Rationality$_2$: No guide for the perplexed?

Shira Elqayam

University of Plymouth

Acknowledgments

- ESRC grant to Handley, Elqayam, Evans
- Comments: Jonathan Evans, Valerie Thompson, Helen Neilens

Speeding to catch the match…

- British traffic law
- SEU
- Adaptive rationality

Overview

- Conflation of normative / computational theory leads to fallacy
- Typology of relation between empirical paradigm and normative system(s)
  - Most situations require arbitration
- No viable decision procedure

The dual concept of rationality

- E.g., Anderson, 1990; Evans, 1993; Evans & Over, 1996
- Rationality$_1$ – personal, adaptive, goal oriented
- Rationality$_2$ – normative
- Often conflated with dual processing
  - Evans (1993) calls rationality$_2$ 'rationality of process'
  - Evans & Over (1996) emphasise 'a reason for what one does sanctioned by a normative theory' – again leading to DP
- Where does the conflation come from?

Normative ≠ Computational

- The muddle is a symptom of conflating normative theory with computational theory
- Much clearer in linguistics
- 'I don't know nothing' – double negation
- Non-normative BUT
- Grammatical in African American Vernacular English (AAVE)
- Competence / computational theory of AAVE will include double negation

Normative ≠ Computational cont.

- Computational: *What* is computed
- Algorithmic: *How* it is computed (Marr, 1982)
- The computational / algorithmic (or competence / performance) should not be conflated with normative / descriptive
- *Computational and algorithmic are both descriptive*
- When theoretical accounts compete:
  - You can arbitrate between computational accounts with empirical data
  - But not between normative accounts (more later)

A typology of normative relations

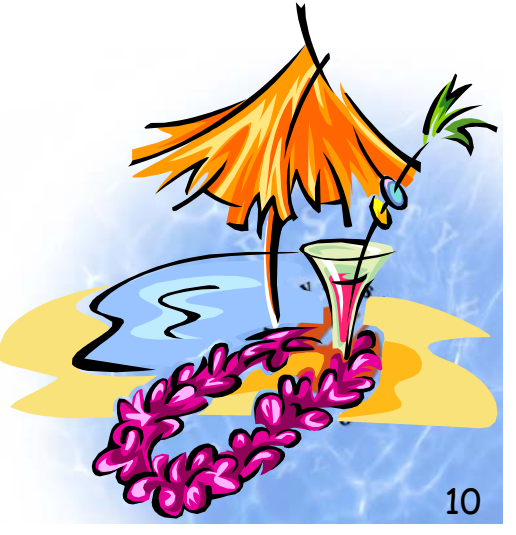
- How do you decide which is the appropriate normative system?
  - Normative system problem (Evans, 1993)
  - Inappropriate norm argument (Stanovich, 1999)
- Even more complex than acknowledged
- Four types of normative theory and experimental data

Type A: One norm, no conflict

- Example: disjunction effect (Tversky & Shafir, 1992).
- Seems to offer the prototypical normative condition
- Actually, nearly extinct after half a century of contentious research
- To find one look for recent paradigms

Type A: One norm, no conflict

- Hawaii vacation problem: Imagine you took tough qualifying exam; will only know results tomorrow
- Attractive discounted holiday package to Hawaii, offer terminates today.
- Majority chose to pay a fee to defer their decision
- BUT: majority opted to buy the package when told to imagine they had failed the exam – or when told they had passed.



Type A: One norm, no conflict

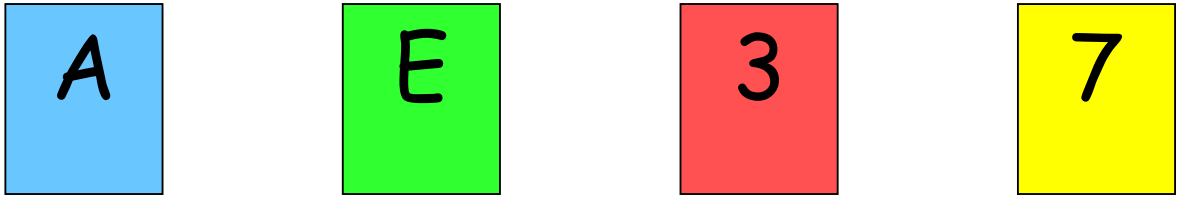
- Violates STP: If one prefers X to Y when condition C obtains, and X to Y when C does not obtain, one should prefer X to Y regardless of C
- Normative theory *not challenged (yet?)*

Type B: Standard-alternative conflict

- Behaviour irrational by standard norm ($norm_s$) but rational by alternative ($norm_a$)
- Far more common
  - Extensively covered in Stanovich (1999)
- The classic example: Wason selection task

Type B: Standard-alternative conflict

If there is an A on one side of the card, then there is a 3 on the other side of the card



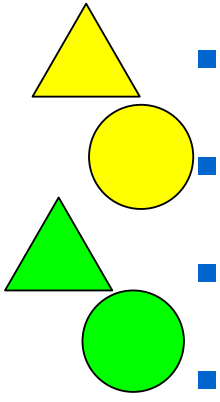
p ¬p q ¬q

Type B: Standard-alternative conflict

- Prevalent response patterns: p, or p&q
- p & q non-normative by standard textbook logic ('non-normative$_s$')
- Normative when task is interpreted as inductive (Oaksford & Chater, 1998; 'normative$_a$' – alternative norm)
- What makes one norm 'standard' and another 'alternative'?
- Probability of conditionals (Evans & Over, 2003; Oberauer & Wilhelm, 2003; Over et al., in press)

Type B: Standard-alternative conflict

- Suppose you know that a deck of cards contains the following cards:



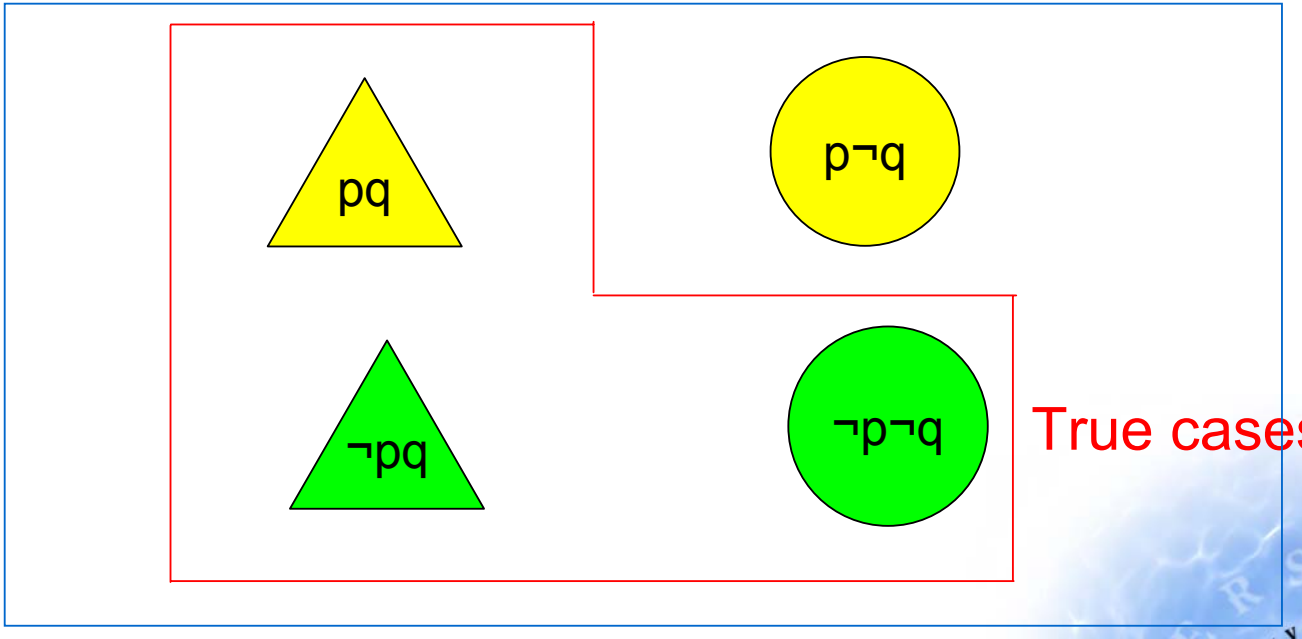
  - 50 yellow triangles
  - 150 yellow circles
  - 500 green triangles
  - 500 green circles

- Suppose a card is taken at random from the pack. What is the likelihood that the card conforms to the following rule?
  - If a card has a yellow shape then this shape is a triangle.

Type B: Standard-alternative conflict

- Competing normative theories sanction different responses

Material conditional

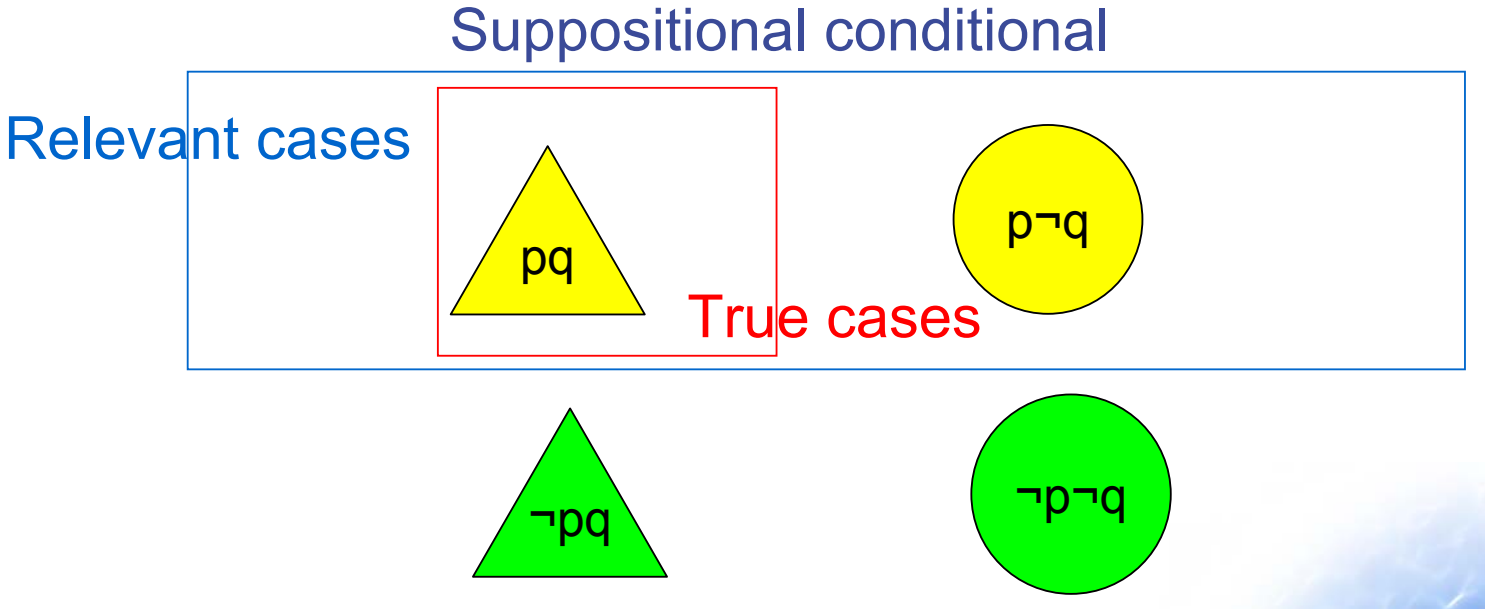


Relevant cases

= 1050 / 1200 = 87.5%

Type B: Standard-alternative conflict

- Competing normative theories sanction different responses



= 50 / 150 = 25%
(conditional probability)

Type B: Standard-alternative conflict

- What makes one norm 'standard' and another 'alternative'?
  - Historical precedence?
  - Kuhnian paradigm?
- MC has historical precedence
- SC has better case as Kuhnian paradigm, at least for everyday conditionals in natural language (Edgington, 1995)

Type C: Multiple norm conflict

- More radical: no normative standard
- Many norms compete
- Metadeduction (e.g., Rips, 1989; Byrne et al., 1997; Schroyens et al., 1999; Elqayam, 2006): on the Island of Knights and Knaves, knaves always lie and knights always tell the truth

- I am a knave and snow is black

- Conjunctions always false when one conjunct is false

Type C: Multiple norm conflict

- Stock answer in most meta-deduction literature BUT:
- Normative situation a lot more complex (Elqayam, 2003, 2006)
- 'I am a knave' is a paradox
  - A knave can't utter it because it would be true
  - A knight can't utter it because it would be false
- 'Liar' paradox knight-knave style
- How do we evaluate sentences with paradoxical / indeterminate constituents?

Type C: Multiple norm conflict

- Many valued truth systems (Gottwald, 2001; Rescher, 1969)
  - Strong / conservative truth tables: sentences evaluated based on determinate constituents if possible
  - Weak / radical: Any sentence with indeterminate constituents evaluated as indeterminate
  - Both presented within several pages of the same work! (Kleene, 1952; van Fraassen, 1966, 1969)
- No standard, no alternative
- All systems equally standard and equally alternative
- Are many type B conflicts actually type Cs in disguise?

Type D: No-norm no-conflict

- Where no normative theory exists at all
- Example: conditional inference with meta-deductive constituents (Elqayam et al., 2005)
- 'If I am a knave then I live in Emerald City'
  - 50 knaves in Emerald City
  - 150 knaves in Diamond City
  - 500 knights in Emerald city
  - 500 knights in Diamond city
- Same distribution as the pack of cards
- Common response pattern: 0%

Type D: Non-norm no-conflict

- Where meta-deductive components are not involved, the suppositional conditional has very clear norm
- Ramsey test: mentally simulate the antecedent and evaluate the consequent in that context
- Can you mentally simulate a paradox?
- →No pre-existing norm

Summary: the four conflict types

Type	Conflict / No Conflict	No. / type of norms involved	Example
A	No Conflict	1	Disjunction effect
B	Conflict	1 Standard + at least 1 Alternative	Wason Selection Task
C	Conflict	Multiple	Meta-deduction
D	No Conflict	0	Metadeductive conditionals

- Only for Type A the relation between normative theory and experimental paradigm is not strained
- Most Type B / C started off as Type A
  - Just a question of time until an alternative norm is proposed
- Type D paradigms by themselves are enough to cast doubt on normative rationality
- With Types B / C we need an arbitration procedure

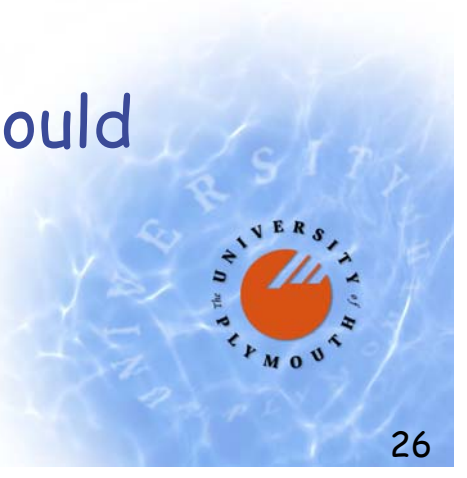
When theories compete

- Where theories compete, we need an arbitration procedure
- When computational theories compete, we can use empirical data to decide between them
- When normative theories compete, using empirical data amounts to a fallacy
- No guide for the perplexed!

Is-ought fallacy

- First identified by Hume
- When premises are descriptive (is), it is logically invalid to derive a normative conclusion (ought)
- E.g.:
  - Human beings have natural fear of heights (is)
  - Therefore, we should not fly in aeroplanes (ought)

- The implicit 'ought' premise: we should avoid whatever we naturally fear



Is-ought fallacy

- Confusing normative with computational triggers is-ought fallacy
- Computational / competence is an 'is' type theory (descriptive)
- Normative is an 'ought' type theory
- You can support an 'is' theory with 'is' evidence, but not an 'ought' theory
- Two influential solutions
  - Rational analysis (Oaksford & Chater)
  - Understanding / acceptance and individual difference (Stanovich; Stanovich & West)

Is-ought Fallacy I: Rational Analysis

- *A* rational norm is one that is *computationally adequate* (Oaksford & Chater, 1998).
  - Computational theory is psychologically complete if it can generate all 'intuitively correct' answers
  - Similar to Chomsky's notion of descriptive adequacy: A linguistic theory is descriptively adequate if it can generate all grammatical utterances

Rational Analysis critique

- 'Intuitively correct' begs the question
- Far removed from intuitive grammaticality
- Different people have different logical 'intuitions' (Stanovich, 1999; more later)
- Do Ps with higher ability speak a different logical native language?

Rational Analysis critique cont.

- Conflates normative / descriptive with computational / algorithmic (rational norm should be computationally adequate)
- Maintaining the distinction is essential to avoid the is-ought fallacy

Back to Rational Analysis

- Is-ought fallacy in Rational Analysis:
  - People behave in a way that approximates Bayesian rules (is)
  - This behaviour is successfully adaptive (is)
  - We should follow whatever normative system that makes our behaviour adaptive (ought).
  - Therefore, Bayesian rules are the appropriate normative system (ought)
- However, this is oversimplified.
- The argument is valid if the 'ought' premise is added

Back to Rational Analysis

- Rational analysis follows this route by explicitly stating that any adaptively rational behaviour should be justified in terms of some normative system
- With this premise in the argument is valid
- BUT:
  1. Can only be added axiomatically
  2. Oaksford and Chater compound the problem by proposing to address the question empirically – see whether adaptive behaviours conform to some normative standard – is-ought fallacy all over again

Is-ought Fallacy II: Understanding / Acceptance

- Stanovich (1999; Stanovich & West, 1998; 2000) adopt Slovic and Tversky's (1974) understanding / acceptance principle:
- The more one understands the normative principles involved, the more likely is one to endorse them
- Hence, more cognitively gifted reasoners are more likely to respond in congruence with 'appropriate' normative model
- And vice versa: accept as normative whatever is congruent with responses of higher ability reasoners

Understanding / acceptance critique

- Even if we accept the rule, the reverse may not be true – Ps may endorse a rule for reasons that have nothing to do with the normative system
- What happens when competing norms are both supported?
  - WST: highest ability: p&¬q; second highest: p only (compatible with Margolis's 'open reading')
  - Which is 'the' normative system?
- What if higher ability Ps reject inferences considered to be valid
  - Higher ability: fewer MT (Newstead et al., 2004; Evans et al., 2005)

Is-ought again

- Is-ought fallacy: From the 'is' (bright Ps response data) derive the 'ought' (appropriate norm)
- Akin to naturalistic fallacy, which derives moral norms from natural phenomena (aeroplane example)
- Stanovich acknowledges the naturalistic fallacy (1999, p. 60), but asserts:
  - 'if the theorists discussed so far are actually committing the naturalistic fallacy, then many of the best minds in cognitive science seem to be doing so'
  - → Appeal to authority

Are there two types of rationality?

- NO
- Normative rationality is a myth
- Asking whether behaviour on, e.g., the WST conforms to a specific norm (textbook logic, Bayesian system) is analogous to asking if AAVE conforms to 'good English':
  - At best, relatively uninteresting
  - At worst, does not make any sense

What about adaptive rationality, then?

- Prima facie, rationality$_1$ – adaptive / instrumental rationality – involves no norms
- Therefore, no is-ought fallacy
- Can we keep the normative question out of the game?
- Stanovich (1999): both rationality$_1$ and rationality$_2$ are cases of normative rationality (both relate to personal goals)
- Is instrumental rationality normative?

Concluding comments

- Linguistics only came of age when it discarded its historical obsession with norms.
- The obsessive back-and-forth dialogue with normative theory is a peculiarity of reasoning and decision making theories (Evans, 1993).
- Dual process theories of reasoning should be liberated from dual theories of rationality.
- In spite of the muddle in much of the literature, the two are *not* mutually dependent, and dual-process theories of reasoning would only benefit from a clear separation.
- Let us leave the 'ought', then, to clergymen and politicians, and concentrate on the 'is' instead.

Thank you!