Reconciling probabilistic approaches and experimental pragmatics: The case of conditionals

Ira Noveck

Laboratoire sur le Langage le Cerveau et la Cognition (L2C2)
Lyon, France

In the old days (circa 1980’s), the reasoning world was pretty straightforward

The question was about representation.

Mental Logic

	Prémisses du schéma	Conclusion du schéma	Aussi connu comme
1.	Not not P	P	Double negation
2.	If P or Q then R ; P	R	Disjunctive Modus Ponens
3.	**P or Q; not P**	**Q**	**Disjunction elimination**
4.	Not both P and Q ; P	not-Q	Negated disjunction
5.	P or Q ; If P then R ; If Q then R	R	
6.	P or Q ; If P then R ; If Q then S	R or S	
7.	**If P then Q ; P**	**Q**	**Modus Ponens**
8.	P; Q	P and Q	Conjunction introduction
9.	P and Q	P	Conjunction elimination
10	P and (Q or R)	(P and Q) or (P and R)	
11	P and not-P	Incompatibility	Contradiction
12	P or Q ; Not P ; Not Q	Incompatibility	Contradiction

- These inferences are easy
- Applied in a programmatic order
- Provide people with inferences automatically.

<u>I tell you that on the blackboard.</u>

There is a C or T.
There is no T.
If there is a C then there is a G.
There is not both a G and a D.

And I ask you to evaluate as True or False:
There is a D.

Disjunction elimination	*C or T; no T// C*
Modus Ponens	$C \rightarrow G; C \,//\, G$
Negated conjunction	*~(G & D); G// ~D*
Contradiction	*D & ~D// Contradiction*

Mental Models

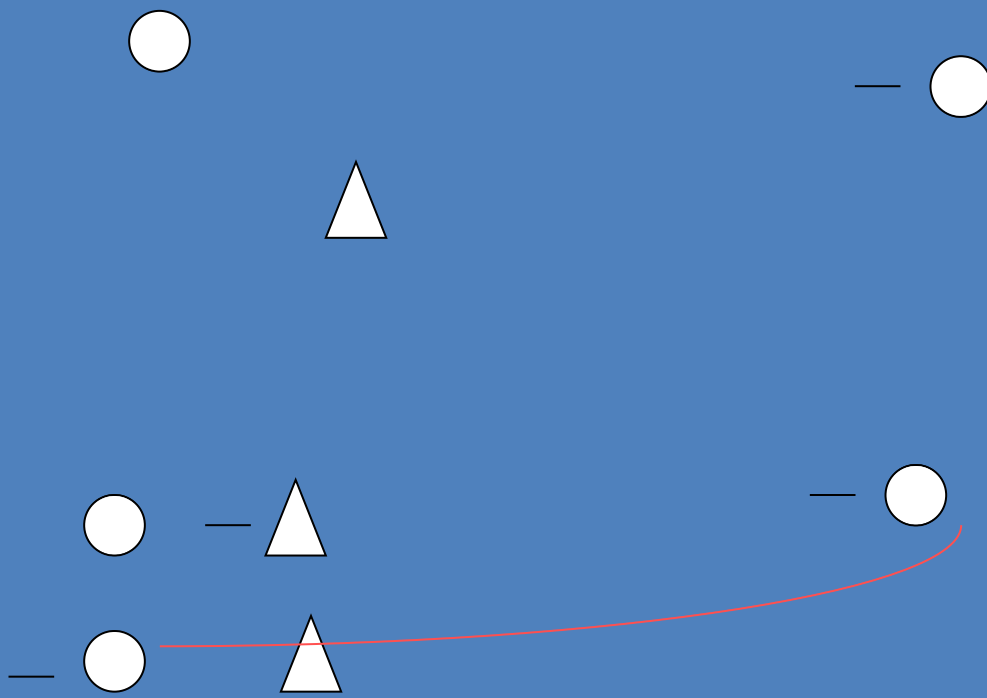
Connective	Initial	Explicit
p and q	*p* *q*	*p* *q*
p or q	*p* *q*	
If *p* then *q*	[*p*] *q* . . .	

Inclusive	Exclusive
p q	
p not-q	p not-q
not-p q	not-p q
Conditional	**Biconditional**
p q	p q
not-p not-q	not-p not-q
not-p q	

There is a circle or a triangle *There is no circle*

1.

2.



Conclusion:

3.

Propositional reasoning by mental models? Simple to refute in principle and in practice.

By O'Brien, David P.; Braine, Martin D. S.; Yang, Yingrui

Psychological Review, Vol 101(4), Oct 1994, 711-724.

Psychological Review
1994, Vol. 101, No. 4, 725–733

Propositional Reasoning by Model?

Luca Bonatti

Two theories of propositional deductive reasoning are considered: Johnson-Laird's mental models and Braine's mental logic. The model theory is said to account for practically all of the known phenomena of deductive propositional reasoning, offer a general theory of conditionals, account for the most important aspects of Braine's theory, and predict new phenomena that rule theories cannot explain. I argue that (a) the model theory is flawed in a way that is difficult to overcome, (b) conditionals are seriously misrepresented, (c) the algorithms proposed to implement it either allow invalid inferences or are psychologically useless, (d) Braine's theory accounts for all of the new phenomena worth considering, and (e) the model theory can predict Braine's results only at the cost of self-refutation. I conclude that the mental model theory of propositional reasoning offers no reason to reject the program of mental logic.

Psychological Review
1994, Vol. 101, No. 4, 734–739

0033-295X/94/$3.00

Why Models Rather Than Rules Give a Better Account of Propositional Reasoning: A Reply to Bonatti and to O'Brien, Braine, and Yang

P. N. Johnson-Laird, Ruth M. J. Byrne, and Walter Schaeken

O'Brien, Braine, and Yang argue that the mental model theory of propositional reasoning is easy to refute, and they report 3 experiments that they believe falsify the theory. In contrast, Bonatti argues that the model theory is too flexible to be falsified. We show that O'Brien et al.'s experiments do not refute the model theory and that Bonatti's claims are ill founded. Formal rule theories of propositional reasoning have 3 major weaknesses in comparison with the model theory: (a) They have no decision procedure; (b) they lack predictive power, providing no account of several robust phenomena (e.g., erroneous conclusions tend to be consistent with the premises); and (c) as a class of theories, they are difficult to refute experimentally.

There were three dominant reasoning research programs

Representational battle:	Mental Logic	Mental Models	Heuristics and Biases

“The picture that emerges from this focus on mechanistic explanation is of the cognitive systems as an assortment of apparently arbitrary mechanisms, subject to equally capricious limitations, with no apparent rationale or purpose.”
-- Chater & Oaksford, 1999

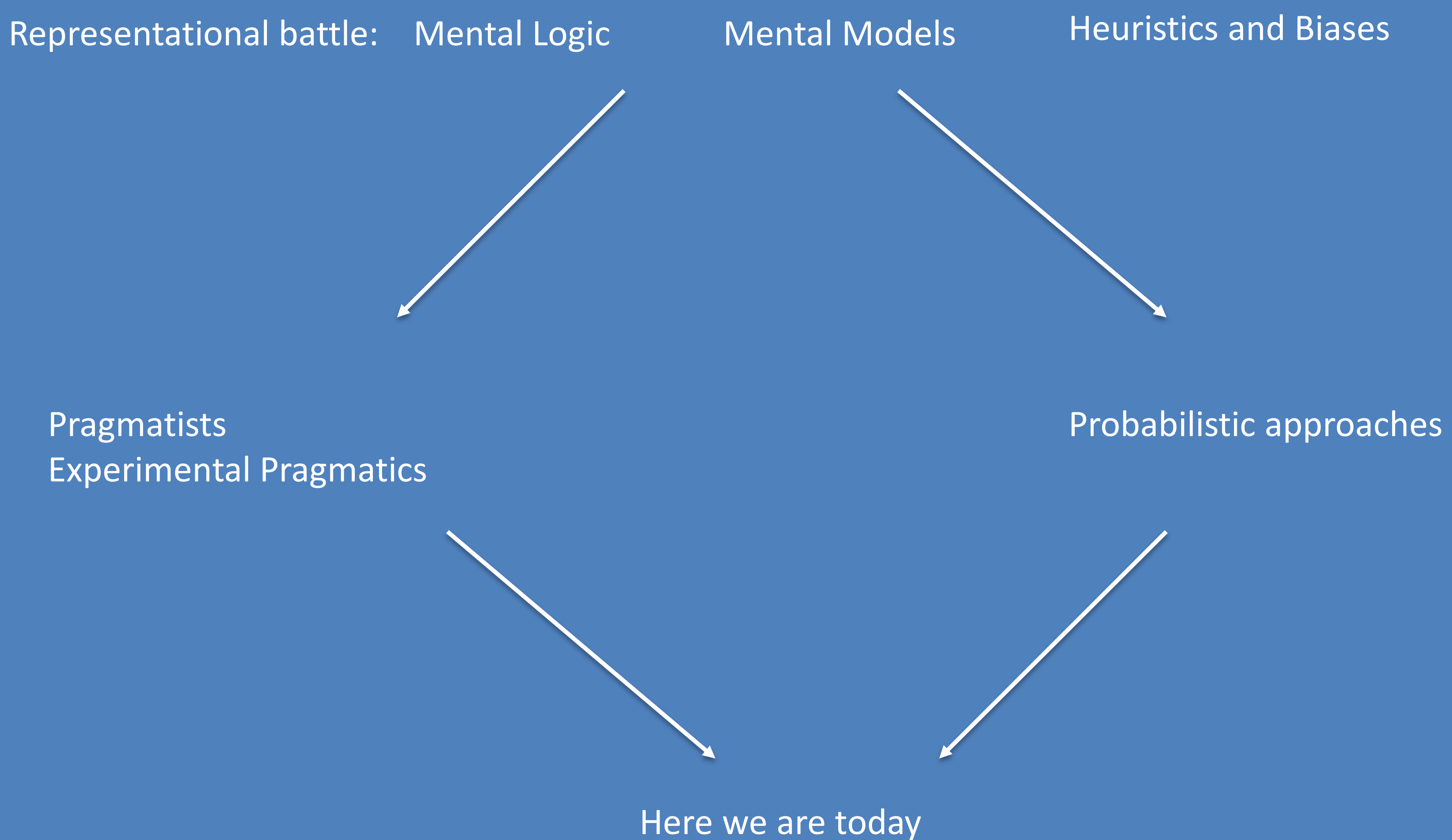
Representational battle: Mental Logic Mental Models Heuristics and Biases

Probabilistic approaches

Representational battle: Mental Logic Mental Models Heuristics and Biases

Pragmatists
Experimental Pragmatics

Probabilistic approaches


Representational battle:
Mental Logic
Mental Models
Heuristics and Biases
Pragmatists
Experimental Pragmatics
Probabilistic approaches
Here we are today

One can superimpose Marr’s 3 levels on this too:

Computational
Algorithmic
Implementational

Representational battle: Mental Logic Mental Models Heuristics and Biases

Pragmatists
Experimental Pragmatics
(algorithmic & implementational)

Probabilistic approaches
(computational)

Here we are today

Representational battle: Mental Logic Mental Models Heuristics and Biases

Pragmatists
Experimental Pragmatics

(Inside out)

Probabilistic approaches

(Top-down)

Here we are today

See Love, 2015

“Focusing solely on the environment, numerous theoretical constraints are discarded, such as those provided by physiology, neuroimaging, reaction time, heuristics and biases, and much of cognitive development.”

Now, during that earlier epoch, the focus was on conditionals. When the computational approach came around, focus was on the fact that

a) The material conditional is not a good model,
b) modus ponens is distinguishable from modus tollens,
c) conditional inferences are defeasible….
d) conditionals lead to inferences differently as a function of negatives…

…etc. (see Oaksford & Chater).

Not much has been devoted to knowing more about conditional processing itself.

Compare this to scalars, which explored in detail the frontier between semantics and pragmatics of key terms such as *some, or, might* etc. (though primarily *some*)

When compared to scalars (a very small set of terms) – which has now been the main example of analyzing the semantics/pragmatics frontier -- this same frontier wrt conditionals has gotten only a cursory glance from experimentalists…In other words, there is actually a paucity of information about conditional processing per se.

There is still much more to know about conditionals before modeling them.

Before modeling conditionals , the semantics of conditionals would benefit from an experimental pragmatic treatment.

To summarize very briefly, the empirical literature has established that pragmatic enrichments, as exemplified by scalars (e.g. enriching an utterance with *Some* as *Some but not all*), are costly to generate based on the following sort of findings:

i) children enrich less often than adults

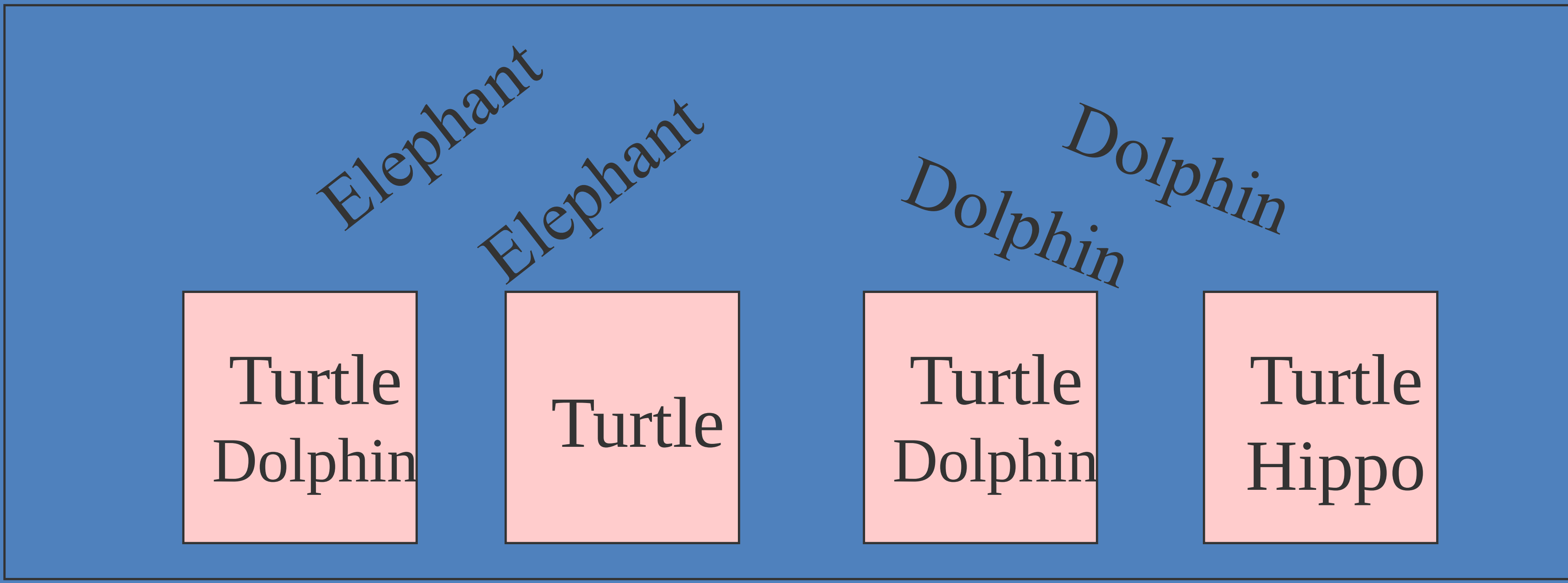
ii) scalars are *not* produced systematically among adults

iii) presence of scalar inference is linked with longer reading times (Bott & Noveck, 2004)

iv) sentence-processing and eye-tracking studies show that an enriched reading (of 'Some') does not immediately disambiguate between two options (Breheny et al., 2007; Huang & Snedeker, 2009)

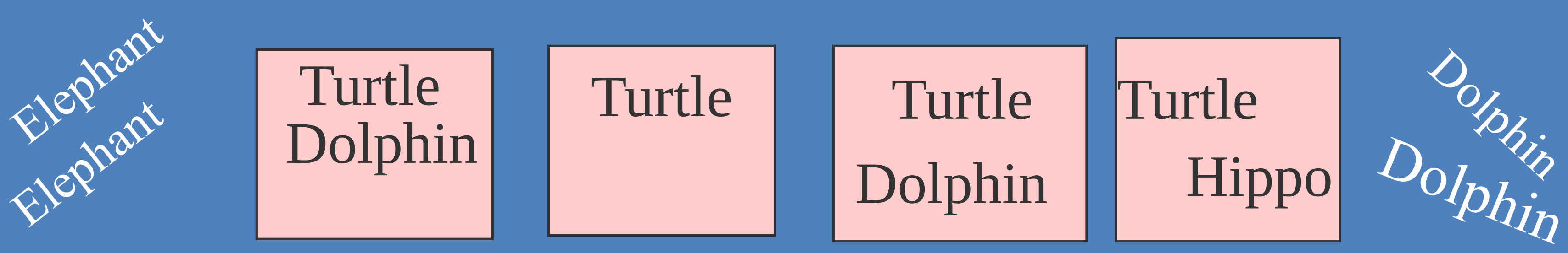
v) Theory of mind abilities arguably figure into scalar inference making (Nieuwland et al., 2010)

From Pouscoulous, Noveck, Politzer & Bastide, (LA, 2007)

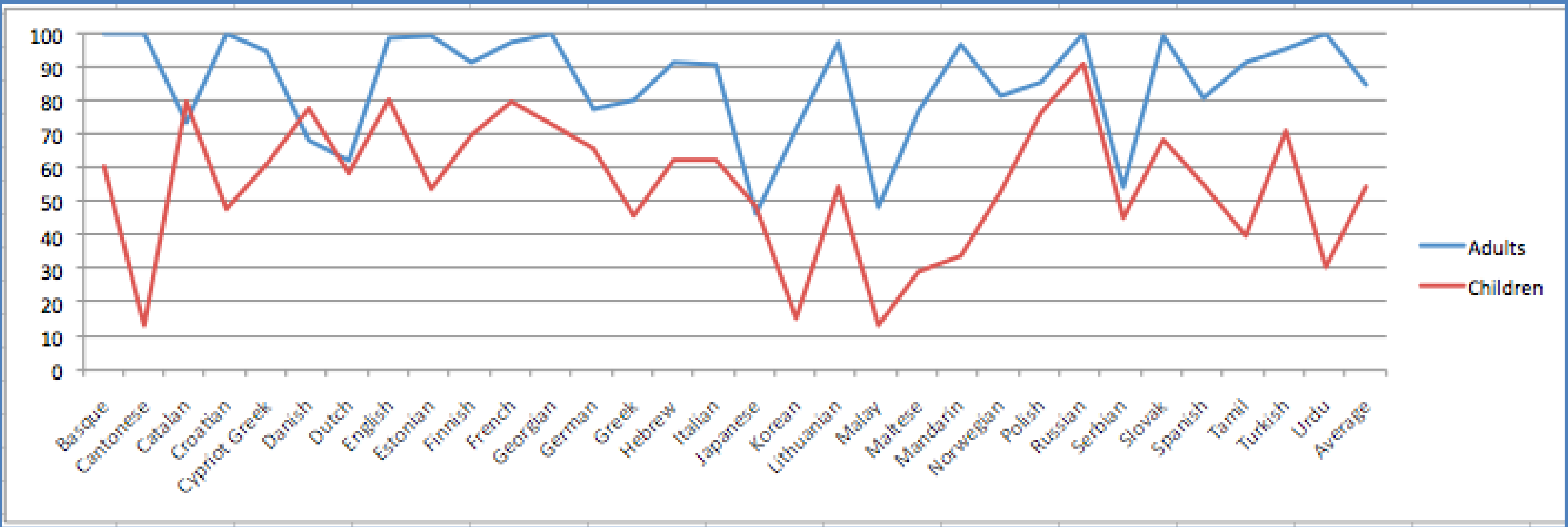


Percentage of children and adults who respond logically to 4 of the 10 critical statements presented in Experiment 1, including the one of central interest (in bold).

Statements	Logical response	Children 9-10 y.o. (N=23)	Adults (N=19)
All the turtles are in the boxes	True	100%	100%
Some turtles are in the boxes	**True**	**91%**	**53%**
Some turtles are not in the boxes	False	100%	100%
All of the dolphins are in the boxes	False	100%	100%



% rejecting underinformative utterances



Giant COST A33 project initiated by Uli;
Napoleon Katsos coordinated the investigation on scalars; Chris Cummins and I also played important roles…(just resubmitted)

What these examples have in common is a gain in informativity, where the pragmatic inference reduces the number of possible true cases.

Adult processing tells a similar story

Will processing data show that initial treatments of *Some* are compatible with the logical or a “default” pragmatic interpretation?

Following up on a developmental study showing children’s reticence to enrich scalar utterances (Noveck, 2001), consider a categorization task where the quantifier is weaker than it ought to be:

For example,
Some cows are mammals (Certaines vaches sont des mammifères).
Some trout are fish (Certaines truites sont des poissons).
Some parakeets are birds (Certains perroquets sont des oiseaux).

Of course...
All cows are mammals.
All trout are fish.
All parakeets are birds.

Thus, an interlocutor has a reason to be tempted to say that "Some cows are mammals" is *false* because it implicates that *Not All cows are mammals.*

Lewis Bott and I (Bott & Noveck, 2004, *JML*) ran 4 experiments based on the following paradigm:

54 items randomly presented by computer.

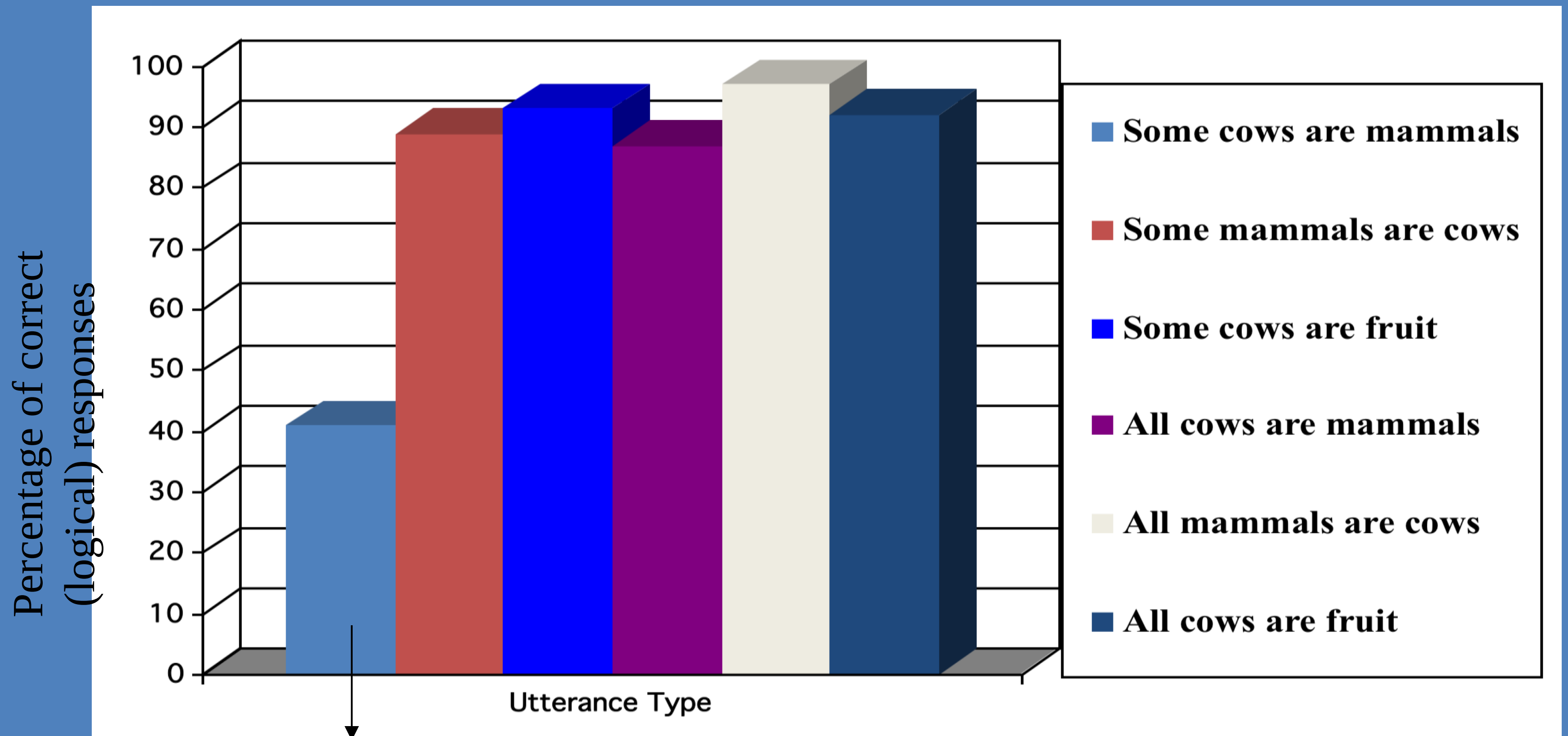
6 Categories -- mammals, fish, reptiles, shellfish,
birds, fruit

1) Some cows are mammals. * (True logically/False with enrichment)
2) Some mammals are cows. (True)
3) Some cows are insects. (False)
4) All cows are mammals. (True)
5) All mammals are cows. (False)
6) All cows are insects. (False)

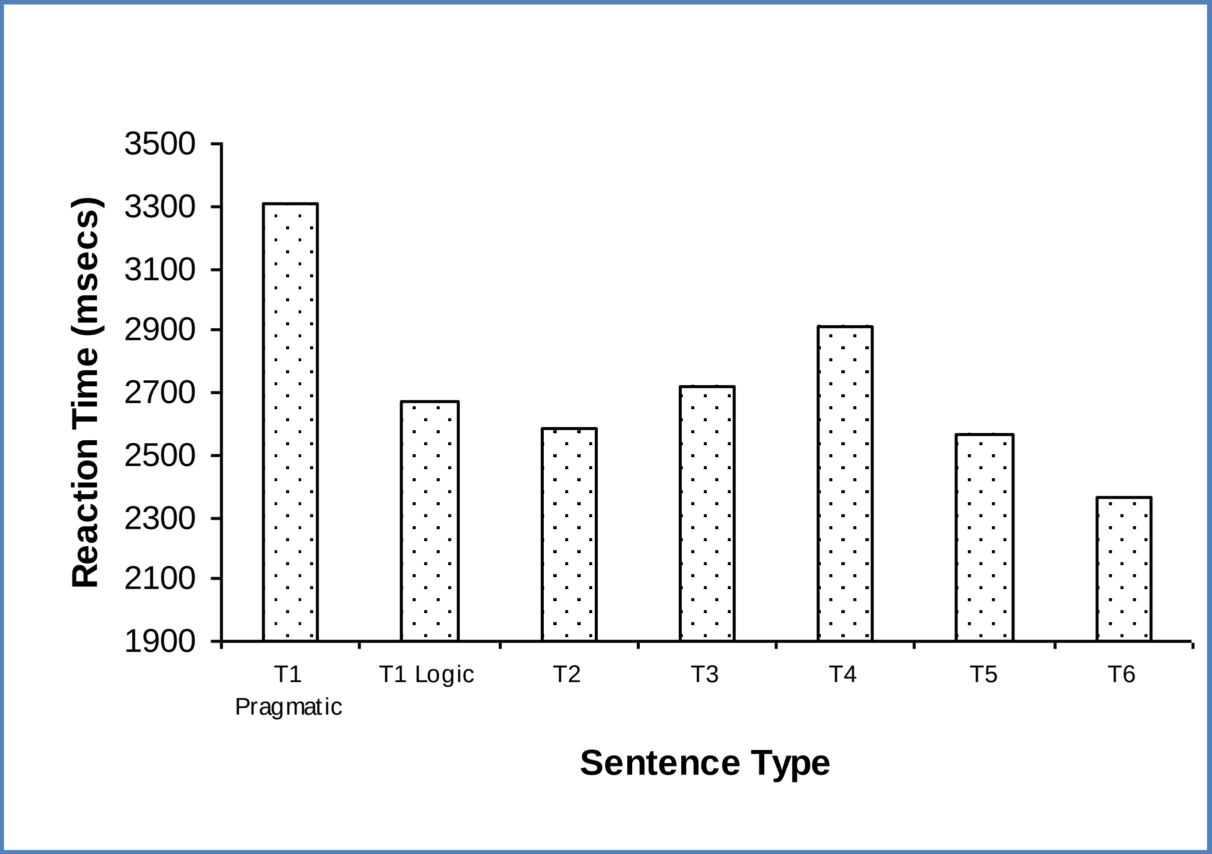
Experiment 3 : Whole sentences, no specific instruction

e.g. Some cows are mammals (Certaines vaches sont des mammifères)

Presented 54 items (6 categories; 9 per condition) and asked 32 participants to respond "True" or "False".



41% True and 59% False (in line with other data; Noveck, 2001)



T1) Some cows are mammals. (Logic - true / Pragmatic - false)
T2) Some mammals are cows. (True)
T3) Some cows are insects. (False)
T4) All cows are mammals. (True)
T5) All mammals are cows. (False)
T6) All cows are insects. (False)

Experiment 4: Controlling the amount of available effort.

The same task as before but with the following features

- One word at a time (to control uptake)

- Two lag times: A) Short amount of time (900 msecs)
  B) Long amount (3 seconds).

Percentage saying "true" to each of the Sentence types in Experiment 4. N= 45.

Utter.	Example	Short Lag (900 msecs)	Long lag (3000 msecs)	Logical response diff.	
T1	Some robins are birds	.72	.56	-.16	*$p < .001$*
T2	Some birds are robins (T)	.79	.79	.00	
T3	Some robins are fish (F)	.12	.09	+.03	
T4	All robins are birds (T)	.75	.82	+.07	
T5	All birds are robins (F)	.25	.16	+.09	
T6	All robins are fish (F)	.19	.12	+.07	

Bimodal distributions are generally the rule.

- Strategies remain largely consistent throughout a session.
- Situational factors such as quickened response times, or secondary tasks, can prevent deeper processing.
- Individual differences figure into prompting different interpretations.

These were all findings before computational accounts came into the picture

THINKING AND REASONING, 2001, 7 (2), 121–172

The processing of negations in conditional reasoning: A meta-analytic case study in mental model and/or mental logic theory

Walter J. Schroyens, Walter Schaeken, and Géry d'Ydewalle
University of Leuven, Belgium

Psychological Review
1991, Vol. 98, No. 2, 182–203

A Theory of *If*: A Lexical Entry, Reasoning Program, and Pragmatic Principles

Martin D. S. Braine
New York University

David P. O'Brien
Baruch College of the City University of New York

1) One cannot generalize from theoretical approaches to scalars…

On the one hand, enrichments of weak utterances appear to restrict meanings…

Informativeness

~~All~~ ***~~Excellent~~*** ***~~And~~***

Some ***Decent*** ***Or***

Conditionals on the other hand, do not….

Informativeness

If and only if p then q

If p then q

2) The conditional is…

- Well known for its pragmatic potential…

Geis & Zwicky (1971) in their seminal squib :

[This] regular association [linking *if* and the situations in which they arise]….asserts a connection between linguistic form and a tendency of the human mind ‘to perfect conditionals to biconditionals’ in words suggested to us by Lauri Kartunnen. This tendency is manifested in two classical logical fallacies, Affirming the Consequent (concluding X from $X \supset Y$ and Y) and Denying the Antecedent (concluding ~ Y from $X \supset Y$ and ~X).

There are many inference forms associated with conditionals

Two pragmatically justified ones

Affirmation of the Consequent

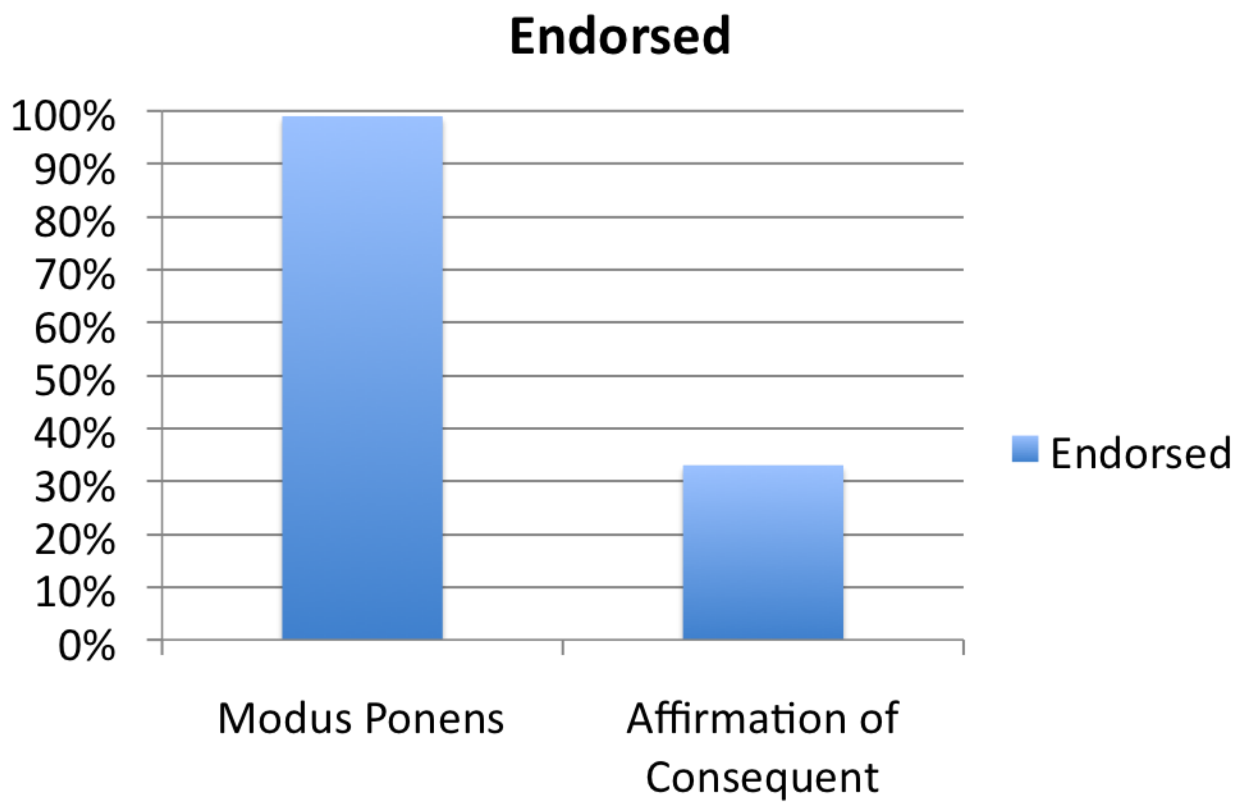
If Jean goes to the cinema, then he travels by bicycle	If p then q
Jean travels by bicycle.	q .
Therefore, he goes to the cinema	p

Denial of the Antecedent

If Jean goes to the cinema, then he travels by bicycle	If p then q
Jean does not go to the cinema	~p .
Therefore, he does not travel by bicycle	~q

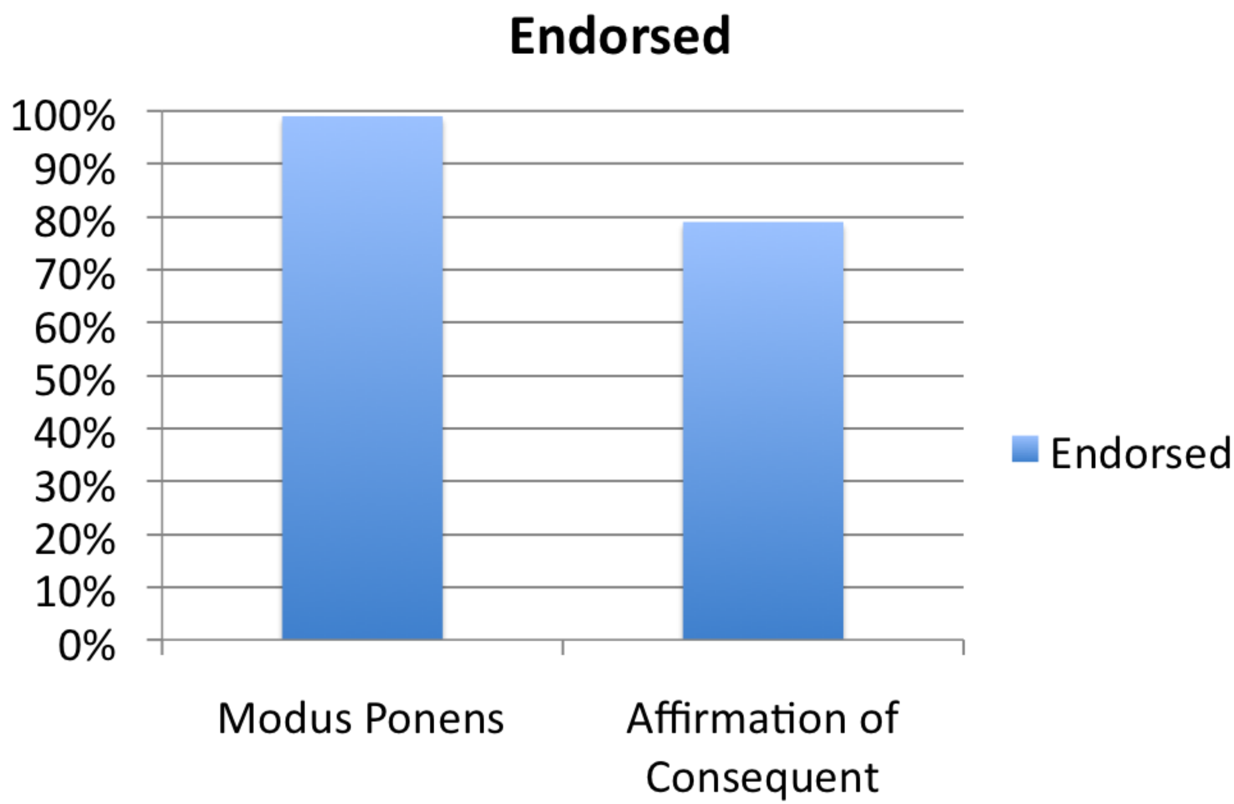
3) Invited inferences vary quite a bit (like with scalars)

Here's an example of a low rate of AC endorsements



e.g. Marcus & Rips, 1975

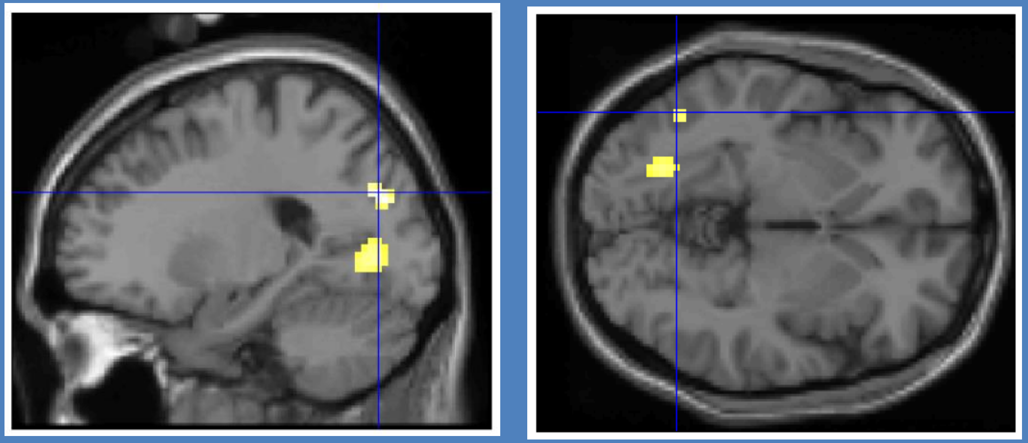
Highest rate of endorsement – 79%



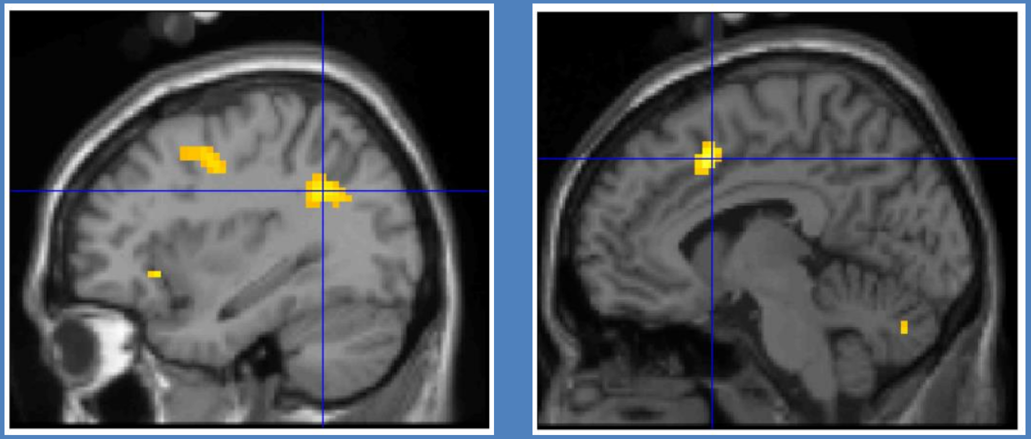
(e.g. Barrouillet et al.2000)

4) **An fMRI study (Noveck, Goel, & Smith, 2004) that focused mostly on the valid inference forms, Modus Ponens & Modus Tollens.**

Modus Ponens - Baseline

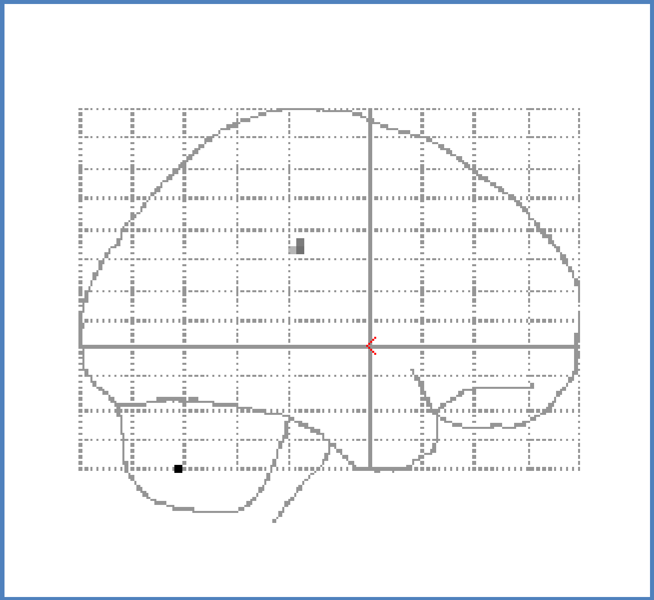


Modus Tollens - Baseline

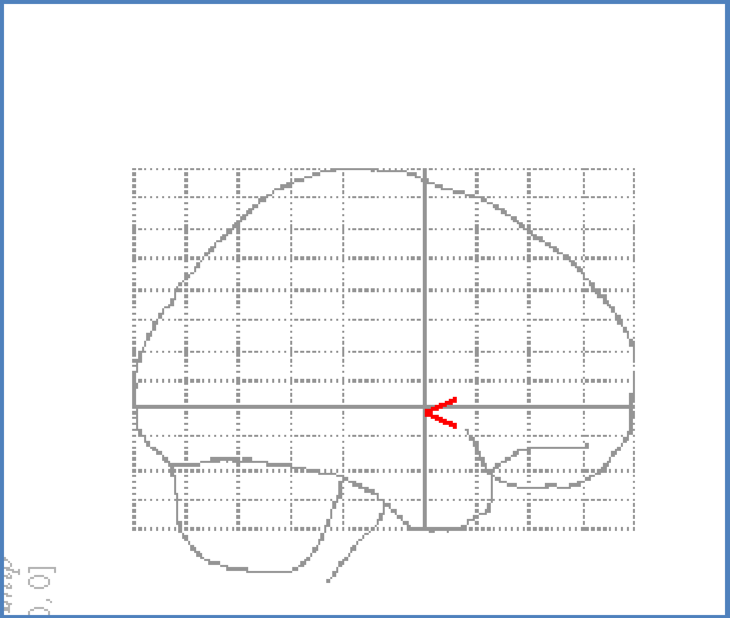


N=16

A) We captured little or no activity for these two forms in the conclusion.



Affirmation of the Consequent



Denial of the Antecedent

B) Rather high rates of correct responses (saying Inconclusive).

-- 69% said inconclusive to AC
-- 60% said inconclusive to DA

5) Existing accounts:

Break them down into two groups:

I. Predominant group: Those that encourage AC interpretations

1. Horn: If is *strengthened* directly into *iff :* likens it to cases such as “drink” for alcoholic beverage or how “I don’t believe that p” becomes “I believe that not-p”

2. Braine and Mental Logic: *If p then q* comes with an invited inference “If q then p.”

. . .

II. Those that discourage an AC interpretation

1. Markovits. If P then Q (at least for causals) prompts *If A then Q* as in *If the electricity goes out then school is cancelled* prompts other causes, such as snow.

If A then Q

If P then Q
Q

2. Von Fintel:

~~Q (no matter what)~~

If P then Q

If that is the case, then
The minor premise in

If P then Q
Q .
prompts a contradiction

Not Q

If there is activity to be captured, it is happening before the conclusion…

1) We used an Experimental Pragmatic strategy of isolating pragmatic effects in order to better understand the lexical contribution made by conditionals

2) We dealt only with Modus Ponens and Affirmation of the Consequent; we avoid negations.

3) We use temporal measures (reading time measures and EEG).

With an eye toward an ERP experiment …

Modus Ponens

If John watches television then he eats cookies.
He watches television.
+ He eats cookies.

MP': + He eats fries.

Affirmation of the consequent

If John takes a metro then he reads a newspaper.
He reads a newspaper.
+ He takes a metro.

AC': + He takes a bicycle.

<u>Details</u>:
In French
18 words total
RSVP: Presented one at a time
Response Options: Conclusion acceptable
Conclusion not acceptable

Initial attempt

Word number:

1	2	3	4	5	6	7	8	9	10
11	12	13	14						
15	16	17	18						

Si Jean regarde la tele alors il mange des gateaux.
Il regarde la tele.

MP Il mange des gateaux.

MP' Il mange des frites.

Si Jean regarde la tele alors il mange des gateaux.
Il mange des gateaux.

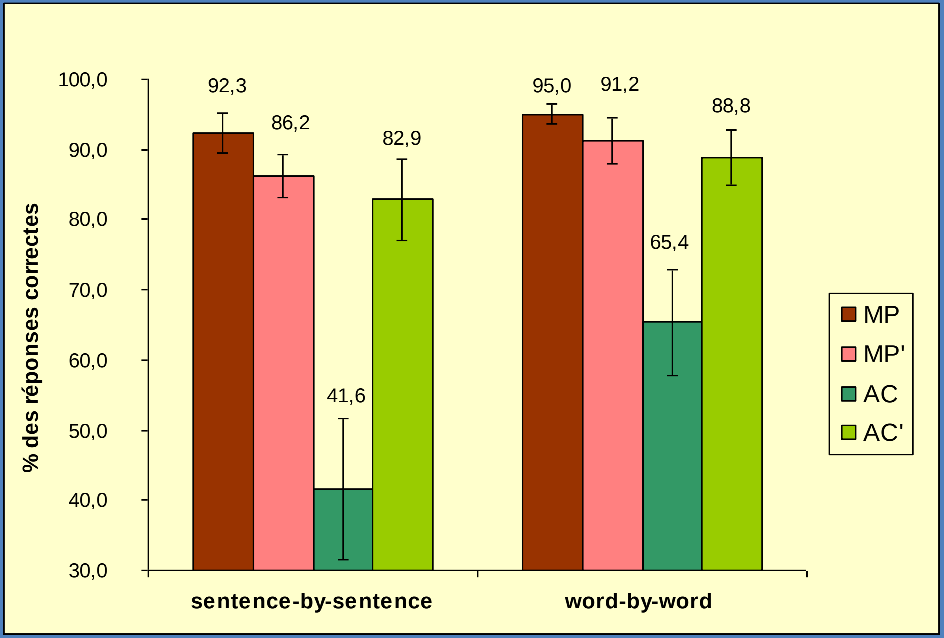
AC Il regarde la tele.

*AC' Il regarde la lune.

(A)

Experiment with the MP/MP'/AC/AC' paradigm:

Behavioral study



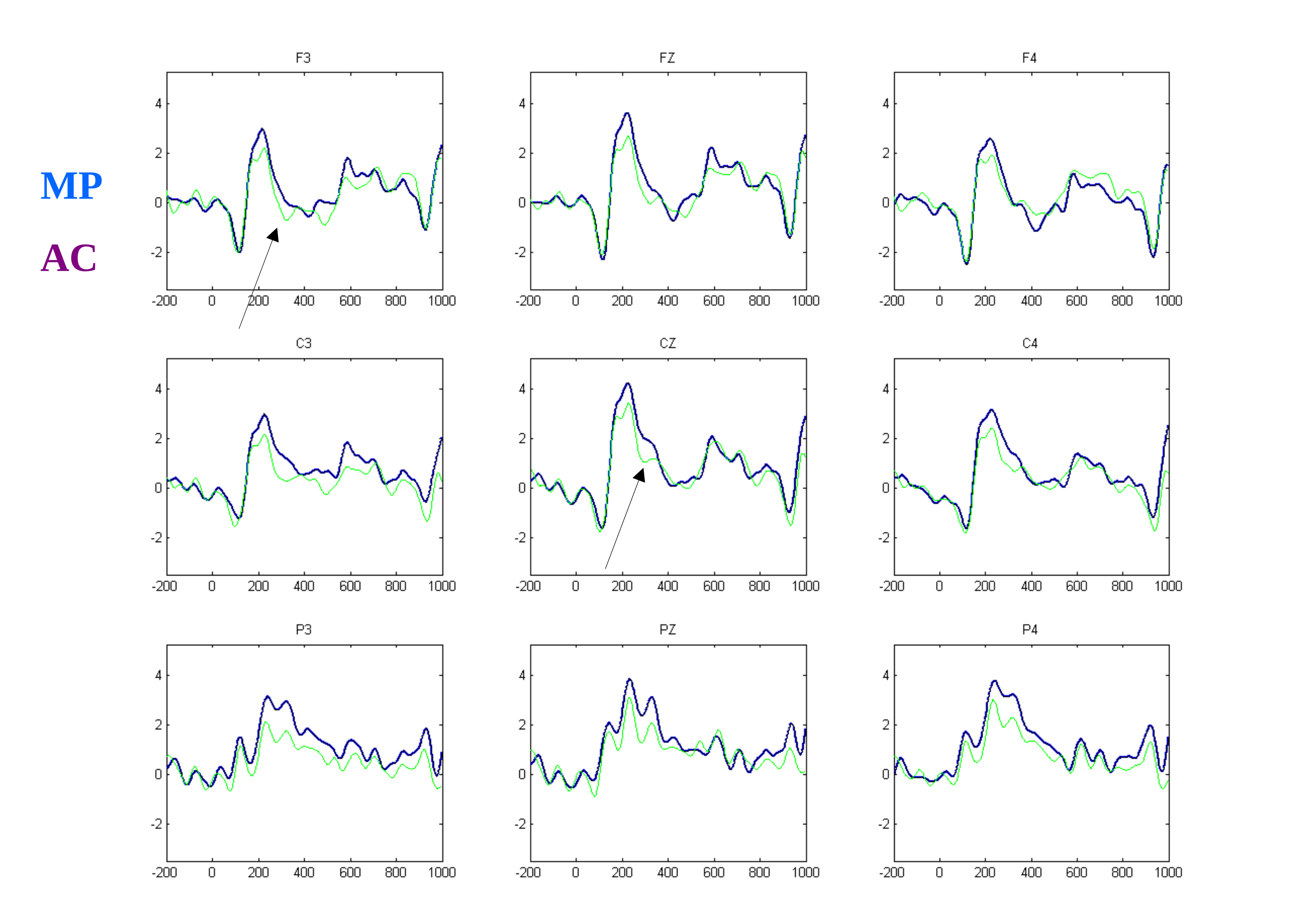
% <u>correct</u> responses

Slowing down is linked to an increased likelihood to respond "normatively" to AC syllogisms….

Van der Henst, Ciceron, Bujakowska & Noveck, 2006

(B)

Modus Ponens vs Affirmation of the Consequent : First word in the minor premise indicating whether MP or AC is called for (Word 12)



With respect to the Minor Premise of Modus Ponens, The AC Minor premise is linked to an N400, an indication of inconsistency

Difference between MP & AC : 200-300 ms

(c)

Neuropsychologia 47 (2009) 3125–3133

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

ELSEVIER

Using letters

What's behind an inference? An EEG study with conditional arguments

Mathilde Bonnefond, Jean-Baptiste Van der Henst*

Laboratoire Langage, Cerveau et Cognition (L2C2), Institut des Sciences Cognitives - CNRS, Uni

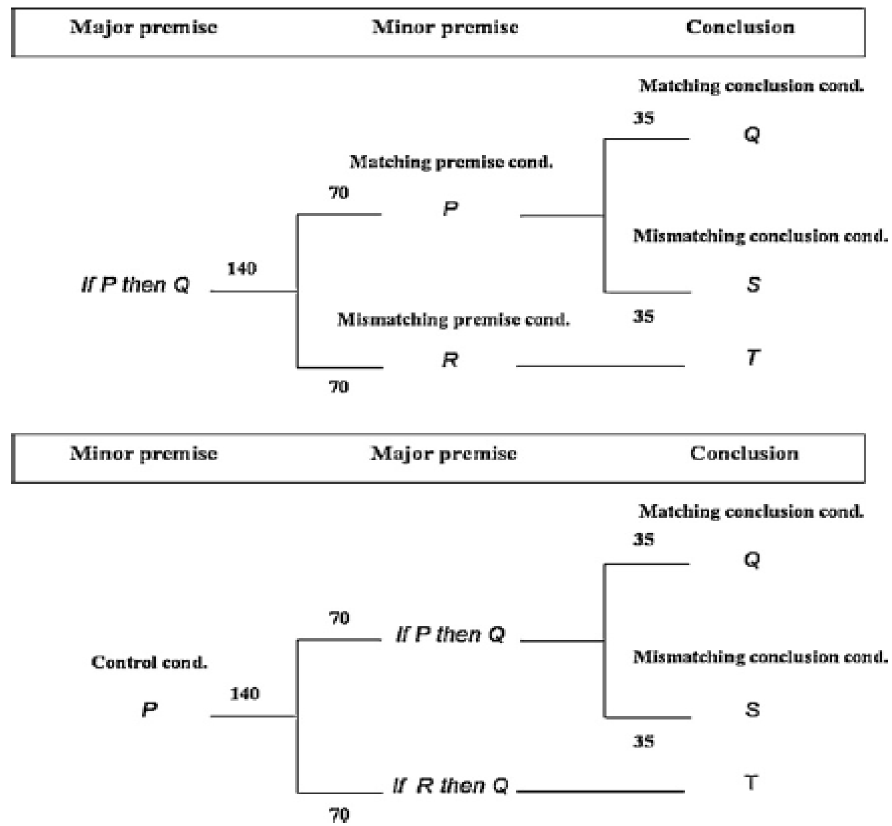
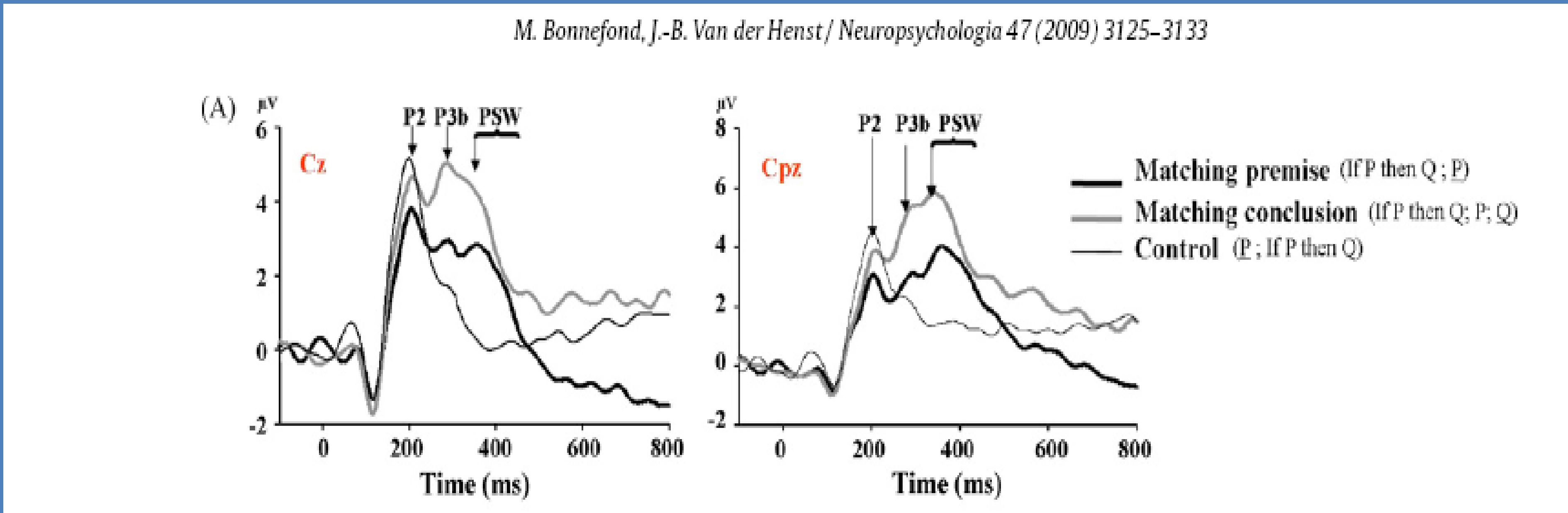


Fig. 1. Experimental design. Top: minor premise presented after the conditional statement; bottom: minor premise presented before the conditional statement. The numbers correspond to the number of trials in each condition.

If P then Q
P

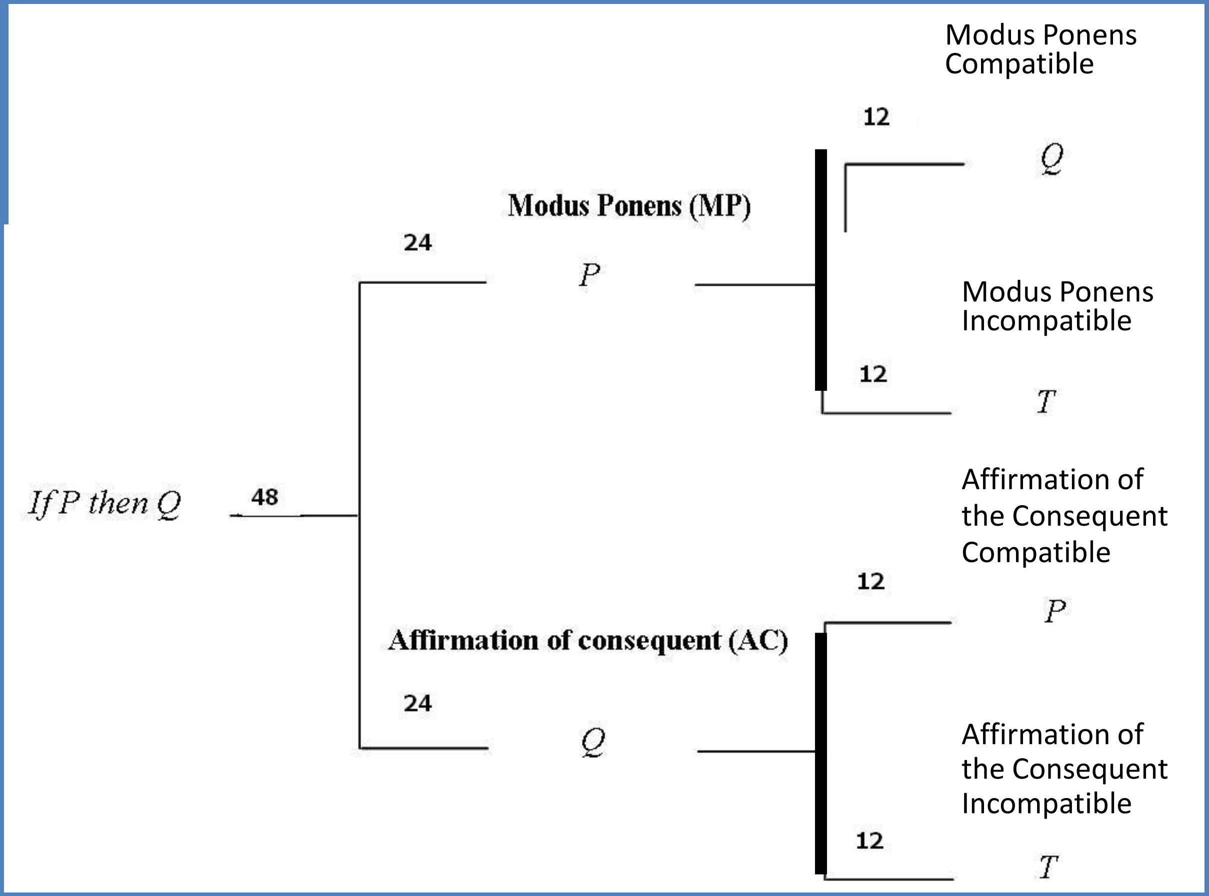
Versus

P
If P then Q



The minor premise of a Modus Ponens conditional elicits a P3b, an indication that an expectation was satisfied.

Our new, sparer paradigm…



Bonnefond et al. (JML, 2012)

Main interest: How does the second premise react,
even as we expect two sorts of responses…

That is, we expect two groups in response to AC : Endorsers and Rejectors

Affirmation of the Consequent

If P then Q
<u>Q .</u>
P

Accept as "logical"   Accept as "not-logical"

Modus Ponens

If P then Q
<u>P .</u>
Q

If the conditional comes with an enrichment that *facilitates* an AC inference, then the reaction to the Minor premise should be like that of Modus Ponens…

If the conditional comes with an enrichment that *discourages* an AC inference, then the reaction to its Minor premise should be slower than the one in Modus Ponens…

Two studies: One behavioral and one EEG

+

If there is a J then there is a N.

There is a J.

There is a N.

Styles (n) = Endorsers (15), Rejectors (13) & Mixed (9)

Premise 1 *If P then Q*

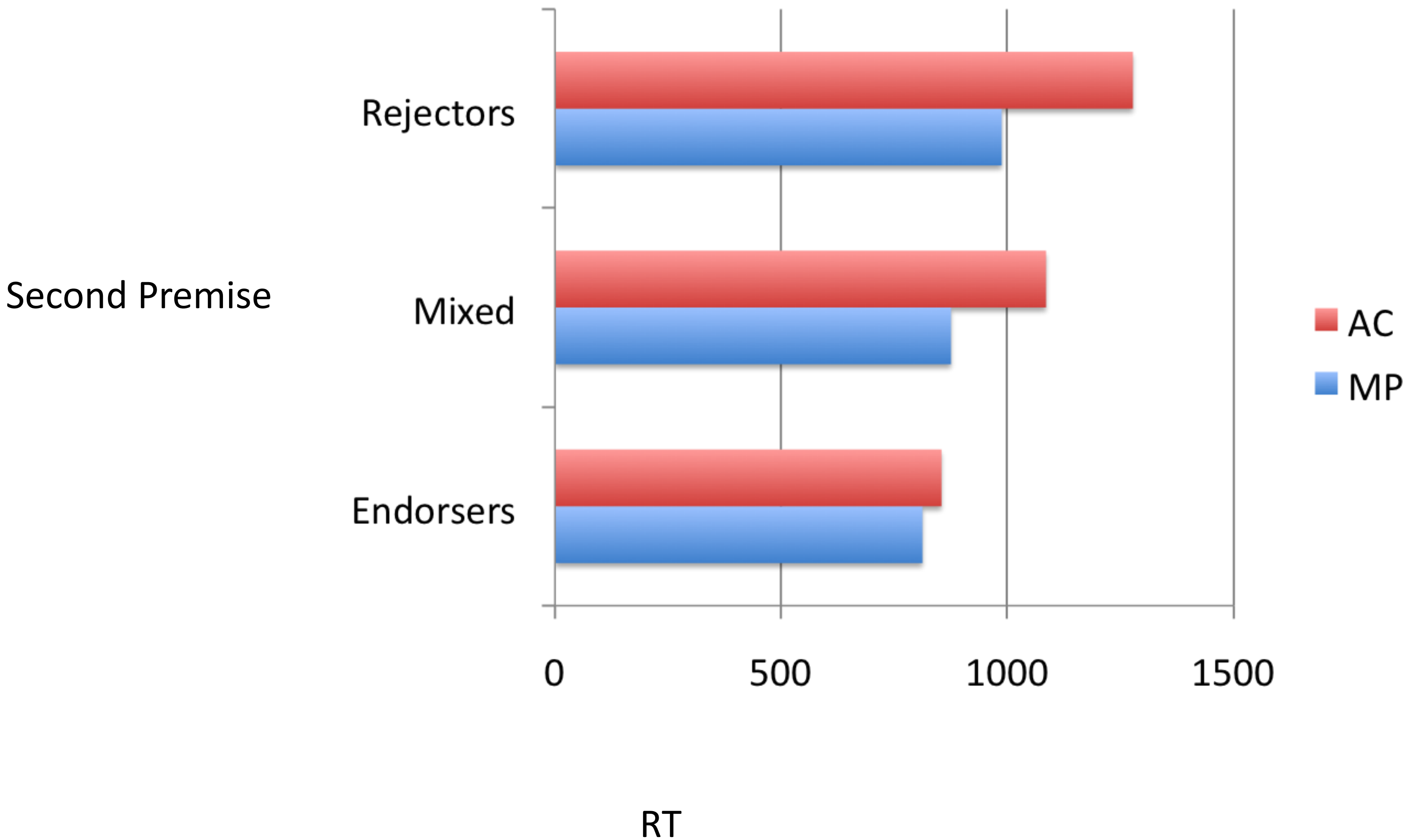
Endorsers	**1805**
Mixed	**2440**
Rejectors	**2552**

Premise 2

	P (MP)	*Q* (AC)
Endorsers	**812**	**854**
Mixed	**875**	**1075**
Rejectors	**987**	**1277**

Conclus'n

	Q MP Compatible	*R* MP Incompatible	*P* AC Compatible	*S* AC Incompatible
Endorsers	**847**	**965**	**835[e]**	**892**
Mixed	**843**	**925**	**1000[e]/888[r]**	**921**
Rejectors	**966**	**1131**	**1230[r]**	**1055**


Second Premise
Rejectors
Mixed
Endorsers
0
500
1000
1500
RT
AC
MP

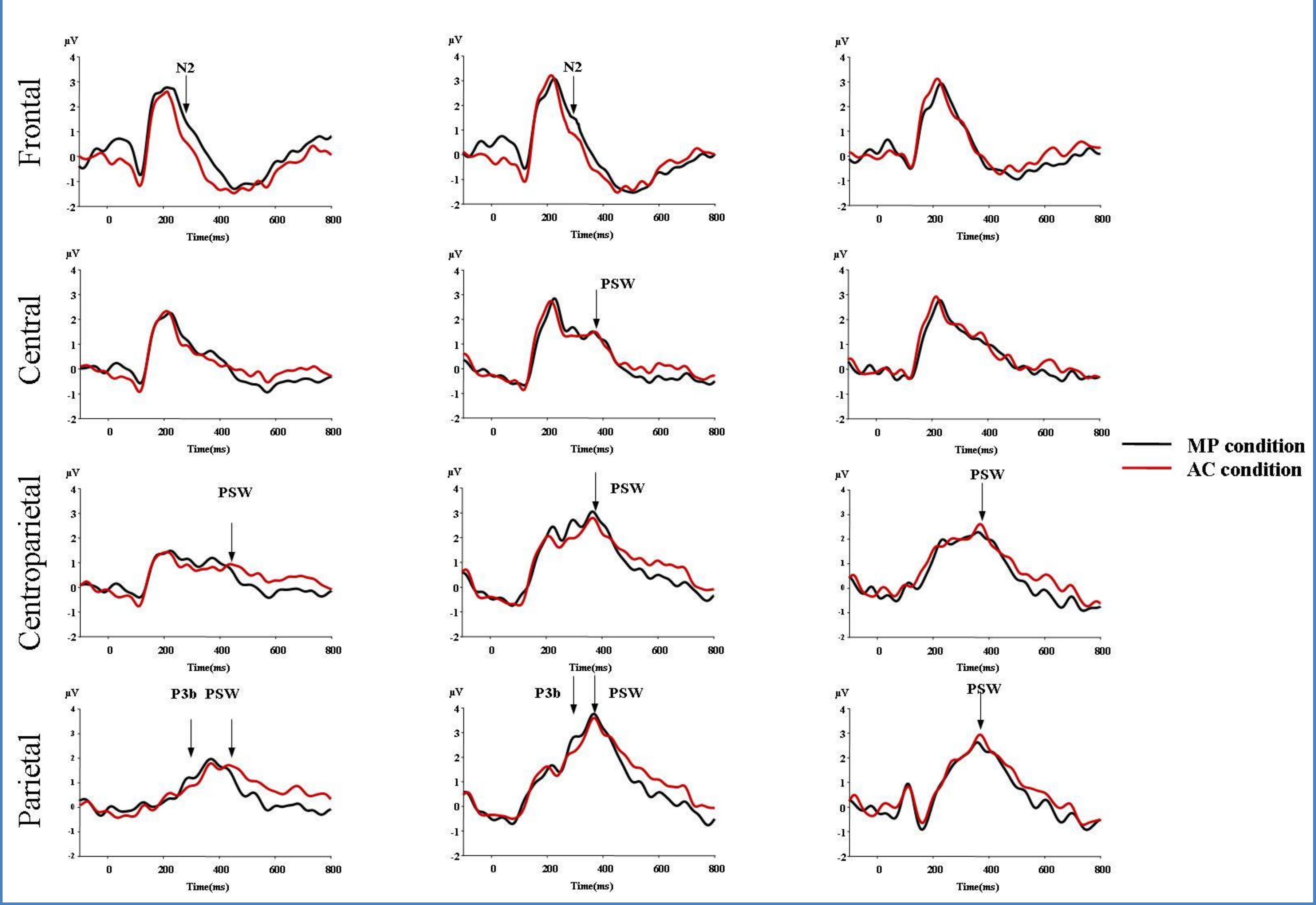
EEG study -- What to look for:

- Second Premise

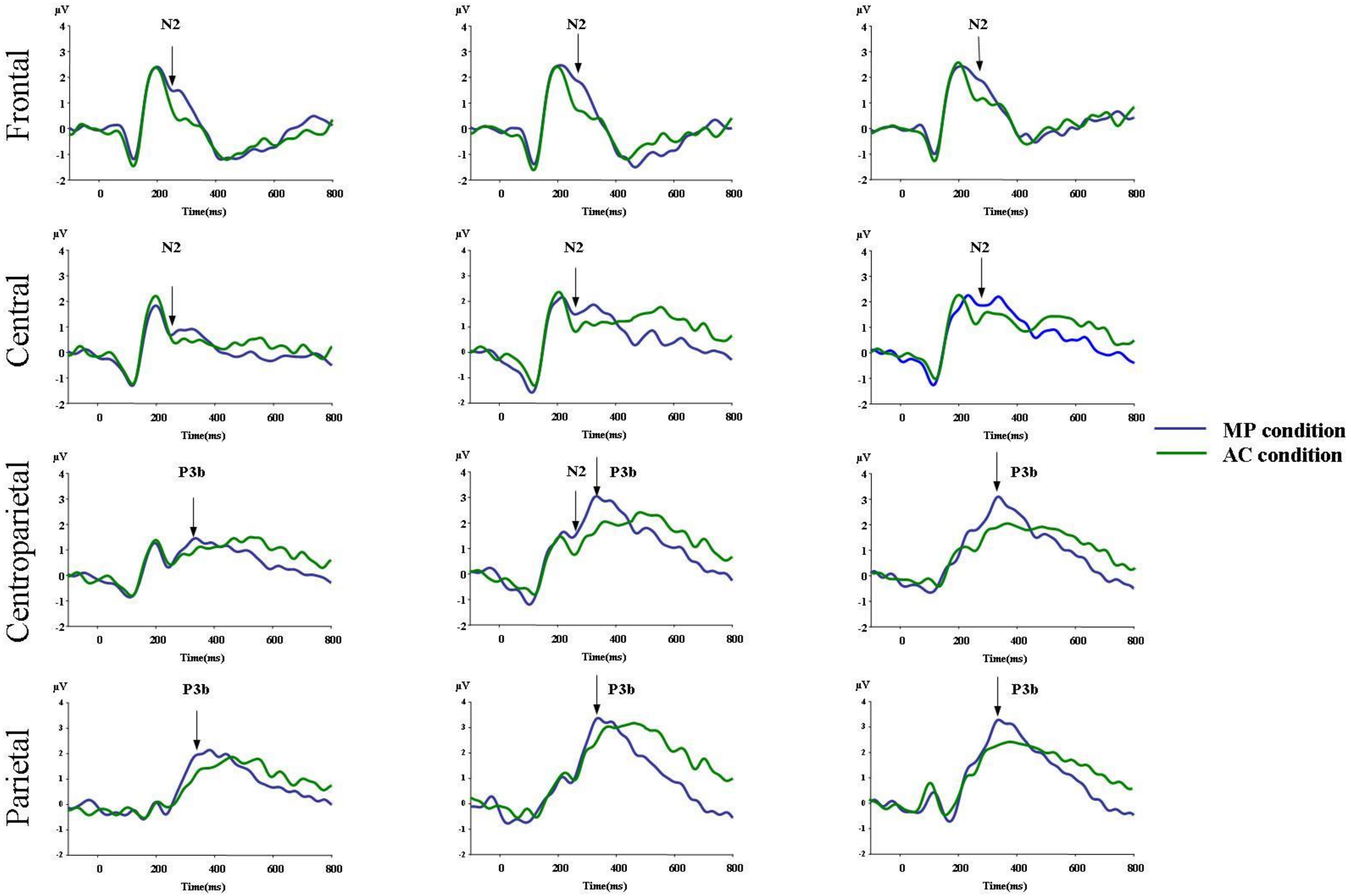
- N2   indicates a violation
- P3b indicates an expectation that was met

- Focus on Endorsers and Rejectors

Second Premise **Endorsers**



Second Premise **Rejectors**



Everyone rejects AC's second premise; ***everyone*** arguably expects MP's minor premise

-- what an individual will do in light of that will vary

-- the distinction between MP & AC is more pronounced among the Rejectors

Varieties of narrowing

"invited inferences" go here

Voluntary		Coerced
The drawing out of a pragmatic inference is optional		The drawing out of a pragmatic inference is obligatory in order to make sense of the utterance
The (hypothesized) intended reading relies on a narrower version of the linguistically encoded one		There is no obvious relationship between the linguistically encoded reading and the (hypothesized) intended one
The linguistically encoded reading is often good enough for discerning a speaker's informative intention		The linguistically encoded meaning is implausible
When the linguistically encoded reading is narrowed, extra effort is generally detectable		Extra effort compared to what?

Does this distinction allow us to cash anything else out?

Yes

A lot of work shows that cases – that are arguably the voluntary kind – are not problematic for ASD folks (e.g.Chevallier, Happé, Wilson & Noveck, 2010), but there are differences of the coerced kind.

J Autism Dev Disord (2010) 40:1104–1117
DOI 10.1007/s10803-010-0960-8

ORIGINAL PAPER

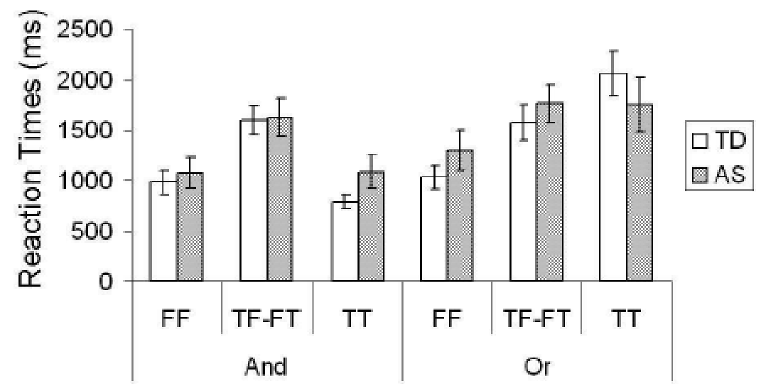
Scalar Inferences in Autism Spectrum Disorders

Coralie Chevallier · Deirdre Wilson ·
Francesca Happé · Ira Noveck

J Autism Dev Disord (2010) 40:1104–1117

Table 3 Percentage of correct answers as a function of connective type (And, Or), condition (FF, FT/TF, TT) and group (TD, ASD); results of the Mann–Whitney *U* tests comparing the groups

Connective	Truth condition	TD ($N = 22$) Mean (SD)	ASD ($N = 22$) Mean (SD)	Mann–Whitney *U* test z; p	Difference in mean ranks (Effect size index)
And	FF	100 (0)	99 (4)	$z = 0.00$; $p = 1.00$	0
	TF-FT	99 (4)	99 (4)	$z = -0.77$; $p = 0.44$	–0.07
	TT	100 (0)	97 (7)	$z = -0.25$; $p = 0.80$	–0.02
Or	FF	100 (0)	99 (5)	$z = 0.83$; $p = 0.40$	0.07
	TF-FT	42 (45)	57 (47)	$z = 0.35$; $p = 0.72$	0.03
	TT	43 (44)	48 (48)	$z = 0.01$; $p = 0.99$	0.001



Arguably, ASD deficits are linked to the coercion label

Table 5
Proportion of responses for the simple task and the suppression task. MP = modus ponens, MT = modus tollens, AC = affirmation of the consequent, DA = denial of the antecedent, add = with additional premise and alt = with alternative premise.

% responses	Autism			Control		
	Yes	No	Maybe	Yes	No	Maybe
MP	89.6	0.0	10.4	96.1	2.5	1.4
MP add	71.0	1.1	28.0	51.1	0.7	48.2
MP alt	92.9	0.4	6.8	97.5	0.7	1.8
MT	1.4	79.6	19.0	2.5	92.8	4.7
MT add	0.7	62.1	37.1	0.7	45.0	54.3
MT alt	0.4	90.3	9.3	1.1	95.0	3.9
AC	45.0	1.1	53.9	67.1	2.1	30.7
AC add	28.1	1.1	70.9	35.7	0.0	64.3
AC alt	12.2	2.2	85.7	9.6	0.0	90.4
DA	1.1	48.0	50.9	0.4	69.1	30.6
DA add	2.9	28.9	68.2	2.5	33.6	63.9
DA alt	3.2	15.7	81.1	1.1	10.4	88.5

From Pijnacker, Geurts, Van Lambalgen et al.

Also see, Happé et al., 1984

Summary:

Using scalars as an experimental model for analyzing the semantic/pragmatic frontier, we investigated conditionals (at that same frontier).

Conditionals (at least *out of the blue* conditionals in our restrained environment), show that listeners anticipate MP inferences over AC inferences.

Conditionals do not fit into scalar (and referential) framework. There is a distinction to be made (between what I call *voluntary* and *coerced* cases). This distinction allows one to cash out empirical findings.

Danke