HUB ZWART

CONSENSUS FORMATION AS A BASIC STRATEGY IN ETHICS

1. INTRODUCTION

Consensus formation is an ethical strategy for reaching viable, tenable judgements or decisions in dealing with problematic cases - one strategy among others. In some cases, we may prefer to use other (perhaps more reliable) strategies, such as testing the logical coherence and empirical validity of judgements, or their concordance with canonical documents and authoritative statements. Whenever these strategies are not available, however, or whenever they fail to produce convincing results, opting for consensus formation becomes the most reasonable alternative.

The term *consensus* first emerged in the history of philosophy as the *consensus gentium*-argument (Suhr, 1971). Notably, it was used as an argument (one among others) to ascertain the existence of God. It was argued that, since throughout the world human communities and cultures seem to believe in God, one way or another, there must be some truth to this idea, even if incontestable proof of His existence (either of a metaphysical or of an empirical fashion) is beyond the limited capacities of human reason (under mundane circumstances). In the medieval epoch, the consensus argument was largely supplanted by strategies of sound reasoning and concordance with canonical documents or authoritative statements (such as the Bible or papal Bulls). In the 16th century, however, as traditional authorities found themselves fundamentally contested, the concept and strategy of consensus made its reappearance. The Protestant churches arranged a series of historical consensus meetings (Consensus Tigurinus (1549), Consensus Pastorum Genevensium (1551), Consensus Sendomiriensis (1570)) in order to contain the centrifugal tendencies, inherent to the protestant movement. Finally, in the present, the concept of consensus formation has proven its significance once again. The term 'consensus' is now used to refer to meetings of experts (such as physicians or medical ethicists) aimed at establishing a common standard of good professional practice. The *Appleton Consensus*, for example, refers to a meeting of bioethicists who tried

H.A.M.J. ten Have and B. Gordijn (eds.), Bioethics in a European Perspective, 281-288.

to reach consensus on a series of biomedical issues. *Consensus* now has become one of the standard terms (and standard techniques) of contemporary ethics (Zwart, 1998a).

Furthermore, the term consensus *a priori* contains a logic of its own. Whenever the word consensus is used, certain prepositions are already set to work. As soon as the word 'consensus' turns up, the moral discourse is bound to take a certain direction, determined by the tacit, inherent logic of the term. What does this logic look like?

To begin with, the term 'consensus' *a priori* refers to a middle condition, situated between two extremes - the one being a situation of complete agreement, the other a situation of complete disagreement. A completely closed community (fictitious no doubt) that unequivocally relies on some canonical document or other, interpreted by an authoritative readership, in accordance with a fixed set of interpretative rules, has no need whatsoever for a consensus strategy. All participants in a debate can be made to agree to a certain statement by more reliable and effective means, such as quoting (and adequately applying) the canonical documents at hand.

The other extreme (the situation of complete disagreement) is fictitious as well. If moral agreement is completely absent, consensus formation makes no sense at all and the parties involved will have recourse either to warfare or to arbitrary regulation. In short, consensus formation requires a considerable measure of agreement, while a limited number of (albeit important) issues is left open to debate. Logically speaking, moreover, the term *consensus* indicates that it is not something which can be *enforced* on others. Eventually, the consensus statement requires the free and deliberate consent of all parties involved. Also, the term *consensus* indicates that its outcome (the consensus statement) has not solely been adopted on rational grounds, but entails something like *sense* or *sensibility* as well.

Several strategies for reaching consensus have been developed. All these strategies will rely on some protocol or other. This protocol may be rudimentary (in which case the consensus process will be fairly open) or rather elaborate (in which case the consensus process will be fairly standardized). Yet, some kind of protocol will always be involved, at least in outline. Moreover, in every consensus procedure, there is a tendency towards more complete, more detailed protocolisation. Finally, it must be stressed that the aim of a moral protocol is not to ignore or disqualify the personal responsibility or sensibility of those involved, or to *force* them into some kind of final statement, but to indicate (as precisely as possible) where instances of serious disagreement are likely to occur and what the reasonable options are.

By focusing on case studies, I will now further clarify the logic of consensus formation in health care ethics, limiting myself to two cases: the case of animal experimentation (or more precisely, the role of the animal ethics committee in a university hospital) and the case of do-not-resuscitate-

decisions. The analysis will not involve concrete, real-life cases, but rather focus on *typical* cases, that is, on the basic moral *scripts* bound to emerge in the real-life cases we may encounter.

2. CASE 1: ANIMAL EXPERIMENTATION FOR MEDICAL RESEARCH

In all university hospitals, animal experimentation takes place, often on a considerable scale. Without experimentation on animals, the practice and development of modern medicine would be impossible.

Now imagine a sensitive human person who, for the first time in his life, enters a university hospital laboratory where experimentation on animals is actually being performed.[1] *Animals like rats, mice and goats are subjected to experimental trials and bound (at least in the majority of cases) to experience suffering or some level of 'discomfort'. Finding himself confronted with instances of animal suffering, the sensitive individual is likely to experience feelings of uneasiness or even disgust. We start, that is, from a dual situation: on the one hand the moral image of the suffering animal, and on the other the (more or less spontaneous and immediate) experience of sensitive moral subject. Responding to the situation in an impulsive manner, the sensitive person might for example flee the laboratory, or try to free the suffering animals from their cages.*

The mere feeling that something is not right here, however, does not suffice as a well-considered moral judgement or as a legitimate ground for action. It is, for example, somewhat unlikely for the professional care-takers or researchers working in this same laboratory to experience similar feelings. Quite on the contrary, they will no doubt be eager to indicate that there are certain moral reasons (apparently convincing to them) for participating in this kind of animal practice. That is, they are able to account for and justify what they are doing. Thus, the consensus process is initiated as soon as the sensitive person articulates his reasons for criticizing the practice at hand, while the professional articulates his reasons for regarding it as admissible or even necessary. The sensitive visitor is called upon to explain *why* he experiences uneasiness or even disgust in the face of animal discomfort. His condemnation of the situation *as a whole* gives way to an effort to *analyze* it as precisely and meticulously as possible. Likewise, the professional care-taker will be called upon to explain why he regards the situation as basically legitimate, why under these circumstances cruelty towards animals is not regarded as immoral, but as morally justified, etcetera.

Should we study a series of discussions like the one described above, between the sensitive lay-person and the conscientious professional, a certain monotony becomes noticeable, a certain basic script is bound to emerge, and

the speech acts of the participants turn out to be more and more predictable. At least implicitly, that is, they all adhere to a basic protocol. And we may try to articulate and formalize it, in order for the process of consensus formation to be furthered.

This is what happens, for example, during the meetings of an Animals Ethics Committee (AEC). On the one hand, it will be acknowledged that inducing discomfort to animals is problematic in itself. On the other hand, it will be acknowledged that a certain level of discomfort may be morally acceptable *if* the intentions behind it are legitimate ones - in terms of the scientific or social relevance of the experimental trial involved. A considerable number of participants in a consensus process is likely to agree, for instance, that it is morally permissible to sacrifice a limited number of rats in order to test an experimental cure for a disease from which a considerable number of human individuals are suffering. It then becomes the task of the AEC to determine at what point *precisely* a particular experiment on animals is regarded as morally justified, if both the animal *discomfort* and social or scientific *relevance* are taken into consideration.

It is possible now to discern a basic scheme to which the consensus formation process adheres. A series of stages and a number of transitions can be distinguished. The point of departure is the *image* of a situation - in this case the suffering animal, an image triggering uneasiness or even abhorrence, voiced in phrases like "Something here is not right!", or "This is not good!". It is the initial stage of the consensus formation process - a moral response triggered by an image of the situation as a whole.

Different people, however, will respond differently to different situations. The inherent logic of the consensus process now forces them to verbalize *why* they regard the situation as problematic, even objectionable. And they are able to do so by relying on certain basic *standard terms*, borrowed from the vocabulary of ethics, terms such as *discomfort* and *relevance*. Thus, the participants in the consensus formation process are provided with a limited set of items, a basic moral vocabulary that allows them to discern the basic structure of the situation and to really set the consensus formation process going.

Finally, as the consensus formation process proceeds further and further, a point will be reached where efforts towards quantification become relevant. For example, a scale will be introduced in order to determine the *precise level* of discomfort - that is, to *measure* it. A similar scale might be developed for relevance as well. Thus, it becomes possible to balance off discomfort and relevance in a less intuitive, more precise manner. Indeed, the consensus formation process is found to be guided by an inherent tendency to proceed from *images* (concrete situations) to *standard terms* (basic aspects or structures), and from standard terms to *numbers* (measurable variables).[2]

At a certain point, however, the consensus procedure runs the risk of becoming a routine and the sense of uneasiness is bound to recur - "Something is not right here!", "This is not good!", etcetera. The AEC and its procedures may now be regarded as biased and partial because of the fact that a fair majority of its members are themselves involved in research and will regard *any* level of discomfort acceptable as soon as *some* scientific relevance is to be expected. Moreover, the philosophical question might be raised whether a quantification in terms of physical discomfort really is an adequate way of capturing the moral phenomenon involved. At this point, we may feel that the standard moral vocabulary stands in need of broadening. A term like *intrinsic value*, for example, may serve to indicate that there is something of a problem in the instrumentalisation of animals *as such*, even in the absence of suffering. The introduction of such a term may encourage us to become more keen on limiting the number of experiments, for example. And finally, if taken to its logical conclusion, a proposal might be forwarded to quantify the concept of intrinsic value in some way - for instance by adding a fixed numeral to the score for *discomfort* by way of standard procedure.[3]

3. CASE 2: THE DECISION NOT TO RESUSCITATE

I will now turn to the analysis of a second *typical* case. Imagine the following situation.

A friendly and compassionate physician [A] pays a visit to an elderly patient and finds him in a deplorable state, which she describes as loss of decorum, loss of dignity, or grace. It is a terrible scene: the aspect of a patient whose physical state has suddenly deteriorated. In view of his physical condition, she decides not to resuscitate her patient, but allows him to die. Is there something wrong with this?

I think there is. Why this is so becomes clear as soon as another physician [B] pays a visit to this same patient and, finding him in the same deplorable state, describes it as a state of *emergency*. Instead of allowing him to die, she will immediately take a series of initiatives with the explicit objective of saving the patient's life. What we are faced with here is clearly a lack of consensus. Still, it is possible to initiate a process of consensus formation - on behalf of (more or less similar) future cases.

Let us analyze the case. To begin with we may note that both physicians are responding to a *scene*, an *image* of the situation as a *whole* [First Stage]. This situation is subsequently diagnosed in *moral* terms [Second Stage], borrowed from the basic vocabulary of ethics, such as *dignity* [physician A] or *emergency* [physician B]. Physician A basically claims that it is problematic, or

even objectionable, to prolong the life of a patient in case of loss of dignity. Physician B, on the other hand, may well claim that, in cases of emergency, it is the physician's duty to save the patient's life, whatever the circumstances. At this point, the transition from image to standard term, from immediate response to a verbal analysis of structural aspects and items, has already occurred. The participants in the debate will now find that moral standard terms like *dignity* and *emergency* have an inherent logic of their own. As soon as the word 'loss of dignity' is introduced, some options will seem more plausible than others, and some trains of thought will seem more convincing than others, due to the moral presuppositions inherent in using this particular term. Phrases like *loss of dignity* allow us to make sense of the initial image or situation, of our immediate response to the situation as a whole, but may imply a bias as well and may even limit our possibilities for action. By elaborating the conflict between 'dignity' and 'emergence', we may be able to analyze the case at hand more carefully.

Thus, a consensus formation process is initiated. In the course of this process, a series of aspects is likely to be acknowledged as highly relevant.

1. It is the duty of the physician to save the patient's life. Whether or not this must be regarded as categorical and unconditional, is open to debate. Should physician B maintain that this principle remains valid at all times, regardless of the circumstances, the level of disagreement is such that we must have recourse to alternative strategies, such as arbitrary regulation - for instance by taking votes - instead of consensus formation proper. The consensus formation process presupposes that the range of moral principles is limited, and it tries to determine (as precisely as possible) where and on what grounds reasonable limits to the physician's tendency to intervene can be set.
2. The phrase loss of dignity indicates that there *is* such a limit. Moreover, a scale may be used to predict health prospects in terms of the patient's future quality of life.[4] Thus, the transition phase from standard terms to numerals is reached. On the other hand, we may have doubts whether this particular score, this particular form of quantification really covers what the physician initially experienced as she first uttered the term 'loss of dignity'. Quantification may well help us to increase the precision of our moral diagnosis, but may also cause us to neglect important aspects that were conveyed by our initial phrases and responses.
3. Finally, the will of the patient will by many be acknowledged as a decisive aspect. Others, however, will question whether patients are really able to anticipate future preferences.

These aspects (and other, additional ones) can be built-in into a moral protocol or 'consensus statement'. Such a statement will not solve all our problems, nor will it predetermine all our decisions (as an authoritative judgement in a closed community would), but it will help us to distinguish

the clear cases from the difficult ones, and to indicate (as clearly as possible) where and how basic disagreements are likely to occur.

4. CONCLUSION

The consensus formation process adheres to a script that basically consists of three stages: *visualisation* (recognizing and responding to certain images, such as the image of the suffering animal or the image of the suffering elderly patient), *verbalisation* (articulating and justifying moral responses by means of a limited set of standard terms that allow us to discern the basic logical structure of the situation), and *quantification* (the effort to balance off relevant items in a more precise manner by introducing parameters, numerals and weights). A consensus statement, moreover, always remains open to contestation. Continuous questioning of the statement reached, of the provisional moral protocol, allows us to improve our efforts at visualisation, verbalisation and quantification. Thus, the consensus formation process may well be regarded as a practice of moral experimentation, indispensable for furthering the scientific quality and reliability of ethics.

NOTES

1. Cf. Zwart, 1998b.
2. This tendency to proceed from concrete, visual phenomena to verbal structures, and from verbal structures to numerical relationships, is of course a tendency which ethics has in common with any other science.
3. In fact, in a recent report on animal experimentation new model for balancing human and animal interests is presented. Items such as relevance and discomfort are to be carefully measured and scored in terms of a 10-point scale. Subsequently, the authors suggest that a fixed score of 2 points must be added in favor of animal interests in order to take their 'intrinsic value' into account (Vorstenbosch, *et al.*, 1997, p. 76).
4. The term quality of life is, of course, yet another standard term with a logic and persuasiveness of its own.

REFERENCES

Suhr, M.: 1971, Consensus gentium. In: Ritter, J. and Gruender, K. (eds.), *Historisches Wörterbuch der Philosophie*. Wissenschaftliche Buchgesellschaft, Darmstadt.

Vorstenbosch, J., Joles, J., Stafleu, F., and Tramper, R.: 1997, *Ethiek, dierproeven en de afweging van menselijke tegen dierlijke belangen*. Centrum voor Bio-ethiek en Gezondheidsrecht, Utrecht.

Zwart, H.: 1998a, Moral deliberation and moral warfare: Consensus formation in a pluralistic society [1998]. In: Have, H.A.M.J. ten, and Sass, H.M. (eds.), *Consensus formation in healthcare ethics*. Series: Philosophy and Medicine, Vol. 58; European Studies in Philosophy of Medicine, Vol. 2. Kluwer Academic Publishers, Dordrecht, p. 73-92.

Zwart, H.: 1998b, Moral legislation in the Netherlands on euthanasia and animal experimentation. *Gegenworte: Zeitschrift für den Disput über Wissen* 1: 24-30.