

Rationality₂: No guide for the perplexed?

Shira Elqayam

University of Plymouth

ABSTRACT

My aim in this paper is to suggest that the dual notion of rationality cannot be sustained. More specifically, I argue against the notion of normative rationality, the Rationality₂ of the Evans and Over (1996) concept of dual rationality, based on having ‘a reason for what one does sanctioned by a normative theory’ (ibid, p. 8). To elucidate the difference between these two type of theories, I draw on an example from linguistics where the distinction is well established, and suggest that the same would benefit the rationality debate in reasoning and decision making. I propose that normative rationality is commonly conflated with a computational level analysis, and that this muddle gives rise to the ‘is-ought’ fallacy. This fallacy means that, while one can empirically arbitrate between competing computational accounts, there is not way to do so for competing normative ones. I suggest a typology of relations between normative systems and an empirical paradigm and demonstrate that nearly all of them pose a problem to normative rationality. I then proceed to examine two prominent rationality agendas, rational analysis of Oaksford and Chater (e.g., 1998) and the use of the understanding / acceptance principle in Stanovich’s individual differences programme (e.g., 1999). Each of these programmes commits the is-ought fallacy in some part of its arguments. I conclude that normative rationality cannot be an issue for empirical research, and that dual process theories of reasoning are better off without the dual normative framework.

A critical football match is about to start and Jack speeds along the deserted highway. One normative system says he shouldn't do that – British traffic law. Another system says he's quite justified – his chances of being caught are small enough and his longing to see the game strong enough to justify risking the fine for speeding, so by Subjective Expected Utility he is being rational. And of course, he's being adaptively rational because he acts towards satisfying his personal goal of seeing the match. Is Jack, then, being rational? There, in a nutshell, we have a classic dual rationality dilemma.

The dual concept of rationality has been around for quite some while, often – although not always – in conjunction with a dual process theory of thinking (e.g., Anderson, 1990; Evans, 1993; Evans & Over, 1996). Dual rationality systems distinguish between personal, adaptive, goal-oriented rationality – dubbed 'rationality₁' by Evans (Evans, 1993; Evans & Over, 1996; Evans, Over, & Manktelow, 1993); and normative or impersonal rationality, or 'rationality₂'. In one of the most influential versions of this approach, Evans and Over (1996) employ the distinction as a keystone in a general dual process theory of reasoning. In this version, the link between rationality₂ and analytic or explicit processes became explicit, by making rationality₂ contingent on having 'a reason for what one does sanctioned by a normative theory' (ibid, p. 8) – for one can only have an explicit reason when explicit processing is involved. The link between normative rationality and analytic processing actually goes back to Evans's earlier work (1993), in which rationality₂ was suggestively dubbed 'rationality of process'.

I don't think that it needs a lot of discussion to show how tenuous this link is, between normative rationality and analytic processing, and indeed, in later work Evans maintains as much (Evans, in press), opting instead for the much weaker

assertion that analytic processing is more likely to produce solutions congruent with normative rationality.

This muddle, however, is but a symptom of another, more far-reaching one, which gave rise to it. In this paper I will try to dig out the roots of the confusion, and show that it is still with us today, and in the writing of influential theorists. My assertion is that the source of the confusion is conflating normative theory with computational-level analysis. Once the confusion is disentangled, I aim to demonstrate that the notion of normative rationality cannot be sustained, and that dual process theories in reasoning and decision making are better off giving up the quest for normative rationality.

First of all, what is the difference between normative theory and computational or competence theory and why should we make it? The distinction is much clearer in linguistics, so this is where we now turn to. For a classical example, consider the sentence ‘I don’t know nothing’. Although in Standard English the use of double negation is non-grammatical, this sentence is nonetheless a grammatically correct expression in non-standard varieties of English, e.g. African American Vernacular English (AAVE). A normative theory of language will condemn double negation as ‘wrong’ (doubtless numerous teachers in primary education do so to this very day). A competence theory, though, is a very different matter. A competence theory of AAVE would strive to describe the rules governing these expressions in AAVE that its native speakers consider correct. A competence theory of AAVE will thus include double negation as a grammatical rule, although it is non-normative in Standard English, describing rather than judging. The accepted wisdom in linguistics – and one that goes back to Ferdinand de Saussure’s seminal ‘Course in General Linguistics’ (1959) – is that an adequate linguistic theory would be far more

interested in a description of grammar rules in AAVE than in condemning it as ‘irrational’ or ‘ungrammatical’ (e.g., Trudgill, 2000. Of course, a sociolinguist may well observe (ibid) that the use of double negation is negatively correlated with socioeconomic status, but that's a different matter altogether). This example makes quite clear the difference between a normative theory and a competence-level theory (or, in Marr’s terms, computational level analysis; see later). It is the latter that, I will argue, constitutes a hard-core scientific question, whereas the former is at best a matter for educational policy.

The muddle between normative rationality and analytic processing boils down to conflating the two levels. Rationality₂, of course, requires a normative theory by definition; but to explicate analytic reasoning one needs a computational theory, a theory of deductive competence. Once the difference becomes apparent, it is clear that dual processing accounts need not have dual rationality presuppositions; on the contrary, such presuppositions are counterproductive, as I will now try to establish.

The difference between the normative and the computational becomes crucial when theoretical accounts compete – either competing normative accounts or competing competence accounts. When one has to arbitrate between different competence / computational theories, one can do so with the aid of empirical data to support or undermine one or the other. I will try to show that using the same strategy to arbitrate between competing normative accounts is logically unsound, and a dubious scientific endeavour.

I will first examine the various ways in which normative systems can relate to a specific empirical database. I will then go on to examine two prominent theoretical proposals, each with its own arbitration mechanism: Oaksford and Chater’s rational analysis, and Stanovich’s individual differences agenda. I will demonstrate how both

proposals suffer from the same fallacy, and conclude that the fallacy is inherent, and why no amount of empirical evidence can arbitrate between competing normative accounts.

A normative conflict typology: some examples

What Evans (1993) calls the ‘normative system problem’ and Stanovich (1999) calls ‘the inappropriate norm argument’ means that deciding on an appropriate normative system for any set of experimental findings is more often than not far from obvious. Indeed, one is hard put to find an experimental paradigm that has just one obvious norm to compare against and no competing alternative norms. Such cases though do exist, as well as more radical cases of normative conflict. In the following I suggest a typology of four normative situations, based on the nature and number of competing normative accounts of a particular experimental paradigm. Of the four types, two involve normative conflict and two involve no conflict.

Type A. Perhaps the simplest (and quite rare), in this one-norm no-conflict situation there is just one applicable norm, typically the one that the originators of the paradigm had in mind.

Type B. Far more common state of affairs, this is the standard-alternative conflict situation, with a standard norm competing with alternative accounts. It is so common that it has become the only acknowledged type in, for instance, Stanovich’s account (1999).

Type C. The existence of a standard norm is not universally the case – there is also type C, the multiple norm conflict type, in which there are several alternative norms, none of which has any claim for ascendancy.

Type D. Finally, type D is the no-norm no-conflict condition: where no norm exists there is obviously no conflict. This condition makes the idea of normative rationality even more difficult to maintain. I will now examine some typical examples to illustrate each case.

Table 1 summarise the four types and their characteristics.

(Insert Table 1 about here)

Type A one-norm no-conflict situations seem to be offer the prototypical normative condition, but after half a century of contentious reasoning and decision making research, such cases have become all but extinct. One has to look for relatively recent experimental paradigms to find some in which the norm that first motivated the authors has remained unchallenged. One such type A paradigm is Tversky and Shafir's disjunction effect (1992), and the norm in question is Savage's Sure Thing Principle (1954). According to STP, if one prefers X to Y when condition C obtains, and one also prefers X to Y when condition C does not obtain, then one should prefer X to Y regardless of C. Thus, if I prefer red wine to white wine regardless of what I eat, I should order red wine in the restaurant before I decide if I want curry or steak for the main course. Tversky and Shafir's (1992) classic paper demonstrates repeated violations of STP. Perhaps the most striking example is the Hawaii vacation problem, in which participants are asked to imagine that they have taken a tough qualifying exam and will only know the following day if they have passed or failed. They now have a chance to purchase an attractive discounted holiday package to Hawaii, but the offer terminates today. The majority of these participants chose to pay a fee to defer their decision until after they know if they have passed or

failed. This, in contrast to the majority of participants who opted to buy the package in a test condition in which they were told to imagine they had failed the exam – and the condition in which participants were told they had passed. I will skip the theoretical analysis Tversky and Shafir offer in terms of prospect theory, since the point in question is the appropriate norm. As far as I am aware, although the disjunction effect itself has been challenged (e.g., Kühberger, Komunská, & Perner, 2001), no one has challenged the assumption that STP is the appropriate norm against which to judge it. There may be other examples of one-norm no-conflict situations but they are not all that common.

A far more prevalent type is B, the standard-alternative conflict situation so extensively covered in Stanovich (1999). In a typical debate of this type, a standard account of a particular observation competes with another, alternative account (or accounts), making an observed behaviour rational₂ according to the latter but not according to the former (and vice versa). Perhaps the most familiar debate in this regard is the one over the Wason selection task. I will describe the task very shortly as I assume that most would be familiar with it. In this widely studied task (for a recent review see Evans & Over, 2004) the participants are presented with four cards and a rule – typically a conditional rule of the form ‘if p then q’; for instance, ‘If there is an A on one side of the card then there is a 3 on the other side’. The cards bear the characters A, D, 3 and 7 (called p, not-p, q and not-q respectively). The participants are requested to choose the cards – and only the cards – that need to be turned over in order to decide if the rule is true or not. To cut a rather long story short, the typical response pattern is the p card alone, or a combination of the p and the q card, although the standard normative answer is a combination of the p and the not-q cards. No more than 10% of intelligent adults (our customary euphemism for first year undergrads)

typically come up with this normative answer. From the point of view of standard textbook logic, the p & q only response pattern is just as erroneous as the p pattern. However, Oaksford and Chater (1998) have argued that the p & q response is the normative one when the task is interpreted as inductive.

What makes one account 'standard' and the other 'alternative' is less than clear in Stanovich's account. Perhaps the simplest criterion is that the 'standard' norm is the one that originally motivated the paradigm. Or one could perhaps propose a sociology of science assertion that the 'standard' norm is a sort of Kuhnian paradigm. Or simply choose the norm that has been around longer. The problem, of course, that while some criteria may coincide at times, others may give rise to different answers. Once one considers this question, many type B cases may turn out to be far from obvious. Take, for instance, the case of probability of conditionals (Evans, Handley, & Over, 2003; Oberauer & Wilhelm, 2003). Suppose you know that a deck of cards contains the following cards:

50 yellow triangles (pq)

150 yellow circles (p not-q)

500 green triangles (not- p q)

500 green circles (not-p not-q)

Now 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.

I will not go into the competing psychological theories (for review see Evans & Over, 2004; Johnson-Laird & Byrne, 2002); however, they are all related to competing normative theories of everyday conditionals. According to the material conditional, a conditional is true in the pq case, both not- p cases, and false in the p not- q case. This makes the conditional rule given true for the 50 yellow triangles plus the 500 green triangles and the 500 green circles, but false for the 150 yellow circles. Hence, the normative response for a material conditional should be $(50+500+500)/(50+150+500+500) = .875$, or 87.5%. In contrast, the suppositional conditional is true for $p q$, false for p not- q , and is neither true nor false otherwise ('truth-value gaps'). Hence, for the suppositional conditional the normative response would be the conditional probability, q given p , in this case $50/(50+150) = .25$, or 25%. (I will get back to the suppositional conditional in our next example.)

Given these conflicting norms, can we easily determine which is the normative_s response (in Stanovich's terms) and which the normative_a – that is, the standard norm and the alternative norm? The answer is far from clear. The material conditional has historical precedence on its side, but few contemporary logicians accept the material conditional as an adequate description of everyday conditional (Evans, Over, & Handley, 2005) – in terms of Kuhnian paradigm, one could argue for the suppositional conditional. So for experiments in probability of conditionals, it is not easy to determine which is the standard norm and which is the alternative.

Some type B cases, then, may actually be type C cases in disguise. To fully appreciate this, we now turn to examine this type. The type C multiple-norm conflict is a more radical case, in which there is no normative standard whatsoever and many systems compete for the same observation. For example, take the reasoning literature on metad eduction (Byrne & Handley, 1997; e.g., Byrne, Handley, & Johnson-Laird,

1995; Elqayam, 2006; Rips, 1989; Schroyens, Schaeken, & d'Ydewalle, 1999). Here reasoners are presented with the Island of Knights and Knaves, in which all the inhabitants are either knights, who only tell the truth, or knaves, who only lie. Participants are given statements and asked to identify whether the speakers are knights or knaves. Consider the following statement:

I am a knave and snow is black

Since conjunctions are false whenever one of the conjuncts is false, it seems that the sentence is false and the speaker is a knave. This is the stock answer in metaductive literature, and answers that deviate from it are typically considered erroneous by most authors regardless of theoretical persuasion (e.g. Rips, 1989; Johnson-Laird & Byrne, 1990). However, this is not the only possible answer. On the Island of Knights and Knaves, 'I am a knave' is paradoxical – a knave cannot utter it because it would be true and knaves lie, and a knight cannot utter it because it would be false and knights don't lie. In fact, this is a version of the Liar paradox (e.g., Martin, 1984). The question now shifts to how we treat sentences with paradoxical constituents, which brings us to many-valued truth systems. Elqayam (2003) has argued that the plethora of such systems (for reviews see Gottwald, 2001; Rescher, 1969) does not allow for one type of solution to be preferred over the other. For instance, the stock answer in metaduction amounts to adopting one particular type of solution – 'strong' (Kleene, 1952) or 'conservative' (Van Fraassen, 1966; Van Fraassen, 1969) truth tables – that is, truth tables in which a proposition's truth value is determined by its determinate constituent, where such exists. Hence, in a strong system 'I am a knave and snow is black' will be evaluated as false, because a

conjunction is false whenever one of its constituents is false, regardless of the other constituents' truth status. However, within several pages of the same work, both authors also offer 'weak' or 'radical' solutions, in which propositions with an indeterminate constituent are evaluated as indeterminate. This means that any sentence that contains a paradoxical component, such as in our example, will be considered as indeterminate, or a 'truth-value gap' – neither true nor false.

Inexplicably, this perfectly legitimate option is considered an error in most of the metaduction literature.

Notice the difference between this type C multiple-norm conflict to a type B standard-alternative norm conflict. In the latter there exists a 'normative standard' response (e.g., p and not-q on the Wason selection task) and a 'normative alternative' one (e.g. p & q only on the WST). In a type C conflict, however, there is no standard, no alternative; there are many systems, each equally standard and equally alternative, each sanctioning a different response patterns, and often suggested by the same authors in the same work. Hence, the argument cannot be that reasoners should conform to some normative system X, or normative system Y rather than X: one should first convince why any system should have any sort precedence.

Let us now turn to the last type, D, the no-norm no-conflict type. As the name implies, this is where no pre-existing normative theory exists at all. One such case is conditional inference with meta-deductive constituents (Elqayam, Handley, & Evans, 2005). In this study participants were presented with various conditional statements made on the Island of Knights and Knaves, such as 'If I am a knight then I live in Emerald City', and with distributions of knights and knaves in the Island's various cities, e.g.:

50 knaves in Emerald City

150 knaves in Diamond City

500 knights in Emerald city

500 knights in Diamond city

Note that this is the same distribution that we had earlier for our pack of cards, and indeed, when the Liar paradox is not involved, the suppositional theory of conditionals (Edgington, 1995; Edgington, 2003; Evans & Over, 2004) has a very clear norm. Consider the following conditional: 'If Pete is a knave then he lives in Emerald City', with the same distribution as above. This is a perfectly good conditional and the antecedent is unremarkable: it is not a paradox or indeterminate for any reason. In this case, the famous Ramsey test (Ramsey, 1931) is applicable: reasoners mentally simulate the antecedent (in this case, that Pete is a knave) and assess their belief in the consequent in that context. The normative solution according to the suppositional theory of the conditionals is therefore the conditional probability, $P(q|p)$. This is also the response pattern that the majority of intelligent adults come up with when object-language level conditionals (i.e., conditionals that refer to anything but semantic concepts) are involved (Evans et al., 2003; Evans & Over, 2004).

However, when the antecedent is the Liar paradox, things get complicated. Can one even mentally simulate a paradox? To the best of my knowledge, there is no discussion in the philosophical literature on suppositional conditionals with paradoxical or otherwise indeterminate antecedents. Yes, we could begin such a discussion, perhaps invent a norm (we came up with several, mutually-excluding ones). No doubt this would be of philosophical interest. But doing so merely to find some norm against which to compare empirical responses borders on the ludicrous.

We could also fall back on the material conditional, and find numerous multi-valued systems with various material conditional solutions, but this would only put us in a type C condition, with a multiple-norm conflict – hardly helpful.

Out of these four types, there is just one in which the relation between the experimental paradigm and the potential normative system(s) is not strained, and that is type A, where there is just one normative system. However, most type B and type C paradigms started off as type A paradigms, with alternative norms added on by subsequent critics. I think it is a safe bet that for the few type A paradigms still around, it is just a matter of time until someone finds an alternative norm.

The case for type D paradigms is even more tenuous. When there is no formal system to conform to, how can one be rational₂? Type D paradigms by themselves are enough to throw serious doubt on normative rationality.

Finally, for the conflict cases B and C, it is clear that we need a guide for the perplexed. How does one choose a normative system? Is there any failsafe principle to select and distinguish between normative alternatives? My answer is that any arbitration criterion that involves empirical evidence would be inherently fallacious, because a normative theory is not a competence theory. To understand this, let us examine in detail two of the major rationality programmes in the field: Oaksford and Chater's (1998), and Stanovich's (1999).

Is-ought fallacy I: Rational analysis

The first proposal I will examine is rational analysis, Oaksford and Chater's (1998) prominent rationality programme. They suggest (ibid, p. 7) that a rational norm is one that is computationally adequate. The distinction is based on Marr's (1982) definition of three levels of theoretical explanation: the computational level –

what is computed; the algorithmic – *how* it is computed; and the implementational – the ‘hardware’ / ‘wetware’ physical level. Oaksford and Chater propose that a computational system would be psychologically complete if it can generate all ‘intuitively correct’ answers (just as Chomsky (1957) proposed that a linguistic theory would be descriptively adequate if it can generate all grammatical utterance).

I can see various difficulties with this. First, there may be cases in which none of the alternative norms are complete in the sense that Oaksford and Chater propose. Metaduction is again a case in point. Some responses tend to be ‘strong’ / ‘conservative’ whereas others are ‘weak’ / ‘radical’, but no system incorporates both.

Moreover, the concept ‘intuitively correct’ itself begs the question. It seems far removed from Chomsky’s notion of intuitive grammaticality. We are all capable of judging grammaticality in our own native language or dialect, but different people can have very different logical ‘intuitions’. Stanovich’s individual differences programme (e.g., 1999) – the next to I will explore – demonstrates how different we can be in terms of logicality intuitions. For instance, different response patterns on the Wason selection task tend to be associated with different ability scores. Do higher ability reasoners speak a different logical native language?

Suppose, then, we take away the intuition part of the equation, and maintain that a theory is computationally complete if it creates all correct answers (rather than all intuitively correct answers). Will this fit the bill? Unfortunately, this will not remedy the situation. ‘All correct answers’ necessitates one to define a norm to judge if answers are correct or not. Thus, we are left with a circular definition.

But the major problem I have with this proposal is that, while adopting a similar strategy to the linguistic agenda, it does not keep to the linguistic distinctions. It fails to separate the normative from the computational, a separation that is, as we

have seen, very clear in linguistics. Another way to put it is that they conflate the normative / descriptive distinction with the computational / algorithmic distinction. However, a computational theory – or a competence theory, in Chomsky’s terms – is not conceived as a normative theory; it does not endeavour to dictate ‘good’ language. I cannot conceive of a contemporary linguist suggesting that an adequate linguistic theory is also a good normative theory of how people *should* use language. Linguistics as a rule tends not to be overly interested in how people should talk, only how they do talk. In other words, a computational level theory is just as descriptive as an algorithmic level theory. Conflating computational with normative is unfortunately quite common, but the analogue with linguistics helps to clear this up.

The lack of clear boundaries between computational theory and normative analysis in Oaksford and Chater is not a problem in itself, but unfortunately maintaining such boundaries is essential if one is to avoid the is-ought fallacy. The is-ought fallacy, first identified by Hume (also see Cohon, 2004; 2000), means that whenever the premises in an argument are merely descriptive (‘is’), it is logically invalid to derive from them a normative conclusion (‘ought’). For instance, the following argument contains an is-ought fallacy:

Human beings have natural fear of heights. (is)

Therefore, we should not fly in aeroplanes. (ought)

The conclusion that we should not fly in aeroplanes only follows if we add the implicit ‘ought’ premise that we should avoid anything that we have a natural fear of.

Confusing a competence theory with a normative theory inevitably triggers the same sort of fallacy. A competence / computational theory is an ‘is’ type of theory,

and empirical evidence is an 'is' type of argument. Supporting the former with the latter involves, therefore, no is-ought fallacy. However, a normative theory is an 'ought' type of theory – how we *should* reason rather than how we *do* reason. One should not derive the 'ought' from the 'is'.

A simplified account of rational analysis arguments for normative rationality can be presented as:

People behave in a way that approximates Bayesian rules. (is)

This behaviour is successfully adaptive. (is)

Therefore, Bayesian rules are the appropriate normative system. (ought)

Note, however, that, the argument is made valid if we add the 'ought' premise:

We should follow whatever normative system makes our behaviour adaptive.

(ought)

Oaksford and Chater follow this route by explicitly stating that any adaptively rational behaviour should be justified in terms of some normative system (1998, pp 291-297); otherwise, they maintain, its rationality is meaningless. Thus, one can maintain that rational analysis avoids the is-ought fallacy. However, this strategy depends on a circular argument: the premise that we should follow the normative system that makes behaviour adaptive. Oaksford and Chater compound the problem by proposing that the answer should be addressed empirically, by looking to see whether adaptive behaviours actually conform to some normative standard. This is, of

course, an is-ought fallacy. So at best, the rationally analysis treatment of normative rationality is circular; at worst, it commits the is-ought fallacy.

Is-ought fallacy II: The understanding / acceptance principle

Another highly influential research programme that falls prey to the is-ought fallacy is that of Stanovich (1999; Stanovich & West, 1998; Stanovich & West, 2000), who adopts the understanding / acceptance principle suggested by Slovic and Tversky (1974): the more one understands the normative principles involved in an inference task, the more likely one is to endorse these principles. This means that the more cognitively gifted reasoners are more likely to respond in congruence with the ‘appropriate’ normative model – such as it is – for a particular problem set. Stanovich reverses the principle and maintains that we should accept as normative whatever is congruent with responses given by higher ability reasoners. In other words: if the participants who respond according to a specific pattern P1 are more intelligent than participants who respond according to another pattern P2, then P1 should be considered normative and the system it conforms to should be considered appropriate.

However, the understanding / acceptance principle in Stanovich’s writing suffers from a host of minor problems. First, even if one accepts the rule, the converse may not necessarily be the case – even if we agree that understanding triggers acceptance, it does not necessarily mean that acceptance is diagnostic of understanding. Participants may behave according to the rule but for their own reasons. Secondly, like many normative arguments, this one is at least partially circular: we know a normative system is ‘right’ because the brighter participants comply with it, but we know they are brighter because they comply with these same

normative systems, whose assumptions are often incorporated into general intelligence tests.

Also, what happens when the competing normative response patterns both receive support in the form of higher ability participants? This is the case with the Wason selection task. Although participants who provide the p & not-q response have the highest ability, second highest are participants who respond with p only – a pattern that is compatible with Margolis's (1987) 'open reading' analysis of the task. So, should we accept both accounts?

Finally, if participants with higher ability are our standard to what should be considered right or wrong, what would we do when these bright kids reject inferences that are generally considered valid? For instance, higher-ability participants tend to draw *fewer* Modus Tollens inferences (Evans, Handley, Neilens, & Over, 2005; Newstead, Handley, Harley, Wright, & Farelly, 2004). Modus Tollens is non-controversial as a valid argument under any conditional theory (Evans & Over, 2004), but if norms should be established according to whatever smarter participants prefer, than Modus Tollens should be rejected on the basis of these data.

However, all these are rather trivial problems compared to the biggest hitch in the understanding / acceptance agenda: the 'is-ought' fallacy. The understanding / acceptance principle involves the is-ought fallacy since it derives normative, 'ought' conclusion from descriptive, 'is' data, the data concerning performance of higher ability participants. It says, in effect: X is the way smart people *do* respond (is); therefore, X is the way we *should* respond (ought).

The is-ought fallacy is akin to the naturalistic fallacy (but see also Frankena, 1939; Moore, 1903), which derives moral norms from natural phenomena (recall our aeroplane example: that was a typical case of naturalistic fallacy). Stanovich (1999)

acknowledges the naturalistic fallacy in the understanding / acceptance argument, but goes on to argue that ‘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’ (p. 60). This answer is no help: it merely replaces the naturalistic fallacy with an appeal to authority one. One could even consider it another case of the is-ought fallacy, as ideas are again judged valid due to their source. Regrettably, we would have to reject the understanding / acceptance principle as guide for the perplexed.

Is-ought fallacy III: Explanatory adequacy is not an option

One last proposal, again congruent with Chomsky’s levels of adequacy criteria. Perhaps we should take the Chomskyan analogy one step further and concentrate of explanatory adequacy instead. A linguistic theory is adequate at the explanatory level if it can account for linguistic universals (for instance, the phenomenon that all languages have nouns). Perhaps what we should look for is logical universals. If we discover that some principle or other appears in different logical systems in various times and in geographically unrelated places, then we could have a good candidate for a logical universal, and perhaps for a rational norm. Tempting as this proposal may be, though, it cannot save normative rationality. It can provide an interesting starting point for a computational level theory, but we have already established that computational level is not the same as a normative account. This way lies the is-ought fallacy. A computational theory is not a normative theory, and any muddle between the two is not likely to be beneficial to reasoning theories.

Can normative rationality be experimentally arbitrated?

A clarifying aside seems to be in order, as there seems to be a superficial resemblance between the arguments presented here and some of Cohen's (1981). In a highly controversial paper, Cohen (1981) argued that human rationality should be considered as a given, and therefore no amount of empirical evidence can demonstrate that humans are irrational. This is a variation of what Stanovich (1999) calls a 'Panglossian' position, the position that humans are a-priory rational. The position seems to be similar to the one presented in this paper because my argument, too, denies the applicability of empirical evidence to human rationality. However, the resemblance ends there. The force of Cohen's argument aims to save human normative rationality by bestowing on it an a-priory status. In contrast, the position in this paper is not Panglossian. The force of my argument is not to except human rationality as a given, but to exclude it from the list of issues amenable to empirical scientific research.

Concluding comments

Can we, then, have a viable *dual* account of rationality? I am forced to conclude that we can't. Yes, there are two types of rationality in the literature; but the second is a fiction, a myth. The way I see it, asking whether participants in, say, a Wason selection task experiment are conforming to a set of normative rules (whether textbook logic or Bayesian makes no difference), is analogous to a linguist trying to ask whether speakers of AAVE conform to rules of 'good English' – whatever those may be. At best, the question is relatively uninteresting; at worst, it does not make any sense.

What, then, about instrumental or adaptive rationality? Prima facie, instrumental rationality involves no norms, no 'ought', and therefore no is-ought or naturalistic fallacy. Insofar as we can keep the normative question out of the game, insofar as we can regard instrumental rationality at a purely descriptive level (whether computational or algorithmic makes no difference), I have no problem with it. However, the question is far from obvious: for instance, Stanovich (1999, p. 243-244) maintains that by his framework, both rationality₁ and rationality₂ are cases of normative rationality, since they both relate to personal goals. The problem whether instrumental rationality should be considered normative is beyond the scope of this paper. However, this does not alter the main conclusion of this paper: *normative* rationality is not amenable to empirical investigation, so we cannot keep the dual notion of rationality.

Linguistics only came of age when it discarded its historical obsession with norms. Linguistics as a science coincides with beginning to pay attention to what speakers actually say, rather than how they should say it. The obsessive back-and-forth dialogue with normative theory is a peculiarity of reasoning and decision making theories (Evans, 1993). Similarly, dual process theories of reasoning should be liberated from dual-process 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 will only benefit from a clear separation. Let us leave the 'ought', then, to clergymen and politicians, and concentrate on the 'is' instead.

ACKNOWLEDGEMENTS

I am grateful to Jonathan Evans for comments on an earlier draft and to Helen Neilens for useful observations.

Reference List

Anderson, J. R. (1990). *The Adaptive Character of Thought*. Hillsdale, N.J.: Erlbaum.

Byrne, R. M. J. & Handley, S. H. (1997). Reasoning strategies for suppositional deductions. *Cognition*, 62, 49.

Byrne, R. M. J., Handley, S. J., & Johnson-Laird, P. N. (1995). Reasoning from suppositions. *The Quarterly Journal of Experimental Psychology*, 48A, 915-944.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, 4, 317-370.

Cohon, R. (2004). Hume's moral philosophy. In E.N.Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2004 Edition), URL: <http://plato.stanford.edu/archives/win2004/entries/hume-moral/>. (.<

Edgington, D. (1995). On conditionals. *Mind*, 104, 235-329.

Edgington, D. (2003). What if? Questions about conditionals. *Mind & Language*, 18, 380-401.

Elqayam, S. (2003). Norm, error and the structure of rationality: The case study of the knight-knave paradigm. *Semiotica*, 147, 265-289.

Elqayam, S. (2006). The collapse illusion effect: a pragmatic-semantic illusion of truth and paradox. *Thinking and Reasoning*, 12, 180.

- Elqayam, S., Handley, S. J., & Evans, J. St. B. T. (2005). The limits of supposing: Semantic illusions and conditional probability. In *Proceedings of the 27th Annual Conference of the Cognitive Science Society* (Hillsdale, NJ: Lawrence Erlbaum).
- Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin and Review*.
- Evans, J. St. B. T. (1993). Bias and rationality. In K.I.Manktelow & D. E. Over (Eds.), *Rationality: Psychological and Philosophical Perspectives* (pp. 6-30). London: Routledge.
- Evans, J. St. B. T., Handley, S. H., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29, 321-355.
- Evans, J. St. B. T., Handley, S. J., Neilens, H., & Over, D. E. (2005). The mental representation of abstract conditional sentences: A study of qualitative individual differences. *Unpublished manuscript, University of Plymouth*.
- Evans, J. St. B. T. & Over, D. E. (1996). *Rationality and Reasoning*. Hove: Psychology Press.
- Evans, J. St. B. T. & Over, D. E. (2004). *If*. Oxford: Oxford University Press.
- Evans, J. St. B. T., Over, D. E., & Handley, S. J. (2005). Suppositions, extensionality, and conditionals: A critique of the mental model theory of Johnson-Laird & Byrne (2002). *Psychological Review*, 112, 1040-1052.

- Evans, J. St. B. T., Over, D. E., & Manktelow, K. I. (1993). Reasoning, decision making and rationality. *Cognition*, 49, 165-187.
- Frankena, W. (1939). The naturalistic fallacy. *Mind*, 48, 464-477.
- Gottwald, S. (2001). *A Treatise on Many-Valued Logics (Studies in Logic and Computation, vol. 9)*. Baldock: Research Studies Press Ltd.
- Hume, D. (2000). *A treatise on human nature (Original publication date 1739-1740)*. Oxford: Clarendon Press.
- Johnson-Laird, P. N. & Byrne, R. M. J. (2002). Conditionals: a theory of meaning, pragmatics and inference. *Psychological Review*, 109, 646-678.
- Kleene, S. C. (1952). *Introduction to meta-mathematics*. Groningen: Wolters-Noordhoff.
- Kühberger, A., Komunska, D., & Perner, J. (2001). The disjunction effect: Does it exist for two-step gambles? *Organizational Behavior and Human Decision Processes*, 85, 264.
- Margolis, L. (1987). *Patterns, Thinking and Cognition: A Theory of Judgement*. Chicago Il., University of Chicago Press.
- Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco: Freeman.
- Martin, R. L. (1984). (Ed.) *Recent essays on truth and the Liar paradox*. Oxford, UK: Oxford Univerity Press.

Moore, G. E. (1903). *Principia ethica*. New York, NY: Cambridge University Press.

Newstead, S. E., Handley, S. H., Harley, C., Wright, H., & Farelly, D. (2004).

Individual differences in deductive reasoning. *Quarterly Journal of Experimental Psychology*, 57A, 33-60.

Oaksford, M. & Chater, N. (1998). *Rationality in an Uncertain World*. Hove, UK:

Psychology Press.

Oberauer, K. & Wilhelm, O. (2003). The meaning(s) of conditionals: Conditional

probabilities, mental models and personal utilities. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29, 680-693.

Ramsey, F. P. (1931). *The foundations of mathematics and other logical essays*.

London: Routledge and Kegan Paul.

Rescher, N. (1969). *Many-valued logics*. New York, NY: McGraw-Hill.

Rips, L. J. (1989). The psychology of knights and knaves. *Cognition*, 31, 85-116.

Saussure, F. d. (1959). *Course in General Linguistics (Original publication 1916)*.

New York, NY: McGraw Hill.

Savage, L. J. (1954). *The Foundations of Statistics*. New York: Wiley.

Schroyens, W., Schaeken, W., & d'Ydewalle, G. (1999). Error and bias in meta-

propositional reasoning: A case of the mental model theory. *Thinking and Reasoning*, 5, 65.

Stanovich, K. E. (1999). *Who is Rational? Studies of Individual Differences in*

Reasoning. Mahway, NJ: Lawrence Erlbaum Associates.

Stanovich, K. E. & West, R. F. (1998). Cognitive ability and variation in selection task performance. *Thinking and Reasoning*, 4, 193-230.

Stanovich, K. E. & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences*, 23, 645-726.

Trudgill, P. (2000). *Sociolinguistics : An introduction to language and society*. (4th ed.) London, UK: Penguin.

Tversky, A. & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological Science*, 3, 305-309.

Van Fraassen, B. (1966). Singular terms, truth-value gaps, and free logic. *Journal of Philosophy*, 17, 481-495.

Van Fraassen, B. (1969). Presuppositions, supervaluations, and free logic. In K.Lambert (Ed.), *The logical way of doing things*. (New Haven and London: Yale University Press.

Table 1

The four types of normative conflict

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	Metaductive conditionals