



Determining the common morality's norms in the sixth edition of *Principles of Biomedical Ethics*

Peter Herissone-Kelly^{1,2}

¹International School for Communities, Rights and Inclusion, University of Central Lancashire, Preston, UK
²University of Bolton, Bolton, UK

Correspondence to

Dr Peter Herissone-Kelly,
Philosophy Section, International
School for Communities, Rights
and Inclusion, University of
Central Lancashire, Preston PR1
2HE, UK;
pnherrissone-kelly@uclan.ac.uk

This paper is a contribution to
the 'Symposium on *Principles of
Biomedical Ethics*', Workshop,
December 16 2008, University
of Bochum, Germany.

Received 4 March 2009
Revised 11 June 2009
Accepted 15 June 2009

ABSTRACT

Tom Beauchamp and James Childress have always maintained that their four principles approach (otherwise known as principlism) is a globally applicable framework for biomedical ethics. This claim is grounded in their belief that the principles of respect for autonomy, non-maleficence, beneficence and justice form part of a 'common morality', or collection of very general norms to which everyone who is committed to morality subscribes. The difficulty, however, has always been how to demonstrate, at least in the absence of a full-blooded analysis of the concept of morality, whether the four principles are foundational, and so globally applicable, in this way. In the recently published sixth edition of *Principles of Biomedical Ethics*, an imaginative and non-question-begging empirical method of determining the common morality's norms is suggested. In this paper, I outline this method, before arguing that it is difficult to see how it might be thought to achieve its purpose.

COMMON MORALITY THEORY AND THE GLOBAL APPLICABILITY THESIS

One of the central tenets of the four principles approach to biomedical ethics, alongside its commitment to pluralism and its related emphasis on the need for agents to weigh and balance conflicting norms, is a claim that we can call the global applicability thesis. This thesis holds that principlism is universally applicable to the sorts of ethical issues that typically arise in the practice of medicine. This is a consequence of the supposed fact that the four principles—respect for autonomy, non-maleficence, beneficence and justice—are norms that form part of the common morality. The common morality, in turn, is to be understood as a collection of very general norms adhered to by everybody, everywhere, who is authentically committed to morality. In short, the global applicability thesis is true, so the thought goes, just because the norms that it claims to be globally applicable are derived from a common moral framework shared by all morally committed persons.

In a paper published in 2003, I suggested that many criticisms of the global applicability thesis rely on the view that the common morality theory is to be conceived of as empirical or sociological in nature.¹ When it is viewed in this way, I argued, two possible charges can be levelled against it. First, we might want to hold that, far from common morality theory's being empirically supported, empirical considerations can readily defeat it. We can produce empirical examples to show that there

is no one, universal morality, but instead many and varied particular moralities.

This last point concerning moral variation has been more than adequately dealt with in the sixth edition of *Principles of Biomedical Ethics*³, without abandonment of either common morality theory or the global applicability thesis. It has, that is, been made clearer than ever before that particular moralities are, in Beauchamp and Childress's view, one and all *specifications* of the common morality. Such specifications will vary across cultures; the general norms thus specified will not.

JUSTIFYING THE COMMON MORALITY THEORY IN THE SIXTH EDITION

The second worry I expressed in my original paper was a rather deeper one. If common morality theory is to be regarded as empirical in nature, we should by empirical means alone be able to demonstrate not only that a common morality exists, but that it comprises certain specific norms. And, I argued, it is difficult to see how such a justification could be achieved in a non-question-begging way. To explain: recall first of all that the common morality is by definition a set of general, and variously specifiable, norms that are subscribed to by anyone who is morally committed. Empirically to determine what the norms of such a morality are, then, we first have to select a reasonable number of morally committed people from a suitably large array of cultural backgrounds. But how are we to identify the morally committed, other than by the norms that they hold? And in identifying them in this way, it seems we have already decided to what norms the morally committed adhere, because we have used those norms to tell us who is to count as morally committed in the first place.

Nevertheless, I went on to suggest that, given a stronger reading of the principlist common morality theory, such criticisms can be made to miss their targets. On that stronger reading, the theory is to be regarded as a *conceptual* rather than an empirical thesis, according to which the norms that make up the common morality are to be construed as a set of very general values that are somehow *constitutive* of morality. The claim here is essentially that which Philippa Foot makes as follows:

[T]here are ... starting points fixed by the concept of morality. We might call them "definitional criteria" of moral good and evil, so long as it is clear that they belong to the concept of morality—to *the* definition and not to some definition which a man can choose for himself. What we say about such definitional criteria will be objectively true or false.²

In the sixth edition of *Principles of Biomedical Ethics*, Beauchamp and Childress make it clear that, as I had suggested, the common morality theory *can* be regarded as a conceptual thesis, and consequently as a fitting object of conceptual justification.³ However, they insist that forms of justification other than the conceptual remain possible. We might attempt a normative justification, showing that a particular set of norms flows from and is undergirded by a defensible normative moral theory. And, significantly given my 2003 argument, they argue that it is possible to develop a non-question-begging and effective empirical justificatory method.

If this last claim is true, it is of some significance. If we were constrained to rely solely on the possibility of a conceptual justification, it appears unlikely that we would ever reach a suitably uncontentious identification of the common morality's norms. The only way to arrive at a candidate for a set of universal norms would be through a fairly major philosophical analysis of the concept of morality, an analysis that would hold certain norms to be essential to that concept. And the trouble with analysing concepts of any real complexity is that one's results are unlikely to be universally regarded as uncontroversially correct. An empirical justificatory method may provide a more widely acceptable means of determining the norms of a common morality. In what follows I want briefly to set out the way in which Beauchamp and Childress think such a method would work, and then to ask at greater length whether it would be fitting to place our confidence in it.

A NON-QUESTION-BEGGING EMPIRICAL METHOD FOR DETERMINING THE COMMON MORALITY'S NORMS

Beauchamp and Childress present two possible methods for empirically determining the common morality's norms. I want to ask how successful the first of those methods is likely to be. Doing so will supply a number of points that could be transferred wholesale to an assessment of the second method, since the second is really only a variation on, and must rest upon the same foundations as, the first.

To start with, we need to be clear about how the proposed method manages to avoid the charge of being question-begging. Beauchamp and Childress write that:

An investigation would include only persons who had already been screened to assure that they are committed to a principle of morality that is reasonable to expect all morally committed persons to accept. We suggest that a reasonable principle to choose would be nonmaleficence, as it is unimaginable that any morally committed person would reject this principle.³

Once a suitably diverse group of agents who adhere to the norm of non-maleficence has been selected, we can go on empirically to establish what other general norms they share. Those further norms, it is suggested, will be the remaining constituents of the common morality.

At first glance, we might think that, in line with the criticism in my earlier paper, this method risks skewing the investigation's results by focusing only on people who subscribe to a norm that, prior to any justification, we just happen to think must be among the principles adhered to by the morally committed. Any proposed survey, we might object, is rendered suspect by our taking as a point of departure a norm from our own personal, pre-theoretical conception of morality. I believe that, given a suitably strong reading of what Beauchamp and Childress say, this objection can be shown to fail. It is entirely possible, and indeed plausible, to construe the sixth edition account of empirical justification as non-question-begging.

Recall Beauchamp and Childress's stated position: the principle of non-maleficence is a norm that we should use to identify the morally committed, just because it is supposedly *unimaginable* that it should be rejected by anybody accurately categorisable in that way. Precisely how we read this unimaginability claim will depend on what sense of 'unimaginable' we take to be intended in it. If the notion of unimaginability employed is that of a contingent fact about our (current) psychology, then it could be that we are simply appealing to our pre-theoretical conception of morality in judging non-maleficence to be a foundational norm, and the charge that the method is question-begging will not have been avoided. Consider that many people in the early-to-mid 20th century might have found it unimaginable that a morally committed person could regard homosexual practices as morally permissible. Most of us would now want to say that that just revealed a fact about the limitations of their moral vision, and not about the nature of morality itself. We are perhaps tempted to say that they ought to have exercised their imaginations a little harder.

On the other hand, we might want to suppose that we are unable to imagine a morally committed person rejecting a norm of non-maleficence, just because such a rejection would be, in the strongest sense, inconceivable, or incoherent. The thought may be that we can no more imagine a system of morality lacking a norm of non-maleficence than we can imagine a triangle lacking three sides. Such notions are inconceivable, not because of the contingent limits of our imagination, but because they are self-contradictory. This way of understanding the unimaginability claim rescues the proposed empirical method from being question-begging. It does this by assuming that the foundational nature of non-maleficence is a conceptual claim; adherence to non-maleficence is a mark of the morally committed, just because a norm of non-maleficence is a constitutive element of any system that can be counted as a morality in the first place. To understand the unimaginability claim in this way is not to adulterate an empirical justificatory method with conceptual elements. The application of the method remains a thoroughly empirical affair. But it does mean that the method is, so to speak, *conceptually informed*.

How can we justify the claim that non-maleficence is internal to morality, without first undertaking a fully fledged conceptual justification of common morality theory? Well, even in advance of a detailed analysis of the concept of morality, we can presumably identify *some* of its constitutive norms. There is a general truth about conceptual analysis here. If, for example, we were to attempt fully to explicate the concept of personhood, we would find ourselves faced with a quite considerable analytical task. But in advance of such a task, there will be certain features of the concept that are just obviously essential to it (say, that anything that qualifies as a person must be self-aware). So it is, we might claim, with morality and a principle of non-maleficence.

CONCERNS ABOUT THE EMPIRICAL METHOD

Beauchamp and Childress's proposed empirical method is, then, non-question-begging. But is it likely to deliver reliable results? I want to start my answer to this question by making the following point: if the empirical method is to be thought reliable, there must be some reason for supposing it will reveal the morally committed to adhere to those norms that are internal to the concept of morality, and which would be unearthed by an ideal conceptual justification. My thought here is that although there may be three entirely separate ways of justifying it, there can be only one common morality. If the proposed justifications

are to work, they must each deliver the same results: they must all identify, in their various ways, the privileged norms internal to morality.

Having established that much, I next want to address this question: why suppose that someone's adhering to one of the constitutive norms of morality (non-maleficence in Beauchamp and Childress's example) will entail, or even strongly support, their adhering to them all? Let us concede that the person who is authentically committed to a norm of non-maleficence ought to be counted as morally committed. Why does this license us to suppose that she will also be committed to the full range of general norms constitutive of morality? If we think it does, then we might seem to be relying on the claim that adherence to one norm that is internal to morality will invariably be accompanied by *knowledge* of which other norms are similarly privileged. It is not obvious why this claim should be true. Indeed, the claim might seem to suggest that anyone who is committed to non-maleficence, and who thereby qualifies as morally committed, has carried out, or somehow been made aware of the results of, a comprehensive analysis of the concept of morality, especially as we can hardly expect her to have carried out her own cross-cultural empirical study (and if she had undertaken such a study, we can still ask what reason she would have to trust its results). Prior to such an analysis, the subscriber to a principle of non-maleficence might just not know that, for example, respect for autonomy is a norm internal to morality. She may consequently not realise that her commitment to the institution of morality demands that she be committed to that value.

A 'UNITY OF THE PRINCIPLES' DOCTRINE?

We might escape this last objection by insisting that propositional knowledge of the content of the concept of morality is not necessary here. We might urge that there is instead some sort of analogue of the Ancient Greek notion of the unity of the virtues at work, and label that analogue 'the unity of the principles doctrine'.

The central tenet of the unity of the virtues doctrine is that if an agent has one virtue, she automatically has them all. At first sight, this might seem a dubious claim, though on closer inspection the doctrine is found not to be without a certain appeal. The thought is that each of the virtues is, so to speak, implicated in all the others. Take as an example the virtues of kindness and courage. We may suppose it incontrovertibly true that a person could be kind, and yet want for courage. However, we need to consider that for a person genuinely to have the virtue of kindness is for her to be someone who can be relied upon to act kindly across the broad range of circumstances in which such action is required. Among these circumstances will be situations in which her acting kindly will demand that she also act with courage. Therefore, no-one is genuinely kind—can be relied upon to act kindly in the full range of appropriate circumstances—unless she is also courageous.

Might a similar unity attach to the principles of the common morality? Are they, too, implicated in each other in such a way that if a person were to adhere to one, she would find herself automatically adhering to all the others? Precisely because such adherence *would* be automatic, we could happily jettison the unpalatable supposition that all those who are morally committed must have carried out a full-blooded analysis of the concept of morality, or been apprised of the results of one. Nonetheless, it is not immediately apparent how we might set about arguing for the unity of the principles doctrine. And crucially, if the argument were to prove successful, its success might be more damaging than helpful to the principlist

approach. That is, it is difficult to see how a claim that all the common morality's principles are implicated in each other could be reconciled with principlism's radical pluralism. Morality, according to Beauchamp and Childress, is supposed to encompass diverse, variegated values. Positing too much unity among the common morality's principles might be thought to imperil that diversity. This, I take it, would be too high a price to pay for an empirical means of establishing morality's privileged norms.

A 'UNITY OF THE VIRTUES' DOCTRINE?

There may, nonetheless, be another, related avenue open to Beauchamp and Childress. Instead of embracing an *analogue* of the unity of the virtues doctrine, they might embrace that doctrine itself. It is held in *Principles of Biomedical Ethics* that the common morality's principles will have corresponding virtues. For example, the person who is committed to the principle of respect for autonomy will possess and exercise the virtue of respectfulness. Similar correspondences will hold between the principle of non-maleficence and the virtue of non-malevolence; between the principle of beneficence and the virtue of benevolence; and between the principle of justice and the virtue of justice, or fairness.³ Might we not say, then, that these *virtues* are implicated in each other, while their corresponding *principles* are not?

Adoption of this strategy would help explain how it is that a person committed to a principle of non-maleficence, will also be committed to all the other principles in the common morality. Since she adheres to a norm of non-maleficence, she possesses the virtue of non-malevolence. And since she possesses the virtue of non-malevolence, she also possesses all the other virtues, including benevolence, respectfulness and so on. Possession of *those* virtues, in turn, ensures that she subscribes to their corresponding principles (beneficence, respect for autonomy and so on), since there is a close conceptual tie between each virtue and its matching principle. At the same time, since the postulated unity holds only between the virtues, and not between the principles with which those virtues are paired, the unity of the principles doctrine can be held to be false, and principlism's radical pluralism consequently be safeguarded.

Unfortunately, I am less than sure that this move away from the unity of the principles and towards the unity of the virtues would be especially useful to the principlist. First, even though the unity of the virtues doctrine might be more acceptable than it appears at first sight, it seems to me that a considerable amount of argument would be needed before it could be regarded as anything like firmly established. Second, and more importantly, it might seem reasonable to suppose that if the unity of the virtues doctrine is correct, and all the virtues are implicated in each other, then that unity will be transferred to and infect the corresponding principles. In other words, if it is correct to say that, for example, the virtue of benevolence is implicated in the virtue of respectfulness, then since there is such a strong conceptual tie between each virtue and its corresponding principle, we may seem forced to conclude that the principle of beneficence is implicated in the principle of respect for autonomy, and indeed in all the other principles of the common morality. Once again, principlism's radical pluralism is placed under threat. If we want to retain that pluralism, we need to drop both the unity of the principles doctrine and the unity of the virtues doctrine.

CONCLUSION

What I have aimed to do in this paper is briefly to assess the reliability of Beauchamp and Childress's proposed method of

empirically determining the common morality's norms, as it is presented in the sixth edition of *Principles of Biomedical Ethics*. If that method is to do its job, it should produce the same results as an ideal conceptual justification; however, I have argued that there seems no good reason to suppose that it will.

Of course, we may hold that all this shows is that the agents we study can make mistakes—that they may not recognise, for example, that a principle of respect for autonomy is every bit as integral to the concept of morality as is the principle of non-maleficence to which they already subscribe.

But from this claim, there seems to follow one of two others, both of which, I take it, would be uncongenial to the principlist. Either (a) in order to check it for accuracy, we are forced to compare the outcome of an empirical survey with the outcome of an ideal conceptual analysis, or (b) we must admit that the fallible agents that feature in our survey are not really committed to morality after all, since the common morality is by definition a set of norms *all* of which are shared by all morally committed persons. Option (a) is unattractive on two counts. First, it would undermine the supposed independence of an empirical justification from a full-blooded analysis of the concept of morality. Second, it would compel us ultimately to rely on the results of a conceptual justification that would be

more contentious than we had hoped an empirical justification would be. Option (b) would be similarly unsatisfactory, as it would leave us with no way of reliably identifying the morally committed other than by first determining all the constituent norms of the common morality via a conceptual justification. That is, an agent's adherence to non-maleficence alone would not be sufficient to establish that she is committed to morality; the fact of agent-fallibility means that only those who are committed to *all* the common morality's norms can qualify as morally committed. Consequently, we need to know what all those norms are in advance of an empirical justification. And this would clearly render an empirical justification superfluous.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Herissone-Kelly P.** The principlist approach to bioethics and its stormy journey overseas. In: Häyry M, Takala T, eds. *Scratching the surface of bioethics*. Amsterdam/ New York: Rodopi, 2003:65–77.
2. **Foot P.** Morality and art. In: Honderich T, Burnyeat M, eds. *Philosophy as it is*. London: Pelican Books, 1979:14.
3. **Beauchamp TL,** Childress JF. *Principles of biomedical ethics*, 6th edn. Oxford: Oxford University Press 2009:395, 393–4, 45–6.

Information in a hurry...

If you need the latest information in emergency care then you need the **Emergency Medicine Journal**. Packed with research, educational papers and debate of all aspects of emergency medicine, the journal will make sure you know everything you need to.

FOR MORE DETAILS OR TO SUBSCRIBE,
VISIT THE WEBSITE TODAY

emj.bmj.com



BMJ Journals