

## **Stavroula Tsinorema**

University of Crete, Philosophy and Social Studies Department, Rethymno Campus,  
GR-74100 Rethymno (Crete)  
tsinorev@uoc.gr

# **The Principle of Autonomy and the Ethics of Advance Directives**

### **Abstract**

Advance directives are conceptualised as a means of increasing “patient autonomy”, as they enforce individuals’ power of choice over a post-competence dying process. There is, however, controversy over their moral force. Rebecca Dresser and John Robertson offer a conceptual argument grounded in epistemological considerations concerning personhood which challenges their authority. Roland Dworkin defends forcefully “precedent autonomy” in planning post-competence medical care. This paper examines the above opposing theses and assesses their main arguments. Limitations are detected in both. Regarding the former, its conceptualisation of the notion of personhood is found to be problematic, and regarding the latter, its conception of individual autonomy is found to be too narrow. An alternative route is explored by reconstructing Kant’s conception of moral autonomy. It provides a framework for moral reasoning, from which certain contemporary understandings of autonomy as a right, as a reflective capacity of the individual, as responsibility and integrity can be properly assessed and justified. Normative conclusions follow regarding the extension of personal autonomy in advance medical choice.

### **Key words**

advance directives, agency, autonomy, Rebecca Dresser, Ronald Dworkin, Immanuel Kant, personhood, respect

Discussions of the ethics of advance directives have intensified in recent years as a result of unprecedented advances in medical technologies and subsequent applications in healthcare, which enable patients to be kept alive beyond the point at which they are competent to express their consent. They are taken to show how individuals can exercise greater control over their treatment, and thus how their power of choice can be reinforced and extended in the probability of their own future marginal competence or complete incompetence.

Advance directives are advocated as a means of preserving patient self-determination at the end of life, as they may particularly improve surrogates’ understanding of patients’ wishes regarding life-sustaining treatment. They may take two specific forms: they either give instructions in advance on the kinds of treatment that should or should not be provided (instruction directives), or designate someone else (a proxy) to make decisions on the author’s behalf in accordance with what she would wish to be done in the event of becoming incompetent (appointment directives).

They are conceptualised – and justified – as enactments of the principle of “patient autonomy”, which has acquired unique prominence in contemporary (bio)ethical discussions. They highlight the importance of being in control of one’s own life, even in cases of future severely incapacitating illnesses or loss of mental powers which hinder one’s capacity to make choices. They

are motivated by fears of being medically over-treated should one end up in a state of incompetence. On the whole, authors tend to specify the situations in which they would want treatment to be administered or discontinued.

Questions arise about the moral authority of advance choice, as well as its scope and application in a wide variety of cases. Executing an advance directive in anticipation of loss of consciousness, for instance, is different from that which is executed in anticipation of dementia. In the former case, its moral justification seems less debatable because, at the time that it is meant to take effect, there is no active agent whose choices need to be taken into account. In the case of severe dementia, however, the author of the advance directive wills to take control of her future at a time when she may still possess some attributes of agency, may be capable of experiencing pleasure and pain or discomfort, and of expressing needs and wants. In the latter case, the patient may be conscious, but partially or completely incompetent.

The worry, in such cases, is that the person who authored the advanced directive has undergone such severe transformation that there exists a patient who is a self very different from the author. Unlike the permanently unconscious, the demented individual may still have experiences, wants and desires, and therefore interests, which may conflict with those expressed in the advance directive. A standard criticism of such cases is that one person's treatment (the patient's) is dictated by another person's wishes (the author of the advance directive). And this is morally unacceptable.

The challenge goes deep and concerns the proper understanding and normative force of individual autonomy extending to the far end of human agency. What are the moral grounds for honouring advance directives and what is the scope of the latter? Objections fall within a large scheme of sceptical questions about the conceptual and normative dimensions of "precedent autonomy" and the idea of autonomous choices extending into the future, with authority extending into a period of an agent's incompetence.

### **"Margins of agency" and precedent autonomy: Two perspectives**

Two major theoretical approaches have dominated the current debate: that of Rebecca Dresser and John Robertson, on the one hand, and that of Ronald Dworkin, on the other. Both invoke a certain view of personhood, but they part company in their understanding of this notion.

#### ***a. The Dresser and Robertson thesis: Is there a problem with personal identity?***

Dresser and Robertson aim to undermine the moral force of instructional advance choice by invoking a conceptual argument about personhood and personal identity, which derives from empiricist epistemic premises. As they put it:

"It is difficult, if not impossible, for competent individuals to predict their interests in future treatment situations when they are incompetent because their needs and interests will have so radically changed."<sup>1</sup>

To substantiate their claim, they appeal to a certain philosophical theory of personal identity associated with Derek Parfit,<sup>2</sup> according to which personhood, one's status as a person, is a matter of one's psychology. Personal identity refers to that aspect of one's psychology which makes one the particular

person that one is. It consists in the continuity of psychological states through time. If one had a radically different psychology, then one would not be the particular person that one is. Similarly, if one's current psychology were to change drastically, then one's particular identity would also change.

On this reading, as the psychological connectedness and continuity between different stages in the life of a person may decrease, this decrease can alter the force of our normative commitments. In the case of deep psychological changes between the former and later self of a demented person, one could say that they are different persons. In such cases, an advance directive issued by the former person cannot have moral force for the course of action to be taken with regard to the person existing after the psychological change.

Dresser and Robertson have systematically defended the claim that, if the degree of psychological discontinuity is so great that we may talk about two different selves, then the demented subject's contemporary interests are the ones that should be legitimately cared for. The patient's earlier preferences should not count because the demented patient may not be the same person as the one who once expressed these preferences.<sup>3</sup> Caretakers and medical practitioners are faced with individuals who, although not in the "full maturity of their capacities", are nonetheless conscious beings capable of pleasure and pain, who can have claims about their current needs and desires. Why should one, morally speaking, ignore these in the name of past values which are now extinct? Dresser puts the issue succinctly as follows:

"Legal decision-makers have accepted the dubious notion that what was vitally important to incompetent patients when they were competent remains vitally important to them in their incompetent states. But incompetent patients differ from competent patients in material ways that invalidate this notion. Incompetent patients are incapable of appreciating the values and preferences they once held dear. As a consequence, standards attempting to honor those values and preferences fail to advance the incompetent patient's present welfare."<sup>4</sup>

There is a moral requirement for medical professionals to do what is in the current best interests of the now incompetent or marginally competent patient, in situations in which her earlier wishes conflict with these interests. Decisions taken by others which affect a demented subject at a given time must be true to that person's point of view at that given time. Abiding by and trying to fulfil the wishes and values which the patient no longer endorses or experiences may harm the patient. Along these lines, Dresser rejects the view

1

Rebecca S. Dresser, John A. Robertson, "Quality of Life and Non-Treatment Decisions for Incompetent Patients: A Critique of the Orthodox Approach", *Law, Medicine & Health Care*, Vol. 17 (1989), No. 3, pp. 234–244 (239).

2

Derek Parfit, *Reasons and Persons*, Oxford University Press, Oxford 1984. Parfit's approach has invigorated the 18th century Lockean tradition. John Locke proposes a relational view of personal identity as opposed to a substance-based, Cartesian, view of it (according to which, identity is preserved between two stages of a person's stages only in virtue of his/her consisting in one and the same substance). Against the latter, Locke argues that X at t1 is identical to Y at t2 just in case X and Y are related via consciousness,

i.e., just in case Y remembers the thoughts and experiences of X. As he puts, it, in case Y's consciousness "can be extended backwards" to X (John Locke, "Of Identity and Diversity", in: John Locke, *An Essay Concerning Human Understanding* [1694], Penguin Classics, London 1998, Book II, Ch. XXVII).

3

Rebecca Dresser, "Life, Death, and Incompetent Patients: Conceptual Infirmities and Hidden Values in the Law", *Arizona Law Review*, Vol. 28 (1986), No. 3, pp. 373–405; cf. R. S. Dresser, J. A. Robertson, "Quality of Life and Non-Treatment Decisions for Incompetent Patients".

4

R. Dresser, "Life, Death, and Incompetent Patients", p. 389.

that, in the event of becoming incompetent, patients retain the right of competent patients to refuse treatment including life-saving measures.

However, the justification of the Dresser-Robertson thesis rests on doubtful premises. As Joel Feinberg has argued,<sup>5</sup> a person's interests may be harmed even when she does not and can never experience the harm. To suggest otherwise is to take too narrow a view of what an interest is. Posthumous harms are a case in point. Such harms may occur when somebody's treasured goods suffer destruction, or her valued reputation as a person is undermined by wicked lies, or when important promises that she was given to be performed after death are broken. For similar reasons, a person's interests may be harmed by the way she is treated after she loses her cognitive capacity and awareness of the fact that such harms are being done to her.

The Dresser-Robertson thesis rests on a questionable conception of personhood which does not square adequately with a structural feature of personhood, namely that of persons as agents. Firstly, their analysis blurs the distinction between competence and personal identity. In saying that psychological discontinuity leads to a change of personality, they seem to ignore the gap between loss of competence for medical decision-making and change or loss of personality. Competence is unravelled in terms of certain psychological characteristics, and these are, in turn, taken to define personhood. Nonetheless, the move is left unargued. Secondly, it is unclear why such a conception of personhood should have any moral relevance at all. Personality is a far richer notion than their empiricist methodology recognises. It is a normative notion, and not a fact-stating matter. Their perspective is unable to answer adequately the question of why one's interests or preferences ought to be cared for in the first instance. One cannot validly conclude that something ought to be the case morally just because it works in a certain way psychologically. Even in cases in which there is psychological connectedness between the different stages of one's life, why should one's preferences or wishes (contemporaneous or advance) bind another's action (the surrogate's or anyone else's)? The transition from factual premises about psychology (facts about preferences and other mental states) to normative conclusions about moral commitment (what ought to be done) is epistemologically questionable.

### ***b. Ronald Dworkin: Autonomy as integrity***

A different response to the moral authority of advance instruction comes from Ronald Dworkin, who furnishes reasons for adhering to the demented person's earlier instructions by appealing to the "right to autonomy" extending into one's future.<sup>6</sup> His conception of personhood stresses the active side of our human nature, that is, the standpoint that we adopt when we view ourselves not merely as subjects of (psychological) experiences through time, but as agents who actively decide what to do and take responsibility for what they do. This standpoint separates our actions from mere behavioural events determined by biological and psychological laws. Dworkin's approach focuses on the idea of freedom to make our own choices, do things for ourselves and take responsibility for what we do. This is what it means to be persons as opposed to mere things. Personhood is tied up with the practical standpoint, i.e., with agency and responsibility.

From the practical standpoint, we view our relationship to our choices and actions as that of an author: we view them as *our own*. When we think of ourselves in this way, our own lives matter to us personally. We conceptualise ourselves as living our lives, including having certain psychological experi-

ences, as something we actively engage in, as something that we do. Dworkin conceptualises this important feature of our sense of personal identity, which the Dresser-Robertson thesis leaves out, through his notion of personal autonomy. Personhood, for him, is intertwined with the capacity for autonomy. He understands the latter as “the ability to act out of genuine preference or character or conviction or sense of self”.<sup>7</sup> To be a person is to be an autonomous agent, and autonomy implies taking responsibility for one’s life.

The decisive criterion of autonomy is the right to govern the course of one’s life, including one’s future incompetence according to a “recognised and coherent scheme of value”.<sup>8</sup> The value of autonomy

“... lies on the scheme of responsibility it creates: autonomy makes each of us responsible for shaping our own lives according to some coherent and distinctive sense of character, conviction, and interest. It allows us to lead our own lives rather than being led along them, so that each of us can be [...] what he has made of himself.”<sup>9</sup>

A human life is to be judged “as we judge a literary work [...] whose bad ending mars what went before”.<sup>10</sup> From the point of view of a person, it is not “zoe”, physical or biological life, that has inviolable moral worth but “biography”.<sup>11</sup> When an autonomous person becomes incompetent, we “worry about the effect of his life’s last stage on the character of his life as a whole, as we might worry about the effect of a play’s last scene or a poem’s last stanza on the entire creative work”.<sup>12</sup> A bad ending may be worse than death as it may leave a person “a narrative wreck [...], a life worse than one that ends when its activity ends”.<sup>13</sup>

We ought to adhere to the patient’s earlier values and wishes, which originated when she was still capable of acting autonomously and capable of judging what was important for her overall well-being. We ought to respect the patient’s “precedent autonomy”, because we ought to respect the abilities and capacities that gave rise to such values and choices. We ought to respect that person’s personhood when it was intact. He specifies “precedent autonomy” as follows:

“A competent person’s right to autonomy requires that his past decisions, about how he is to be treated if he becomes demented, be respected even if they do not represent, and even if they contradict, the desires he has when we respect them, provided he did not change his mind while he was still in charge of his own life.”<sup>14</sup>

Dworkin furnishes two arguments in support of the above claim, one grounded in personal autonomy and the other in beneficence,<sup>15</sup> both of which lead to

5  
Joel Feinberg, *The Moral Limits of the Criminal Law: Vol. 3 – Harm to Self*, Oxford University Press, New York 1984, pp. 83–95.

6  
Ronald Dworkin, *Life’s Dominion*, Alfred A. Knopf, New York 1993, pp. 218–237.

7  
*Ibid.*, p. 225.

8  
Ronald Dworkin, “Autonomy and the Demented Self”, *The Milbank Quarterly*, Vol. 64 (1986), suppl. 2, pp. 10–14.

9  
*Ibid.*, p. 8.

10  
R. Dworkin, *Life’s Dominion*, p. 27.

11  
*Ibid.*, pp. 82–83.

12  
*Ibid.*, p. 199.

13  
*Ibid.*, p. 211.

14  
R. Dworkin, “Autonomy and the Demented Self”, p. 13.

15  
R. Dworkin, *Life’s Dominion*, pp. 222–229, 229–231.

the moral prioritisation of deeply held goals and values ahead of one's future incompetent mental condition. The reasons for honouring advance directives stem from considerations of respect for the right to individual autonomy and (their author's) right to our beneficence. He understands beneficence as the right which a person entrusted to the care of another has that the latter make decisions in the best interests of the former. Just as it would be a mistake to conceive a demented person as a person with the right to autonomy and interests of *her own*, which are distinct from those of her previous competent stage, so it would be an error to think that we would satisfy the patient's best interests if we ignored her previous choices.

Unravelling the notion of best interests, he draws a distinction between one's "experiential" and one's "critical" interests. Victims of dementia, *qua* demented, may still have experiential interests in their lives, i.e., good or bad experiences, such as enjoying comfort or feeling pain or fear. But they have lost the capacity to think about how to make their lives more successful as a whole:

"They are ignorant of self – not as an amnesiac is, not simply because they cannot identify their pasts – but more fundamentally, because they have no sense of a whole life, a past joined to a future, that could be the object of any evaluation or concern as a whole. They, therefore, have no contemporary opinion about their own critical interests."<sup>16</sup>

Although they cannot have a view of their life as a whole, and so they cannot make a judgement about their critical interests, nor can they assess what makes their life a success or a failure, they, nevertheless, have to be treated as persons who have become demented. The autonomy and interests of the whole personality – whose life should be viewed as encompassing far more than the period of their dementia – should be taken into consideration. And these can be affected by what happens to individuals in this final, demented, stage. Respecting both the patient's right to autonomy and her best (critical) interests (her "right to beneficence") requires adhering to her previous will, giving weight to her critical interests as she conceived them when she was competent to do so. There is no conflict between autonomy and beneficence in deciding what to do in such cases. Demented patients' advance directives ought to be honoured.

In a nutshell, Dworkin's analysis leads to the unambiguous conclusion that advance directives should be respected, because they are expressions of the critical interests that a person has and constitute that person's well-being, manifesting at the same time the exercise of her autonomy. His reconstruction of autonomy and beneficence leads to the claim to non-interference and to respecting the will of competent individuals concerning future treatment, even if a demented person expresses a will to the contrary. The rights of competent and conscious individuals to continue treatment, refuse treatment, or end their lives through more active means or interventions (so long as they did not change their minds while competent) are to be respected.

At the core of his understanding of agency and autonomy is the decision-making capacity of the individual and the ability to be in control and plan one's life as a whole, which warrants protection from unacceptable paternalism. This reconstruction of personal autonomy has its roots in John Stuart Mill.<sup>17</sup> He strikes a Millian note when he argues:

"Recognizing an individual right of autonomy makes self-creation possible. It allows each of us to be responsible for shaping our lives according to our own coherent or incoherent – but, in any case, distinctive – personality. It allows us to lead our own lives rather than be led along them,

so that each of us can be, to the extent a scheme of rights can make this possible, what we have made of ourselves.”<sup>18</sup>

Dworkin draws attention to important aspects of individuality and the good life, including autonomous living. However, two premises in his argument need further defending, namely, (a) the view that the demented patient can no longer generate “critical interests”, and (b) that the crucial dimension of autonomy is that of a capacity for being able to set one’s own values and goals and follow them.

Regarding (a), Dworkin assumes that critical interests stem from “convictions about what helps to make a life good on the whole”.<sup>19</sup> However, critical interests may plausibly be said to be generated from something less: simply from convictions concerning what is good to have, without requiring the ability to grasp, or review, one’s life as a whole. This not only characterises dementia patients, but may also be ascribed to many of the rest of us. Critical interests need not be understood in terms of a person’s grasp of what is good for her life as a whole, but can be traced to somebody’s convictions about what would be good for her – what is valuable for her to have or achieve.

A forceful criticism along these lines comes from Agnieszka Jaworska,<sup>20</sup> who – while agreeing with Dworkin about the normative dominance of critical interests over experiential ones – understands persons with dementia as still capable of generating new critical interests, including those about the value of their life. In order to substantiate her thesis, she appeals to a distinction between values and desires. While desiring is a more basic, first-order notion, valuing involves reflection. We do not always value what we desire. We often try to give up a bad habit, for example. To value something is not merely to desire it, but to think that it is good; and this is a second-order appraisal. For Dworkin, to form critical interests one must be capable of having convictions about what makes one’s own life good as a whole. She objects that critical interests can stem from simpler second-order desires, and that convictions about what is good to have do not require the ability to grasp or reflect upon one’s life as a whole.

Along these lines of argument, she contends that a demented person who has the capacity to value (even if that person has lost the capacity to implement her values) is capable of generating new critical interests. And the latter should be taken into account, thereby relativising the force of advance choice. It is one thing for someone to write an advance directive planning treatment in case of permanent unconsciousness, but quite another in case of dementia. In the first instance, the directive has full authority because later there will be no active agent to change some of her values, while in the second, later there will still be a person who is a valuer.

A demented person may be capable of generating contemporaneous critical interests as a person capable of valuing things. And her values may change and become different from the ones she endorsed when she was healthy. If

16  
Ibid., p. 230.

17  
John Stuart Mill, “On Liberty”, in: John Stuart Mill, *On Liberty* [1859], Penguin Classics, London 1982, Ch. 3.

18  
R. Dworkin, *Life’s Dominion*, p. 224.

19  
Ibid., pp. 201–202.

20  
Agnieszka Jaworska, “Respecting the Margins of Agency: Alzheimer’s Patients and the Capacity to Value”, *Philosophy and Public Affairs*, Vol. 28 (1999), No. 2, pp. 105–138.

this is so, then the conflict which occurs in our dilemmas about caring for that person may be best described as a conflict between the patient's ongoing experiential interests and her ongoing critical interests.<sup>21</sup>

However, in crucial respects, Jaworska and Dworkin do not address each other's point. Their views are grounded in different conceptions of critical interests, which lead them to different moral assessments of dementia. For Jaworska, anyone capable of second-order desires is a valuer, thus capable of generating new critical interests by changing one's mind. For Dworkin, on the other hand, critical interests involve more than just any values or second-order desires. They involve particular evaluative characteristics, such as life-long commitments and views of "what makes life good as a whole"; they involve the integrity of one's life in its entirety. In the end, Jaworska does not show that Dworkin is mistaken about the ability of demented patients to generate new critical interests in his sense of critical interests. Rather, what she formulates is a different conception of critical interests, one that does not address the kinds of reasons that Dworkin offers for insisting that critical interests override experiential interests in the overall assessment of best interests, namely their grounding in life's integrity, the value that people attach to their lives as a whole.

As for (b), it seems to me that, here, Dworkin's approach is at its most vulnerable. It concerns the fact that he invokes too narrow a conception of autonomy in ethics. He argues that it matters morally to individuals that they each be able to lead a self-directed life, which each personally finds fulfilling. He focuses on autonomy as an individual's capacity, and value-formation as the basis for the exercise of this capacity. Autonomy is thus reconstructed as an individual's capacity for value-formation. The value of personal autonomy emerges as the value of an individual's rationally self-directed life in the above sense. Within this scheme, the ethics of personal autonomy rests on the acceptance of a fundamental right of individuals to make choices with regard to their own bodies. However, this reconstruction is weak at the point at which it restricts justification for respect for one's chosen action to a capacity to have a distinctive character and act out of a sense of identity with one's values (individuality). Only where this capacity exists over a continuous period of time, according to Dworkin, does one have the required abilities so as to claim a right to autonomy.

However, why should a personally autonomous life in the above sense create an obligation that others should respect or assist an individual in attaining it? Going back to classical moral theory, for both Kant and Mill, the capacity for autonomous agency acquires normative significance because it is related to a moral endeavour which goes beyond the life of the individuals concerned. Particularly, in Kant, it is humanity in one's own person as well as in the person of everybody else that has moral worth (it is an "end-in-itself"<sup>22</sup>). Equally for Mill,<sup>23</sup> the normative ideal of human progress places individual self-development within the scheme of a higher social ideal, which is the welfare of the greatest number of agents possible. In classical moral theory, there is an individual-transcending dimension which sees morality as that which characterises our relations to others. This element is lacking in Dworkin's analysis, as indeed in most contemporary writings on autonomy in bioethics, despite their self-proclaimed allegiance to the above tradition.

### *c. Kantian autonomy*

Were we to offer a classification of the various conceptions of the notion of "patient autonomy", so vehemently invoked in contemporary bioethics, we



would detect at least three different senses according to which a person's autonomy can be said to be morally valuable: (i) as mere, sheer choice, that is, as some form of individual independence thought to be desirable; (ii) as a valued reflective capacity of the individual, a capacity to reflect on one's desires and adopt the rationally acquired ones; (iii) as Kantian or "rational autonomy", that is, as acting on principles valid from the standpoint of all.

The first conception is a minimalist one.<sup>24</sup> It counts all choosing unhindered by external sources as autonomous choosing. Respect for it is merely respect for patient choice (choices that may endanger life or health are not excluded). The only restriction that some of its advocates would accept is the prevention of harm or risk to others. The second conception views autonomy as choice that meets some additional standard, e.g., as a choice which is informed or reasoned as a second-ordered desire to have particular desires or preferences (cf. Harry G. Frankfurt<sup>25</sup>), or as reflective or authentic choice, coherent with other choices in one's life as a whole (cf. R. Dworkin), or as second-order valuing (cf. A. Jaworska). However, issues arise and controversies develop regarding what these standards should be. Debates are inconclusive. They only appear to be significant, if we start our moral reasoning by favouring some such characteristic of the individual agent, leaving otherwise open the question of why and how this characteristic should generate obligations for all others.

The third sense of autonomy mentioned above characterises principles rather than individuals. It is related to universal self-legislation and depicts that binding, individual-transcending aspect in our moral relations. It is firmly grounded in Kant's conception of morality, characterising principles as autonomous if and only if the reasons for acting on them could be adopted by all (rather than by reference to some particular authority of limited scope – which, for Kant, would be acting heteronomously). Kantian autonomy provides reasons why concern for others or dependence on others fits into the moral scheme of personal autonomy. It shows that the core of the capacity for autonomy consists in the ability to lay down the principles that will govern one's actions, which, *qua* moral, are capable of being shared by all. It is not merely independence of judgement or self-direction in one's life that constitutes moral autonomy. On the contrary, acknowledging the force of independently binding principles constitutes the core constituent of the personally autonomous agent. In Kant's words:

21  
Ibid, pp. 111, 134ff.

22  
Immanuel Kant, *Grundlegung zur Metaphysik der Sitten* [GMS], 1785, in: *Kant's Gesammelte Schriften*, Königlich-Preußische Akademie der Wissenschaften (later: Deutsche Akademie der Wissenschaften zu Berlin), Walter de Gruyter, Berlin 1903–, Vol. 4, pp. 428–429; English translation: Immanuel Kant, *Groundwork of the Metaphysics of Morals* [GMM], in: Mary Gregor, *Immanuel Kant. Practical Philosophy*, Cambridge University Press, Cambridge 1996, pp. 78–80.

23  
J. S. Mill, *On Liberty*.

24  
See, e.g., John Harris, "Rights and Reproductive Choice", in: John Harris, Soren Holm (eds.), *The Future of Human Reproduction: Ethics, Choice, and Regulation*, Clarendon Press, Oxford 1998; Julian Savulescu, "Autonomy, the Good Life and Controversial Choices", in: Rosamond Rhodes, Leslie P. Francis, Anita Silvers (eds.), *The Blackwell Guide to Medical Ethics*, Blackwell, Oxford 2007, pp. 17–37; John Robertson, *Children of Choice: Freedom and the New Reproductive Technologies*, Princeton University Press, Princeton 1994.

25  
Harry G. Frankfurt, "Freedom of the Will and the Concept of a Person", *The Journal of Philosophy*, Vol. 68 (1971), No. 1, pp. 5–20.

“Autonomy of the will is the property of the will by which it is a law to itself. [...] The principle of autonomy is, therefore: to choose only in such a way that the maxims of your choice are also included as universal law in the same volition.”<sup>26</sup>

We thus obey the categorical imperative (the moral law) by declining to adopt any maxim that could not rationally be endorsed by all others, including the party affected by our action – in the present case, by the patient. What is morally worthy is the individual’s capacity for self-legislation, and it is worthy in a particular way: we accord respect to it. We respect the patient’s autonomy by regarding her participation in our decision-making as a constraint on what decisions we permit ourselves to make.

Respecting persons’ autonomy in the Kantian sense is not just a matter of giving them effective options and assisting them in achieving some of them, but making decisions in which they (and all others) could rationally join in. The substantial principles of non-coercion and non-deception follow from the above notion. Kant’s conception of autonomy does not merely propound the value of a person’s end-setting capacity. It is “the will’s moral capacity to determine itself” independently of subjective drives and in accordance with principles which all others could adopt (principles of practical reason). It is our law-making ability which is called “autonomy”. It is “the ground of the dignity of human nature and every rational nature”.<sup>27</sup> This is what is captured by the Kantian notion of personality. It may sometimes require setting aside the individual projects that one may have for the sake of morality.

Integral to Kant’s notion of autonomy is the common human perspective. Morality as autonomy is about our relations with others (abiding by self-legislated principles valid from the standpoint of all). If this is so, if morality is interpersonal in the above sense, then to speak of the moral importance of personal autonomy implies that it has a role in interpersonal relations. The individual-transcending dimension of morality is captured by Kant as follows:

“... every person’s own will directed to himself, is restricted to the condition of agreement with the *autonomy* of the rational being, that is to say, such a being is not to be subjected to any purpose that is not possible in accordance with a law that could arise from the will of the affected subject himself; hence this subject is to be used never merely as a means but as at the same time an end.”<sup>28</sup>

Unlike its Dworkinian counterpart, Kantian autonomy is not a “value” or a “right”. Autonomy is necessarily attributed to the will of each and every moral agent, *qua* member of the moral community. It is not an empirical matter to be discerned, that is to say, it is not the case that some individuals have more or less than others, as indeed is the case with independence of judgement or mind and the ability to control events in one’s life as one chooses. It is a model of moral agency that applies to all human beings, *qua* moral agents. Autonomy thus constitutes the ground of the dignity of rational nature, and such autonomy is expressed primarily through the act of giving laws to oneself.

Put differently, autonomy means responding to moral reasons. It is not merely an individual right or a value, but a structural feature of moral agency and is, as such, presupposed by all rights and duties. Rights stem from (moral) autonomy. The latter offers a solid normative ground for claims of rights, giving them coherent content and justification.

Kant does not ground moral requirements on some prior value, or some valuable feature to be found in other human beings. Rather, he turns the relation the other way round: something has value because it is morally required. So, we owe respect to others not because they possess some kind of value, but

because it is a direct demand of our moral rationality. It is from this demand that all other values follow. This claim is encapsulated in the “formula of humanity”, which expresses the demand not to treat others merely as means, but always at the same time as ends in themselves. The demand is a categorical imperative, that is, a prescription law-like in form and universal in scope. These latter two characteristics cannot be inferred or acquired from experience, but are structural characteristics of an a priori conception expressing moral status:

“This principle of humanity, and in general of every rational nature, *as an end in itself* (which is the supreme limiting condition of the freedom of action of every human being) is not borrowed from experience; first because of its universality, since it applies to all rational beings as such and no experience is sufficient to determine anything about them; second because in it humanity is represented not as an end of human beings (subjectively), that is, not as an object that we of ourselves actually make our end, but as an objective end that, whatever ends we may have, ought as law to constitute the supreme limiting condition of all subjective ends, so that the principle must arise from pure reason.”<sup>29</sup>

Kant’s “Formula of Humanity” states that one ought to act in such a way that one treats humanity in one’s own person and in each and every other person as an end and never merely as a means. Human beings, *qua* rational agents, have dignity in virtue of their humanity (their rational moral agency). It is not the bare life of an individual human being that has dignity. Kant is not a vitalist. He does not defend life at all costs. He puts it this way: “morality and humanity insofar as it is capable of morality, is that which alone has dignity”.<sup>30</sup> We respect persons by respecting their humanity and thus their rational agency that constitutes humanity. It is this status that confers dignity upon them, unconditional worth beyond any price. Dignity cannot be traded off or measured against any other values.

All the main aspects of Kant’s formulations of the categorical imperative come together in the idea of a “kingdom of ends”, from which specific practical principles follow. A “kingdom of ends” is a systematic union of rational agents under common laws, being both legislators and subject to them. They “legislate” laws as rational beings with autonomy of the will and mutual recognition of their status as ends in themselves (persons).

Important consequences regarding the treatment of beings at the margins of agency follow from the above conception of moral autonomy. Kant’s conceptual argument is that autonomy is constitutive of moral agency, and this is independent of our beliefs about which human creatures are, empirically speaking, moral agents. Although we cannot consider severely brain-damaged or demented human beings as responsive to moral reasons, they are, nevertheless, due our respect because we relate to them as members of the

26

I. Kant, GMS, Ak. 4:440; I. Kant, GMM, p. 89. It is known as the “Formula of Humanity” of the Categorical Imperative.

27

I. Kant, GMS, Ak. 4:436; I. Kant, GMM, p. 85.

28

Immanuel Kant, *Kritik der praktischen Vernunft* [KpV], (1788), in: *Kant’s Gesammelte Schriften*, Königlich-Preußische Akademie

der Wissenschaften (later: Deutsche Akademie der Wissenschaften zu Berlin), Walter de Gruyter, Berlin 1903–, Vol. 5, p. 87; English translation: Immanuel Kant, *Critique of Practical Reason* [CPR], in: M. Gregor, *Immanuel Kant. Practical Philosophy*, p. 210.

29

I. Kant, GMS, Ak. 4:431; I. Kant, GMM, p. 80.

moral community in virtue of what they have been and in virtue of that which they exhibit (the “image” of a human face, metaphorically speaking).

One does not have to find out whether others possess this or that (valuable) feature as a matter of fact in order to be bound by respect to them. When Kant invokes the notion of dignity and attributes it to persons by virtue of their rational (and to that effect, moral) nature, thus placing them beyond any price, he does not invoke anything that requires the ability to walk unassisted, or to control one’s movements, or to eat without help, or to be self-reliant. He does not mean physical independence or strength. Moral autonomy does not involve empirical abilities to function independently, or lack of dependence on continuous medical intervention, or freedom from physical or cognitive deterioration. The conception of dignity invoked in bioethical discussions often seems to refer to a kind of condition threatened by physical deterioration or dependency. But this is not the status that can ground the demands on our moral actions which are so fundamental in Kant’s ethics.

### **The unity of personhood and the authority of advance directives**

On Kantian grounds, mentally disabled or incapacitated persons are to be treated as full members of humanity, as subjects of rights and interests. The fact that they are unable to tell us what they want does not make them strangers to our moral community. Even regarding mature rational agents, the latter may express their rational autonomy in action in various ways. What is more, Kant repeatedly stresses that we cannot judge their inner motive with assurance. This implies that we too remain uncertain about the degree to which others, mature adult individuals, express their personal autonomy. We need the power of imagination in order to understand others in their individual separateness, or strangeness or even mental impairment, and ought to respect them even if we do not fully understand them.

If persons are human beings with moral status, then the loss of the ability to display certain empirical properties does not make them non-persons. It does not deprive them of being members of humanity, the community of moral agents. Even when someone is in a persistent vegetative state, there are still certain interpersonal moral attitudes that we extend to them: they are due our respect of the sort that forbids us to maltreat, harm, humiliate or degrade them.

The reluctance we feel to degrade the comatose demonstrates that their body has to be treated with respect as an integral part of their personality, because of the “human form” it manifests. Each of us is an “image” or a manifestation of humanity. To the extent that the “kingdom of ends” evokes respect, each of its instances also evokes respect.

There is reason to respect advance directives furnishing instructions on how one should be treated in case of permanent loss of consciousness or cognitive capacity, just as there is reason to respect the directives of the deceased furnishing instructions on what is to happen after their death concerning their property and its distribution to their loved ones. Respect for rational agency requires both sustaining conditions of rational agency and respecting decisions made by rational agents. Advance directives have moral force because a person’s determination of the will and moral realisation include a claim in shaping her future image and recollections.

Despite their incapacity to sense, to exhibit cognitive or linguistic competence, or even appreciate violations of interests, or have experiences, etc., incompetent individuals possess moral status and have to be treated with respect. It is important given that people seek to shape their image and the way they will be remembered, something which will survive their competent stage of life and their life itself. Respect for a person morally justifies respect for the person's decisions surrounding a post-competence dying process.

So long as the directive reflects a considered and unambiguous awareness of its content and effect, and so long as there is no indication that the person's choices changed while she was competent, it ought to be honoured. A moral community ("kingdom of ends") in which rational decisions were not respected after the loss of decisional competence or the death of those issuing them would be less respectful of humanity and hence not a kingdom of ends, morally speaking. A world in which rational decisions were not respected after the loss of (empirical) competence or even the death of the person in question would show less than respect for humanity and would be less than a "realm of ends".

At the same time, however, a note of caution should be sounded. It is imperative that advance directives be carefully considered in their multifarious forms. Above all, the reasons upon which they are based have to be carefully scrutinised and assessed. (For instance, advance directives cannot be morally accepted if they are the outcome of prejudice, ignorance or an unreflective assumption, or any kind of irrational preference.) As has above been mentioned, to those who subscribe to a minimalist conception of individual autonomy – including, most paradigmatically, Ludwig Minelli, the founder of the Swiss organisation *Dignitas*, which caters to foreigners travelling to Switzerland for assistance in dying – it does not matter why the person wants to die. The only value which is taken to be morally relevant is the individual's self-determination. But, on our account of moral autonomy, the reasons which motivate authors to issue advance directives do make a difference and are relevant. As Paul Menzel has put it, quoting Nancy Rhoden,

"When they start saying, 'If I can't do higher mathematics, kill me', we will have to worry in earnest about the limits of precedent autonomy."<sup>31</sup>

Just as, morally speaking, there is a limit to what a competent person can do to her competent self, so there is a limit to what a person can dictate for her future incompetent stage of life. This limit is not defined, nor is it constituted by a person's mere free individual choice. Not everything can count as a morally acceptable course of action simply because it is freely chosen. Degrading conduct towards a mentally deceased patient is not an ethical option, regardless of any consent that may have been given. There are moral limits to what can be tolerated in the name of individual autonomy, regarding both contemporaneous and advance choices. Just as the right to self-determination cannot morally include consent to slavery or consensual cannibalism, so the prospective imposition of an utterly degrading or dehumanising status on an incompetent patient is beyond all moral bounds. Similarly, the imposition of

30

I. Kant, GMS, Ak. 4:435; I. Kant, GMM, p. 84.

31

Paul T. Menzel, "Advance Directives, Dementia, and Eligibility for Physician-Assisted

Death", 58 *N.Y.L. Sch. L. Rev.* 321 (2013–2014), p. 339; see also Nancy K. Rhoden, "The Limits of Legal Objectivity", *North Carolina Law Review*, Vol. 68 (1990), No. 5, p. 860.

significant and gratuitous suffering or deliberate degradation on a mentally impaired patient, despite the patient's ignorance or inability to understand reasons or her inability to deny consent, is absolutely morally unacceptable.

In addition, it is absolutely essential that no evidence exists that the patient has changed her mind or has hesitated after authoring the directive and before losing competence. Moreover, when obvious harm to the contemporary patient's well-being will be a consequence of the application of the advance directive, those making the decision ought to scrutinise carefully whether and to what extent the author of the directive anticipated and considered the effects now occurring. As Onora O'Neill puts it, quoting Bernard Williams, "we should not put too much weight on the fragile structure of the voluntary".<sup>32</sup>

When obvious harm to the now incompetent patient is anticipated by the enforcement of the directive, the decision makers have to make sure that the previously competent patient did envision adequately the situation that has now occurred. In such cases, the advance directive has to be explicit about such developments. Any sign of a will to live on the part of a person who is no longer competent relativises the authority of an advance directive in which treatment is refused, unless the exact medical situation is described in the directive in clear and specific terms, explicitly mentioning that any sign of a will to live is to be treated as insignificant and is not to be taken into consideration in the decision-making process.

We should also be sensitive to complex practical issues relating to questions of predicting future facts, estimating future appreciation of such facts, as well as unforeseeable future advances in medical technologies and therapeutic methods. Practical decisions have to be taken with extreme care and caution, by paying careful attention to the complexities of willing something in the future, as well as the conditions under which the decision has been made. Amongst these crucial parameters are freedom from coercion and heteronomous, external determination, freedom from internal coercive powers, such as depression, impaired cognitive powers, stress and fear that one may be a burden to one's family, and so on.

To recapitulate, we need to avoid the temptation to invoke, on the one hand, sceptical empiricist conceptions of personal identity extending into the future (*à la* Dresser) and, on the other, misleading, hyper-idealised conceptions of individual autonomy. The morally relevant notion of autonomy is making a choice based on principles that one judges that everybody could choose to act on. The trait of the personally autonomous agent is that of acknowledging the constraints of objective, sharable principles.

Lastly, discussions about the morality of end-of-life choices should not be reduced merely to claims about advance directives and "patient autonomy". Advance directives have moral force in their own right. But their moral authority cannot compromise or minimise the doctors' and health providers' responsibility to protect the "humanity", basic rights and interests of the terminally ill, incompetent or unconscious patients. It cannot compromise the obligations of the state, either, to establish institutions which comfort the life of those exiting it, including palliative and hospice care. These are but some of the vital moral and political obligations of a democratic state.

**Stavroula Tsinorema**

**Princip autonomije i  
etika smjernica za postupanje na kraju života**

**Sažetak**

*Smjernice za postupanje na kraju života (advance directives) zamišljene su kao sredstvo povećavanja »autonomije pacijenta« jer one osnažuju moć izbora pojedinca u post-kompetentnom procesu umiranja. No njihova je moralna snaga sporna. Rebecca Dresser i John Robertson nude konceptualni argument utemeljen u epistemološkim razmatranjima osobnosti koja osporavaju autoritet smjernica. Ronald Dworkin snažno brani »prethodnu autonomiju« u planiranju post-kompetentne medicinske skrbi. Ovaj rad istražuje gore navedene suprotstavljene teze i ocjenjuje njihove glavne argumente. U oba su slučaja ustanovljena određena ograničenja. S obzirom na prvu tezu, problematičnom se smatra njezina konceptualizacija pojma osobnosti, a što se tiče druge, smatra se da je njezin koncept individualne autonomije suviše uzak. Alternativni se put traži kroz rekonstruiranje Kantova shvaćanja moralne autonomije. Time se nudi okvir za moralno rasuđivanje iz kojega se može primjereno izvesti i opravdati određeno suvremeno razumijevanje autonomije kao prava, kao refleksivne sposobnosti pojedinca, kao odgovornosti i integriteta. Normativni zaključci proizlaze iz proširenja osobne autonomije u napredni medicinski izbor.*

**Ključne riječi**

smjernice za postupanje na kraju života (*advance directives*), djelovanje, autonomija, Rebecca Dresser, Ronald Dworkin, Immanuel Kant, osobnost, poštovanje

**Stavroula Tsinorema**

**Prinzip der Autonomie und Ethik der Patientenverfügung**

**Zusammenfassung**

*Die Patientenverfügung (advance directives) ist als Mittel zur Vergrößerung der „Patientenautonomie“ gedacht, weil sie die Macht der individuellen Wahl innerhalb des post-kompetenten Sterbeprozesses kräftigt. Jedoch ist deren moralische Kraft umstritten. Rebecca Dresser und John Robertson unterbreiten ein konzeptuelles Argument, das sich auf epistemologischen Betrachtungen der Personalität gründet, die die Autorität der Patientenverfügung anfechten. Ronald Dworkin verteidigt mit aller Kraft die „vorherige Autonomie“ in der Planung der post-kompetenten medizinischen Betreuung. Diese Arbeit untersucht die oben angeführten entgegengesetzten Thesen und bewertet ihre Hauptargumente. In beiden Fällen wurden gewisse Beschränkungen festgestellt. In Anbetracht der ersten These wird ihre Konzeptualisierung des Personalitätsbegriffs für problematisch befunden, und was die andere angeht, wird ihr Konzept der individuellen Autonomie als übertrieben schmal beurteilt. Ein alternativer Weg wird durch die Rekonstruktion von Kants Auffassung der moralischen Autonomie gesucht. Dadurch wird ein Rahmen zum moralischen Ermessen angeboten, aus dem sich ein bestimmtes zeitgenössisches Verständnis der Autonomie als Recht passend ableiten und rechtfertigen lässt, ein Verständnis der Autonomie als reflexive Fähigkeit des Individuums, als Verantwortung und Integrität. Normative Schlussfolgerungen gehen aus der Ausdehnung der persönlichen Autonomie auf eine fortgeschrittene medizinische Wahl hervor.*

**Schlüsselwörter**

Patientenverfügung (*advance directives*), Tätigkeit, Autonomie, Rebecca Dresser, Ronald Dworkin, Immanuel Kant, Personalität, Achtung

**Stavroula Tsinorema**

**Le principe d'autonomie et  
une éthique des directives anticipées**

**Résumé**

*Les directives anticipées (advance directives) ont été conçues comme un moyen pour augmenter « l'autonomie du patient » puisqu'elles renforcent son pouvoir décisionnel une fois le processus de fin de vie entamé, c'est-à-dire une fois les compétences diminuées. Toutefois, leur force morale est discutable. Rebecca Dresser et John Robertson proposent des arguments tirés de considérations épistémologiques de la personnalité qui contestent l'autorité des directives. Ronald Dworkin défend vigoureusement « l'autonomie antérieure » dans la planification des soins médicaux « post-compétents ». Ce travail examine les thèses adverses mentionnées ci-dessus et évalue leurs arguments principaux. Dans les deux cas, des limitations déterminées sont établies. Concernant la première thèse, la conceptualisation de la notion de personnalité est jugée problématique, et quant à la seconde thèse, son concept d'autonomie individuelle est estimé bien trop étroit. Une voie alternative est recherchée à travers la reconstruction de la compréhension kantienne de l'autonomie morale. De cette manière, un cadre est proposé pour un jugement moral à partir duquel il est possible de déduire et de justifier de manière adéquate la compréhension contemporaine d'autonomie en tant que droit, en tant que faculté réflexive de l'individu, en tant que responsabilité et intégrité. Les conclusions normatives dérivent de l'élargissement de l'autonomie personnelle relative au choix médical avancé.*

**Mots-clés**

directives anticipées (*advance directives*), activité, autonomie, Rebecca Dresser, Ronald Dworkin, Emmanuel Kant, personnalité, respect