The Methods of Ethics. Conflicts Built to Last

Björn Eriksson
Stockholm University

Introduction

An impressive amount of evidence from psychology, cognitive neurology, evolutionary psychology and primatology seems to be converging on a 'dual process' model of moral or practical (in the philosophical sense) psychology according to which our practical judgments are generated by two distinct processes, one 'emotive-intuitive' and one 'cognitive-utilitarian'. In this paper I approach the dual process model from several directions, trying to shed light on various aspects of our moral and practical lives. I first outline this model and hint at some of the evidence for it (section 1). I then show how the dual process model may explain how we are able to handle some practical decision problems (section 2). After that follows a discussion of some of the psychological mechanisms by which the dual processes involved in the model interact (section 3). After these preliminary matters follows a discussion of the normative relevance of the dual process model. A perspective according to which our long lasting normative disagreements are the output of a practical decision system in perfectly good working order is proposed and defended (section 4). Lastly the question of the model's possible implications for explanation of intentional action is given a brief treatment. (section 5).

1. The dual-process model: a sketch. Resources and problems.

According to the dual process model our moral decisions are driven by emotion *and* cognition. These different kinds of processes do however not work in unison or harmony, but to the contrary they compete, often yielding conflicting results. I start with a characterization of the emotive side of the practical decision systems. I'll come back to the cognitive side in a while.

Emotive moral judgements. Evolutionary psychology and primatology stress the role of emotion in moral decision making. From primatology we know that moral behavior (or at any rate behavior which we are prone to regard as very moral-like - a fact which is just as telling) is widespread among primates: they engage in complicated 'moral' interaction in spite of lacking our cognitive functions. The suggestion is that primate morality is driven by social emotions targeting salient kinds of interaction such as cooperation, reciprocation, violence and helping behavior. We, being the descendants of our progenitors, should be expected to show some of these social-emotive capacities. And since humans are very apt at social learning and we should expect that in us these emotive capacities are culturally calibrated thus allowing for and explaining moral variance even in emotively driven moral judgement.

Cognitive neurology adds to the picture. From Phineas Gage (Damasio 1994) and onwards the role of brain areas and processes implicated in emotion has consistently been seen to play an indispensable role in moral judgment (Damasio 1994). So called frontal patients (with damaged ventral and medial portions of the frontal lobes) are known to be emotionally blunted and they are

also poor practical reasoners. Their IQs and other cognitive functions seem however not to be affected. Those that are damaged later in life also retain a social and moral knowledge that they, however, find strikingly difficult to translate into action and decision. Patients with early damage to the relevant areas lack even this 'dead' social competence (Bechara 1996, Anderson 1999). So the emotive processing normally effected in the ventral and medial portions of the frontal lobes seems necessary both for social-moral learning and for putting the result of such learning to work in moral and social decision. This focus on the emotive aspect of the practical decision is well in accord with observations made using fMIR imaging during processing of various kinds of practical problems. When subjects are mulling over morally laden proposition (*they hung an innocent*), activity is increased in the areas implicated in emotional processing mentioned. These areas are not active to the same degree when the subjects are pondering non-moral propositions (*stones are made of water*) (Moll 2001).

Recent psychological studies on moral judgment and reasoning also provide evidence that our 'moral minds' have a clearly emotive aspect. Normally we produce moral judgements immediately, effortlessly and without any conscious reasoning taking place (Haidt 2001). Sometimes we do engage in reasoning of a distinctly cognitive flavor, but then often post hoc and normally to convince others and reach consensus among our peers. So, in this sense too, reason seems to be 'the slave of passion:' it is the ambassador of emotion. The judgments delivered by the emotive processes so far sketched are triggered by specific kinds of information, normally with some social relevance and processed below the level of conscious cognitive processes. Cognitive processes enter, if at all, later to spread the word to those whom we find it worthwhile to convert. Haidt calls this model of our moral-practical mental processing the 'social intuitionist model' underlining the fact that the judgements so produced have a distinctly 'intuitive' flavor and are triggered by socially entrenched emotions. Haidt argues that not only is this emotive-intuitive aspect of our moral 'thinking' common, it is dominant. What little cognitive moral thinking there is, is post hoc and serves communication and negotiation rather than actually affecting the results of moral deliberation. Such a one-sidedly emotion driven picture may seem too simpleminded. It gives too little room for the role of conscious cognitive reasoning in moral matters.

The dual process model. John Bolender has recently argued for a somewhat more balanced picture of our practical decision making where cognitive processes are allowed to play a more active role (Bolender 2001). And Joshua Green and collaborators have conducted several studies that underscores the role cognitive and what Bolender calls 'informationally unrestricted' processing of moral matters plays in certain kinds of decision situations (Green 2008). Such cognitive processes are systematically associated with utilitarian judgments in a broad sense.

Though the work already done seems to provide valuable insight in the workings of the processes behind our moral and practical decision making, much work remains to be done. In particular it seems that no one has a very good idea about which kinds of stimuli elicits typical emotive-intuitive judgments, in contrast to cognitive-utilitarian moral responses. Green started off in 2001 by proposing as a first approximation that 'personal' dilemmas elicit 'emotive' moral responses whereas 'impersonal' stimuli elicits 'cognitive' moral responses. Personal dilemmas are those where direct personal harm is imminent, such as pushing a bystander in front of a runaway trolley. Impersonal dilemmas are those where there is no harm in the offing or else the harm is caused only indirectly - by pulling a lever, say - and/or where the victims are remote and/or unknown or undetermined at the time of action. The very same problem at hand can elicit responses of both kinds, thus possibly creating a mental conflict that has to be resolved in order to reach a decision, something that takes time (that can and have been measured by Green and collaborators [2004]).

Bolender is more sketchy in his account of what triggers the intuitive processes of the moral mind, but seems to take a somewhat more computational stand towards the intuitive and informationally restricted aspects of moral problem processing than Green. Recently Mikhail (2008) has proposed an alternative to Green's clearly not fully adequate distinction between 'personal' and 'impersonal' dilemmas. His alternative is also computational and along the lines of a 'moral grammar' that categorize stimuli and applies normative principles of the categorized stimuli thus providing judgmental output. It seems to me that what is valuable in this Mikhail's computational-grammatical approach can be adopted for improving on Green's first approximation. The crucial things are that the 'grammar' works with categorizations that make sense from an evolutionary perspective, since *ex hypotesi* the emotive-intuitive aspect of the decision processes was formed by evolution. Also the output must be emotive in kind rather than cognitive and the grammatical computations should be of the effortless kind characteristic of 'intuitive' processes.

I what follows I will indicate the as yet poorly understood distinction between stimuli that typically elicits emotive-intuitive responses and stimuli that elicits cognitive responses by talking about 'old' and 'new' moral problems. 'Old' ones being problems of a kind that evolution has had a say on, 'new' ones being problems we have to solve without the help of Mother Nature. Problems can be complex, having both old and new aspects, thus provoking moral conflict. (Below I propose that the distinction between 'old' and 'new' problems should be developed on the basis of a fuller range of 'moral emotions' and their triggers rather than on the meagre conception of direct personal harm.)

The nature of the cognitive processes are roughly speaking utilitarian, paying heed to what happens as a result of alternative actions and to how good or bad that is. If the emotive-intuitive processing is silent, we process our problems by evaluating the consequences of our alternatives. This is also what is observed in patients with frontotemporal demential which makes them emotionally impaired: they become more utilitarian in their judgments. (Mendez, Anderson & Shapira 2005) But the very problems at hand can presumably also be of *kinds* that don't awake our emotive-intuitive processes.

One aspect of the evolutionary origins of our emotive-intuitive moral processing that Green doesn't mention is that one important job for our emotive-intuitive moral capacities is to act as a bulwark against cognitive judgment. There is always the temptation to do as Hobbes' 'Foole' and think that one should go for maximum pay-off instead of keeping agreements and cooperating. Bolender argues, rightly I think, that our decision processing systems are the result of an evolutionary trade-off between the *effective* and *fast* emotive-intuitive judgments and the *flexible* and *opportunistic* cognitive-utilitarian judgements. One important function of the emotive-intuitive processing is thus to disarm over-eager calculation. This competition from utilitarian calculus has in latter days - after human cognitive capacities began to approach present levels - also been one of the most salient aspects of the environment for our evolving emotive-intuitive responses. The ability of the emotive-intuitive responses to defeat the verdicts of utilitarian calculus provides one of their most striking benefits as we shall see below. It also goes the other way: one evolutionary *raison d'être* for our cognitive-utilitarian capabilities may be to silence the rough and ready emotive-intuitive responses. This aspect of the competition has presumably been the more important since the emotions in question probably has been present for a longer time than our higher cognitive functions.

Thinking about our practical reason in terms of competing processes invites a speculative hypothesis contributing to the question 'Why are we humans so much smarter than everyone else?' The phenomenon of arms-races in evolution is well-known and generally taken to be a powerful kind of evolutionary process. Gazelles and leopards are very, very fast. Why? They have tried to

outrun each other for thousands of years. Now, why are we so smart? Perhaps because our emotive-intuitive and our cognitive-utilitarian decision processes have 'tried to' out-smart each other for quite a while. Of course, our two kinds of processes are no evolutionary 'units of selection' so the analogy with fully fledged organisms like leopards not serious. But given that there is an advantage in having 'balanced' practical mind where the powers involved have the right relative force against each other, *and* given that there is some selective pressure towards increasing the force of any one (or both separately) of the two processes involved, there will be pressure for the 'other' process to keep up. (Perhaps the evolutionary dynamics of the balance works like this: equilibrium is always better, and adjusting upwards is always better.) This would then lead to, let's say, 'heightened consciousness' as regards handling of both 'old' and 'new' problems. This in turn is surely an aspect of smartness in the wider sense. Anyway, this illustrates that the dual process model invites speculations that may be turned into serious hypotheses, fit for of serious testing.

The general picture of practical decision making that emerges from the dual process model is that when confronting a moral problem, it is first determined whether there is anything 'old' about it. If it is, then the emotive-intuitive processes kick in and produces the material for an emotionally underwritten moral judgment. If there is nothing 'old' about it, our moral minds skip forward to the cognitive processing and produce a 'utilitarian' judgement based on an evaluation of the the consequences of the alternatives at hand. If the problem shows both 'old' properties - so that the intuitive-emotive processing kicks in - *and* salient 'new' properties, both systems work on the problem and eventually produces a 'winning' judgment of either the emotive-intuitive kind or the utilitarian kind. The 'competition' between the two processes leave empirically accessible traces. Some are only too painfully accessible to introspection, witness the familiar phenomena of having residual doubts, feelings of guilt and the sense of there being genuine moral dilemmas. All these phenomena may plausibly be explained by the dual process model. The 'competition' is also visible in fMRI imaging. During the deliberations that eventually terminates in a judgment where the emotive-intuitive responses have been subdued by cognitive-utilitarian concerns, brain areas implicated in 'cognitive control' show heightened activity similar to the brain activity patterns shown in subjects faced with the color-naming Stroop task. (Greene *et al*. 2004)

The true picture is bound to be much more complicated than this. In particular it seems that the idea that there are *two* distinct processes is a simplification. This simplified perspective may also account for the difficulties Green *et al.* have had in specifying two different 'job-descriptions' for the two processes, what I call 'old' and 'new' problems. Work in evolutionary psychology, in particular the work of Cosmides and Tooby, shows that it is likely that there is 'a number of of different evolved specializations, each of which is functionally specialized for negotiating a particular domain of social life.' (Cosmides and Tooby 2008: 58). In this context, however, I shall bypass this complication and interpret the dual process model as in effect saying that there are two *kinds* of processes, one characterized by having a social focus and being emotionally driven, and one kind characterized by not having any such restriction in focus and not being emotionally driven. For now I'll go on and tuck away this under the blanket terms 'old' and 'new' problems. I'll also continue say 'two processes,' meaning 'two kinds of processes.'

I conclude that the total evidence for something along lines of the dual process model is strong enough to merit discussion even at the present time, when many questions are still unanswered.

Thus, the background for the discussion to follow will be this.

1 There are two separate information processing systems that are involved in producing moral output, moral judgments and perhaps practical judgment generally.

2 One process is created by evolution for specific problem solving in social contexts and handles 'old' problems by yielding emotive-intuitive responses to specific stimuli. The emotive-intuitive responses are translated to emotive-intuitive moral judgments.

3 The other process is informationally unrestricted and can process any kind of problem yielding 'cognitive-utilitarian' judgements based on considerations of evaluations of consequences of alternative actions.

4 These processes can yield conflicting results in concrete situations.

5 Normally the defeated process in a conflict don't disappear, but leaves traces in the form of excuses, rationalizations, confabulations and in general the stuff that moral philosophy is made of.

2. The dual processes and practical dilemmas.

In this section I argue that the existence of two distinct and to a large extent independent processes of the kind outlined above partly explains both the existence of certain practical decision problems and our capacity to 'solve' them by transforming them. Though there are numerous problems and decision situations that are of relevance to this discussion I shall focus on three: The Prisoners' Dilemma, The Ruined Lawn (and Global Warming) and The Rigid Deontologists.

The prisoners' dilemma. This is really to well-known to be described, so I will not do that. The dilemma of the dilemma is, as we know, that the prisoners will go for inferior outcomes (each one individually) if each one doesn't pursue a clearly irrational strategy. The relevance for my discussion is that we seem to be quite apt at steering clear of PDs. The present hypothesis is that what allows us to do that is that the emotive-intuitive tire of the two-tire model of practical reasoning may kick in and transform a threatening PD to a harmless, indeed possibly profitable, occasion for cooperation. Where we *only* equipped with a rational utility calculator as some students of economics, patients with frontotemporal dementia and similar impairments seem to be, we would be trapped in our PDs.[1] But sometimes we escape by heeding our emotive-intuitive judgements about fairness and just contribution. If this is along the right lines we should expect that PDs that are set so that the concrete problem at hand is 'old' rather than 'new' are more consistently evaded by the 'players'. This could be tested, if we had a more robust interpretation of 'old' and 'new' problems than the one I have given.

Global Warming and ruined lawns. There are probably as many utilitarianisms as utilitarians. In the present section I mean individual act-utilitarianism by 'utilitarianism' and 'consequentialism'. (Otherwise I use these words in a wider sense, mostly indicating a rough contrast to deontological intuitions.) Individual act utilitarianism says that each agent on each occasion ought to do the act that produces most value overall of the acts available to her on that occasion.

---

[1] Actually there are two distinct issues at play here. One is how it can be that we are equipped with anything at all that allows us to avoid PDs. This kind of question is discussed and answered for example by Skyrms (1996) and Okasa (2007). Here I address another question: Given that we are equipped to avoid PDs, *how* do we do that. The same *mutatis mutandis* goes for the other two decision problems discussed.

Imagine that some of the individual act-utilitarians have formed a club for promotion of the greater good. Between the Utilitarian Club House and the parking lot is a beautiful lawn, much appreciated by the members. However, they all correctly judge that the disvalue of the harm to the lawn caused by any one walk across it is far less than than the benefit to the member crossing the lawn on her way to the parking lot. So they all agree that each member ought to cross the lawn. It is soon ruined, all in good utilitarian spirit. The members mourn their lawn and curse their cherished moral theory that got it ruined as a result of successfully following its principle. They also note that this is only a triviality compared to where their theory will land their descendants in the not so distant future: in a heated world with much worse conditions for pleasurable living. They will contribute to such a future because what each one of them do individually does not matter for global warming, even in the long run. But it does matter for their present quality of life. So in good utilitarian spirit they drive their SUVs and produce considerable amounts of greenhouse gasses. The relevance of the ruined lawn and global warming to the present discussion is that unrestricted albeit successful application of a value-maximizing decision rule to particular decisions may result in a sub-optimal aggregate of decisions. It seems that for hard core utilitarians of the act-utilitarian flavor this is a constant threat. I shall argue that the dual process model outlined above contains the resources to explain why and how utilitarians manage to escape tricky situations like these. To glance ahead one explanation is that output from the emotive-intuitive processes can tilt the utilities so that what otherwise would have been the value maximizing alternative turns out to be inferior. Another is that the emotive-intuitive output may fool the utilitarian into thinking that it really *is* for the best to mend his ways. In the first case the gain from crossing the lawn is diminished by the costs of counteracting the intuitive prohibition against participating in a collectively sub-optimal activity. In the second case the intuitive prohibition is directly translated to a biased value-calculus that inflates the disvalue of crossing the lawn.

The prisoners' dilemma and the lawn/SUV cases are quite similar. Correctly aggregating and aiming for maximizing the *overall* value does the trick if one can choose the right sets of actions to aggregate, as it would do in a prisoners' dilemma. If only the prisoners/utilitarians could go for the highest overall pay-off obtainable by aggregating the value of the optimal set of actions, they would be off the hook. So one problem is the agents' separateness: the fact that they each act independently of the others. If their actions were tied together somehow so that each agent in fact decided not only what she should do now, but rather what should *happen* as a result of some collective action, things would look much brighter. The ruined lawn and SUV-driving examples present thresholds that are passed by no one in particular, so the disvalue of passing them is not retraceable to any one of the actions that produced them. In this sense the disvalue is not a result of the individual actions, but only of the aggregate of actions. That is why the separateness of the agents matter.

Glover and Parfit and moral mathematics. It may be thought that problems such as the ruined lawn and SUV-driving are solved by modern utilitarianism. Jonathan Glover (1975) and Derek Parfit (1984) has treated problems like these and shown that enlightened utilitarians can show (and see) that lawn crossing and SUV-driving is wrong even though the negative effects are very very small or indiscernible. That is wrong. To the extent that Glover and Parfit have shown how utilitarians can get out of problems like these they have shown how to cleverly make exceptions to individual act-utilitiarian theories or how to revise the values at stake so the problems disappear. That is also how my own analysis will turn out as we shall see, so let me for now only as a parenthesis argue that Glover and Parfit in something resembling *mauvais foi* does not really succeed in rescuing utilitarians from global warming and lawn ruining. They provide means for utility-revisions that

make the problems disappear or they provide motivations for disregarding the utilitarian principle in certain cases.

Glover defends the 'principle of divisibility' that says that even actions that only cause harm that is sub-discernible are wrong. They are wrong 'to the extent that they cause harm' so that if one lawn-crosser causes a sub-discernible deterioration of the lawn that amounts to a 100th of a discernible deterioration, then the lawn-crosser has caused harm and acted wrongly to that extent. His story about the bandits who eat all of the villagers' food while causing no discernible harm to them is intended to prove his point. Each one of the bandits only take one baked bean from each one of the villagers' plates. That 'harm' is imperceptible, so no-one causes any discernible harm to anyone. But the villagers are left starving, since there are just as many bandits as there are villagers and beans on each plate. According to the principle of divisibility each of the bandits cause one fraction of a discernible harm to each one of the villagers and acts wrongly towards each villager, as it where, to that extent. He also seems to hold that all the wrongs done to all the villagers by any one bandit should be aggregated to one wrong done by that bandit: the sum of the fractions of discernible wrong done. This seems attractive if the only other way to judge the case is that the bandits do no wrong at all.

However, when Glover goes on to discuss whether the principle of divisibility is applicable in cases where even the *aggregated* effects are below the discernibility-threshold, he argues that for a utilitarian the principle may be unacceptable in *those* cases. Discussing an example where someone over-uses electricity in times of power rationing he writes: '…it seems absurd for a consequentialist who is certain that the threshold will not be reached to refrain from using electricity although he knows that this will in no way avoid any detectable discomfort or inconvenience to anyone' (1986: 129). Indeed, it does seem absurd. But it seems equally absurd in cases where the threshold *is* reached, I submit. In such cases too, the reasons for regarding it absurd applies: the consequentialist 'knows that this will in no way avoid any detectable discomfort or inconvenience to anyone.' Utilitarians must stick to this judgement if they are to remain utilitarians. Still the principle of divisibility seems attractive, even for utilitarians, I think. Which means that it seems attractive in cases of this kind to attend to other things than consequences of particular actions and their values when deciding how to act. In short: it seems attractive to find a way of paying heed to the voice of our emotive-intuitive moral processes. The principle of divisibility is a way to do just that, and it may seem consistent with utilitarian thinking, but that is not right. Only because the matter is so convoluted we may think it is.

The straightforward utilitarian analysis of the bandits is that when each only causes an imperceptible harm to the villagers, they do, individually, no wrong. But collectively they do. *They* act wrongly but no-one among them do. The trick we have to manage is to see the situation, individuals that we are, from the point of the collective. Individual utilitarians cannot do that *qua* individual utilitarians, but their capacity for emotive-intuitive moral processing is (in fact, I think) built for helping them to do just that and similar things.

I think Parfit's discussion of mistakes in moral mathematics can be viewed in roughly the same light. Parfit's discussion is quite complex, so I will just pick one part of it and fearlessly claim that what I say about this goes for the rest of his discussion of moral mathematics too. Consider Parfit's

> *Case Two: X* tricks me into drinking poison, of a kind that causes me a painful death within a few minutes. Before this poison has any effect *Y* kills me painlessly. (1984: 70)

Since *Y* does not harm me in Case Two, Parfit holds, the wrongness of his act must be explained without reference to *Y*'s harming me. How? By appealing to

> *C7:* Even if an act harms no one, this act may be wrong because it is one of a set of acts that *together* harm other people. Similarly, even if some act benefits no one, it can be what someone ought to do, because it is one of a set of acts that together benefit other people. (*loc. cit.*)

Again, as was the case with Glover, the straightforward utilitarian response is rejected: *Y* acts wrongly in *Case Two*. But not because the consequences of what he does; he doesn't harm anyone according to Parfit. The wrongness is explained by *C7*: '…in *Case Two*, *Y* does not harm me . *Y* acts wrongly in *Case Two* because he is a member of a group who together harm me.' (*loc. cit.*) This may be the best answer. But it is not a very utilitarian answer. However, Parfit's 'C' in '*C7*' stands for 'consequentialism.' To my mind seems better interpreted as a proposed exception from strict consequentialism, Parfit obviously sees it as an integrated part of an amended consequentialism. I think that *C7* and other principles defended by Parfit in his discussion of mistakes in moral mathematics should be seen as departures from utilitarian thinking motivated by deliverances from our emotive-intuitive practical decision processes. So, to conclude this parenthesis I hold that in spite of their seeming utilitarian leanings both Parfit and Glover actually proposes exceptions to utilitarian problem-solving in cases like the ruined lawn and the innocuous SUV-driving, rather than utilitarian ways out of the problems. Now to the third kind of decision problem.

The unreasonable Deontologists. Having seen what utilitarianism leads them to do, the utilitarian club members dissolve the club and form a new Deontological Club. No one walks across the lawn in front of the club house anymore and they all take the bus since they have disposed (for recycling) of their SUVs. Indeed, what they have done is to systematically disregard their individual act-utilitarian judgments and started trusting their intuitions. This constantly lands them in trouble. They frequently forgo opportunities to do good that seem counter-intuitive and immoral to them. For example they become wary even of giving promises (because they *must* be held), so many opportunities for fruitful cooperation are foregone. They are *driven* by their rigid moral intuitions. We, however, seem to be quite good at avoiding to become rigid deontologists. The explanation of this, as the reader will guess, is that the 'utilitarian' side of our practical reason is difficult to silence. Just as the intuitive processes can kick in and save us from too much smart utilitarian or rational calculation, the smart calculators in us kick in and temper the verdicts of our intuitive thinking.

3. The psychological economy of our dual moral minds

Bolender presents his version of this dual-process model of the moral mind as a 'two-tire model'. The picture that emerges from these sketches of how we solve practical dilemmas by playing out the two parts of or moral minds against each other is, however, more a picture of a split personality than something resting in equilibrium on its two tires. It is more like two *persons* riding a single bike than like the tires of the bike.

The question in this section is how the two kinds of practical decision processes are integrated in us and sometimes by us so as to deliver *one* decision, and sometimes the wise decision. There seems to be various ways for the two tires of our practical reason to interact so as to help us navigate the pitfalls of calculation and intuition. Starting with the emotive-intuitive processes there seems to be

two main ways for them to hold sway over decision under the threat of utilitarian calculation: revision of utilities and effacing.

Revision of utilities. Our emotive-intuitive decision processes can effect a revision of the utilities in various ways. First there are emotional responses that make what would otherwise have been attractive seem less so or even disgusting. Pocketing a dropped wallet starts to seem a rotten thing to do when you consider the fact that the wallet's owner always has been kind and even generous towards you. It becomes more costly for you to take the wallet. This may be tied to the emotions surrounding one's conception of oneself: if I take the wallet I get, so to speak, a dent in my moral character and that is a disutility. In a more positive vein we may also have our utilities revised 'upwards' as when an action can be seen as a step in a character building project. Returning the dropped wallet to its owner may increase its utility if the agent sees herself as engaged in a project of reforming her character from being an opportunistic exploiter of others to a responsible member of the community. There are many variations on these themes, but the central point is that they all work through emotions and emotionally charged descriptions of our alternatives. This goes for the proactive emotional responses like anticipation of pride or shame. The reactive responses, such as anger or remorse, aimed at correcting past actions and their effects, work in the same way. Being angry at the person I saw pocketing my dropped wallet increases the utility of taking action against him, as well as tending to increase my chances of success in rectification. Feeling ashamed may drastically increase the utility of making amends. Reactive emotional responses too are dependent of viewing the alternatives at hand under the relevant descriptions.

Connecting this to the hand-waiving terminology above where 'old' problems (prone to emotive-intuitive processing) and 'new' problems (prone to utilitarian-cognitive processing) were distinguished, the gist of the above can be stated as follows: our ability to counteract the deliverances of utilitarian-cognitive processing by having our emotive-intuitive processes revising the values of the outcomes is a function of our ability to focus on 'old' rather than 'new' aspects of the problems at hand. I suggest that improvement on the 'personal - impersonal' distinction introduced by Green could be done by connecting the emotive-intuitive processes to 'moral emotions' in general, rather than those specifically evoked by personal direct harm. Such emotions as anger, shame, self-esteem and compassion seems to be no worse off than aversion towards direct personal harm as candidates for fitting into a relevant evolutionary explanation.

So having access to utility revision is partly a matter of having access to the right kinds of descriptions of what one does. These kinds of utility revisions are effected by emotions such as guilt, anger shame and pride which are in principle very forceful but negotiable emotions. One can accept pangs of guilt if the reward is big enough. There are also utility revising emotions of a more decisive kind. Coming to see an act or its consequences as disgusting or heavenly may very radically alter the 'utility landscape' so that there is nothing to negotiate about: the case is clear, there is no temptation. These kinds of stronger emotional utility revisions borders on the second way of our intuitive decision mechanisms to influence decisions that would otherwise driven by our calculating decision mechanisms alone: effacing of alternatives.

Effacing of alternatives. Sometimes an alternative action will strike one as unthinkable or undoable. Indeed, sometimes it doesn't strike one at all because one does not even consider it as an alternative. I could save burial costs and earn some insurance money if I were to burn down the house were my newly deceased mother's corps lies. (Only respect for the reader's squeamishness prevents me from even more pertinent examples.) However, that thought doesn't occur. In these cases the very thought of the alternative is blocked through the continuing working of taboos of

various kinds. Though there is no emotion involved in *not* thinking about a disgusting action, such taboo-driven effacings of alternatives have an emotional underpinning. The very taboos that direct and restrict our attention are upheld emotionally. They are emotionally sculpted habits of practical thinking.

Apart from emotionally driven or upheld utility revisions and alternative effacings I want to mention another and quite peculiar way for the emotive-intuitive processes to influence decision: promises. Promises and their effects have not to my knowledge received any notable attention from psychologists, which is a bit strange. But the non-existence of empirical studies of promises and their effects leaves the field open for enlightened speculation. As I see promises they are 'portable' and 'instant' utility revision devices. Roughly the way they work is like this: by promising you change the relative utility pay-off to yourself of the promised course of action either by simply increasing it or by decreasing the utility of the salient alternative (the temptation). You do this normally by saying that you promise, but the real work is done by the resulting utility change: that is what makes it a *sincere* promise. I guess that the promising device works by exploiting the proactive emotional mechanisms mechanisms just mentioned. The workings of this familiar psychological device is more complicated than this, but for now this very rough sketch will have to do.

I thus submit for consideration that proactive and reactive emotional responses and the promising device exploiting the proactive responses are paramount in shaping our emotive-intuitive tempering of our utility calculating decision processes. But what has the cognitive-utilitarian calculating mechanism to put up against that competition?

The answer is emotive reinforcement on the one hand and emotive weakening on the other. As utility calculators we may be able to stick to our guns in the face of countervailing emotions by either hijacking emotive strength so that the output of our calculations becomes armored by emotive force. We may also weaken the emotive strength of the countervailing considerations so that they loosen their grip on us. The former may happen when, for example, we make the suffering of anonymous multitudes 'vivid' so that our thoughts about them get access to our emotive resources, as it were (Kagan 1989). We turn the 'new' problem into an 'old' one by vividly representing the persons affected thus in a sense bringing them closer. We may then get our emotive-intuitive responses to line up behind the judgement that we should donate more to Oxfam rather than spending more time at our favorite sea-side resort. The second route to cognitive disciplining of our emotions may happen when we 'see through' the automatic workings of our emotions. As when we come to see that eating pork isn't really anything different (morally or in terms of health or quality of taste) than eating lamb. We may thus overcome the heartfelt resistance against pork eating or, say, women not wearing the right kind of veils or whatever. The result would not normally be that pork eating becomes emotionally positively valenced but rather that the aversion gets emotionally duller or disappear altogether. How this works I do not know. It seems that mere dwelling on matters may sometimes do the trick, as in the Oxfam case. Also habituation seems to play a role, as when medical students get habituated to handling dead bodies and students in the humanities get habituated to handling the writings of Jacques Lacan.

The upshot of these speculative musings on the tempering of our calculations and our emotive responses is, it seems, that we have a certain influence over our motivational and judgmental lives, so that we may influence what we, in the final analysis, decide to do. And this seems to be quite beneficial: hard core utilitarians, 'rational agents' in the sense of game theory and deontologists may in reality all get their edges blunted because of the workings of 'the other' decision process.

For utilitarians and 'rational agents' emotive-intuitive concerns comes to rescue and for deontologists the voice of utilitarian rationality provides needed flexibility. This is the good news. The bad news is that the two kinds of processes have their own, not fully understood, agendas: They have their own respective kinds of problems that they are prone claim authority over. So we should beware the way we see the world, and consider that emphasizing the 'old' or the 'new' aspects of a problem is a forceful way of influencing the decision reached in the end. Too bad that we don't know *yet* how to spell out this distinction.

I now turn to the more philosophical part of this discussion. The question is: given that the dual process picture of practical decision painted so far is correct, what are the philosophical implications, if any? I shall restrict my discussion to two kinds of implication: implications for normative issues and methodological implications having to do with the structure of explanation of action. I start with the normative issues.

4. Normative issues

Since the dual process model is purely descriptive or non-normative, it seems that no normative conclusions follow from it; one wouldn't want to be caught breaking Hume's law. But it can't be that simple. The dual process model is a theory about of the very processes that yield our normative decisions. The correct picture of how these processes work can in principle have relevance for the status of the normative decisions themselves, and thus indirectly for normative questions.[2] First we have the debunking arguments. If it can be shown that the decision-processes are flawed in some way, the conclusion would be that normative decisions stemming from them are undermined by that very decent. And conversely, if the mechanisms could be shown to be reliable, their output would then earn some credibility from originating from them. Also there is the notorious 'evolutionary ethics' actually trying to bypass Hume's law and reaching full-fledged moral conclusions from evolutionary premises. I'll briefly consider the reliabilist argument, the debunking argument and the evolutionary ethicist's project in turn.

Reliability. Is there a case for proclaiming one of the processes (or both in their respective spheres) reliable? No argument for such a claim seems to be available for the simple reason that there is nothing to measure reliability against but the output of our decision mechanisms themselves. In order to assess reliability independently one must be able to independently assess correctness or adequacy of relevant decisions. After having assessed the correctness of a particular decision one could then go on to see whether the decision in question also was condoned by one or both of our decision processes. But the two distinct steps required in this procedure are not available: the assessment of the correctness of the decision is performed by the very same processes the reliability of which is at issue.

---

2 Attemtps to break Hume's law are nowadays commonly equated with comitting the 'the naturalistic fallacy.' This could be called 'the naturalistic fallacy fallacy' if it were not for the fact that it strickly speaking is no fallacy, its just not correct. On the issue of supporting norms with statements of fact it seems that every philosopher worth his salt has to have a say. This in spite of there being a rough agreements of how the issue turns out: You can't derive norms from facts but you can support normative claims with statements of fact. Richards (1986) is one of a few dissidents who actually holds that you can derive norms from facts (because there are valid inference rules from facts to norms, he thinks). 'Duke naturalism' has recently been proposed as the name for en evolutionary ethical naturalism that leaves no room for norms (in distinction to facts) at all, it seems, and so deprives Hume's law of application. (Flanagan, Sarkissian & Wong 2007) 'Debunking arguements' like the one I discuss below do not strictly speaking have normative conclusions and so does not really infringe on Hume's law's territory.

When we confront practical or normative decision problems we have no other resources than our practical decision processing. This seems equally true when we ascend to a meta-level and judge the quality of the decision mechanisms themselves. Decisions with a deontological emotive-intuitive profile may be challenged from a utilitarian perspective and *vice versa*, but there is no respectable vantage point from which one could reasonably judge that e.g. the utilitarian perspective is right or better or less biased. One may think that normative moral theory could be called upon to resolve this issue. But it seems that normative moral philosophy is the very game our dual moral processes are playing against each other. It is not the neutral referee in the match between them. Anyway, the match is still going on and nowhere near ending, it seems. So no determined answer is presently at hand even if moral philosophers could in principle settles these matters.

Debunking. Is there a case for a debunking argument, proclaiming one or both of the processes non-reliable? Indeed I think there is a strong case for a debunking of morality and other non-hypothetical practical judgements to the effect that something is right or ought to be done. I will not, however, pursue that now. That is a wider issue. [3] What I want to discuss now is the more limited question of whether the dual process model gives rise to any new twists on the old debunking issue. One of the master minds behind the dual process model, Joshua Green himself, argues that the evolutionary origins of the emotive-intuitive processes specifically provide strong reason to reject the deontological ethical judgements they produce. The cognitive-utilitarian processes on the other hand he thinks are untouched by any such considerations. All in all, then, he thinks the dual process model speaks in favor of a more utilitarian and less deontological ethics.

Our emotive-intuitive responses, on which deontological judgments are based according to Green, are built to get us to convey our genes into future generations. They are not built to yield *right* or *warranted* practical decisions. They *are* built to yield *fitness promoting* kinds of behavior in the environment where our ancestors evolved. So we have no reason to trust that they yield right or warranted decisions if we have no reason to trust that tokens of historically fitness promoting kinds of behavior are right or warranted in any sense we should now honor. And we have no reason to believe that. *If* the interests of our genes and what is the right or warranted moral decision diverge, we know that our emotive-intuitive responses will stay loyal to our genes, that's what they are built for, among other things.

On the assumption that his empirical claims about the origin and function of our emotive-intuitive moral responses are correct, Green writes

> The *only* reason why we are motivated to make a moral distinction between nearby drowning children and faraway starving children are that the former push our emotional buttons and the latter do not. And the *only* reason we exhibit this pattern of emotional response is because we did not evolve in an environment, like our current environment, in which we could have meaningful interactions with faraway strangers. (Green 2008: 116)

Green admits the logical compatibility of the beliefs (1) that the *only* explanation why one is not moved by the suffering of faraway children refers only to morally irrelevant factors of our evolutionary history and (2) that this complacency is nevertheless perfectly justified. Though no contradiction is involved, Greene finds that combination of assertions 'uncomfortable.' Reformulated as an principle about warranted judgments and generalized, this lack of comfort could

---

[3] Many attempts at debunking morality has been made in modern times. The pioneer was Gilbert Harman (1970). Others worth mentioning are John Mackie (1977), Michael Ruse (1986) and Sommers and Rosenberg (2002). Richard Joyce (2006) provides the most recent and ambitious attempt. I also think Joyce is quite successful.

be expressed like this: If person *P* asserts the judgment that *X* only because *P* has certain emotive responses that are caused only by morally irrelevant factors, then *P*'s assertion of *X* is unwarranted. This seems attractive, but also highly problematic. First, it should be noted that there are *two* judgements at issue: (1) that one should rescue the nearby drowning child and (2) that one need not care very much about faraway starving children. *Both* these judgments are plausibly held, given Green's empirical results, because of 'certain emotive responses that are caused only by morally irrelevant factors'. So it seems that either both should be rejected or the principle about warrant should be rejected. In fact, since presumably all that deserves to be called part of human nature is caused by morally irrelevant factors, the principle about warrant seems to imply that there are *no* warranted judgements based on our human nature. The question how to stop this slippery slope from landing us in a through and through rejection of all that we value is according to Green 'among the most fundamental questions we face in an age of growing scientific self-knowledge'. (2008: 76) Elsewhere Green has tried to answer this question. He argues that the utilitarian parts of our moral minds (at least when matters of public life is concerned) is untouched by the discrediting argument above and so may retain some legitimate authority. He also holds that a more utilitarian outlook *follows psychologically* from seeing the truth about our moral psychology and its origins, meaning that seeing the truth makes us more prone to actually take a utilitarian stand (Green 2002).

The claim that exposure to the truth of the dual process model makes us more utilitarian, psychologically speaking, may be true about many of us. But I suspect that some may find that they are instead moved in the direction of down-playing the importance of moral concerns more generally so that they feel free to act on their deeply felt emotive-intuitive responses. But even if reflecting on the evolutionary debunking of judgments that are caused by our emotive-intuitive responses would in fact make us more prone to taking a utilitarian stand, this would mean nothing for the question of vindication of utilitarian thinking. Seeing the truth may make me more prone to depression, for example, or to take an self-centered amoralist stand. So the fact that some belief or attitude is the psychological effect of seeing the truth is not in general warrant for those beliefs or attitudes. Only if seeing the truth would make us more prone to take a utilitarian stand *in the right way*, we would have something like a vindication of utilitarianism. Since all I have to say about what counts as 'the right way' is that it would indeed provide vindication, that argument ends here for me.

Another problem with Green's argument against deontological thinking based on the origin of the emotive-intuitive responses is that it seems that the evolutionary origins of our utilitarian-cognitive thinking about practical matters is equally shabby and morally irrelevant: it seems unwarranted to hold that our cognitive-utilitarian decision processes are left untouched by the debunking argument from normatively irrelevant evolutionary origins. I guess that our general capacity to evaluate (or prefer or desire) and process judgements about pay-offs from alternative actions, has evolved from more or less self-centered concerns that have been somewhat expanded by such things as kin-selection and the dynamics of reciprocity. More fully utilitarian impersonal concerns seem to be based, psychologically speaking, on those morally irrelevant ego-centric concerns. The circle may be expanding, but I am at the center of it. This means that the explanation for any impersonal value-promoting views I hold refers only to morally irrelevant factors in the environments of my ancestors *plus* unwarranted expansions of the scope of those concerns. In short: that I care about myself is the result of morally irrelevant factors, that I care about others impersonally, if I do, is dependent of generalizing those self-centered concerns. It seems that there is an even stronger case for ruling out the authority of these beliefs and attitudes than there is for ruling out the authority of our emotive-intuitive responses. I conclude that the *partial* debunking argument Green advances fails. The suspect origins of our decision processing in general is reason to mistrust their claims in general. On

the other hand they are all we have for making decisions, so one important question is how we should best put them to use, when we have come to doubt the authority of their outputs. I'll come back to this question shortly, but first a quick look at evolutionary ethics.

Evolutionary Ethics. By 'evolutionary ethics' I mean ethical systems that are putatively vindicated in the epistemological sense by arguments where statements of evolutionary facts function as essential premises. There are many problems associated with this project. Some of the evolutionary ethicists only delivers instrumental justification, not epistemic. Dennett (1995) and Campbell (1996) seem to do just this. Or they deliver the wrong kind of normativity: the 'ought' of expectation or the 'ought' of biological function in stead of the moral 'ought'. Richards (1986) and Casebeer seem both to make this mistake (Joyce 2006: ch. 5). To add to these problems it seems that the dual process model spells trouble for evolutionary ethics in general. The standard project among evolutionary ethicists is to try and pinpoint a function for morality, and then to argue that moral decisions that are in line with that function, that for which morality evolved, are the right decisions. The most commonly attributed function is probably co-ordination in its various forms and settings (Lillehammer 2003: 569). The corresponding judgments that these decisions are indeed right are thus vindicated, it is argued. A moral theory that truly captures this function in a set of norms is the true or correct moral theory. However, the dual process model seems to imply that there is no one function of moral reasoning and decision. Our practical reason is poly-functional and, crucially, the various functions seem to be in perpetual conflict and perhaps even built for counteracting each other. Our practical reason is a complex mechanism for negotiating decisions and not one for reaching a predetermined goal, the empirical evidence is saying. It is conceivable that the sub-set of practical decisions that are moral decisions may not show this poly-functional and conflictual character. But since what counts as moral is quite a wide category, determined by such things as having to do with societal and individual flourishing; smooth and productive interaction; justice, reciprocity and reliability; purity, ceremony and tradition, the possibility seems remote that this wide and heterogenous category should succeed in singling out a mono-functional and co-realizable sub-set of practical judgments. So I take it that the conflictual character of practical reason trickles down to moral reason proper.

The standard evolutionary ethical argument tries to make the leap from statements about the proximal goals of our moral dispositions to vindication of those goals. Proximal goals in this context are goals we recognize and the recognition of which by us is the *means by which* evolution makes us promote 'its goals' (gene propagation). The standard argument runs like this:

1 It is morally right/obligatory to advance that which it is the function of our moral responses to promote as proximal goals.

2 It is the function of our moral responses to promote proximal goals *g*1,…, *g*n.

---

3 Advancing *g*1,…,*gn* is morally right/obligaory.

The poly-functional and conflictual nature of or practical decision processes ensures that there will be no co-realizable set of *g*s such that 2 is true. There may well be an non co-realizable set of *g*s that makes 2 true, but that will bring out 3 as an inconsistent moral principle, implying that one and the same action is right and wrong or obligatory and forbidden. So either 2 is false or 3 is. If morality is good for any one thing it seems to be something like getting along or counteracting the effects of limited sympathy (Mackie 1977). And that doesn't readily boil down to any set of moral principles.

Perhaps, the best that can be achieved along the lines of evolutionary ethics is an argument to the effect that morality as such - with its dual processes and inherent conflicts - gets *some* support from evolutionary thinking. Rather than vindicating any specific moral system of norms or any one of the two kinds of processes, the poly-functional conflict prone set-up of our moral psychology may get some vindication. If I am right, this is indeed where Campbell's argument should land him, whether he likes it or not. This would also reconnect to one of the other notorious problems for evolutionary ethics mentioned above: this somewhat fluffy vindication would in any case be of the wrong kind, only instrumental (Joyce 2006: 160 ff.).

Conclusion concerning normative matters

The general normative conclusion from an evolutionary perspective on morality has to be nihilistic, I believe: We have no reason to regard the deliverances of even our careful moral thinking as *evidence* for moral truth, and there is no other source of evidence. This I have not argued for. I have argued that none of the two processes of the dual process model fares better than the other at the hands of the debunking argument; if the emotive-intuitive process does not provide evidence for moral judgments, then the cognitive-utilitarian one doesn't.

Where does this leave us, if it is right? The normative position or attitude I think best suited to the combination of the dual process model and moral nihilism is, first, not to choose sides in the deontological-intuitive vs. utilitarian-cognitive competition and, secondly, to change focus from finding correct answers to moral questions to questions about how we can deliberate over moral problems as well as possible. Focus turn to what characterizes 'responsible deliberation,' collectively and individually.[4] I suspect that a correct diagnosis of the history of moral philosophy and the present state of the art, including the dual process model, will prove to be that the stubborn conflicts between broadly deontological and broadly utilitarian arguments are the natural (quite literally) outcome of our practical reasoning *at its best*. Among other things, as we saw above, we need both sides to avoid PDs and ruining our lawns and at the same time retaining our fruitful flexibility and opportunism in practical decision-making. Also, the very fact that we need the two independent decision-generating processes to negotiate such situations supports the hypothesis that each one has a certain build-in stubbornness: they are each built to counteract the other in certain kinds of situation. Considerations having to do with broadly utilitarian concerns are not easily silenced by intuitive-emotive responses and *vice versa.* Thus, I suspect that there is no sense in which either side, so to speak, can hope to come out on top. And that is not a thing to be lamented: the ongoing intra- and inter-personal conflicts and our handling of them in our attempts to further our interests broadly conceived – normally including considerable concern for others, of course – is the best we can hope for. The consequences of this perspective on moral philosophy is to move from a traditional focus on the *result* of the processes of deliberation to the very processes themselves. The main question is thus not 'which action is right?' but rather 'how should I go about deciding what to do?' Rather than discovery of moral truths or invention of moral theories, I think moral reasoning should be focused on *negotiation* between intuitionistic and utilitarian views. By negotiating practical matters we come to *decide* them, we don't *detect* correct answers to them. I'll leave this topic with a slogan by which to remember the position I find most promising: Not investigation and detection, but negotiation and decision!

---

[4] The phrase 'responsible deliberation' comes from the Norwegian philosopher Harald Ofstad (1920-1994), who in the eighties advanced the idea that one important but by philosophers normally neglected topic is how we go about making moral decisions (Ofstad 1982). Ofstad was my first philosophy teacher at Stockholm university. In view of how Ofstad's ideas were rapidly drowned by niceties in utilitarian theory and the fascination I found in that, I am surprised and a bit dismayed to note that my present ideas on the topic are approaching Ofstad's.

On this view of morality, *beliefs* that there are moral truths to detect and false moral claims to be argued against serves negotiation and decision in that they makes it seem a serious matter, concerned with arriving at the right decision. So beliefs in the authority, the objectivity and even truth of moral claims is instrumental in the upholding of what really matters: responsible negotiation and decision. So a certain amount of self-deception or fictionalism seems both to be part of how we think about moral matters and also quite useful for getting us to take them seriously. Indeed this is part of the standard evolutionary account of our moral psychology. Too rigid moralistic attitudes are, however, probably not conducive to successful negotiation, so realist and objectivist attitudes should probably be held in check in order not to block moral interaction and negotiation.

5. Methodological issues. Explanation of action and the dual process model

This brings me finally to the question of explanation of action in the light of the dual character of our practical reason. Intentional action can sometimes be explained by reference to the agent's beliefs and desires. Intentional action can be explained in many ways, but when we try to explain intentional action by giving the agent's *reason* for acting, the belief-desire model is the standard approach. I'll call this familiar mode of explanation 'reason explanation'. (Bergström 1990) But the agent's reason for acting is often the outcome of the interaction or negotiation between the emotive-intuitive and the cognitive-utilitarian decision processes, or so I am assuming. How could these two perspectives on decision be fitted into the framework of reason explanation, is it even possible? That is the last issue I want to discuss.

The most immediate way to merge the belief-desire model of reason explanation with the dual process model is to focus on the practical dualism first and then call in the belief-desire model to provide the last part of the explanation. If I for example decide to take the train to Hannover rather than flying that decison or action could according to the present proposal be explained by first detailing the workings of my two decision mechanisms in order to clarify how they interacted in yielding the decision as output.

We would need to know if I engaged in calculation (or less ambitiously: dwelled on the consequences of my action and its alternatives) before I reached the decsion or if I just saw the action as required. Or perhaps I saw myself in different and differently valenced light depending on which decision I imagined myself taking. In short, we would want to know *how* the choosen action and its alternatives were categorized, or interpreted: in a manner suited for processing by the emotive-intuitive or by the cognitive-utilitarian processes. In this we should thus hope for some help from further developments of the theoretical framework of Green *et al*. Roughly we should ask whether the issue at hand (travelling to Hanover) is seen as an 'old' problem or a 'new' one. Say that in this case the presumption would be that it was a decision of the latter kind, that is a kind that normally is processed by the cognitive-utilitarian process. But on the other hand it is a kind of decision where the cognitive-utilitarian process is prone to get us into 'error', remember SUV-driving and lawn deterioration. So perhaps, since I actually decided to take the train, the presumption should be that I managed to get my intuitions to get a grip on the situation in spite of its being of a kind normally processed by utilitarian thinking: I managed to look at it as involving an 'old' problem, as avoiding free-riding for example. Of course this conjecture could be corroborated by asking me how I saw the matter, but my answer should not be taken as decisive. We have ideals and preferences regarding our own reasoning and that makes us vulnerable to wishful thinking about our own thinking, so that I may well unwittingly embellish my decision as a purely

'rational' cognitive-utilitarian decision, when in fact it is the emotive-intuitive response after having vividly considered my moral standing as a upright member of the moral community.

This example shows that in order to finish the first part of an explanation of my action we would need to consult not only me and my interpretation of the situation, but also neurology, experimental psychology and decision theory in order to provide evidence for an hypothesis regarding which one of my decision mechanisms was responsible for the decision. The total evidence will often be contradictory because of our capacity to divert the decision mechanisms from their normal ways and put them to work on problems they normally don't process. Eventually we could reach a respectable hypothesis and that could then be further processed into the belief-desire framework. For example, we could say that I believed that I would best protect my moral standing as an upright member of the moral community if I took the train, and I wanted to do that, so I took the train. Alternatively the story could be that I wanted to maximize utility, and having thought about the long term effects of pollution by planes and the indirect effects of influencing the behavior of others (having read Glover's 1986 paper) I believed that taking the train would maximize utility, so I did that. So one lesson from the dual process model is that there are crucial factors relevant for decision that are far too interesting to be taken as given and stowed away in utility functions, even though they can be treated that way. Fuller explanations of action need to take into account the duality of practical reason and the agent's interpretation of the situation that is crucial for the workings of the dual processes.

There is a case to be made for a more radical conclusion regarding reason explanation in the light of our dual practical reason. Perhaps the workings of the emotive-intuitive decision process doesn't fit very well in the belief-desire framework at all. The picture that would then emerge is one where Humean belief-desire explanations are given for decision processed by the cognitive-utilitarian processes and decisions and actions that are the outcome of the emotive-intuitive processes need some other kind of explanation. Consider the paradigmatic working of the emotive-intuitive decision processes as it emerges in the work of Haidt (2001). A judgment that an act of consensual incest (say) is morally wrong even when there is nothing (else) to hold against it is issued on the spot without any conscious deliberation. The judgment is also highly resistant to revision by argument. How could such a judgment be fitted into the belief desire-model of reason explanation? Suppose I decide that I will refrain from, and when possible prevent, acts of consensual incest. What is the explanation? One explanation fitting the belief-desire model would be that I wanted to prevent and refrain from consensual incest and I thought that deciding to refrain from and prevent such acts would lead to me refraining from and preventing such acts. This seems not very explanatory. A 'real' belief-desire explanation cites one belief about means and ends where these are distinct. This putative explanation cites a means-ends belief where the means seems to be the end or in any case is much too close to them for explanatory comfort. There are two routes to follow from here. On the one hand we may say that the belief-desire model is OK even for intuitively processed decisions. Then we must improve on the above explanation of my counter-incestual decision. On the other hand we may say that there is a better model for intuitively processed decisions than the belief-desire model. Then we should provide that better model. I shall briefly comment on these two possibilities.

First, surely there is *something* I want to obtain from my counter-incestual decision, besides being against consensual incest. For example I may want an incest-free world, and if that is not within reach I may want to minimize incest incidents (consensual or not). I may reasonably believe that the best I can do to further that goal is to work consistently against incest, so I decide to refrain from and counteract incest. This seems a lot better than the claustrophobic explanation given above. So,

amendment within the belief-desire model should not be ruled out. However, this may seem a bit strained, and what's worse, it seems not true to what happens when the intuitive mechanism is in the drivers seat. The most interesting point of Haidt's incest-adverse subjects is precisely that they retain their judgment even if *all* arguments for the judgement are undermined, and this includes crucially arguments built on means-ends relations. So what we see is people who hold on to a decision even when all available resources for support are exhausted, that includes *any* beliefs to the effect that the decision furthers any goal or satisfies any preference at all. It's like in the scene from the movie I have forgotten the name of where the unfaithful husband returns to his home in the morning and meets his wife who asks where he has spent the night. 'In the hammock in the garden' he answers. 'But we burnt that hammock three years ago,' she retorts. 'Well, that's *my* story and I'll stick to it,' he finishes the discussion. This seems closer to the truth in the Haidt-cases than any means-ends story, I think: 'Why are you against consensual incest even when there is nothing wrong with it?' 'Well, it's my story and I'll stick to that!'

So, there is a case for examining the possibility of a methodological dualism in the humanities, regarding the explanation of action, reflecting the duality of practical reason. Sometimes, when the emotive-intuitive processes calls the shots, we do things, not to obtain anything, but to do just what we are doing. We do things, so to speak, for their own sakes. Such scenarios can often, perhaps always, be described correctly in a belief-desire model, but those descriptions will at least sometimes be explanatorily vacuous.

Of course these methodological remarks have an ontological side. The debate about the existence of what is sometimes called 'arational' action was initiated by Anscombe's complicated discussion of intention in the book by that name. Anscombe recognizes a class of intentional actions that are done for no (further) reason but still intentional and she also claims that the agent has non-observational knowledge of what he is doing when he acts intentionally (Anscombe 1957: 9 ff.). Both the no (further) reason part and the non-observationality part sits well (up to a point) with the present emotive-intuitive perspective, it seems. This discussion has continued and arguments for the idea that some intentional actions are not fit for explanation of the belief-desire kind have been advanced (Hursthouse 1991, Spicer 2004). The arguments have for the most part been built on 'philosophical' premises in the form of armchair examples of actions that sit uncomfortably with the belief-desire model. The present discussion points to the existence of further evidence for that kind of view.

How would a model for explanation of emotive-intuitive action look? I think it would have to hook into the biology and the sociology of the mind more directly than explanations of actions of a more cognitive-utilitarian kind. In explaining my incest-adverse judgment, we may for example have to go into the natural history of incest and its effects, the sociology of culture and ceremony and tell a story to the effect that being adverse to incest, no matter what, is indeed not at all mysterious. It is what one should expect from members of a cultured species of our kind. Explanations along these lines would not be rationalizations in Davidson's sense. Quite the contrary, they would show why the explanandum obtains even though it's not rational in the sense that it helps the agent get anything she wants.

6. Summing up

It looks as if there is a practical dualism: that our decisions are driven by two distinct and sometimes contrary processes. That could go some way towards understanding how we avoid some

obvious traps of practical reason. Furthermore, even though it is difficult to see that such a practical dualism have any direct normative consequences, it seems that conflicts in moral philosophy well could be seen as the end point for our practical reasons, harboring, indeed nursing, conflicts we are built to manifest and resolve as the way to navigate our ways though our lives. A little less speculatively the dualism seems to spell trouble for evolutionary ethics, that rely on the idea that our moral sensibilities have developed to further any specific or at least non-contradictory proximal goal. Also this dualism invites some novelty in the field of explanation of action. First by introducing the need for digging into the works of the two distinct and sometimes competing mechanisms when providing belief-desire hypotheses, and secondly by inviting a departure from the belief-desire model in favor of an evolutionary and social psychological emotive-intuitive model. This last also promises to provde some evidential input a and analytical tools to the debate concerning so called arational action.

References

Anderson, S. W. *et al.* 1999. "Impairment of social and moral behavior related to early damage in human prefrontal cortex." *Nature Neuroscience* 2: 1032-1037.

Anscombe. G. E. M. 1957. *Intention.* Oxford: Basil Blackwell.

Bergström. L. 1990. "Explanation and interpretation of action," *International Studies in the Philosophy of Science* 4: 3-15.

Bechara, A. *et al.* 1996. "Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex." *Cerebral Cortex* 6: 215-225.

Bolender, J. 2001. "A two-tired cognitive architecture for moral reasoning." *Biology and philosophy* 16: 339-356.

Campbell, R. 1996. "Can biology make ethics objective?" *Biology and Philosophy* 11: 21-31.

Casebeer, W. D. 2003. *Natural Ethical Facts: Evolution, Connectionism, and Moral Cognition.* Cambridge, MA: The MIT Press.

Cosmides, L. and Tooby, J. 2008. "Can a general deontic logic capture the facts of human moral reasoning?" In *Moral Psychology. Vol. 1*, ed. W. Sinnott-Armstrong. Cambridge, MA: MIT Press.

Damasio, A. R. 1994. *Descarte's Error: Emotion, Reason, and the Human Brain*. Putnam Adult.

Dennett, D. 1995. *Darwin's Dangerous Idea.* Simon and Schuster.

De Waal, F. 1996. *Good Natured. The Origins of Right and Wrong in Primates and Other Animals.* Harvard University Press.

Glover, J. 1986. "It makes no difference whether or not I do it." In *Applied Ethics*, ed. Peter Singer Oxford: Oxford University Press. (First published in *Proceedings of the Aristotelian Society*, Supp. Vol. XLIX, 171-90.)

Green, J. *et al.* 2004. "The neural bases of cognitive conflict and control in moral judgment." *Neuron* 44: 389-400.

Green, J. 2002. "The terrible, the horrible, no good, very bad truth about morality and what to do about it." Doctoral dissertation. Princeton University.

Greene, J. 2008. "The secret joke of Kant's soul." In *Moral Psychology. Vol. 3: The Neuroscience of Morality*, ed. W. Sinnott-Armstrong. Cambridge, MA: MIT Press

Haidt, J. 2001. "The emotional dog and its rational tail: a social intuitionist approach to moral judgment." *Psychological Review* 108: 814-834.

Harman, G., 1977. *The Nature of Morality. An Introduction to Ethics.* Oxford: Oxford University Press.

Hursthouse, R. 1991. "Arational Actions." *The Journal of Philosophy*, 88: 57-68.

Joyce, R. 2006. *The Evolution of Morality.* The MIT Press.

Kagan, S. 1989. *The Limits of Morality.* Oxford: Oxford University Press

Lillehammer, H. 2003. "Debunking morality: evolutionary naturalism and moral error theory." *Biology and Philosophy* 18: 567-581.

Mackie, J. L. 1977. *Ethics. Inventing Right and Wrong.* Pelican Books.

Mendez, M. F., Anderson, E., & Shapira, J. S. 2005. "An investigation of moral judgment in frontotemproral dementia." *Cognitive and Behavioral Neurology* 18: 193-197.

Moll, J. *et al.* 2001. "Frontopolar and anterior temporal cortex activation in a moral judgment task: preliminary functional MRI results in normal subjects." *Arquivos de neuro-psiquiatria* 56: 657-664.

Ofstad, H. 1982. *Ansvar och handling.* (In Swedish.) Stockholm: Prisma.

Okasha, S. 2007. "Rational choice, risk aversion, and evolution." *The Journal of Philosophy* 104: 217-235.

Ruse, M. 1986. *Taking Darwin Seriously: A Naturalistic Approach to Philosophy.* Oxford: Basil Blackwell.

Richards, R. J. 1986. "A defence of evolutionary ethics." *Biology and Philosophy* 1: 265-293.

Flanagan, O. J., Sarkissian, H. and Wong, D. 2007. "Naturalizing ethics." In *Moral Psychology. Vol. 1*, ed. W. Sinnott-Armstrong. Cambridge, MA: MIT Press.

Sommers, T. and Rosenberg, A. 2003. "Darwin's nihilistic idea: evolution and the meaninglessness of life." *Biology and Philosophy* 18: 653-668.

Skyrms, B. 1996. *The Evolution of the Social Contract.* Cambridge: Cambridge University Press.

Skyrms, B. 2003. The *Stag Hunt and the Evolution of Social Structure.* Cambridge: Cambridge University Press.

Spicer, F. 2004. "Emotional behaviour and the scope of belief-desire explanation," in *Emotion, Evolution, and Rationality,* eds. D. Evans, and P. Cruse, Oxford: Oxford University Press.