Title
Deductive Reasoning Competence: Are Rule-Based and Model-Based Methods Distinguishable in Principle?

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 19(0)

Author
Yule, Peter

Publication Date
1997

Peer reviewed

Deductive Reasoning Competence: Are Rule-Based and Model-Based Methods Distinguishable in Principle?

Peter Yule
Department of Psychology
Birkbeck College
University of London
Malet Street
London WC1E 7HX, U.K.
p.yule@psyc.bbk.ac.uk

Abstract

Much argument has been generated concerning the problem whether human deductive performance can best be viewed as rule-based (e.g. Rips) or model-based (e.g. Johnson-Laird). This paper argues that the distinction is ill-founded, and demonstrates that an ostensibly model-based syllogistic reasoning method can easily be implemented in a natural deduction calculus, which moreover makes fully explicit reference to the different possible interpretations of the premisses. More generally, it is unclear that other model-based methods cannot be given similar natural-deduction treatments, raising doubts about the distinguishability in principle of rule-based and model-based methods.

Introduction

The "rules versus models" debate in the psychology of reasoning has generated a considerable amount of argument over the best way to model human competence in deductive reasoning tasks such as syllogistic and propositional reasoning. The question bears a superficial resemblance to the distinction in logic between proof theory and model theory: proof-theoretic issues are usually addressed with reference to axiomatic or natural deduction systems, whereas model theory is concerned with individuals and sets of individuals. Thus the "rule theorists" such as Rips (1983) hold that human deductive competence is implemented in a theorem-proving system which uses computational analogues of natural-deduction rules (and a control module) to derive proofs. By contrast, Mental Models theory (see e.g. Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991) uses more concrete tableaux which represent individuals in a spatial array, and specifies procedures to transform these tableaux in ways which exhaust the (relevant) logical possibilities, without the explicit use of inference rules.

Superficially at least, then, Mental Models appear to be semantic whereas rule-based methods are syntactic. However, insofar as Mental Models theory specifies an effective procedure for its logical domain, it is also necessarily a proof system, and subject to the usual limits of computability and incompleteness on proof systems (see e.g. Boolos & Jeffrey, 1980). As a consequence, it must be possible to recast the Mental Models system in rule-based, proof-theoretic terms.

Of course, the rules versus models debate is primarily concerned, not with distinguishing rule-based and model-based systems in general, but with the empirical evaluation of particular systems of each type, with respect to human reasoning data. To the extent that the different systems make different empirical predictions, they can be compared for empirical adequacy. While this goal is reasonable in principle, it has proved difficult to achieve in practice: both theories can account for much of the data, and as Evans & Over (1997) and Roberts (1983) observe, both camps allow themselves sufficient free parameters that the question which is the best may be hard if not impossible to resolve using conventional psychological methodology.

Despite these worries, Evans & Over (1997) conclude that model-based methods are more plausible, since (they claim) they are better suited to modelling hypothetical reasoning, or the consideration of possible alternative situations. They make it clear that their idea of mental models is not necessarily equivalent to any of the versions proposed by Johnson-Laird and his co-workers, but rather they have a much more minimal account in mind: rationality in a deductive task consists in the consideration of multiple alternative situations which are consistent with the premisses. They hold that the question could be settled by analysing reasoners' proof protocols, and argue that if reasoners mentioned alternative interpretations of the premisses, this would be sufficient to establish that they were using a model-based method.

The aim of this paper is to show that this is no solution, since an ostensibly model-based method can be implemented as a rule-based, natural deduction method, albeit a slightly unconventional one. This method makes explicit the same information as do model-based methods, and handles the alternative interpretations of the premisses explicitly, in a manner closely analogous to the model-based method. Thus the rule-based method could underlie the protocols if the model-based method could. More generally, this paper suggests that similar treatments could be given for many of the domains used in the psychology of deductive reasoning, and seeks to emphasise the indistinguishability of rule-based and model-based methods in these domains.

Explicit and Implicit Models

One one interpretation, to say that reasoners consider all possible models of premisses is to say no more than that they have a grasp of logic — all sound deductive mechanisms, whether overtly proof-theoretic (e.g. axiomatic or natural de-

1	(1)	$(\exists x)(Ax \,\&\, Bx)$	A
2	(2)	$(\forall x)(Bx \longrightarrow Cx)$	A
3	(3)	$Aa \,\&\, Ba$	A
3	(4)	Aa	3 & E
3	(5)	Ba	3 & E
2	(6)	$Ba \longrightarrow Ca$	2 ∀ E
3,2	(7)	Ca	5,6 MPP
3,2	(8)	$Aa \,\&\, Ca$	4,7 & I
1,2	(9)	$(\exists x)(Ax \,\&\, Cx)$	1,3,8 ∃ E

Figure 1: A simple natural deduction proof of the syllogism *Some A are B, all B are C, so some A are C.*

duction systems) or model-based (e.g. Euler Circles or Mental Models), either implicitly or explicitly ensure that their conclusions hold in all logical models of their premisses — this is just what soundness means.

Of course, the proponents of model-based theories intend that explicit representations of models are used, but it is worth observing that most ostensibly model-based methods do not represent all logical models explicitly. In the domain of syllogistic reasoning, although Erickson's (1974) method using Euler Circles interprets each separate diagram as corresponding to an individual logical model, making this method maximally explicit, this approach leads to a combinatorial explosion, so more recent methods do not represent models this way. Mental Models theory (e.g. Johnson-Laird, 1983; Johnson-Laird & Byrne, 1991) adopts the expedient of condensing several logical models into one representation (a "Mental Model") by explicitly marking the distinction between necessary and merely possible individuals. Consequently the number of distinct representations to be considered is drastically reduced, to between one and three.

The same device can be used to make Euler Circles tractable (Stenning & Oberlander, 1995; Stenning & Yule, 1997 forthcoming) — if we mark regions of the diagram that correspond to necessary individuals, then the number of diagrams required for each premiss is reduced to one. Moreover, provided the diagrams are interpreted correctly, only one compound diagram for the premiss pair need be constructed. The Euler Circle method is summarised below, but it should already be clear that on a scale from minimum to maximum explicitness, neither Mental Models nor the Euler Circles method represent logical models maximally explicitly.

Nevertheless there seems to be a strong contrast between these and the minimally explicit natural deduction proof in Figure 1, which I assume is a typical example of a rule-based proof. The crucial step to notice in this proof is the application of *Modus Ponens* in line (7). Consideration of the truth table for $\longrightarrow$ reveals that $P \longrightarrow Q$ can be rewritten as $(P \,\&\, Q) \vee (\neg P \,\&\, Q) \vee (\neg P \,\&\, \neg Q)$ — this is known as *Canonical Disjunctive Normal Form* (see Lemmon, 1965) — and making this change is sufficient to form the basis of a natural deduction method which closely parallels the Euler

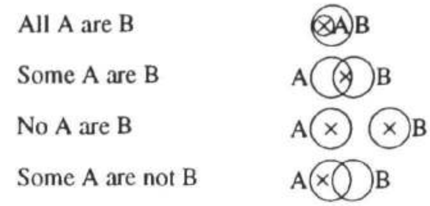

Figure 2: Euler diagrams for each premiss type. Regions representing necessary individuals are marked with a ×.

Circle method, and which will be demonstrated later.

The Method of Euler Circles

Since the Method has been described elsewhere (Stenning & Oberlander, 1995; Stenning & Yule, 1997 forthcoming), only a brief summary will be given here.

The method uses diagrams composed of circles to represent the premisses of the syllogism (see Figure 2). Each circle represents the set denoted by a term (A or B), and each region in a diagram represents an individual whose existence is *consistent* with the truth of the premiss. Regions which represent individuals whose existence is *entailed* by the premiss are marked with an ×.

Any well-formed syllogism has one term that appears in both the premisses, the *middle term*, and two others which each appear in only one of the premisses, the *end terms*. To solve a syllogism, a compound diagram is constructed from the two premiss diagrams by identifying the circles corresponding to the middle term, while preserving the topological relationships of the premiss diagrams. For some problems, this may be achieved in different ways (these correspond rather closely to the multiple-model problems in Mental Models theory), but in these cases the diagram with the maximum possible number of subregions should be constructed — it is sufficient to overlap the end-term circles if possible. Finally, if any ×-marked region has been bisected by the arc of another circle during this process, its mark is removed; otherwise, its mark remains in the compound diagram.

The semantics of the resulting diagram are the same as those of the premiss diagrams[1]: any marked regions in the compound diagram correspond to individuals whose existence is entailed by, and the rest represent individuals which are consistent with, the premisses taken together. Only if there is a necessary individual, is there a conventional quantified conclusion, although there are necessary individuals which do not form the basis of quantified conclusions because they are inexpressible as syllogistic sentences (Stenning & Oberlander, 1995).

Particular conclusions can be read off necessary individuals directly, by dropping the middle term and picking a positive end term as the subject of the conclusion, but for uni-

[1]Mental Models do not have this property

	Existential Implications	*CDNF Inferential Constraints*
All A B	$(\exists x)(Ax \& Bx)$	$\& \ (\forall y)((Ay \& By) \vee (\neg Ay \& By) \vee (\neg Ay \& \neg By))$
Some A B	$(\exists x)(Ax \& Bx)$	$\& \ (\forall y)((Ay \& By) \vee (\neg Ay \& By) \vee (Ay \& \neg By) \vee (\neg Ay \& \neg By))$
No A B	$((\exists x)(\neg Ax \& Bx) \& (\exists y)(Ay \& \neg By))$	$\& \ (\forall z)((\neg Az \& Bz) \vee (Az \& \neg Bz) \vee (\neg Az \& \neg Bz))$
Some A not B	$(\exists x)(Ax \& \neg Bx)$	$\& \ (\forall y)((Ay \& By) \vee (\neg Ay \& By) \vee (Ay \& \neg By) \vee (\neg Ay \& \neg By))$

Figure 3: Representing premisses in Monadic Predicate Calculus. Each premiss is represented as the conjunction of its *existential implications* and its *Canonical Disjunctive Normal Form (CDNF) inferential constraints.*

versal conclusions, a further condition applies: the marked region should be an unbroken circle corresponding to an end term, and this end term is the subject of the conclusion.

Implementation in Monadic Predicate Calculus

It will now be shown how the Euler Circle method can be implemented in Monadic Predicate Calculus (MPC) using natural deduction. Since space is short, only the derivation of necessary individuals is covered here; see Stenning & Yule (1997 forthcoming) for more on drawing quantified conclusions in a variety of implementations.

In order to understand how the implementation works, it should be recalled that in the proof shown in Figure 1, the premisses had different roles: one of them (the particular premiss) was treated as an existential conjunction, $(\exists x)(Ax \ \& \ Bx)$, while the other was treated as a universally quantified conditional, $(\forall x)(Bx \longrightarrow Cx)$. *Modus Ponens* was then used, with the former premiss[2] providing the antecedent and the latter providing the conditional. Thus the role of the first premiss was to establish the existence of some individual, and the role of the second was to provide information which was used to make an inference about that individual. Similarly, in the Euler Circle implementation, we distinguish between the existential implications of a premiss, and the set of consistent individuals. The set of consistent individuals can be viewed as a set of inferential constraints; the universal premisses, having only three regions (including the background), place constraints on the total set of possible individuals, whereas the particular premisses, having four regions, impose no such constraints (see Figure 2). Also, we have seen that a conditional can be translated into a sentence with three disjuncts, its Canonical Disjunctive Normal Form representation, which as we will see, corresponds directly to the set of consistent individuals. The strategy in implementing the MPC version of the Euler Circle method is to explicitly represent both components of each premiss type — its existential implications and its inferential constraints.

In MPC, each premiss can be represented as the conjunction of two sentences (see Figure 3). The first sentence, a conjunction of existentially quantified conjunctions, expresses the existential implications of the premiss under its standard syllogistic interpretation, and corresponds to the set of $\times$-marked regions in the Euler Circle system. We can call it the *Existential Implications.* The second sentence is a universally quantified disjunction of conjunctions, and expresses the list of individuals which are consistent with the premisses, in Canonical Disjunctive Normal Form, which explicitly constrains the set of consistent individuals. We can refer to it as the *CDNF Inferential Constraints.*

In the proofs that follow, the natural deduction system is a variant on Lemmon's (1965) system, but for convenience and to shorten the proofs, some modifications have been made to the system. $\&$-elimination and $\vee$-elimination have been generalised to handle multiple conjuncts and disjuncts respectively. Also the *Reductio Ad Absurdum* rule has been replaced with *Ex Falso Quodlibet* (EFQ), which has the form $P \& \neg P \vdash Q$ — this sequent is provable in propositional calculus, so the soundness of the system is unaffected.

Ex. 1: A Valid Conclusion

Figure 4 shows a proof in Monadic Predicate Calculus for the example *Some A are B, All B are C*. This problem has the valid conclusion *Some A are C*, and we will prove the existence of the necessary individual from which this consequence follows, by analogy with the Euler Circle system. The strategy is to take the sentence which expresses the existential implications of one of the premisses, and instantiate it, and then the main body of the proof uses a large $\vee$-elimination on the inferential constraints of the other premiss, to show that the fully specified three-term individual exists.

Lines (1) and (2) are the assumptions, representations in our chosen form for the two premisses. In line (3) the sentence which expresses the existential implications of the first premiss is derived; since this is existentially quantified, (4) assumes the corresponding arbitrarily instantiated sentence. (5) and (6) derive the sentence expressing the inferential constraints of the second premiss, and instantiate it with the chosen arbitrary name.

We are now ready to perform the main inferential step, corresponding to the *Modus Ponens* step in the proof in Figure 1. In this case, however, it is necessary to use $\vee$-elimination to perform this function, so each of the disjuncts in (6) must be assumed in turn, as in lines (7), (11) and (15). The strategy is to find one of these disjuncts which has the same formula containing the B predicate as in (4) (the instantiated existential implications of first premiss), then to unify it with the formula in (4), to give a formula denoting the necessary individual (9). When the disjunct contains a contradictory formula

[2]Strictly, an *instantiation* of the former premiss was used to provide the antecedent.

1	(1)	$(\exists x)(Ax\&Bx)\ \&\ (\forall y)((Ay\&By) \vee (\neg Ay\&By) \vee (Ay\&\neg By) \vee (\neg Ay\&\neg By))$	A
2	(2)	$(\exists x)(Bx\&Cx)\ \&\ (\forall y)((By\&Cy) \vee (\neg By\&Cy) \vee (\neg By\&\neg Cy))$	A
1	(3)	$(\exists x)(Ax\&Bx)$	1 &E
4	(4)	$(Aa\&Ba)$	A
2	(5)	$(\forall y)((By\&Cy) \vee (\neg By\&Cy) \vee (\neg By\&\neg Cy))$	2 &E
2	(6)	$(Ba\&Ca) \vee (\neg Ba\&Ca) \vee (\neg Ba\&\neg Ca)$	5 ∀E
7	(7)	$(Ba\&Ca)$	A
7	(8)	Ca	7 &E
4,7	(9)	$(Aa\&Ba\&Ca)$	4,8 &I
4	(10)	Ba	4 &E
11	(11)	$(\neg Ba\&Ca)$	A
11	(12)	$\neg Ba$	11 &E
4,11	(13)	$(Ba\&\neg Ba)$	10,12 &I
4,11	(14)	$(Aa\&Ba\&Ca)$	13 EFQ
15	(15)	$(\neg Ba\&\neg Ca)$	A
15	(16)	$\neg Ba$	15 &E
4,15	(17)	$(Ba\&\neg Ba)$	9,16 &I
4,15	(18)	$(Aa\&Ba\&Ca)$	17 EFQ
4,2	(19)	$(Aa\&Ba\&Ca)$	6,7,9,11,14,15,18 ∨E
1,2	(20)	$(\exists x)(Ax\&Bx\&Cx)$	3,4,19 ∃E

Figure 4: Drawing a valid individual conclusion from the premisses *Some A are B*, *All B are C*.

for the B term, then we derive a contradiction (as in (13) and (17)), and using *Ex Falso Quodlibet* we are then permitted to derive anything ($P\&\neg P \vdash Q$), so we derive the same formula as we did by unification, as in (14) and (18). Since each of the disjuncts in (5) entails the same conclusion, ∨-elimination is permitted, discharging the auxiliary assumptions (7), (11) and (15) so that the conclusion now follows from (2) and (4).

All that remains to be done is to show that the conclusion in fact follows from the original premisses. Since the name in (4) does not occur in (2), ∃-elimination is permitted, so in (20) the existentially quantified conclusion now follows from (1) and (2), completing the proof.

Ex. 2: Failure to Draw a Valid Conclusion

We have seen how to draw a valid conclusion; next we examine a case where a valid conclusion cannot be drawn.

Figure 5 shows the proof for the problem *All A are B*, *Some C are not B*, which has a valid conclusion *Some C are not A*, but is a multiple-model problem in terms of Mental Models theory. In this example we use the existential implications portion of the MPC representation for *All A are B*, and the inferential constraints portion of the representation for *Some C are not B*, and we cannot derive a necessary individual, since the existential implications sentence is unifiable with more than one disjunct of the inferential constraints sentence. The best we can do is to derive a disjunctive conclusion, which holds for each disjunct of the inferential constraints.

Lines (1)—(6) proceed analogously with Example 1, selecting one existential implications sentence and one inferential constraints sentence, and instantiating them with an arbitrary name. But in the main ∨-elimination, we cannot derive the same formula for each disjunct — note that in lines (9) and (14) we have different formulas, so that by line (19) we have reached an impasse. However, lines (21)—(23) show that for each disjunct we can derive the same *disjunctive* formula, and thus complete the ∨-elimination in line (24), but the final conclusion (25) does not specify a single necessary individual.

This outcome is the direct analogue of the removal of an ×-mark in a bisected region of an Euler Diagram. The valid conclusion *could* be derived by using the existential implications portion of the *Some... not* premiss with the inferential constraints portion of the *All* premiss. This is not illustrated here owing to space limitations, and can be left as an exercise for the interested reader.

The fact that a valid conclusion can be missed in this way demonstrates the importance of selecting the right premiss to supply the existential implications sentence; in general it is necessary to try both possible assignments in order to derive all necessary individuals. However, when only one premiss is particular, the existential implications of that premiss should be used with the inferential constraints of the universal sentence, since particular sentences never license conditional inferences.

Discussion

It should be clear that there is a high degree of correspondence between the graphical Euler Circle method and the natural deduction formulation; however there are some potentially important differences. In the Euler Circle method, owing to the specificity of graphical representations (Stenning & Oberlander, 1995), the attempt to find all the necessary individuals cannot avoid producing the side-effect that

1	(1)	$(\exists x)(Ax\&Bx)\ \&\ (\forall y)((Ay\&By) \vee (\neg Ay\&By) \vee (\neg Ay\&\neg By))$	A
2	(2)	$(\exists x)(Cx\&\neg Bx)\ \&$ $(\forall y)((Cy\&By) \vee (\neg Cy\&By) \vee (Cy\&\neg By) \vee (\neg Cy\&\neg By))$	A
1	(3)	$(\exists x)(Ax\&Bx)$	1 &E
4	(4)	$Aa\&Ba$	A
2	(5)	$(\forall y)((Cy\&By) \vee (\neg Cy\&By) \vee (Cy\&\neg By) \vee (\neg Cy\&\neg By))$	2 &E
2	(6)	$(Ca\&Ba) \vee (\neg Ca\&Ba) \vee (Ca\&\neg Ba) \vee (\neg Ca\&\neg Ba)$	5 ∀E
4	(7)	Ba	4 &E
8	(8)	$Ca\&Ba$	A
8	(9)	Ca	8 &E
4,8	(10)	$Aa\&Ba\&Ca$	4,9 &I
11	(11)	$\neg Ca\&Ba$	A
11	(12)	$\neg Ca$	11 &E
4,11	(13)	$Aa\&Ba\&\neg Ca$	4,12 &I
14	(14)	$Ca\&\neg Ba$	A
14	(15)	$\neg Ba$	14 &E
4,14	(16)	$Ba\&\neg Ba$	7,15 &I
17	(17)	$\neg Ca\&\neg Ba$	A
17	(18)	$\neg Ba$	17 &E
4,17	(19)	$Ba\&\neg Ba$	7,18 &I
4,8	(20)	$(Aa\&Ba\&Ca) \vee (Aa\&Ba\&\neg Ca)$	10 ∨I
4,11	(21)	$(Aa\&Ba\&Ca) \vee (Aa\&Ba\&\neg Ca)$	13 ∨I
4,14	(22)	$(Aa\&Ba\&Ca) \vee (Aa\&Ba\&\neg Ca)$	16 EFQ
4,17	(23)	$(Aa\&Ba\&Ca) \vee (Aa\&Ba\&\neg Ca)$	19 EFQ
4,2	(24)	$(Aa\&Ba\&Ca) \vee (Aa\&Ba\&\neg Ca)$	6,8,20,11,21,14,22,17,23 ∨E
1,2	(25)	$(\exists x)((Ax\&Bx\&Cx) \vee (Ax\&Bx\&\neg Cx))$	3,4,24 ∃ E

Figure 5: Failure to draw a valid conclusion from the premisses *All A are B*, *Some C are not B*

the complete set of consistent individuals (the unmarked regions) is also represented. By contrast, in the natural deduction method, these individuals are not represented, since the method is "focussed" on only the candidate necessary individuals. This makes no difference to its effectiveness, since the possible existence of these "irrelevant" individuals has no bearing on the validity of any conclusions. Mental Models theory exhibits a similar neglect of irrelevant individuals; in early versions (e.g. Johnson-Laird, 1983) several possible individuals may be summarised by a single row of the tableau, and in more recent versions (Johnson-Laird & Byrne, 1991) the model is never "fleshed out" to the extent that they become explicit. In this respect Mental Models theory is intermediate between the Euler Circles and natural deduction methods.

However, one apparent difference is the distinction in Mental Models theory between single-model and multiple-model problems. In Mental Models theory, when two Mental Models are treated as alternatives, it is as an argument against some specific conclusion, and in fact this is the basis for the individuation of Mental Models. Only the differences which are relevant to the truth or falsity of some given conclusion are made fully explicit. But it is exactly this type of alternation that is summarised when a conclusion is refuted in the present natural deduction system, and only a disjunction follows: the disjunction summarises the only salient difference between the two sets of logical models. It should be clear on this basis, *contra* Evans & Over (1997), that protocol analyses would be unlikely to be sufficient to distinguish the model-based method from this natural deduction method, if we assume that participants are most likely to mention the most salient differences between the cases under consideration.

Altough the natural deduction method in its present form is primarily intended as a sceptical argument, it is easy to sketch how it could be developed into a serious psychological model. Perhaps the most important developments would concern the production of appropriate error patterns, since it is sometimes argued that model-based methods provide a natural account of errors, especially invalid conclusions, in a way that rule-based methods cannot (Christoph Schlieder, personal communication). In the present method it is actually trivially easy to generate invalid conclusions — just assume that a reasoner might omit to consider all the disjuncts in the ∨-elimination. Failure to draw a valid conclusion is easily implemented in conventional rule-based implementations, such as Rips' (1983) system, and potentially in the present one, by making rule application probabilistic, but also, as we have seen in Example 2, in the present method failure to consider the existential implications of *both* premisses can lead to such errors of omission.

There is no need to assume that the rules in use would be implemented exactly as in the 'paper' version. For example, much of the proof structure (i.e. the ∨-elimination) could if necessary be parallelised, and treatment of contradictory cases could be changed in numerous ways. Hyperproof (Barwise & Etchemendy, 1994) provides a good example of the way a multiple-cases proof structure can be implemented without the need for much of the fine detail used in the present implementation: Hyperproof's ∨-elimination permits cases to be eliminated immediately when a contradiction is found, without the need for the *Ex Falso Quodlibet* rule used here, and disjunctions can be read off the remaining cases without the need for an explicit rule of ∨-introduction. Changing such details certainly would not make the method any less rule-based.

Although this paper has focussed on only the case of syllogistic reasoning, it is straightforward to extend the same approach to other logical fragments commonly studied in the reasoning literature. Propositional reasoning is certainly susceptible to such treatment, since any propositional formula can be rewritten in Canonical Disjunctive Normal Form (see Lemmon, 1965), so any model-based propositional reasoning method could be implemented using natural deduction with maximally explicit representation of models.

Conclusions

In view of the ease with which model-based methods can be implemented as rule-based methods, it is difficult to see what substance there is in the rules versus models debate. Evans & Over's (1997) approach to the issue, avoiding commitment to particular versions of Mental Models theory, owing to the difficulty of resolving the debate by conventional empirical means, inadvertently leads them to endorse model-based methods in general, leaving them open to the criticism that there is no distinction between rule-based and model-based methods in general.

If the debate can be resolved, it will be by virtue of one or other well-specified theory's empirical superiority. But ultimately any model can be implemented in numerous different ways, and it is features expressed at a relatively abstract level that serve to distinguish the class of empirically adequate models from the rest. Stenning & Yule (1986) have shown that a variety of diverse implementations of the Euler Circles method have similar empirical consequences, since more abstract logical features of the class of algorithms account for the empirical data. Thus imaginably, many implementationally distinct but logically similar methods could exist in the population, and the main psychological trends (such as the "multiple models" effect and the effect of Figure on term order in conclusions) would still emerge from the statistical aggregate, since they reflect important logical constraints on any good solution strategy. We can understand these higher-level constraints only by examination of a diverse range of alternatives; the danger is to take a single method too seriously, and thus to fail to understand why it works the way it does.

The problem of finding an appropriate level of abstraction at which to express theoretical constraints is not unique to the study of reasoning; in cognitive modelling generally, it is hard to separate 'mere' implementational detail from the theoretical commitments of a model (Cooper, Fox, Farringdon & Shallice, 1996). It is easy to be misled by the technicalities of implementing a model as a computer program, or even on paper, such that properties of these implementations come to dominate the psychological debate. Ironically perhaps, a better understanding of the reasons for behaviour may come from a more thorough examination of alternative models than has usually been conducted.

Acknowledgements

Thanks to Richard Cooper and John Fox for their comments on this paper, and to Keith Stenning. This research was supported by EPSRC grant GR/L03637.

References

Barwise, J. & Etchemendy, J. (1994). *Hyperproof.* CSLI Lecture Notes. Chicago University Press.

Boolos, G.S. & Jeffrey, R.C. (1980) *Computability and Logic (2nd edition).* Cambridge University Press.

Cooper, R., Fox, J., Farringdon, J. & Shallice, T. (1996). A Systematic Methodology for Cognitive Modelling. *Artificial Intelligence*, **85**, 3-44.

Erickson, J.R. (1974). A set analysis theory of behaviour in formal syllogistic reasoning tasks. In R.Solso (ed) *Theories in cognitive psychology: the Loyola Symposium.* Potomac: LEA.

Evans, J.St.B.T. & Over, D.E. (1997). Rationality in reasoning: the problem of deductive competence. *Current Psychology of Cognition*, **16 (1)**.

Johnson-Laird, P.N. (1983). *Mental Models.* Cambridge University Press.

Johnson-Laird, P.N. & Byrne, R. (1991). *Deduction.* Hove: LEA.

Lemmon, E.J. (1965) *Beginning Logic.* Walton-On-Thames: Nelson.

Rips, L.J. (1983). Cognitive Processes in Propositional Reasoning. *Psychological Review*, **90 (1)**, 38-71.

Roberts, M. J. (1993). Human Reasoning: Deduction Rules or Mental Models, or Both? *Quarterly Journal of Experimental Psychology*, **46A (4)**, 569-589.

Stenning, K. & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: logic and implementation. *Cognitive Science*, **19**, 97–140.

Stenning, K. & Yule, P. (1997 forthcoming). Logic and Implementation in Human Reasoning: A Syllogistic Illustration. *Cognitive Psychology.*