

# THE NATURALISTIC FALLACY. A NOTE ON A NOTE.

Jaap Hage

Maastricht University

Department of Metajuridica

P.O Box 616

6200 MD Maastricht

email: [jaap.hage@metajur.unimaas.nl](mailto:jaap.hage@metajur.unimaas.nl)

## 1 Introduction

Having written a book on the biological and psychological basis of behavioural norms (Crombag 1973), Hans Crombag may have felt liable of being accused to have committed the 'naturalistic fallacy', the illegitimate derivation of normative or evaluative conclusions from purely factual premises. In a brief paper, called 'A Note on the Naturalistic Fallacy' (Crombag 1986), he made an attempt to parry such accusations. The argument in this paper is fundamental to the whole enterprise of evaluating norms from the perspective of empirical sciences and occupies for that reason an important place in Crombag's scientific oeuvre. In this paper I will restate Crombag's argument and examine it to determine which conclusions can, and which ones cannot be based on it. My conclusions will be that Crombag does not give sufficient reasons to remove the naturalistic fallacy from the philosophical map, but that he makes it sufficiently plausible that empirical sciences can play a useful role in the rational evaluation of norms.

## 2 Crombag's argument

The central question that Crombag addresses in his paper is whether the naturalistic fallacy is really fallacious. As his main target Crombag takes the classical text from Hume which is often taken to contain the first delineation of the naturalistic fallacy. In this text (Hume 1739/40, Book III, Part i, section 1), Hume notices that authors on moral philosophy often move imperceptibly from purely factual propositions to deontic propositions, holding what ought or ought not to be done. Hume then remarks that it seems altogether inconceivable how these deontic conclusions can be deduced from the factual ones, which are entirely different.

Having quoted Hume, Crombag continues with posing several questions of his own. The first question is where norms come from, if they cannot follow from facts. He decomposes this question into two related ones: How did the norms that we have or think to have, come into existence? and Why are these norms binding?

To answer these question Crombag discusses four arguments. The first one is by Bentham. He starts from the assumption that mankind is under the governance of two masters, pleasure and pain. These two both point out what we ought to do, and determine what we will do. Crombag interprets pleasure as that which promotes individual survival, and pain as that which hinders it. Under this interpretation Crombag believes Bentham's starting point to be shown true by biologists. Seeking individual survival is trivially crucial for survival.

But then, how does it follow that survival value is relevant for what we ought to do? As an answer to this question, Crombag gives an argument that is worth quoting:

Because there is no alternative for the species. For the species the only alternative would be non-existence, which is obviously counter-factual. The fact that *homo-sapiens* (still) exists, implies that, although adaptation may not always have been perfect, a critical number of its members so far have been adaptive more often than not. For individual members of the species this need not be true, but they do so 'at their peril'.

Briefly summarised, the need to adapt to the circumstances imposes constraints on human behaviour that individuals should obey in order not to risk their own lives and the survival of the species. Without this survival the question what ought to be done would not even arise. This leads Crombag to the conclusion, ascribed to Bentham, but implicitly adopted, that what ought to be done springs from the factual source which is the existence of human life. As a consequence the correct method of acquiring moral knowledge is empirical (namely the establishment of what contributes to life - JH).

The second argument that Crombag discusses stems from Skinner. Having discussed Bentham's argument, we can be brief about Skinner's. His conclusion is that norms are conditional statements of

fact: If you want to be positively reinforced, you ought to do this or to avoid that. Survival has been replaced by positive reinforcement, but the upshot of the argument remains the same: Norms are answers to the demands posed by factual situations. In order to obtain particular effects, survival in the case of Bentham (as interpreted by Crombag), and positive reinforcement in the case of Skinner, humans ought to behave in a particular way. Crombag notes that these norms are hypothetical imperatives in the sense of Kant. They do not prescribe unconditionally, but merely on the hypothesis that some goal is pursued. Kant assumed that there also categorical, that is unconditional, imperatives. But Crombag hastens to add that he does not know of those; 'even the most sacred norm of the inviolability of human life must yield in a situation of self-defence'.

The third argument, or rather a point, stems from Ullman-Margalit, who stated that 'certain types of norms are solutions to problems posed by certain interaction situations'. Again, norms are presented as means to an end.

As a fourth argument Crombag himself points out that norms are senseless if they prescribe behaviour of which humans are not capable. This shows, according to Crombag, that certain factual states of affairs exclude certain norms. This is the negative counterpart of the conclusions from the previous three arguments, which hold that some norms are demanded by facts about the world.

### **3 Has Crombag refuted the naturalistic fallacy?**

Let us assume that Crombag's fourfold argument is correct in the sense that it establishes that factual states of affairs make certain norms necessary, or at least highly desirable, and other norms senseless. Can we then say that, under this assumption, Crombag has shown the naturalistic fallacy to be not fallacious at all? Remember that Crombag formulated two questions which he attempted to answer:

1. How did the norms that we have or think to have, come into existence?
2. Why are these norms binding?

It seems to me that Crombag's fourfold argument has provided at least the beginning of an answer to the first question. The factual situation in which mankind lives explains, at least to a large extent, the norms that exist. In particular it seems not necessary to add other than factual premises to make the explanation complete.<sup>1</sup>

It is less clear whether Crombag has even attempted to answer the second question. At most he has given reasons why it may be rational to obey particular norms, on the assumption that one wants to survive (as a species), to be reinforced positively, or to solve particular problem situations. In other words, Crombag has given reasons for the truth of particular hypothetical imperatives, or technical norms as they are also called. Such technical norms are comparable to recipes for the preparation of certain dishes, and it has, to my knowledge, never been disputed that the truth of such recipes can be ascertained by empirical means.

To answer the second question, Crombag should have provided an argument for the bindingness of what Kant called categorical imperatives, or - to say it in more plain language - for the bindingness of certain prescriptions that hold irrespective of the goals one pursues. In his discussion of Skinner, Crombag explicitly denied the existence of such norms. This explains why he did not even attempt to answer the second question that he posed himself. Crombag gave good reasons why empirical sciences can explain the existence of norms and can establish the truth or falsity of technical norms, but he did not even argue why these sciences provide reasons for the existence unconditional norms, let alone for their bindingness.

In the next sections I will give a more detailed analysis of what the naturalistic fallacy amounts to. From this analysis it follows that, in order to refute the naturalistic fallacy, Crombag should have shown that empirical sciences give decisive reasons for the bindingness of unconditional norms. Therefore, my preliminary conclusion is that Crombag did not even attempt to refute the naturalistic fallacy.

---

<sup>1</sup> It should be noted that these factual premises should include lawlike generalisations, the precise status of which is not fully clear. This complication does, however, not detract from Crombag's point that empirical sciences can contribute to the understanding and the critical evaluation of existing or proposed norms.

#### 4 What does the naturalistic fallacy amount to?

At the beginning of this century the Cambridge philosopher G.E. Moore introduced the 'naturalistic fallacy' into ethical theory (Moore 1903). Moore's own description of what this 'fallacy' amounted to was not very clear, but Frankena (1939) and Taylor (1961) improved upon Moore's work in this respect. The basic intuition behind the notion of the naturalistic fallacy is that purely factual sentences on the one hand, and purely evaluative or deontic sentences on the other hand, have no meaning components in common. As a result it is:

1. impossible to give the meaning of evaluative and deontic sentences in terms of purely factual sentences;
2. impossible to deduce evaluative and deontic sentences from purely factual ones; and
3. possible to agree about all purely factual sentences and still to disagree rationally about some evaluative or deontic sentence (Taylor 1961, p. 242f.).

I will illustrate this by means of the norm that one ought not kill human beings. Clearly, this ought-sentence does not *mean* any of the following:

- a. Nobody kills human beings.
- b. Persons who kill human beings are punished (or otherwise subjected to negative reinforcement).
- c. Killing a human being enhances the chance of the killing person to lose his life.
- d. Killing a human being diminishes the chance of the killing person to procreate.
- e. Killing a human being enhances the chance that the human race dies out.
- f. God (or any other lawgiver) has forbidden to kill human beings.
- g. There exists a norm to the effect that it is forbidden to kill human beings.

Some may believe that one of the sentences a to g describes a sufficient reason why one ought not kill human beings, but that it is not the issue at stake. The issue at stake is what these sentences mean, and not one of them *means* that one ought not to kill human beings. In general no set of purely descriptive sentences means that something ought or ought not to be done.<sup>2</sup> In fact, this immediately follows from the basic assumption behind the naturalistic fallacy that purely descriptive sentences and purely deontic sentences have no meaning components in common.

Deductively valid arguments are usually defined as arguments for which it is impossible that if their premises are all true, their conclusions are false. This definition leaves unclear what kind of possibility is at stake. In general, what is possible depends on the constraints that one takes into account. Physically it may be possible that somebody is both a thief and not punishable, but legally this may be impossible.<sup>3</sup>

It is often assumed that the impossibility at stake in issues concerning the validity of deductive arguments is based on language conventions. The most important in this connection are the conventions that govern the meanings of the so-called logical words, such as 'and', 'not' and 'or'. However, the meanings of words such as 'ought', 'forbidden', and 'permitted' may also be taken into account. On this assumption the characterisation of deductively valid arguments becomes that an argument is deductively valid if, given the meanings of the words that occur in the premises and the conclusion of the argument, it is impossible that all the premises are true while the conclusion is false.

Since purely descriptive and deontic sentences are *ex hypothesi* assumed to have no meaning components in common, the constraints that follow from the meanings of the words in an argument cannot necessitate the truth of a deontic conclusion in an argument with purely descriptive premises. The second aspect of the naturalistic fallacy follows immediately from the underlying assumption of this fallacy, given the characterisation of deductively valid arguments in terms of the meanings of the words used in these arguments.

---

<sup>2</sup> There might still be some dispute about sentence g, which some might take as logically equivalent to the sentence that one ought not to kill human beings. However, I take this sentence as meaning that the norm in question exists as a matter of social fact, and that does not mean that one ought to do anything. Unless, of course, one takes the ought-sentence to describe a matter of social fact too. But this would be against the spirit of the naturalistic fallacy, which deals with ought sentences that indicate what ought to be done, in opposition to what is the case. I will return to this digression in section 5.

<sup>3</sup> I abstract from exceptions to rules.

Whether the third characteristic also follows, depends heavily on the way in which one gives content to the notion of rationality. The underlying assumption is that rational disagreement is always possible if one does not violate the inference rules of deductive logic as based on meaning conventions. Given this rather broad notion of rationality, it is possible to agree about all purely factual sentences and still to disagree rationally about any deontic sentence. The question remains, however, whether we should adopt such a broad notion of rationality.

Before addressing this question, we should look back at our analysis of the naturalistic fallacy and consider what it implies for attempts of refutation as the one launched by Crombag. When we do this, we find that the naturalistic fallacy, as analysed by Taylor, is a direct consequence of three assumptions. The first assumption is that purely descriptive sentences and deontic sentences have no meaning components in common. The second is that deductively valid arguments guarantee the truth of their conclusions, given the truth of their premises, *on the basis of conventions governing the meanings of the words employed in the argument*. If these first two assumptions are taken for granted, the first two aspects of the naturalistic fallacy are beyond refutation. The third aspect can still be doubted if one does not adopt the third assumption, that is the assumption that rationality is nothing else than logical consistency according to the standards of deductive logic as defined here.

Since it was Crombag's concern to establish the relevancy of the empirical sciences for normative issues, the third aspect is the most important for him. To refute this aspect of the naturalistic fallacy, two ways are open. First he can challenge the assumption that underlies all three aspects of the naturalistic fallacy, the assumption concerning the absence of meaning relations between purely descriptive and deontic sentences. And second, he can propagate a narrower notion of rationality, by arguing for constraints of rationality that go beyond those of deductive logic.

Theoretically there is a third option, namely to reject the assumption that deductively valid arguments guarantee the truth of their conclusions, given the truth of their premises, on the basis of the meanings of the words employed in the argument. However, this assumption is probably best read as a stipulative definition of what a deductively valid argument is, and it is in general not useful to argue about stipulative definitions.

## 5 Meaning relations

Are there any meaning relations between purely descriptive sentences and deontic sentences? In the previous section I have given a number of examples to show that there are no. However, in particular the seventh example may be less than convincing. I wrote that the sentence one ought not to kill human beings does not mean that there exists a norm to the effect that it is forbidden to kill human beings. But is that true?

In deontic logic it is traditionally assumed that 'forbidden' means the same as 'ought not'. So the question reduces to whether the phrase 'there exists a norm to the effect that one ought (not) to do x' means the same as 'one ought (not) to do x'. On this question the opinions differ, and the differences can be traced back to different interpretations of 'there exists a norm to the effect that ...' and of 'one ought to do ...'.

The phrase 'there exists a norm to the effect that ...' can both be read as stating something about social reality, and as a redundant phrase like 'it is true that ...'. On the first reading, which is a purely factual one, the sentence 'there exists a norm to the effect that one ought (not) to do x' obviously means something different than the sentence 'one ought (not) to do x' interpreted in a normative way. It might mean the same as the last sentence interpreted in the descriptive way, an interpretation which is to be discussed later in this section.

On the second reading, the redundant one, the sentence approximately means the same as the content of the norm.<sup>4</sup> But on this reading the sentence is a deontic one, and not purely descriptive. Therefore, on this reading, it does not provide a counter-example to the assumption underneath the naturalistic fallacy.

---

<sup>4</sup> Arguably there is still a difference, because the one sentence makes a statement about the existence of a norm, while the other makes a statement about killing humans. However, if the one sentence is true, the other must also be true *because of the meanings of the two sentences*. This suffices for the purpose of the naturalistic fallacy.

The sentence 'one ought (not) to do x' can also be read in two ways. In the one reading, acceptance of the sentence as true implies at least the motivation (not) to do x. This is the deontic reading. The second, descriptive, reading means that according to some (the factually existing) set of norms, one ought (not) to do x. It is possible to accept such a sentence as true, while remaining utterly unmoved by it regarding the question whether one will do x. This descriptive reading means more or less the same as the social reading of the sentence stating that the norm exists. However, this second reading makes the sentence into a non-deontic one, and the existing meaning relations provide no counter-example to the assumption underneath the naturalistic fallacy.

Summarising we can say that both the sentence 'there exists a norm to the effect that one ought (not) to do x' and the sentence 'one ought (not) to do x' can be given a descriptive and a deontic reading. Only if the one is given a descriptive reading and the other a deontic reading, the two sentences might provide a counter-example to the naturalistic fallacy, but precisely, and not coincidentally, then there are no meaning relations between them. If both are read either descriptively or deontically, there are meaning relations, but then they cannot provide a counter-example. Therefore we may conclude that the assumption that underlies the naturalistic fallacy, namely that there are no common meaning components between purely descriptive and deontic sentences still stands. Crombag should aim his arrows to the other assumption, concerning the broad notion of rationality.

## 6 Rationality

The broad notion of rationality that is needed to arrive at the third aspect of the naturalistic fallacy, namely that it is possible to agree about all purely factual sentences and still to disagree rationally about some evaluative or deontic sentence, boils down to it that it is only irrational to be inconsistent according to deductive logic based on meaning conventions. According to this notion it is not irrational to believe that some person is more than twenty-thousand years old, unless one also believes that humans cannot be older than a number of years (considerably) below twenty-thousand. Neither would it be irrational to believe that it is raining, while disbelieving that persons that walk outside without umbrella or something similar, become wet.

The broad notion of rationality outlined in the previous section will probably be too broad for most of us. There are more constraints on rational behaviour, including rational belief, than those imposed by meaning conventions. It would, for instance, be irrational to have beliefs that are in conflict with well-established laws of nature. Arguably, it would also be irrational to believe both that there exists a legal rule to the effect that thieves are punishable, and that thieves are normally not punishable. Yet, it is not so easy to establish which other constraints there are on rational behaviour. Any choice of constraints that are to be taken into account is liable to being accused of partisanship.

A possible solution is to relativise the notion of rationality. Behaviour is rational or irrational, not in itself, but relative to some background of constraints. The backgrounds of the several variants of deductive logic are then the corresponding sets of meaning conventions that are taken into account.<sup>5</sup> Broadening the set of constraints in the background correspondingly narrows the notion of rationality as relativised to that background.

Including the right constraints into the background of some particular notion of rationality may make it irrational to believe both that some action detracts from the chance of survival of the human species and that this action is permitted or even right. Or, to take a more Benthamian example, inclusion of the right constraint may make it irrational to believe that some action diminishes pleasure and increases pain and also that this action is right. The Skinnerian counterpart would make it irrational to believe that some action leads to negative reinforcement and also that this action is right, and so on .... Depending on one's background assumptions it may be irrational to agree on some set of facts, and to disagree nevertheless on what ought to be done.

An obvious objection is to be expected here. This objection runs that the normative premise, necessary to make the step from the factual premises to the normative conclusion, is moved from the

---

<sup>5</sup> For instance, propositional logic only takes the meanings of a set of logical operators into account, while predicate logic is also based on the meanings of quantifiers. Deontic logic, finally, takes also the meanings of the deontic operators into account.

premises of an argument to the background. Maybe it is less visible there, but it is still present. Without this normative premise the transition from is to ought would not be valid.<sup>6</sup>

This objection is only in part to the point. To make the step from 'This action will be positively reinforced' to 'This action ought to be performed' a rational one, it is necessary to include in the background that it is rational to do what is positively reinforced. But this is not fully comparable to having 'silent' premises. 'Silent' premises can make an argument that is invalid in a particular system of logic into a valid one. Such a completion of an argument by adding 'silent' premises presupposes a particular system of logic, which in its turns presupposes a number of constraints that determine which inferences are rational, or- what comes in this connection down to the same thing - logically valid. Changing background assumptions, on the contrary, is rather like changing the logic, to make an argument that was invalid under the old logic valid under the new one.

Let me give an example to illustrate this point. Suppose we have the argument:

It is forbidden to steal.

Therefore: It is not permitted to steal.

This argument is invalid under predicate logic, because predicate logic does not take meaning relations between the words 'forbidden' and 'permitted' into account. The argument can be made valid *under predicate logic* by adding the premise

If something is forbidden, then it is not permitted.

However, adding this premise changes the argument. The original argument made a transition from only one premise to the conclusion. The new argument needs two premises to obtain the same conclusion. Given one and the same logic (predicate logic) the first argument is invalid, while the second is valid.

It is also possible to leave the argument unchanged and to modify the logic. By adding the *rule of inference* to the logic that actions that are forbidden are not permitted the original argument becomes a valid one *under this new logic*. In fact this new logic is a rather common one, which goes under the name of deontic logic.<sup>7</sup>

Analogous to this move from predicate logic to deontic logic we can make the move from predicate logic to 'behaviourist logic'. Behaviourist logic is characterised by the background assumption that actions that are (on the balance) positively reinforced ought to be performed and that actions that are (on the balance) negatively reinforced ought not to be performed. This background assumption is translated into the rules of inference that from the premise that some action is positively reinforced one may conclude that this action ought to be performed and that from the premise that some action is negatively reinforced one may conclude that this action ought not to be performed. Under this logic the naturalistic fallacy that consists of an argument from reinforcement to what ought to be done is not fallacious anymore.

It might seem that the move proposed here, that is replacing silent premises by corresponding rules of inference, has 'all the advantages of theft over honest toil'. Apparently it is possible to replace a silent premise by an inference rule, thereby making a previously invalid argument valid without adding to the premises. But, one may object, this does not make any change in the naturalistic fallacy, because the normative premise, necessary to validate the transition from 'is' to 'ought', must still be smuggled in, whether in the form of a premise or in the form of a new inference rule.

This objection misses the point, however. It assumes that there is something like the 'real' logic, which does not allow the inference from 'is' to 'ought' and that this real logic can be expanded into a treacherous logic that has inference rules that allow argument steps which are not 'really' valid. The crucial point of the previous argument is, however, that there is no such thing as the 'real' logic. There are many logics, each of which is based on a number of background assumptions concerning what it is rational to infer. There is no Archimedean set of background assumptions corresponding to the 'real'

---

<sup>6</sup> I have moved here from the general notion of rationality to the 'logical' notion of validity. This move is justified on the assumption, which I make here, that logic is the theory of rational reasoning, and as such a special part of the general theory of rational action.

<sup>7</sup> For logicians: It is the move from predicate logic to a form of deontic logic which uses the definition that  $F(\text{action}) \equiv \sim P(\text{action})$ .

logic. Any premise can be made into a background assumption, and any background assumption can be made into a premise.<sup>8</sup> What one considers to be a premise and what part of the logic depends on what one wants to make explicit. The logic is presupposed, the premises are explicit.

It should be noted, however, that the phrase 'wants to make explicit' is deceptive in that it suggests that it is in general a matter of choice what is considered to be premise and what to be logic. One of the main points of speaking about a background is that usually one does even realise that there is room for choice. Background is presupposed and therefore most of the times unconscious.<sup>9</sup>

Having come this far, we can conclude that Crombag might be successful in attacking the third aspect of the naturalistic fallacy. Whether it is possible to agree on all the relevant facts and nevertheless to disagree rationally about normative conclusions supposedly based on these facts depends on the background assumptions one has concerning rational belief. For those to which it is, for instance, obvious that one ought to do what is positively reinforced, it is impossible to agree about the fact that action A is positively reinforced and nevertheless to disagree rationally about whether this action ought to be performed. Given the way he writes about categorical imperatives, now to be interpreted as imperatives without background assumptions, it seems obvious that Crombag makes background assumptions that are at least comparable to the one mentioned above.

## 7 Conclusion

In a sense Crombag's argument against the naturalistic fallacy misses the point. Given the almost axiomatic assumption that purely descriptive sentences and deontic sentences have no meaning components in common, and the traditional definition of deductive validity in terms of the meanings of the sentences and words involved in an argument, it is practically by definition true that deontic judgements are not definable in terms of non-deontic ones, and that they cannot be deduced from them either.

Crombag's argument can be read on a different level, however. It can be read as the argument why certain assumptions about human behaviour are so obvious that they should be considered as background assumptions upon which our logic is to be based. Such a logic would validate more inferences than only those based on meaning conventions, and these additional arguments would make a utilitarian ethics logically true. Whether Crombag has succeeded in this argument is a question for a separate discussion. It is clear, however, that such an argument would not commit the naturalistic fallacy, because the naturalistic fallacy presupposes, rather than establishes, that Crombag's argument would be wrong.

---

<sup>8</sup> It is, however, not possible to change all background assumptions into premises, because then there would no logic be left over to argue with the premises. This is the upshot of Lewis Carroll's paper 'What the Tortoise Said to Achilles' (Carroll 1895).

<sup>9</sup> There is an interesting parallel with computers, here. Both rules of inference and premises may be considered as software involved in the computation of conclusions. However, where premises are real 'soft' software, rules of inference are like hard software, built into the chips that do the information processing. It is possible to have a chipset that can handle less instructions, but then the software must express the deleted functions in terms of the remaining ones. This is the equivalent of replacing rules of inference by premises.

This metaphor also clarifies why not all rules of inference can be replaced by premises. See note 8. There must remain a processor that processes the enhanced software. It also explains why the mind is not only software, but needs a physical basis. (Cf. the, inadequately formulated, argument of Searle 1984 against 'strong AI').

## References

- Carroll, L. (1895). What the Tortoise Said to Achilles. *Mind*, 4, pp. 278-280.
- Crombag, H.F.M. (1983). *Een manier van overleven: Psychologische grondslagen van moraal en recht*, Tjeenk Willink, Zwolle.
- Crombag, H.F.M. (1986). A note on the naturalistic fallacy. *Methodology and Science*, vol. 19, pp. 83-90.
- Frankena, W.K. (1939). The naturalistic fallacy, *Mind* 48. Ook in Ph. Foot, *Theories of Ethics*, Oxford University Press, Oxford 1967, blz. 50-63.
- Hume, D. (1739/40). *A Treatise of Human Nature*. Several editions.
- Moore, G.E. (1903). *Principia Ethica*. Cambridge University Press.
- Searle, J. (1984). *Minds, brains and science*. Harvard University Press, Cambridge Mass.
- Taylor, P.W. (1961). *Normative discourse*. Greenwood Press, Westport.