other hand, to simply use the generalisation as a premise, thus taking a 'straight' reading of the question, seems to result in either being uninformative, or rash. If this is indeed the case, we would expect reactions to include further suggestions or queries about specification of the domain; or about the relation of the named particular to the domain. This type of reaction is indeed present in a good many responses.

An alternative means to capture the oddity of the syllogistic question is in terms of the difference between an instance being *known* to conform and *predicted* to conform. Using the given premises as intended implicates that the conclusion is foregone because the instance is *known* to conform, but this is uninformative. On the other hand, assessing whether the instance is *predicted* to conform results in an informative response, but involves assessing the domain and type of quantification, and thus going beyond the question as asked.

This mismatch between domain and interpretation can explain the frequent asking for or volunteering specification of the limits of the domain of quantification. This occurred in all groups but was less common in the schooled group. Witness, for example, Nonkululeko (group 2):

'All people who own houses pay house tax. Sabelo does not pay house tax. Does he own a house?'

S: He doesn't have a house if he's not paying.

E: And now suppose that none of the people in Cape Town pay house tax. Do they own houses or not?

S: They have houses.

E: Why?

S: *They can have houses because there are places where you don't pay tax, like the squatter camps.*

E: So they can have houses and not pay?

S: They may, they can live at the squatter camps.

---

[18] In Xhosa custom a married woman indicates her status by wearing a specific kind of dress: she always covers her head, and wears an apron. There is also special language only for married women: some objects (e.g. cattle, stones) have two names, one for use by married women, one for use by everybody else.

The subjects' response to the question, 'Do [the people in Cape Town] own houses or not?' might be taken as a Luria-style 'ignoring the premises' response, or even as inconsistent with her previous turn, if it were not for her next response, in which she justifies her conclusion, that people in Cape Town do have houses, by explaining that the given generalisation does not hold in Cape Town—that is, we've gone beyond a restricted domain—the context set—in which the quantification was strict. She is in essence saying, in her second turn, 'I am free to assert they have houses in Cape Town because the given statements do not apply there.'

So if the generalisation is so problematic for some subjects, how is it supposed to function from the experimenter's point of view? The domain restriction is intended to be 'self-contained', i.e. given only by the NP, with no further domain restriction given by a previously determined context set. Now we can connect this with the insights of both Westerståhl and Goodman with regards the nature of domain-relations. In anaphoric use of quantification this means that all elements of the domain are 'known'—or at least they have been previously indicated by the context. 'Forward' quantification, so long as it introduces the elements to be quantified over only in the generalisation itself, leaves open the extent of the domain, thus allowing for previously unmentioned or unknown elements to still fall under it. It seems plausible that interpreting the generalisation on an unspecified context set (or, as a putative default, the universal domain), which is required for correct response in the syllogistic task, might be more common in scientific or literate discourses, and the assumption that the unschooled subject understands it as a law is thus problematic. For instance, witness Luria's observation that: "In all instances, when a subject repeated the premises he did not give them the character of universal assertions" (1976, p. 106). As we have seen, taking the strict (unrestricted) universal reading of *all*, where a *ceteris paribus* clause suffices for domain restriction, is atypical for everyday spoken language use of universal quantifiers. In fact, it might be that the whole idea of a universal domain is a literate idiosyncrasy. Yet this is the default domain in syllogistic tasks.

If this is indeed the case, then the difficulties that unschooled subjects have with quantified premises should not be attributed to their faulty reasoning ability, but rather to the experimenter's chauvinistic view of what interpretations are available and plausible for such premises. The most obvious prediction that follows from the account given here of difficulties with *all* premises is that such premises would become unproblematic when the nature of the generalisation is clarified. So, for instance, if the context set was specified then subjects should be able to answer questions about particular instances of the generalisation. Recent work (Haan 2007) addresses this question and verifies that this is indeed the case. Haan conducted two reasoning tasks with an explicit context set, designed to test whether an explicit domain aids unschooled subjects. The results from the tasks confirm that this is the case.

In the first task, the 'box task' (versions one and two), subjects were presented with a tray containing three red boxes and shown that each box contained a ping-pong ball. The content of the boxes was then hidden and the subjects were asked the following questions:

 i.   'Is it true that all red boxes contain a ball?'
 ii. (On one of the boxes being produced again): 'What is in this box?'

and in a third version which rules out *yes*-bias the subject was first shown three blue boxes, only two of which have a ball in them, and asked

iii.   'Is it true that all blue boxes contain a ball?'

The rates of correct answers for the tasks were respectively 100%, 69%, and 100%, thus confirming our prediction that using universal generalisations in a typical fashion (that is, contingently on a specified domain) removes difficulty for unschooled subjects in reasoning with them. Such a task set-up looks to be very fruitful for further investigation because the multi-modal sources of information avoid many of the problems of information structuring often present in purely linguistic tasks.

4 Reasoning with *if*-Sentences

In the foregoing, we identified some aspects of the use and interpretation of *all* which may explain the difficulty unschooled subjects have with universally quantified premises. In this section I investigate the use and interpretation of conditional sentences in everyday spoken discourse, with the aim of finding candidate explanations for the relative ease with which subjects reasoned with conditional premises. Unschooled subjects were much less inclined to give 'non-answers' (what Luria termed 'refusals') when presented with conditional premises, in the so-called suppression effect task. Perhaps even more interestingly, within the schooled group subjects were more inclined to give elaborations to the premises in the conditional tasks than they were with quantified premises. In other words, the gap between the two groups narrows, from both sides, when subjects are presented with conditional premises. This section provides a suggestion for why this is so.

4.1 '*if*' in Everyday Language

Unlike the situation for quantifiers, there is existing research on the the use of conditionals in practice. Firstly, conditionals have consistently found to be more common in spoken than in written discourse: Hwang's (1979) analysis of a corpus of spoken and written English found 4.2 *if*-conditionals per 1000 words in speech against 2.7 per 1000 in text, and Ford and Thompson (1986) found 7.2 per 1000 words in speech versus 4.6 per 1000 in text. Secondly, initial conditional clauses—that is, a preposed *if*-clause—outnumber final conditional clauses by about three to one, or even more in spoken language; Ford and Thompson (1986) found initial *if*-clauses in 82% of spoken occurrences of the conditional. Comrie (1986), in a typological study, found no counterexamples to a preference for initial conditionals. As for what *kind* of conditionals are frequently used, Hwang found the structure 'if+pres., pres' to be most commonly used, closely followed by 'if + pres., will/going to'. Comrie found the latter construction, i.e. future or predictive conditionals, to be the most common.

The majority of the conditionals used in reasoning studies can straightforwardly be classified as habitual factual conditionals, recognised by the fact that 'if' can

unproblematically be substituted with 'when' or 'whenever' (Celce-Murcia and Larsen-Freeman 1983). For example, I used:

If Ntombi wants to see her boyfriend then she goes to East London.

The second type used in reasoning studies is the future conditional, characterised by the use of the future tense in the consequent clause, and generally expressing predictions:

If Thembi works hard then she will pass her exams.

Byrne's original (1989) suppression effect material used future conditionals, such as:

If it is raining then she will get wet.
If she has an essay to finish then she will study late in the library.

The first example given can be doubly classified as future and inferential, where the latter express inferences about time-bound relationships and often contain modal auxilliaries (Celce-Murcia and Larsen-Freeman 1983). The second example from Byrne can, despite the future tense, be read as habitual, as above, as it can be seen to instantiate a stable relationship rather than express a singular connection. In fact, other studies based on this material (Lechler 2004, and the current one) used a 'factual' formulation of the conditional premise in some of the conditions, viz:

If Rosa meets her friend Liz then she goes to the cinema.

and

If Maria studies hard then she gets high grades in the exams.

In summary, the materials used in conditional reasoning tasks are classifiable as expressing either factual or future conditional relationships. Factual conditionals expressed habitual relationships and the future conditionals expressed predictions, which can often also be read as instances of habitual, generic or inferential relationships. They are also used in the preferred manner—that is, with preposed *if*-clauses. Generic conditionals expressing laws have not been used in the reviewed studies. Further, all these types are to be distinguished from imaginative conditionals, comprising hypothetical or counterfactual conditionals, which have not been tested either. To be addressed next is: how are these conditionals used in discourse, and how does this explain the responses given by subjects in tasks involving conditional premises? What is it about these conditionals that makes them seemingly such natural premises?

4.2 Why conditional sentences make natural premises

In this section we employ the widely-accepted idea that conditional clauses serve the discourse function of providing 'topics', as first outlined by Haiman (1978) and supported by further corpus studies such as e.g. Ford and Thompson (1986) since then. In Haiman's words: "A conditional clause is (perhaps only hypothetically) a part of the knowledge shared by the speaker and his listener. As such, it constitutes the *framework* which has been selected for the following discourse." Haiman also points out that in

English, "*if*-clauses are hypothetical most of the time" (1978, p. 581)—see this by a comparison with *given that* clauses.[19]

If-clauses do indeed provide topics, how does this explain the ease with which subjects reason with them? The main reason is that having a conditional sentence as an opening sentence gives the premise set a normal start, with the discourse topic provided in the first clause. A conditional first premise, then, introduces the new information in a way which mimics one way in which new information is introduced in spontaneous discourse.[20]

But why does the second premise, a simple declarative, fit easily together with this? We propose that the crucial factor is the type of conditional clause used in reasoning experiments, namely dominantly habitual or habituality-based predictive conditionals. As we have already seen, habitual conditional relationships are not time-bound, rather they hold over extended periods of time. This is not to say that they are continuously being fulfilled. They are precisely conditional because the antecedent condition is not always fulfilled. When, however, the antecedent condition is momentarily fulfilled, as in the case of the *modus ponens* inference, the conditional premise tells us that the consequent will also be so. This moment of fulfillment is exactly what the second, declarative, premise of the MP argument gives. The relation between the conditional premise and the categorical premise can be understood as one of 'instantiation' or 'manifesting' (of the conditional relationship). The conditional gives the habitual connection between antecedent and consequent; the simple statement tells the subject that the topic is now just such an occasion in which the habitual connection applies. Making the inference is a matter of reiterating the given relation, except as actual and not as hypothetical. Note however that the discourse is still all hypothetical in the sense that it is about a fictional person, Thembi, whom neither of us know, so we are hypothetically referring to her.

Habitual conditionals are to be contrasted with, for instance, a generic conditional which expresses true and unchanging, law-like relationships, such as the following (from Celce-Murcia and Larsen-Freeman 1983)

If you boil water, it vaporizes.

With such conditionals we might expect a different relation between the first and second premises. So note this aspect of the explanation would lose purchase with a change in the materials used. Presenting subjects with tasks such

If you boil water, it vaporizes. You boil water. What happens?

might well generate different response profiles across different groups.

---

[19] Note that Haiman arrived at his proposal on the basis of extensive typological study; it is not an artifact of English usage. Haiman himself considers two definitions of the pair of terms 'topic' and 'comment': (i) distinguishes what the speaker is talking about (topic) from what he says about it (comment) and (ii) distinguishes given or old information (topic) from new information (comment). He argues for the second definition. Here we can afford to remain uncommitted to a particular definition of topic and just make use of Haiman's own characterisation of the 'topic' nature of a conditional clause, because whether we label the discourse function of conditional clauses 'topic' or not is irrelevant. What is relevant is what they *do*.

[20] Note that in explaining the full suppression effect task data, in which some conditions have two conditional premises, you would be able to make use of this idea provided you have a story about subordination or co-ordination of the first and second antecedent clauses.

Two features of conditionals do the work in the explanation offered here as to why conditional premises are easy to reason with: firstly, their classification as habitual conditionals, and secondly, their discourse function as topic providers. The topic-providing character explains why starting with a conditional premise is more natural than for example starting with a universally quantified premise, while the habitual character of the conditional projects onto its relation to the second premise. Most generally this explanation provides a perspective on inferences which grounds them in constraints on information flow and coherence of discourse, as manifested in typical language use.

5 Summary and Discussion

In the first section of the paper we presented one quantitative aspect of the findings from an interview study of reasoning behaviour. Corroborating previous research, we found that unschooled subjects find logical reasoning tasks in general more difficult than schooled subjects. But a novel finding was that both schooled and unschooled subjects found reasoning easier with conditional premises than with quantified premises. Consequently, with the aim of explaining this finding, we investigated the use of both quantified and conditional sentences in spontaneous spoken language. It was found that quantified sentences are commonly anaphoric or deictic to previous discourse or to the context. We argued that they chiefly used with a pre-determined domain, and that they are thus unsuitable bases from which to draw inferences about new instances. With regard to conditional sentences, we found that the most commonly usage of conditionals mimics how they are used in reasoning tasks. As for their discourse function, we employed the broadly-accepted idea that conditional clauses serve the discourse function of providing 'topics', and argued that this makes conditional sentences natural bases for drawing inferences. We used these findings to explain some qualitative aspects of the data.

So, do 'mistakes of reason' rule our minds? The argument presented here suggests that sometimes the mistakes of reason are on the behalf of the experimenter. In the case of logical reasoning tasks, negative results are at least partly a result of inadequate semantic analysis of the material used in reasoning studies. When we do pay attention to the semantic structure of test material, we see that there are all kinds of unresolved semantic factors which complicate interpretation of the test material for subjects. If subjects are not well-versed in the discourse of tests, in which language use is often contrived, their reasoning is even more susceptible to influence from these unresolved semantic factors. The intended interpretation of the premises in a reasoning task is not just a function of the subject's ability to reason to the *intended* interpretation, but also of the relation of this to the *usual* interpretation of the premise materials in everyday language. What constitutes a 'mistake of reason' is thus only partly a subject's fault.

References

Bobaljik, J. D. (2002). Floating quantifiers: Handle with care. In L. L. S. Cheng & R. P. E. Sybesma (Eds.), *The second glot international state-of-the-article book: The latest in linguistics*. Mouton de Gruyter.

Brisson, C. (2003). Plurals, all and the nonuniformity of collective predication. *Linguistics and Philosophy, 26*(2), 129–184.

Byrne, R. (1989). Suppressing valid inferences with conditionals. *Cognition, 31*(1), 61–83.

Celce-Murcia, M., & Larsen-Freeman, D. (1983). *The grammar book: An ESL/EFL teacher's course*. Rowley, Massachusetts: Newbury House.

Clark, H. H., & Bangerter, A. (2004). Changing ideas about reference. In I. A. Noveck & D. Sperber (Eds.), *Experimental pragmatics*. Palgrave Macmillan.

Comrie, B. (1986). A typology of conditionals. In E. C. Traugott, A. ter Meulen, J. S. Reilly, & C. A. Ferguson (Eds.), *On conditionals*. Cambridge: Cambridge University Press.

Counihan, M. (2008). *Looking for logic in all the wrong places*. Ph. D. thesis, University of Amsterdam.

Ford, C., & Thompson, S. A. (1986). Conditionals in discourse: A text-based study from english. In E. C. Traugott, A. ter Meulen, J. S. Reilly, & C. A. Ferguson (Eds.), *On conditionals*. Cambridge: Cambridge University Press.

Goodman, N. (1947). The problem of counterfactual conditionals. *The Journal of Philosophy, 44*(5), 113–128.

Haan, G. (2007). How illiterates interpret syllogistic problems. Masters thesis, University of Amsterdam.

Haiman, J. (1978). Conditionals are topics. *Language, 54*, 564–589.

Henle, M. (1962). On the relation between logic and thinking. *Psychological Review, 69*, 366–378.

Hwang, M. (1979). A semantic and syntactic analysis of if-conditionals. Masters thesis, Los Angeles: University of California.

Lechler, A. (2004). Interpretation of conditionals in the suppression task. Masters thesis, Edinburgh University.

Luria, A. (1976). *Cognitive development: Its cultural and social foundations*. Cambridge, Massachusetts: Harvard University Press.

Oakhill, J., Johnson-Laird, P., Garnham, A. (1989). Believability and syllogistic reasoning. *Cognition, 31*(2), 117–140.

Partee, B. (1985). Situations, worlds and contexts. *Linguistics and Philosophy, 8*, 53–58.

Pelletier, F., & Asher, N. (1997). Generics and defaults. In J. van Benthem & A. ter Meulen (Eds.), *Handbook of logic and language*. Elsevier.

Piattelli-Palmarini, M. (1994). *Inevitable Illusions: How mistakes of reason rule our minds*. New York: Wiley.

Politzer, G. (2004). Reasoning, judgement and pragmatics. In I. A. Noveck & D. Sperber (Eds.), *Experimental pragmatics*. New York: Palgrave MacMillan.

Politzer, G., & Macchi, L. (2002). Reasoning and pragmatics. *Mind and Society, 1*(1), 73–93.

Schwarzschild, R. (1996). *Pluralities*. Springer.

Scribner, S. (1997). *Mind and social practice: selected writings of Sylvia Scribner*. Cambridge: Cambridge University Press.

Sloman, S. (1998). Categorical inference is not a tree: The myth of inheritance hierarchies. *Cognitive Psychology, 35*, 1–33.

Sportiche, D. (1988). A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry, 19*(3), 425–449.

Stenning, K., & van Lambalgen, M. (2001). Semantics as a foundation for psychology: A case study of wason's selection task. *Journal of Logic, Language and Information, 10*, 273–317.

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

Stenning, K., & van Lambalgen, M. (2008). *Human reasoning and cognitive science*. Cambridge, MA: MIT Press.

Tsolwana, H. (1996). *New Xhosa phrase book and vocabulary*. J L van Schaik.

Westerståhl, D. W. (1985). Determiners and context sets. In J. van Benthem and A. ter Meulen (Eds.), *Generalized quantifiers in natural language*. Dordrecht: Foris Publications.