

Intention and Virtue

Boris Hennig, March 2008

Draft for the ETMP conference

This is a somewhat hastily compiled excerpt from what used to be a much longer paper. In order to preserve as much of the original argument as possible, I have cut down on context, thereby excluding especially discussions of Anscombe's Intention and Modern Moral Philosophy, and John Doris' Lack of Character. Let me therefore acknowledge here that the following paper is substantially based on and thus deeply indebted to the writings of Anscombe, as well as Michael Thompson's Naive Action Theory. The lack of context might also make it more difficult to see what the point and aim of the following is. But to understand an argument is one thing, to see its implications is another, and understanding comes first. I must leave the implications for another occasion.

Virtue ethics is often said to stress the importance of agent evaluation as opposed to the evaluation of actions. Its proponents seem to suggest the following:

Imagine an agent to have a fully informed and fully virtuous counterpart. The good or right thing to do is what that counterpart would want the agent to do in the agent's actual circumstances. (Copp and Sobel 2004:546)

The idea is, presumably, that we may first settle the question whether an agent is virtuous, before knowing the moral value of her actions, such that we may then ask what actions she would recommend. But how can one evaluate an agent as virtuous independently of what she does and recommends doing? A virtuous agent is a person who does virtuous things. In order to see whether a counterpart agent is fully virtuous or not, one must consider her particular actions, intentions, and recommendations, and look at *their* moral value. The account seems to be circular: in order to represent an agent as virtuous, we need to already know what actions are virtuous; but we are

supposed to find out the latter by imagining a fully virtuous agent.

This however is a problem for virtue ethics only when it is thought to rest on the assumption that the merits of an agent may be assessed independently of what she does. I submit that virtue ethics does not rest on this assumption. In order to see this, one needs to understand the logical status of virtues.

I will establish my claim by first arguing that actions are most naturally described and classified in terms of the agent's intentions, and that the reason for this is that the intentions of an agent tell us something about what this agent may be expected to do in similar contexts. The correct classification of an action therefore has implications regarding the character of the agent. We cannot describe and classify an intentional action as such without at the same time assessing the character of the agent. Conversely, intentions are always intentions to do something; and to say what character an agent has is to say what she is likely to do. Therefore, we cannot assess agents without at the same time evaluating their actions.

I will further argue that virtues are not character traits that cause certain kinds of behavior as an external result. If virtues are character traits at all, they consist in the fact that the person who has them acts in certain ways. In this respect, virtues are similar to intentions. An intention is that in terms of which an action is most naturally classified. Virtues and vices, I will maintain, are also terms in which kinds of behavior may be classified. More specifically, a virtue will be shown to be that in terms of which a generic action is described, in cases where it is possible to justify this generic action merely by describing it in a certain way. But if virtues are terms in which actions are described and classified, then the description of a virtue cannot be separated from the description of the actions it serves to classify and describe.

What virtue ethicists emphasize is thus not that we need to look at agents *instead* of their actions, but that we need to classify and describe actions appropriately, namely in terms of intentions, and as expressions of virtues and vices, rather than looking at their accidental consequences and circumstances. Virtue ethics is in fact more act-centered than its consequentialist or deontological alternatives: it focuses on the proper classification of actions themselves, whereas consequentialism gives more weight to the consequences of actions, and deontological accounts primarily consider the rules and laws with which an action might comply or interfere. But consequences and laws are external to actions.

1. Types of Actions

I am going to claim that to describe an action in terms of an intention is the most natural way of classifying it as an instance of a *type*. Let me first explain what I mean by “type” here. (I do not intend a general account of what types are; only an account of what types of actions are, in contrast to mere classes or sets of actions.)

First, the different types under which a particular item may be brought are always systematically related, such that their relations may be depicted in a tree structure. For instance, the types to which this particular animal belongs—beaver, rodent, and mammal—constitute a hierarchy such that this beaver is also a rodent and a mammal; and conversely, there are several types of mammals and rodents.

Second, the nature of the instances of a type is often not exhausted by their definition. In order to say what it is to be a beaver, for instance, we need to observe and study actual instances of this type.

Third and most importantly, types allow for atypical instances. A type may well be instantiated by something that lacks essential features of its typical instances. This is, to be sure, an idiosyncratic feature of what I call types and essences here. It is usually thought that essences, that in terms of which types are defined, are features that all instances of the respective type necessarily exhibit. This is not the place to argue against this assumption; I am simply not going to share it. Types, as I use this term here, are indeed defined in terms of features. But these features are not features that all instances must share; they are features that *typical* instances will exhibit insofar as they are typical. To the extent that an instance of a type has these features, it is typical, and to the extent to which it lacks them, it is atypical. Types are not defined in terms of necessary features, but in terms of standards of typicality.

Only certain kinds of entities admit of a classification into types, and where they do, there is always a reason why. Since the instances of a type need not all share a set of properties, this cannot in general be the reason why they belong to the same type; often, the reason is more complicated. The reason why we may classify living beings into types is not that all beavers, for instance, share certain features. The reason is that living beings reproduce and that different types of living beings have evolved from common ancestors. There is a story we can tell, rather than a definition we could

state, that ties the instances of a life form together.¹ The story we may tell about how the instances of a type belong together will not always involve reproduction. There are types of artifacts, and the reason why is that humans (or other living beings) produce artifacts with a certain use in mind, or something along these lines. (I ask the reader to insert his favorite account of the essence and identity of artifacts here.)

The reason why actions may be classified into types is that they are performed with intentions. Let me first point out that when we classify actions in terms of intentions, we do indeed characterize them as instances of types.

As for the first feature of types, one may always give more or less wide descriptions of the same intentional action. (This has been called the “accordion effect.”) This is so because often, an action that is done in order to do *X* may also be classified as doing *X*. In this case, the proper genus of an action is what the agent might answer to the question why she is performing it. For instance, I might cycle down Centre Avenue in order to get groceries. To get groceries is the intention with which I cycle down Centre Avenue. But at the same time, it is simply what I am doing by cycling down Centre Avenue. I may thus answer the question why I am cycling down Centre Avenue by simply re-describing what I am doing as an instance of the type *getting groceries*. The answer to the question “*Why* are you doing ...?” may also be given to the question “*What* are you doing?”

Two qualifications are in order regarding this first feature of types. First, the question “*Why?*” is not always properly answered by giving a re-description of the action in question. One may also refer to an external result of the action, or to circumstances that appear to make it appropriate. For instance, I may answer the question why I am cycling down Centre Avenue by saying that this where the grocery store is and that I am out of food. Second, it is of course true that not all instances of cycling down Centre Avenue are instances of getting groceries; whereas all beavers are mammals. The hierarchy of types of actions is not a hierarchy of classes, but a hierarchy of types that may be simultaneously instantiated by one particular instance.

¹ It is often thought that we classify living beings according to their genes. I do not think this is true. An animal that is born from beavers and behaves in every respect like a beaver, with the only exception that it has different genes, would be properly classified as a beaver, albeit a very atypical one. On the other hand, an animal that is not born from a beaver but is like one in all other respects will also qualify as a beaver.

But still, there *is* a hierarchy of types that a particular action instantiates, and this hierarchy may be depicted by a tree diagram. Getting groceries, for instance, may involve cycling down Centre Avenue and then locking one's bike in front of the Giant Eagle, or the Whole Foods market. Both are actions that may, in the given circumstances, be described as instances of the more general type *getting groceries*. (I am locking my bike in front of Whole Foods. You ask me what I am doing. One correct answer is: "I am getting groceries.")

As for the second feature of types, it is also easily seen that the classification of an action as an instance of a more general type only points to certain crucial features, while the actual performance will always have several other, possibly unknown features. In order to find out more about the type *getting groceries*, it is never enough to look at its definition—if there is one. Like animals and unlike mathematical objects, actions exist in the real world and are not brought into existence by definition (although they may depend on certain constitutive rules). Hence, in order to study the nature of an action, one must observe real instances of its type.

Finally, in classifying an action according to the intention with which it is done, we do not imply that the action actually took the intended course. It may be an incomplete or atypical member of its type. This is the most striking feature that actions and living beings share. A particular animal may be a quadruped without having four feet, and a particular action may be an instance of crossing a street without it being the case that in the end, the street has been crossed. I might be cycling down Centre Avenue and get hit by a truck, such that even though I *was getting groceries*, I *did not get* any in the end.

2. What Intentions Are Good For

So far I have argued that when we classify actions in terms of intentions, we do in fact divide them into what I have called types. But I have promised to show more, namely that the *most natural* way of classifying and characterizing actions is in terms of intentions. In order to see this, we must ask why one should want to classify actions in terms of intentions.

Suppose you have told someone to water your plants, and she has killed them. She may have done so for a number of reasons. For instance:

- (a) It might have happened by sheer accident, or
- (b) because she lacked certain skills, or
- (c) because she intended to kill your plants.

(a) If she killed your plants by accident, you have no basis for determining whether and under what circumstances she might do the same again. (An accident is not an instance of any type, and therefore cannot, *as* an accident, happen again.) (b) If the reason is lack of skill, you know what to do: either you do not let her water your plants again, or you instruct her how to do it. To the extent to which an agent has a certain capacity, one may reasonably expect her to succeed in doing certain things; given that she wants to do them. What these things are depends on the precise nature of the skill in question. For instance, if the person who failed to properly water your plants did so because she failed to remember, she will probably fail at other tasks that also involve memory. If she is bad at estimating amounts of water, she will fail at tasks that involve this capacity. In order to correct her, one must find out what the problem was. (c) Likewise, if she killed the plants intentionally, what you may expect her to do in the future depends on what her intention was. For instance, if her intention was to damage your possession (and consequently, her action was an instance of the type *damaging your possession*), you may safely assume that she will do other things that satisfy this description. What she intended was, because it was done intentionally, of a certain type. By doing what she intended, she has revealed herself to be a person who does things of that type.

We are interested in the reasons for which agents act because reasons are *general*. This holds true whether we refer to (in)capacities or intentions in our explanation of an action. A capacity to perform a particular action is a capacity to perform any number of actions of its type, and an intention to perform a particular action is an intention to quite generally perform instances of its type. When we ask what the intention of an agent was, we are interested in the type of action this agent performs intentionally. We are interested in the intentions of an agent because they tell us something about her character. There is a beautiful passage in Wittgenstein's *Philosophical Investigations*, quoted by Anscombe in *Intention*, where he says:

Why do I want to tell him about an intention too, as well as telling him what I

did? ... because I want to tell him something about *myself*, which goes beyond what happened at that time. (*Philosophical Investigations* §659; *Intention* §25)

We want to know the intentions of the agents we are dealing with because we want to know what to expect from them. This is why we look at the agent's intentions when we morally evaluate and sanction actions. (As Abelard says in his *Ethica* 17,3: *ne quae fiunt sed quo animo fiunt iustitia pensat.*) We blame agents for their intentional actions because we want to thereby alter their intentions. When we blame an agent for what she did, we do not blame the past *action*, for we could not change past actions anyway. We blame agents and not actions, because we do not want more actions of the same type. Likewise, we praise actions in order that there be more actions of the same type. The moral worth of an action is ultimately the worth of that which gives rise to more actions of the same type, and this is the agent and her character.

Further, if we are interested in predicting or influencing the behavior of an agent, we should be interested in her intentions rather than merely in her dispositions. Consider how difficult it is to reliably influence an agent's behavior without caring at all about her intentions. To put people in prison or to drug them are not only inefficient means for changing their behavior as compared to changing their intentions; they also disregard their status as persons. Such measures may work in some cases, but humans are far too complicated for that to be possible in general. For the purposes of human interaction, someone who does not act as a person and for her own reasons is nigh on unpredictable. For humans, to be reliable is to be rational and responsible. That is, what we want in people when we want to rely on them is that they intend and have reason to do the right things. Hence, where we may hope to influence someone's behavior by changing her intentions (rather than forcing her), we should always prefer to do so. Intentions provide the most direct and reliable handle on the present and future behavior of an agent, and this is why a classification of actions with a view to intentions is the most natural one. We classify actions with a view to that which is most important to us: the reasons an agent has, and will continue to have, for performing actions of a certain type.

3. From Intentions to Virtues

So far I have argued that intentions are the terms in which we properly classify

actions in order to bring out their moral significance. This view is meant to replace a fairly common view that I take to be mistaken: that intentions are events in the mind of an agent, which cause her to perform certain actions. I cannot afford to directly argue against this latter view here; I mention it only in order to make clear the negative thrust of my account. Intentions are not events that cause actions; they are that in terms of which we best classify actions.

I will now suggest a similar account of what virtues are. I will argue that just as intentions are not events that cause actions as their separable effect, virtues are not traits of an agent that cause a kind of behavior as their separable effect. Again, this is meant to replace a widespread view that I take to be mistaken: that virtues are dispositions in an agent, which cause her to consistently act in certain ways. As in the case of intentions, I will here concentrate on developing my own account, rather than setting it off, in detail, against its rivals.

In order to see where the virtues fit in, the first thing we need to do is to consider a rather special case: the possibility of justifying a generic action by simply re-describing it. When asked why an agent performs an action, the answer may be given

- (i) by reference to some relevant circumstances,
- (ii) by pointing out that the action has desirable external effects, or
- (iii) by re-describing it in terms of an intention.

It is important to distinguish between (ii) instrumental explanations and (iii) re-descriptions. I may explain why I am cycling down Centre Avenue by re-describing what I am doing as getting groceries. Asked further why I am getting groceries, I might simply say that I want to have paella for dinner. The latter is not a re-description of the action I am performing. I am getting groceries by cycling down Centre Avenue, but I am neither having nor wanting to have paella by getting groceries. By saying what I want, I give an explanation and justification of my action in terms of something that is external to it. Conversely, a re-description of an action, when given and accepted as an answer to a question “Why?”, is an explanation or justification of this action in terms of something internal to it.

Now an explanation of an action shows that it is reasonable, and a justification

shows that it is justified. As far as an explanation or justification by re-description can be given, the action is shown to be *intrinsically* reasonable or justified. The action of getting groceries is not intrinsically reasonable or justified; which of course makes it no less reasonable or justified. It is perfectly reasonable and justified, but only with a view to the end that is external to it: having paella for dinner.

Now consider a case where an action is, in a given context, fully and intrinsically justifiable by classifying it in terms of an intention with which it is done. You ask me why I am doing what I am doing, I reply by describing what I am doing in a certain way, and you are entirely satisfied with my answer. The re-description by which my action is justified will then leave no room for a further question “Why?” Depending on who you are, it might be a description such as “I am defending my country,” “I am helping the poor,” or “I am having sexual intercourse.”

In a further step, take this case and transpose it to the level of *generic* actions. Instead of considering my current behavior, the behavior of a particular agent in a particular context, consider the generic behavior of a generic agent in a generic context.

First, transpose the reference to my particular *action* to the generic level. The particular action that I am performing when I am cycling down Centre Avenue in order to get groceries is something that is properly attributed to me in the present progressive: I *am doing* it. A simple way of switching to a more generic level is to replace the present progressive by the simple present tense. Instead of asking why I *am cycling* down Centre Avenue, that is, the question should be why I *cycle* down Centre Avenue. This is a question about a habit or a way of acting, rather than about a particular action token.

Second, generalize the subject that performs the action. In something like the same way, we can move from a particular *subject*, such as myself, to a generic subject. The question is then, not why *I* cycle down Centre Avenue, but why *one* cycles down Centre Avenue.

Finally, take the very act by which the action in question is explained or justified to be a generic one. The question is, then: what does *one* accept as an explanation or justification of this kind of behavior? What does *one* think why *one cycles* down Centre Avenue? We are thus not any longer considering a case in which I explain to you why I am cycling down Centre Avenue, but a situation where question is what

one would accept as a justification that one cycles down Centre Avenue.²

This is, I submit, where we hit upon the virtues. Those re-descriptions of generic actions that *one* generically accepts as justifications are re-description in terms of virtues.

I am putting this in very general terms here, because I do not want to get involved in the question whether any particular kind of behavior actually is virtuous. I do not want to claim that courage, justice, or charity are virtues; not even that there are any virtues. All I want to maintain is that *if* it should be possible to re-describe a generic action, something that *one does*, in such a way that *one accepts* this re-description as a fully satisfactory justification of the action in question, then the terms in which one re-describes the action are virtues. Or, in a more condensed way: virtues are terms which, when used to describe a generic action, provide a generic justification of it. When we say of an action that it is virtuous, we say that it belongs to a type of actions that may be described in such a way that the description itself generically justifies it. When we say of someone that she is virtuous, we say that she is someone who does that kind of thing.

4. Implications and Conclusion

Let me draw some conclusions. Virtues are something like generic intentions: they perform the same function in explanations why *one does* a certain thing that intentions perform in explanations why *a particular agent is doing* something. Just as an intention is not just an inner episode in the mind of an agent, the virtues are not “fine inner states” of an agent (Swanton 2003:26). That on the basis of which we attribute virtues to agents is not inner and unobservable as opposed to presently observable; it rather is the generic behavior of an agent as opposed to her behavior right now. By calling someone virtuous, we say *what she does* as opposed to *what she is doing*; and this is not the same as a statement about what she has in mind. Just as in the default case, intentions are manifest in intentional actions, virtues are manifest in general ways of acting.

² In a community with a catholic background, for instance, this will rule out justifications of actions by simply describing them as instances of sexual intercourse. For *one* assumes, as a member of such a community, that sexual intercourse is only justified as a means for a further, external end. A description of the same act as procreation might do.

Further, just as intentions are not simply events that cause certain actions, virtues are not just dispositions to act and feel in the right way. Virtually all accounts of what virtues are classify them as qualities, dispositions, or traits of an agent that have certain effects. This may seem to be accurate, since the fact that someone is a virtuous agent may indeed be said to make her act in virtuous ways. But to call a virtue a disposition, tendency, or trait is misleading in that it may easily cause the impression that virtuous behavior is an *external* consequence of the agent's having some certain feature. However, the character trait that a virtuous person possesses is not some independently describable feature that causes her to act virtuously now and then. It is the quality or feature that entirely *consists in* acting virtuously now and then, and in nothing else. (This is why Aristotle describes virtues by simply telling us what their possessors do.) It is impossible to determine whether an agent is virtuous before determining whether she acts well.

Virtue ethics does not rely on the idea that one may first, by independent means, give an account of what it is to lead a good human life, and then determine whether an action is good by asking whether it contributes to that kind of life. (This seems to be what Rosalind Hursthouse, 1999:200f., has in mind.) Again, there is some truth in this idea, since a virtuous action does contribute to a good human life. But the good life that it contributes to is good only insofar as it is a life led in virtue. To lead a good human life is to act in the right ways. Hence, it would be a mistake to suppose that virtue ethicists attempt to derive ethical norms from some kind of metaphysical or even pseudo-biological assumption about the essence of human life. The point is, conversely, that to call an action good or desirable for its own sake is to establish part of what any account of a good human life should look like. We are not supposed to derive ethical norms from an independently established general account of human nature. Virtue ethicists do not propose actions or values on the basis of an account of human life. They say that to propose actions and values *is* to give an account of human life. The point that a virtue ethicist is trying to make is one about the status of ethical judgments. When we call an action good for its own sake, on whatever grounds, we do not thereby say that it complies with certain norms or that it benefits some significant number of sentient beings, but we say that performing this action is doing well as a human being. The evaluation of an action as good in itself is not implied by an account of human nature, but it should be taken to imply one.

References

- Abelard, Peter (1971). *Ethics*. Ed. Luscombe. Oxford: Clarendon Press.
- Anscombe, Elizabeth (1963). *Intention*. Second Edition. London: Basil Blackwell.
- Anscombe, Elizabeth (1981). *Collected Philosophical Papers*. Oxford: Basil Blackwell.
- Copp, David and Sobel, David (2004). "Morality and Virtue: An Assessment of Some Recent Work in Virtue Ethics," *Ethics* 114:514–554.
- Doris, John M. (2002). *Lack of Character. Personality and Moral Behavior*. Cambridge University Press.
- Hursthouse, Rosalind (1999). *On Virtue Ethics*. Oxford University Press.
- Laird, John (1946). "Act-Ethics and Agent-Ethics," *Mind* 55, 113–132.
- Thompson, Michael (forthcoming). "Naive Action Theory," in *Life and Action*, Harvard University Press.
- Swanton, Christine (2003). *Virtue Ethics. A Pluralistic View*. Oxford University Press.