

ethics in progress

Ethics in Progress

Vol 9, No 1. 2018.

Peer-reviewed online journal.

**A multidisciplinary forum for research bridging
between academic ethics and social practice.**

EDITORIAL TEAM

Editor-in-Chief

Prof. Dr. Ewa Nowak, Adam Mickiewicz University, Dept. of Social Sciences,
Poland

Executive Editor

Tomasz Raburski, Adam Mickiewicz University, Dept. of Social Sciences,
Poland

Academic Editor

Prof. Dr. Georg Lind, University of Konstanz, Dept. of Psychology,
Germany

Prof. Dr. Jason Matzke, Univ. of Mary Wahsington, Dept. of Classics, Philosophy, and Religion,
USA

Prof. Dr. Roma Kriauciuniene, Vilnius University, Dept. of English for Physical and Biomedical Sciences,
Lithuania

Dr. Roberto Franzini Tibaldeo, Université Catholique de Louvain-la-Neuve, Dept. of Philosophy,
Belgium

Technical Editor I

Mag. Joanna Dutka, Adam Mickiewicz University, Dept. of Social Sciences,
Poland

Design Editor

Mag. Alicja Skrzypczak, Adam Mickiewicz University, Dept. of Social Sciences,
Poland

Internet Resources Editor

Dr. Marcin Jan Byczyński, Adam Mickiewicz University, Dept. of Social Sciences,
Poland

Table of contents:

David Jaitner (Köln); <i>Reconstructive Habits: John Dewey on Human Functioning</i>	pp. 4 - 24
Christos Douskos (Athens); <i>Deliberation and Automaticity in Habitual Acts</i>	pp. 25 - 43
André Schmiljun (Berlin); <i>Why Can't We Regard Robots as People?</i>	pp. 44 - 61
Karolina Kowalewska (Poznań); <i>Underclass, Culture of Poverty and the Research in Sociomoral Competence among the Poor</i>	pp. 62 - 98
Jianjin Liu (Guangzhou); <i>Chinese Adolescents' Conceptions of Teacher's Authority and Their Relations to Rule Violations in School</i>	pp. 99 - 117
Zhu Hailong (Guangdong); <i>An Empirical Study on Chinese and American College Students' Moral Value Recognition</i>	pp. 118 - 127
Paweł Mazur (Poznań); <i>Remarks About Legal Situation of Animal Injured By Prohibited Act in Polish Criminal Process</i>	pp. 128 - 139
Małgorzata Adamska (Poznań); <i>The Categorical Imperative in Defence of Animal Rights. Christine Korsgaard's Ethical Extensionism</i>	pp. 140 - 173
Lan Wen, Peijue Zheng, Shaogang Yang (Guangdong); <i>Changes on the Qualities of University Counsellors in China</i>	pp. 174 - 182
Małgorzata Bogaczyk (Poznań); <i>On Philosophy of Barbara Skarga: An Approach to Get to Know Her Life and Work</i>	pp. 183 - 191

Reconstructive Habits: Dewey on Human Functioning

David Jaitner

ORCID: <https://orcid.org/0000-0002-7839-9643>
(German Sport University Köln, d.jaitner@dshs-koeln.de)

Introduction

The academic strife to parse, investigate and adjust human functioning¹ establishes at least three key concepts: behavior, action, activity. Considering human functioning as *behavior* it usually is any visible stimulus response of a human organism. Here, stimulus means any external influence that causally activates – immediately or cognitively mediated – specific modes of functioning. Reactions are all those modes of functioning activated due to specific external influences, no matter whether it is about doing or not doing something. Human functioning that is made intentionally, motive-based, and with a specific target is ideal-typical *action*. Here, the focus is on the realization of possibilities that are principally controlled by the human being. The human being is agens, with intentional impact on their environment to shape the future. Human functioning is not determined; it is a doing that faces environmental demands with a basically unlimited autonomy to act or refrain from a certain action, even if this factually should be limited by prevailing conditions. As *activity* human functioning is both active and passive. Activity-theoretical accesses retain the human possibility to shape life, but switch over from arbitrary to intrinsically determined choices. Humans are active when they no longer are universally self-determined, but still capable of shaping situations. Thus, the human being remains subject of their process of existence, however, within the scope of intrinsically and not merely secular limited degrees of freedom. Humans control their activity even in moments when something happens to them as well and when other than their mentally will is responsible.

¹ Because behavior, action, activity, doing, or conduct are terminologically designated in the essay, human functioning is a neutral placeholder to designate “menschliches Tun” in a not yet terminological way.

In the context of human functioning the chosen approach that one takes into account is therefore essential: Depending on the approach chosen, human functioning is defined in a certain way and understanding of freedom. Here, the pragmatic considerations of John Dewey (1859-1952), an American philosopher, psychologist and educational theorist, offer a sophisticatedly formulated theory of human functioning that, undoubtedly, takes action-theoretical paths but formulates underlying assumptions in a significantly unusual way. This is where the essay really begins. The aim is to outline the theory in such a way that clearly shows the unusual as part of the usual and the usual as part of the unusual. Thus, in a first section action is defined as the basic category of human functioning where sensory stimuli, registering elements, and motor responses play a leading role, but according to Dewey questions the today still popular model of behaviorist psychology, that positions isolated and a-cultural stimulus-response-procedures in the human organism. A second section affirms the theoretical inclusion of deliberative elements that constitute human action, but according to Dewey witnesses their substantial and rather sporadic significance in a predominantly habitual human functioning. A conclusive section outlines the possibilities and limits of transforming habitually inured patterns of human conduct by means of reconstructive habits.

Coordinated Actions As Basic Units of Experience

To understand or unfold philosophically based considerations, it is usually worth asking for the philosopher's fundamental project. To this, the life-long project of Dewey is a practically relevant *philosophy of experience* that aims to transcend all that traditional epistemological and consciousness-theoretical approaches placing humans in the world and for themselves. Dewey uses experience to denote all kinds of human to the world and reality. Man is uncircumventably in contact with the world. They factually have approach to a world that they have always been enmeshed in. This inevitably mediated relationship to the world is initially and mostly immediate (cf. primary experience). People have their world, are on the move in a universe of having, "[t]hings are objects to be treated, used, acted upon and with, enjoyed and endured" (LW 1, 28).² Certainly, this also means that in problematic situations human beings can transition into a self-referential form of interaction. Looking at this second level, Dewey speaks of reflexive experience, the humanly possible level of knowledge acquisition and theory formation: universally applicable from the resolution of everyday conflict situations through to academic research or philosophical reflection, and ideally carried out as inquiry, a gradual method of obtaining knowledge and thinking based on the experimental process

² Standard references to John Dewey's work are from the critical (print) edition, *The Collected Works of John Dewey, 1882-1953*, edited by Jo Ann Boydston (Carbondale: Southern Illinois University Press, 1969-1991), published in three series as *The Early Works* (EW; 1882-1898), *The Middle Works* (MW; 1899-1924) and *The Later Works* (LW; 1925-1953). „LW 1, 28“, for example, refers to *The Later Works*, volume 1, page 28.

of empirical sciences in order, in the words of Dewey's *Logic. The Theory of Inquiry* (1938), to transform an indeterminate situation into one "that is so determinate in its constituent distinctions and relations as to convert the elements of the original situation into a unified whole" (LW 12, 108).

In *Experience and Nature* (1925), *Half-Hearted Naturalism* (1927), or *Nature in Experience* (1940), Dewey calls the associated self-labeling "empirical naturalism" or "naturalistic empiricism" (LW 1, 10). Both terms are indeed laid out *sui generis*. In Dewey's reconstruction, naturalism relates, esp. in connection with Charles Darwin (1809-1882; e.g. MW 4, 3-14; MW 12, 114; LW 12, 97), not to an understanding of nature as "something essentially given", but "rather to an open, dynamic and contingent process in which identities and relationships emerge as the actualization of natural potentialities in the context of evolutionary interactions" (Neubert 2009, 20). Accordingly, human experience is a natural affair and is emerging from natural interactions (LW 1, 5). Empiricism, too, does not classically mean a merely subjective, passive, sensually absorbed integration of an objective world (e.g. MW 9, 276-285; MW 10, 6), but rather designates experience through the criteria of 'continuity' and 'interaction'. Both criteria are inseparable from each other, they intercept, unite and "are, so to speak, the longitudinal and lateral aspects of experience" (LW 13, 25): Experiences invariably refer to experiences, "every experience both takes up something from those which have gone before and modifies some way the quality of those which come after" (*ibid.*, 19). Experiences are always interactions of natural events that affect human organisms, "certain modes of interaction, of correlation, of natural objects among which the organism happens, so to say, to be one" (MW 10, 26).

The elementary unit of experience is *action*, designed as a link between doing and undergoing wherein instrumental meanings are distilled out of existential events (e.g., EW 3, 212; MW 9, 147; MW 10, 9; LW 10, 50; LW 11, 214; LW 12, 144). Here, a theoretically consequent *idea of circularity* primarily manifested in an idea of a feedback of context and effect is leading the way. Action, as part of experience in nature, is always interaction between an active organism and the environment that continuously opposes the organism. In a changing and adaptable world these interactions are by no means attributable to one-dimensional effective directions and simple causal attributions, but are characterized by mutual change and adaptation. An individual person takes up selected environmental stimuli, and responds to both his environment and himself by his action addressed to the environment, thus changing the starting conditions for any further action. One cannot go twice in the same river. As part of experience in nature, action is simultaneously characterized by continuity. Interactions are far from being isolated events, but always parts of a wide stream of experiences. Interactions between humanity and environment are inevitably embedded into a context of various adaptation and change processes. There are no definite beginnings and origins. Every effect is, at the same time, cause – and vice versa.

From the very beginning Dewey's work has been guided by this circular train of thought, and continuously differentiates itself concurrently to his theory of experience. Man is the "live creature" (LW 10, 9) that, on the one hand, is like any other living organism in nature and on the other, is the most complex organism, since humans are the self-conscious and socially living organisms in nature (cf. MW 14, 200; LW 1, 194-209; LW 3, 41-54; LW 12, 48-65).

Biological Interactions of Organisms and Environments

First, biological interactions of organisms and environments expressively dominate in Dewey's writings. The basis of that approach can already be found in *The Theory of Emotions* (1894), formative and highly respected to the present is the argumentation in *The Reflex Arc Concept in Psychology* (1896). The essay, first published in *Psychological Review*, identifies the criticism of a specific stimulus-response model, but does not criticize the idea of stimulus and response as such. Thus, Dewey agrees and disagrees with the empirically oriented physiological psychology developing at that time: To distance from the established introspective focus of psychology, Dewey starts with behaviorist approaches with an organic construction of human functioning, but dissociates himself clearly from its far too reductionist designing. Because of the fact that, according to Dewey, behaviorist approaches estimate the reflex arc as basic unit of psychological analysis³ and reduce human functioning to the objective observation of peripheral processes, humans are passive recipients of external impulses, and there is no inner relationship between the sensory and motor unit. The behaviorist contribution to the search for a standardized psychological research variable simply helps neither overcome the traditional separation of sensual perception, thinking and action, nor does justice to the reality of human nature: As a passive agent, humans are arbitrarily stimulated externally to initiate response. Sensory stimulus, central processing and motor responses continue to be isolated and incoherent existences, whose coordination further on can best be explained with transcendental mediating authorities or external pressures from the environment. Thus, the old dualism of body and mind is simply dressed in new clothes (EW 5, 96-99; cf. EW 2, 11-15; MW 9, 150)⁴.

³ The reflex arc denotes a physiological control loop model to explain the path followed by nerve impulses to produce reflex actions. The model is based on a mechanical input-output scheme: A stimulus impulse originates in a peripheral receptor, passes through an afferent nerve to the reflex center in the brain or spinal cord, and then out through an efferent nerve to an effector organ, where the appropriate motor response occurs. Around 1900, the model was willingly adapted by the emerging experimental psychology to establish a scientifically oriented 'exact' academic psychology (cf. Watson 1913).

⁴ This criticism makes Dewey appositely sensing the consequences and pitfalls of the behaviorist stimulus-response model at a very early stage of development (cf. Rorty, 1979, 228; for further critical reflections see MW 10, 53-63; MW 11, 10-17; LW 3, 45; LW 6, 27; LW 14, 185), but does not at all remain a lone voice in the wilderness. Further critical reflections to the reflex arc model are particularly found in *Kurt Goldstein* whose work on brain-injured First World War soldiers

To cope with these dilemmata, Dewey psychologically takes the physiological unity of coordinated action seriously that takes the neural circuitry of the reflex arc as a basis, and integrates the reflex arc into an organic circle of functional coordination of interactions. For Dewey, life is homeostasis, an active effort to preserve or recover the equilibrium that is indispensable for the maintenance and function of an organism or organ. Thus, right on the basis of being alive, the human organism is out to shape its life at least life-sustaining (ibid., 104; cf. MW 2, 290; MW 6, 437-440; MW 10, 7ff.; LW 7, 289). Hence, the unifying principle in psychology is to be coordinated action and the basic units of human functioning are to be sensory-motor coordinations. Stimulus, cerebral activity and response are collaboratively integrated components of a single, functional unity of coordinated human action. This has two important consequences. First, context matters: The human organism has always been in contexts where it perceives a few of a multitude of sensations, then selects and interprets some of them, responding actively. A sensory stimulus not and never comes into contact with an organism unconditionally. A motor response to a stimulus is not and never universally programmed. Stimulus and response are always a result that has been and has to be constructed of all those interactions the human organism had been involved in to date. Second, stimulus and response are phases of a holistic entity: Thus, sensations and movements relate to each other in integrated action patterns, where they receive their functional form and importance, and can exclusively be explained meaningfully. Whether one, e.g., answers a loud bang with a weary smile, prying anticipation, paralyzing fear, ordered retreat, or panic-stricken flight depends, on the one hand, inevitably on the respective context of acting where the acoustic event takes place, and on the other, the respective acquired and currently available coordination patterns. The motion in action always constitutes the sensation, and the sensation in action defines the motion. Only through flight does an acoustic stimulus change into a threatening one. Only the act of hearing makes a gunshot on a dark night perceptible and motor-relevant.

concludes, inter alia, that there are no isolated stimulus-response processes in an organism, but that always the whole organism reacts (cf. 1934, 45ff.). *Jean Piaget*, too, vehemently underlines the unity of movement and perception, but like Goldstein does not attribute it to Dewey, but to Viktor von Weizsäcker's „Gestaltkreis" (cf. 1947, 99; 1970, 86). *Karl Lashley* notably questions the application of the reflex arc to complex human behaviors from a neuropsychological point of view (1951; cf. Rosenbaum, Cohen, Jax, Weiss, & van der Wel 2007). A rather popular philosophic contradiction to the underlying relationship of cause and effect is to be found in *Alan Watts* example of an observer watching a line of passing cats through a narrow slit in a fence: Because the observer always sees the heads first and the tails last, and there are never tails without heads, he reasons that the event ‚head' is the invariable and necessary cause of the event ‚tail', which is the head's effect. From such a perspective, it is simply impossible to see that head and tail go together, are parts of the event ‚cat' (cf. 1966, 30f.). Current efforts, references, and resumptions to Dewey's integration of perception and action in a unifying framework can be found e.g. in James Gibson's ecological approach to visual perception (1979; cf. Burke 1994, 83-96), Bernhard Hommel's „Theory of Event Coding" (cf. Hommel, Müsseler, Aschersleben, & Prinz 2001) or in functional-dynamic alternatives in closed loop neurophysiology (cf. Wallach, Marom, & Ahissar 2016).

The fact that stimulus and responses are merely set as “distinctions of function, or part played, with reference to reaching or maintaining an end” (EW 5, 104; cf. MW 7, 346), i.e., as teleological distinctions without independent existential nature, questions whether these can be distinguished at all in the complex tangle of integrated action. For Dewey, the answer is clear. The distinction of stimulus as stimulus and response as response remains recognizable, but is connected to particular conditions. As long as coordination of actions is fully developed and not affected by any external disturbance, the distinction is impossible and pointless. However, if existing coordination of actions are temporarily disturbed, their flow inhibited and procedures confused, the distinction becomes psychologically effective and functional. Whenever this is the case, the outer scenery as well as the modalities of one’s own action are uncertain and the restitution of coordination of action is required. Then, the broken organic circle needs to be closed in such a way that the relation of the uncertain outside and inside world can be examined and modified into a preferable, vitally suitable sensorimotor coordination (EW 5, 104-109; cf. Mead 1903).

Cultural Interactions of Social Individuals and Habitats

Subsequently, Dewey increasingly focuses on *cultural interactions of social individuals and habitats* without, however, denying or depreciating the biological ideas of human functioning in any matter. Along with a continually matured theory of experience that, piece by piece, is fundamentally extended by social and cultural levels of meaning, an extended framework develops towards designing and measuring human functionality. The early conception of action as an organic circuit is supplemented by social and cultural contexts. In summary, this extension becomes visible in *Conduct and Experience* (1930). The essay, first published in the anthology *Psychologies of 1930*, in a first part very clearly iterates and reinforces the reconstruction of the behavioristic interpretation of the reflex arc and provides, in contrast, an integrated and circular model of human functioning. Thus, it remains clear that human functioning, or conduct – Dewey’s proposed conceptual alternative for behavior – is never a mere sequence of independent and universal segments. Furthermore, it is a fatal self-misunderstanding of behaviorist psychology to believe that they move over neutral ground with their experimental-statistical methods and conduct research impartially. Indeed, the opposite is true. According to Dewey, all parts of action are functions in the context of extensive sequences of action. Human functioning develops in a circular way. Observations without observers, theoretical presuppositions and empirical variables are impossible. Regarding adequate experimental designs, all that inevitably belongs to human functioning has to be captured: the serial continuity of actions within comprehensive contexts of conduct. No act can be understood from the series to which it belongs. Thus, the motifs, intentions and situations of the examined subjects as well as the activating problems, arrangements

of apparatus and selections of conditions set by the experimenters are scientifically relevant and need to be taken into account (LW 5, 218-224; cf. MW 11, 16f.; LW 1, 34f.; LW 6, 8).

In a second part, Dewey defers the primacy of conduct and connects the previous considerations with the idea of experience, or culture in continuity with nature (LW 1, 361; cf. Rölli 2009). In other words, the serial nature of psychological occurrence and life-historical interdependence of concrete acts of conduct are to be theoretically and empirically recognized. Hence, psychology is to be the scientific exploration and account of experience. The obvious subject matter of psychology is to be “the life-career of individualized activities” (LW 5, 224; cf. Hughes 1928). This makes the integration of mental issues mandatory. Given that sensation, perception, consciousness, reflection, or emotion, etc. are relevant in human conduct, a non-reductive empirical psychology has to include them. This means, the categories of the formerly rejected introspective foil have to be empirically revived, and at the same time the pitfalls of the declined empirical behaviorist approach are to be overcome. The given answer to the question of an appropriate relation between conduct and experience is – regarding the underlying interpretation of experience as double-barreled word (cf. LW 1, 17f.) – hardly surprising:

The envisioned psychology receives its feedstock, mission and avenue in commonly essential distinctions of a plethora of qualities of everyday experience. These are by no means academic inventions, but emerged in real-life processes of social interactions. Sensing, perceiving, imagining, thinking, judging, reasoning, loving, or admiring, etc. are serving functions of human life that are evolutionarily established decision support in day to day pressure of constantly having to deal with themselves and the world. By living their life, humans acquire culture-specifically pre-structured modes of experiencing and learn how to distinguish and behaviorally use them. These qualities of experience are not already psychology but – like, e.g., the given differences between oil and water, or iron and tin in physics and chemistry – given categories of common sense that have factually to be recognized. With these the psychological science, according to Dewey’s understanding of inquiry, has to start and to end as well. Hence, the science of *experience psychology* is located in the space between practical entrance and exit. In that context, it is psychologically quite clear and thus less interesting what acts of conduct human beings experience, and at a fundamental disposal, *how* human beings experience these acts of conduct. The aim of psychological studies is to provide access to all those attitudes, preferences, sensitivities, habits and techniques that can be observed when dealing with selected experiences.

In the melting pot of experiences conscious and mental forms of acting play a significant and in a certain way momentous role for human life and daily routine. It has an essentially different effect if human functioning takes place physical, physiological,

or conscious and mental resp. what kind and level of mental activity are involved in its emergence. In acts of consciousness and mind logic, rationality, meaning, symbolic interaction, and communal participation become possible, existences and essences can be bridged, and a reciprocal nexus of means and ends to attain consequences that solve accrued problems can be achieved (e.g., MW 8, 37; LW 1, 133; LW 8, 233; LW 12, 17). With regard to research methodology, that is also slightly surprising when looking at the collective works, Dewey clearly subordinates psychological examination of this inner surface of human conduct to the scientific pattern of inquiry. Even though modes of experiencing in their immediate quality are, to a certain extent, phenomenologically explorable, for Dewey it is the direct observation of the outer surface of individual modes of experiencing which makes psychological research more objectifiable and qualifies it as genuine science (LW 5, 224-235).

Positioning of the Deliberative Moment

Even though in the previous considerations aspects of willing, thinking, reflecting, reasoning, judging, or deliberating repeatedly appear, their concrete position in Dewey's theory of human functioning as action is still not yet systematically defined. Dewey comprehensively discusses this issue by giving a systematic presentation in *Human Nature and Conduct* (1922). The focus of the three lectures, delivered at Leland Stanford Jr. University in May of 1918, is the search for the real foundation of human physiology and psychology (MW 14, 6; cf. LW 6, 29-39). Here, an action-theoretical argumentation is the leading concept that converts a consciousness-philosophical approach into a theory of reconstructive habits and develops an understanding of action that firmly distances itself from any utilitarian design. In doing so the human being remains a subject without presupposing a primarily rationally acting essence. Acting is reasoned response, and thus does not arbitrarily react to activating surrounding aspects, but is not significantly conscious-driven either. Man is an interactive being that, for the most part, actively and unconsciously occupies the surrounding world. Cooperation of nature, human being and society takes place, however, the consciousness that keeps in mind that man is about to engage in something with certain consequences usually lacks. Man does not act without reasons, but generally the respective result is not one that had been the explicitly cited reason but is based on pre-reflexive practical dexterity. Thus, Dewey widely limits the classical action-theoretical access but basically maintains the paradigm. The cornerstones of his argumentation are three key concepts and their interrelationships: *habit, impulse, intelligence*.

Habit

According to Dewey, the major determination of human nature is located in *habits*: Man is „a creature of habit, not of reason nor yet of instinct” (MW 14, 88). Hence, the nature of habits is the very nature of man. The mode of habit forming is the mode of

forming humans. Human existence and development become understandable by decoding habits. Dewey introduces the nature of habit by comparing it with different physiological functions (e.g. breathing, digesting, seeing), but also restrictively points out that the latter are involuntary, and habits necessarily acquired. Common, and the analogy legitimizing feature remains the function in coordinated action. Habits are functional for the access to the world and the interaction of an individual person and the surrounding world. Individual designates the internal aspects of interaction, the individual prerequisites and subjective possibilities. The surrounding world is all that opposes, the objective conditions of the setting, the entire non-self a human being contacts and acquires habits with, embedded in habits of habit forming (ibid., 133; LW 13, 25): „Through habits formed in intercourse with the world, we (...) in-habit the world. It becomes a home and the home is part of our every experience” (LW 10, 109).

Contrary to physical or physiological functions, however, Dewey formats habits not behavior-theoretically but culture-scientifically. Habits are meaningfully acquired in individual interrelationships of elements of human nature as well as natural and social surroundings. They are bodily rooted, culturally shaped and socially multiplied unique features of human functioning. By their habits the human species create themselves a functional long-term memory against a lifetime-related transience of cultural achievements, a form of embodied histories to continually keep up with their own present and future. Habits are the reason that with every new generation the conduct of human life does not start from scratch regarding cultural development. Habits are intergenerative, socially-transferring and developing inescapable mediators of a social group to utilize and incorporate an objectively placed setting. The mode of transfer is part of the social nature of the human being and inevitable: „It is not an ethical ‚ought’ that conduct *should* be social. It *is* social, whether bad or good” (MW 14, 17; cf. EW 1, 231; EW 5, 86; LW 7, 253). As a condition of participation, newborns right from the moment of their birth continually and inevitably assume the objective habits, constitutive patterns of functioning and moral orientations of their social setting that in the respective group is expressed by and manifested in customs, arrangements, morals, standards, values. The remaining result is a connection between individual and objective habits, a similar group thinking and acting, „a common understanding – like-mindedness as the sociologists say” (MW 9, 7; cf. MW 6, 413).

Up to here, the major role of habit in human functioning suggests the thought of „an inherent tendency to action and also a hold, command over” (MW 14, 21). Dewey, however, defines the intended meaning of the term with marked differentiation from and in extension to „unthinking routine” (ibid., 50): to mere remote control, programmed repetition, and behavioral response to environmental stimuli. Habits are limiting tendencies that filter and control perception, thinking, feeling, action, but can principally be changed and influenced. Habits are equipped with energy and drive for power; they are the will of man that generally takes place in acts that cannot be

consulted but might have been questioned in the past, and – at least in the event of a problem – offers the possibility of preventing an immediate relationship of will and action (ibid., 22):

Repetition is in no sense the essence of habit. (...) The essence of habit is an acquired predisposition to *ways* or modes of response, not to particular acts except as, under special conditions, these express a way of behaving. Habit means special sensitiveness or accessibility to certain classes of stimuli, standing predilections and aversions, rather than bare recurrence of specific acts. It means will (ibid., 32).

Hence, the selected interpretation shows „that what can be called distinctively individual in behavior and mind is not, contrary to traditional theory, an original datum” (ibid., 60; cf. MW 10, 58; MW 12, 193; LW 2, 61; LW 11, 30), but a smorgasbord of habits that change, are changeable, and have changed in a continuous interaction of nature, human being and society. In general, Dewey anchors this impact of surroundings not sociologically, but in a naturalistically-based *psychologic social psychology*: as „an attempt to understand and explain how the thought, feeling and behavior of individuals are influenced by the actual, imagined, or implied presence of others” (Allport 1954, 5). The focus is on the individual, personal habits, the person’s mind. The goal of the findings is *not* how individual or collective minds form social spaces, but rather how different groups and their overly-individual characteristics create different minds.

Impressive use of this model is shown in *Interpretation of the Savage Mind* (1902). The essay, first published in *Psychological Review*, addresses the question of how the dominant activity of a social group produces generalized mental patterns and can influence the entire lifestyle significantly and the course of the human species over the long term. In terms of social theory and research method, he orients his argumentation on the demands and conditions of practice: the „mind”, as an action-guiding organ, „is at least an organ of service for the control of environment in relation to the ends of the life process” (MW 2, 41). Relevant goals are not set by pure reason, but rather through the existing tasks in daily life, through that which is done and will be done in a group in order to fundamentally shape life. Each analysis of social groups finds action-guiding patterns of thought and solely in and through the analysis of the dominant activity. A person is what he does and, to a certain extent, also conditioned by socialization processes, what he did for generations: „Occupations determine the fundamental modes of activity, and hence control the formation and use of habits” (ibid., 42, cf. LW 2, 263). Dewey exemplifies this line of argumentation on human developmental in nomadic hunting, but allows for a generalizing and perpetual significance. Characteristic for hunting, as the primary food-obtaining form of life, is its immediacy, the direct relationship between needs, activity and satisfaction of needs, of effort and reward: „Want, effort, skill and satisfaction stand in the closest relations to one another. The ultimate aim and the urgent concern of the moment

are identical" (MW 2, 44). These statements are illustrated ethnographically in an example of Australian aborigines who demonstrate the epochal significance of hunting, the fact and the way that habits form mental patterns, those mental patterns define the important areas of life, and those mental patterns historically, at least partially and in new forms, are passed down through generations. The life of aborigines is so fundamentally defined by hunting that not only the act of hunting as such prompts mental relevance, but many parts of life can only be interpreted against this backdrop. Irrespective of whether it is art, religious ceremonies, funeral rituals, the selection of partners or sexuality, the characteristics of the prevailing activity lead to the preference for those possibilities of cultural endowment that offer a familiar degree of correspondence. According to Dewey, sports and games continue this mentally anchored striving for immediacy in and for modernity (ibid., 45).

Impulse

Thus, the priority significance of social habits is set in detail. For Dewey, however, this cannot be the only variable to develop anthropological insights. The human being is simultaneously though not synonymously an *organic existence*. Man owns numerous native predispositions how to respond to surrounding events: „instinctive modes of action, based in the original connections of neurons in the central nervous system" (MW 9, 67f.; cf. MW 14, 63-116; LW 10, 64). Taking this for granted, the intended classification of the human being has to give the biological reality a meaning and define the relationship of nature and nurture without wanting to use creation images. Impulses are the non-learned origin of habits, a bundle of out-going energies: „In the first place, all conduct springs ultimately and radically out of native instincts and impulses" (EW 5, 77). Stimuli are an essential puzzle piece in genesis and transformation, development and progress of individual and social habits. However, Dewey does not place them as the formative force in human functioning, but on a temporal preliminary basis: „They are secondary and dependent" (MW 14, 65). Genetics equips the embryo with unfocused tendencies. Birth puts a baby in a social setting. The interactions with a compulsorily shared setting result in the transmission of social meanings, for a respectively different direction of similar basic configuration:

The same original fears, angers, loves and hates are hopelessly entangled in the most opposite institutions. The thing we need to know is how a native stock has been modified by interaction with different environments" (ibid., 67).

Humans are born undetermined, but determinable. The undirected character of the natural constitution provides every new and existing generation with an opportunity to arbitrarily spawn certain orientations. The importance of stimuli within the theory of human functioning culminates in the fact that habits exist, that these can change in interaction with settings, and that change can happen even if what

exists is extremely powerful. Social change is innate as well. Stimuli are the organic foundation for progress, “agencies of deviation, for giving new directions to old habits and changing their quality” (ibid.). However, due to the fact that human functioning is primarily limited by surrounding conditions, a change of habits can only be realized by the conversion of those social spaces that shape, align, adjust, or change meanings of stimuli.

Thus, the relationship of impulse and habit provides a roadmap for social improvement in a desirable, that is, according to Dewey, always *democratic* manner (cf. Westbrook 1991, 366): the orientation towards growth and the chance to develop, implement, and examine solutions for an intelligent experimental reorganization of habits. Growth needs flexible instead of rigid habits and often expects these as well. The point is to release impulses in such a way and by following an intelligent ideal that they can serve as starting points for particular habits. Here, the ambition is reflexive habits instead of a firm idea: Habits that empower a person to sense existing habits in the event of a problem and to reorganize them. Docility of man shall not be interpreted as encouragement to imitate, but as “power to re-make old habits” (MW 14, 70). Dewey grounds the appropriate way on two fundamental criteria: the impact on the virginity of the upcoming generation, since in childhood and youth there still exist numerous untouched stimuli that can be limited in a democratically intended manner, *and* transformation of existing habits due to inconsistencies. Habits interactively respond to each other, the more rigid the more friction-prone. Whenever diverging impacts meet, there is the chance to productively work on irritations in a joint solution process. Nevertheless, conflicts between social areas, classes, groups, attitudes do not arise spontaneously but need a trigger: the impulsivity.

Intelligence

Human nature and human functioning are socially limited and organically predisposed, though not programmed, but can generally also be changed and influenced. A third constituent is needed to redesign or change habits. Habit and stimulus provide an essential social and physiological basis, but it is the contribution of reason, or *intelligence* to find ways out and productively give events a direction:

Impulse is needed to arouse thought, incite reflection and enliven belief. But only thought notes obstructions, invents tools, conceives aims, directs technique, and thus converts impulse into an art which lives in objects. Thought is born as the twin of impulse in every moment of impeded habit. But unless it is nurtured, it speedily dies, and habit and instinct continue their civil warfare (ibid., 118).

Usually, humans identify their surrounding world in incorporated patterns of the habitual. Knowledge is *knowing how*, a habit that manages the life in and with the world, a practical way of thinking that does not consciously remind, or perceive

without pausing to think, remember, observe. Only irritations of habits are a possible trigger to consciously processing solutions, and to think in a traditional pure way: “habits are not wholly informed by reasoned purpose and deliberate choice – far from it – but they are more or less amenable to them” (LW 2, 240). For such a conscious search a certain linking of habit and impulse is necessary. Driven by impulses, a particular unknown inside the well-known needs to be uncovered in such a way that well-structured and integrated conduct is achievable again. The expected mode set is experimental inquiry, a common structure of the process of non-random problem solving: Inquiry is a modification tool for existing forms of conduct by reflection of problematic situations of primary experience. The aim is to learn from consequences, and to control and prevent consequences, and better assess them. Dewey outlines the associated way of finding, working on and resolving a problem as a generalized process of a piecemeal hermeneutical adjustment. In a schematic order the human totality thereby is equally involved and needed. Thus, thinking is – according to Dewey – never “the armchair thing it is often supposed to be” (MW 10, 328), an activity exclusively taking place in the human mind’s cerebral black box, but is constituted as an iterative procedure of head, heart and hands (e.g., EW 3, 83-89; MW 4, 82-90; MW 6, 188-192; MW 9, 157f.; MW 13, 61-71; LW 8, 196-209; LW 12, 105-122). In this context, impulses are the elements that push the irritation forward towards a solution. For an experimental solution, however, it is necessary that old habits join in to fill the progressing impulse with content. The impulse is triggered by conflicts, but cannot give itself a direction; it needs help from existing habits to generate new habits. The effect becomes visible in practice, by better or worse solutions of a problem. Knowing how can become, imparted by the experimental method, *knowing that*: “knowledge of and about things, knowledge that things are thus and so, knowledge that involves reflection and conscious appreciation” (MW 14, 125).

This means that habits are more than just a determination of thinking – but they are and continuously generate habits. There is an available creativity inherent in the nature of intelligence, a moment of intellectual stimulated transition. However, the concept of how to theoretically understand this intelligent creativity radically changes within Dewey’s works. The incipient discussion about the idea can prominently be found in *Psychology* (1887). There, habit is “a conception of ideas or acts that, if one be presented, the rest of the series follow without the intervention of consciousness or the will” (EW 2, 100; cf. LW 17, 298-309), a cohesive system that needs an external authority to realize change and adaptation. Habit and thinking – although presumably not equally balanced – coexist, exist independently or necessarily together. In *Human Nature and Conduct* Dewey adjusts this separation fundamentally. Habit is no longer blind automatism, static routine, but predisposition: “It means will” (MW 14, 32). All that had initially been rooted as pure reason, is being unmasked an incorporated practice of thinking in action, as *practical reason*. Thinking happens as habits of

thinking, not as soul, mind or any other autonomously-controlled retrievable dimension of reason. There is no independent organ of knowledge that motivates thinking, but knowing happens in the interaction of organism and the surrounding world and requires a certain experiential something to refer to. Reflection itself is a natural event and can be found in nature because of its own characteristics and those of nature (cf. LW 1, 80). All thinking is conditioned and always happens within the possibilities of existing habits and usual manners because, otherwise, it would be disoriented and without any changing solution: "Habit does not preclude the use of thought, but it determines the channels within it operates. Thinking is secreted in the interstices of habits" (LW 2, 335).

By this, it is consistent to democratically turn away from arbitrary thinking habits. For Dewey, the preferred method of choice is inquiry, the reflexive gradual approach from practice to the practice. Inquiry acts as an instrument for growth for the problem-based modification of arbitrary habits to specific thinking habits, accompanied by the epistemic benefit to learn from the consequences, and anticipate, control and better assess further consequences, and – as highly desired side effect – to be able to think to learn: "Still more important is the fact that the human being acquires a habit of learning. He learns to learn" (MW 9, 50). For and by a democratic experimentation community, exactly this process of singular phases of reflection needs to be transferred in an intelligent and intelligence perpetuating habit of creativity, a problem-based feature of thinking to focus results of actions in a certain way, a *habitus of reflexivity*: "What is necessary is that habits be formed which are more intelligent, more sensitively percipient, more with foresight, more aware what they are about, more direct and sincere, more flexibly responsive than those now current" (MW 14, 90).

Communication

Hence, the argumentation in *Human Nature and Conduct* is variously shaped and seems to be watertight: The human being is a creature of habit that right from the moment of their birth enter a wide and complex space of social habits that inevitably becomes the core of their nature as a social being. The development of individual habits is socially transmitted acquisition of meaning. Right from the start the human being is introduced into an established canon of fixed gateways to the world that are accessible and functionally proven in their social environment (cf. MW 1, 69). Any human being is thus provided with a particular smorgasbord of habits, his or her *character*. In that process, the individual habits are not right next to each other, but instead penetrate one another and ensure a continuous and mixed mutual change of habits (MW 9, 30). The dependence of habits on an agile and changeable surrounding world leads in cases of inappropriate pattern of interaction to the possibility of development and change of well-established habits, and consequently of the character. Indeed, the aspiration of argumentation to plausibly open up the human genesis remains in this state. The individual character, or identity, is somehow determined from the outside to the inside intransparently as implicit imitation

of a cultural repertoire, as imitating repetition of exemplified patterns of functioning, and is, in this rendition, provided with critical requirements by George Herbert Mead (1863-1931) shortly after publication:

The implications of social psychology in this book are found in the treatment of the material under three heads, that of custom and habit, that of impulse and change, and that of deliberation. Under the first he shows how inevitably these habits and customs are social, how the child enters into a vast complex of habits, and how these become the very structure of his nature as a social individual. The psychological account of the child's acquiring these habits is sketchy and inadequate. Mr. Dewey has fallen back too much upon the outworn doctrine of imitation, and for this reason he has no satisfactory account of the arising of the self in conduct. I am confident that the psychology of this is to be found first in the little child's assuming the roles of others in his continual play, and thus taking over their habits, and addressing himself as an other, in other words acquiring what is commonly called self-consciousness (Cook 1994, 378).

Approaches towards an advanced argumentation are to be found in Dewey's work as of *Experience and Nature* – here, however, in an implicit referential context and without any evidence for apparent borrowings from socio-psychological ideas of Mead. In the chapter “Nature, Communication and Meaning” (LW 1, 132-161) Dewey reveals his communication-theoretical essentials, the communicative meaning of social interactions and their capability of contributing to the process of habit forming. To this extension, the formation of human identity is added as an *indirect process*. Condition of the possibility to develop human identity is communication, as anthropological unique feature primarily, but not completely in an organized linguistic form: “Meanings run in channels formed by instrumentalities of which, in the end, language, the vehicle of thought as well as of communication is the most important” (LW 2, 210).

The development and meaning of such a human linguistic form, for Dewey, is a “story of the *use*” (LW 1, 139): Right at the start, vocals and sounds had not predominantly been expressive and meaningful, but unconscious, unintentional, useful by and by depending on consequences: differently perceptible physical events that always then became functional when they created useful reactions in a fellow human being. The initial crying of every newborn, for example, is unplanned, but appropriate because of the generated reactions of care, attention, or affection. The transition to organized language happened by symbolically filling of acoustic signals, “used within a context of mutual assistance and direction” (ibid.), when an event can be followed by a distant, reasonably objectified consideration of what has been said. Then, and only then, sounds became words, incidental usabilities in and from the nature were repeatably preserved (e.g. fireplaces, load capacity, leverage), humans can realize meanings of sounds and react to disclosed meanings of stimuli.

Human socialization is, according to Dewey, a natural organic connection, arbitrarily identical to atoms, physical masses, plants or animals. Thus, participating in human language does not only mean that – even for a lonely hermit, a stranded castaway or in an inner monologue – at least two parties are needed; but also that, by and by, an automated

level of reflection is installed which reveals the participation in symbolic calls to action to the human being, and that by the interpretable meaning of calls to action one's reaction can respectively be intentionally verified. Typically human is not only the view of being in a situation of interaction, but how and with what intention and expected consequences. For Dewey, the particular human structure of the inevitable association and the consequences arising from these particular structures of human communitization are more important than the fact that the association exists. Language is at the same time *mean and end* of the creation of social coherences and coordination of action, and is functional for the process of human cooperation and participation:

Language is always a form of action and in its instrumental use is always a means of concerted action for an end, while at the time it finds in itself all the goods of its possible consequences. For there is no mode of action as fulfilling and as rewarding as is concerted consensus of action. It brings with it the sense of sharing and merging in a whole. Forms of language are unrivalled in ability to create this sense, at first with direct participation on the part of an audience (LW 1, 144f.).

Language acquisition and thus the capability of truly human interacting and being a self is always mediated by social contexts. The self only exclusively develops as meanings created by interaction processes, it is always a social self that reaches its 'I' exclusively inside a 'you', and not just a naturally existing phenomenon that somehow, somewhere, sometimes finds its expression in language: "We form our ideas, our estimates of ourselves and of our self-respect in terms of what others think of us, in terms of the way in which they treat us" (LW 5, 239). Condition of the possibility for interaction and identity is *role-taking*, deposited in a specific characteristic of acoustic sounds. Language is a *vocal gesture* and transfers the speaker – contrary to facial expressions, gesture or posture – into a state of subject and object. The speaker talks and simultaneously hears, like any other person, what is being said. The possibility of being the object of one's own vocal gesture and hearing oneself without perspective enables humans to empathize with their interaction partners, initiate a retrospective to themselves, mutually anticipate reactions to their actions and compare them with real reactions, and establish – in mutual agreement and analysis of meanings – a system of significant symbols. Action becomes mutually predictable through significant symbols and can expediently be coordinated as mutual anticipation of functioning. Thoughtfully delving into the role of another person means estimating prior to an action how the opponent will react to one's functioning by reflecting on one's own functioning and reaction from the other person's perspective, provided that similar conditions of socialization and a rough intersection of significant symbols are given. Only via a differentiated language system does it become possible to creatively assume different roles and outline oneself in the looks of the interactive others. Thus, the mutual anticipation of the expectation of an interaction partner is not only condition for successful interaction, but at the same time prerequisite for a unique identity and for a subject that, additionally, does not limit itself in the imitation of habits

from an objective viewers perspective, but – of course, still without exit from the bias in “habits of speech” (LW 1, 145) – adds socially agreed-upon meanings to events and acquires habits.

Conclusion

The collected works of Dewey are characterized in that the anti-skeptical fallibilism that he epistemologically stands up for does not stop at his own thinking. Initiated by sometimes more valid arguments, sometimes recent scientific data, commonly experiences, or personal relationships, smaller or larger upheavals continuously can be traced. Sometimes, these turn the previous thoughts upside down (cf. LW 5, 147-160), and sometimes ‘just’ an extension or reweighting of positions happens. The aim of this essay was to outline Dewey’s theory of human functioning. The result is a construct that, from the beginning, formulates basic characteristics but that, little by little, extends and reshapes design and weighting. The basic structure is of permanent nature: experience is and remains the basic benchmark. The basic trait is to be alive and, as a living creature, to interactively create a balance to sustain and push life forward. The basic unit is and remains unified action. The basic principle is and remains a specific idea of circularity. The basic mission of academic study is and remains to take the context and continuity of human conduct empirically seriously. In particular, the substantive content undergoes modification: Sustainably influenced by William James’ *The Principles of Psychology* (1890) at the beginning the idea of an active biological organism to be studied through the methods of physiological psychology is in focus. Human functioning is the biological interaction of organism and environment and primarily takes place via sensori-motor functions and coordination. Later, especially in the socio-psychological exchange with George Herbert Mead, social patterns of conduct and meaningful series of action rather predominate without giving away the naturalistic perspective in any manner. The organism is amplified by the social individual, environments with habitats, and organic contexts by cultural ones.

The connector that integrates this extension into the idea of circularity is the concept of habit.⁵ For Dewey, habits are inured structures of conduct, customary values, traditions and institutions that combine social functions and elements of personality. There is a double meaning inherent in the concept. On the one hand, habits are the enormous flywheel of human life, its most precious conservative facilitator: Habits, “overlaid and saturated with the products of the reflection of past generations and bygone ages” (LW 1, 40), have a stabilizing effect and are – for particular cultural groups – predictable regarding the respective daily and smooth functional process of individual and habitat.

⁵ In American pragmatism the steering towards exactly that concept is traditional (cf. MacMullan 2013). Early discussions that Dewey certainly knew about are to be found in Charles Sanders Peirce (1877, 5), or William James (1890, 121f.). For further examinations see, e.g., Brinkmann (2011), Cutchin (2007), Garrison (2002), Koschmann (2000), Pratten (2015), or Schäfer (2012).

Habits “possess us, rather than we them, [t]hey move us, they control us” (MW 9, 34). For the most part, the individual and collective life works *because* it takes place in a culturally specific learned manner, unchallenged (cf. LW 1, 22f.). On the other hand, this functionality bears the risk of restricted cultural reproduction: of routine, dissatisfying stagnation, “dead habits” (MW 14, 51). Life is taking place on socially predefined paths. One stews in his or her own juice. Demands and uncertainties are worked on according to existing patterns, the more solidified the less consciously (LW 1, 235). Change and innovation are only possible in problematic situations and with help of a biological basis. Whenever an organic individual falls out of step with the march of its surrounding, there is the possibility and urgency to recover unison with it (LW 10, 19). Now, this return to pre-problematic equilibria can occur happenstance or via directed effort. The most controlling, hence the instrument of choice to be inevitably preferred – and en passant the bridge to Dewey’s educational and political implications of habits (cf. Campbell 1998) – is experimental inquiry.

Experimental thinking enables humans to overcome their thinking, feeling and action caught up in habits in a hitherto best possible way. According to Dewey, experimental adjustment is the human ability to maintain all those habits from retrospective experiences that are prospectively useful in actively and expediently shaping the environment and habitat for individual and collective purposes, *and* the ability to extract those patterns of conduct from these experiences that offer – regardless of the extent of its previous usefulness – suboptimal and vacant solutions for topical problems and irritations. Experimental thinking is *the* way of adjustment by use of reconstructive non-conformism, and is *the* instrument to work on biological and cultural limitations as well as to achieve freedom and individuality. And, in this context, freedom and individuality do *not* mean solitude and riddance of environment and habitat, but the practical knowledge of their mechanism, and (as extensively as possible) shaped reconstructive habits of reflective thinking and intelligent action:

Experience is not a rigid and closed thing; it is vital, and hence growing. When dominated by the past, by custom and routine, it is often opposed to the reasonable, the thoughtful. But experience also includes the reflection that sets us free from the limiting influence of sense, appetite, and tradition (LW 8, 277).

References

- Allport G. W. 1954. „The Historical Background of Modern Social Psychology“, in G. Lindzey (Ed.), *Handbook of Social Psychology*. Reading: Addison-Wesley (3-56).
- Brinkmann S. 2011. „Dewey’s Neglected Psychology: Rediscovering His Transactional Approach.“ *Theory & Psychology* 21:298-317.

- Burke T. 1994. *Dewey's New Logic: A Reply to Russell*. Chicago: University of Chicago Press.
- Campbell J. 1998. „Dewey's Conception of Community,” in L. A. Hickman (Ed.), *Reading Dewey. Interpretations for a Postmodern Generation*. Bloomington: Indiana University Press (23-42).
- Cook G. A. 1994. „George Herbert Mead: An Unpublished Review of John Dewey's Human Nature and Conduct.” *Journal of the History of the Behavioral Sciences* 30:374-79.
- Cutchin M. P. 2007. „From Society to Self (and Back) Through Place: Habit in Transactional Context.” *OTJR: Occupation, Participation and Health* 27:50S-59S.
- Dewey J. EW 1-5. *The Early Works: 1882-1898*, edited by J. A. Boydston [1969-1975]. Carbondale: Southern Illinois University Press.
- Dewey J. MW 1-15. *The Middle Works: 1899-1924*, edited by J. A. Boydston [1976-1983]. Carbondale: Southern Illinois University Press.
- Dewey J. LW 1-17. *The Later Works: 1925-1953*, edited by J. A. Boydston [1981-2001]. Carbondale: Southern Illinois University Press.
- Garrison J. W. 2002. „Habit as Social Tools in Context.” *OTJR: Occupation, Participation and Health* 22:11S-17S.
- Gibson J. J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Goldstein K. 1934. *Der Aufbau des Organismus: Einführung in die Biologie unter besonderer Berücksichtigung der Erfahrungen am kranken Menschen*. The Hague: Martinus Nijhoff.
- Hommel B., Müsseler J., Aschersleben G., & Prinz W. 2001. „The Theory of Event Coding (TEC): A Framework for Perception and Action Planning.” *Behavioral and Brain Sciences* 24:849-937.
- Hughes P. 1928. *An Introduction to Psychology: From the Standpoint of Life-career*. Bethlehem: Lehigh University Supply Bureau.
- James W. 1890. *The Principles in Psychology*. New York: Dover.
- Koschmann T. 2000. „The Physiological and the Social in the Psychologies of Dewey and Thorndike: The Matter of Habit,” in B. Fishman & S. O'Connor-Divelbiss (Eds.), *Facing the Challenges of Complex Real-World Settings*. Mahwah: Erlbaum (314-19).
- Lashley K. S. 1951. „The Problem of Serial Order in Behavior,” in L. A. Jeffress (Ed.), *Cerebral Mechanisms in Behavior*. New York: John Wiley and Sons (112-36).
- MacMullan T. 2013. „The Fly Wheel of Society: Habit and Social Meliorism in the Pragmatist Tradition,” in T. Sparrow & A. Hutchinson (Eds.), *A History of Habit: from Aristotle to Bourdieu*. Lanham: Lexington Books (229-55).
- Mead G. H. 1903. „The Definition of the Psychical.” *Decennial Publications of the University of Chicago* 3:77-112.

- Neubert S. 2009. „Pragmatism: Diversity of Subjects in Dewey’s Philosophy and the Present Dewey Scholarship,“ in L. A. Hickmann, S. Neubert, & K. Reich (Eds.), *John Dewey Between Pragmatism and Constructivism*. New York: Fordham University Press (19-38).
- Peirce C. S. 1877. „The Fixation of Belief.“ *Popular Science Monthly* 12:1-15.
- Piaget J. 1970. *Abriß der genetischen Epistemologie* (German ed. 1974, trans. by F. Kubli). Freiburg: Walter.
- Piaget J. 1947. *Psychologie der Intelligenz* (German ed. 1976, trans. by H. Aebli). München: Kindler.
- Pratten S. 2015. „Dewey on Habit, Character, Order and Reform.“ *Cambridge Journal of Economics* 39:1031-52.
- Rölli M. 2009. „Natur und Kultur. Oder: Wie der Pragmatismus John Deweys einen traditionsreichen Gegensatz aufhebt“. *Allgemeine Zeitschrift für Philosophie* 34:305-29.
- Rorty R. 1979. *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press.
- Rosenbaum D. A., Cohen R. G., Jax S. A., Weiss D. J., & van der Wel R. 2007. „The Problem of Serial Order in Behavior: Lashley’s Legacy.“ *Human Movement Science* 26:525-54.
- Schäfer H. 2012. „Kreativität und Gewohnheit. Ein Vergleich zwischen Praxistheorie und Pragmatismus“, in U. Göttlich & R. Kurt (Hg.), *Kreativität und Improvisation. Soziologische Positionen*. Wiesbaden: Springer Fachmedien (17-44).
- Wallach A., Marom S., & Ahissar E. 2016. „Closing Dewey’s Circuit,“ in A. El Hady (Ed.), *Closed Loop Neuroscience*. London: Elsevier Science (93-100).
- Watson J. B. 1913. „Psychology As the Behaviorist Views It.“ *Psychological Review* 20:158-77.
- Watts A. 1966. *The Book. The Taboo Against Knowing Who You Are*. New York: Pantheon.
- Westbrook R. B. 1991. *John Dewey and American Democracy*. Ithaca: Cornell University Press.

David Jaitner
(Köln)

Reconstructive Habits: John Dewey on Human Functioning

Abstract: The academic strife to parse, investigate and adjust human functioning establishes varieties of at least three key concepts: behavior, action, activity. Depending on the general approach chosen, human functioning is therefore defined in a certain way and in a certain understanding of freedom. Within this paper, the pragmatist considerations of John Dewey (1859-1952) offer a sophisticatedly formulated theory of human functioning that, undoubtedly, takes action-theoretical paths but formulates underlying assumptions in a significantly unusual way. The main focus is to outline the theory in such a way that clearly shows the unusual as part of the usual and the usual as part of the unusual. For this purpose, the first section defines action as the basic category of Deweyan human functioning where sensory stimuli, registering elements and motor responses play a leading role, but according to Dewey questions the today still popular model of behaviorist psychology, that positions isolated and a-cultural stimulus-response-procedures in the human organism. The second section affirms the theoretical inclusion of deliberative elements that constitute human action, but according to Dewey witnesses their substantial and rather sporadic significance in a predominantly habitual human functioning. The conclusive section outlines the possibilities and limits of transforming habitually inured patterns of human conduct by means of reconstructive habits.

Keywords: human functioning; habit; reconstruction; pragmatism; John Dewey.

Ethics in Progress (ISSN 2084-9257). Vol. 9 (2018). No. 1, Art. #1, pp. 4-24.

Creative Commons BY-SA 3.0

Doi: 10.14746/eip.2018.1.1

Deliberation and Automaticity in Habitual Acts

Christos Douskos

ORCID: <https://orcid.org/0000-0002-1776-1100>
(Hellenic Open University, cidou@upatras.gr)

1. Introduction

Most current philosophers and psychologists assume that habitual acts do not ensue from deliberation, but are direct responses to the circumstances: habit essentially involves a variety of automaticity. My objective is to show that this view is unjustifiably restrictive. I argue that a habit can explain an act in various ways. Pointing to the operation of automaticity is only one of them. The view that habitual acts do not ensue from deliberation unduly narrows the explanatory scope of the idea of habit.

In section [2] I introduce a skeletal conception of habit that implies minimal commitments, and a corresponding conception of habit explanation. I suggest that the commitments of this minimal conception are generally acknowledged by accounts of habit in philosophy and psychology. This will enable me to use this conception as the basis of the ensuing discussion. In section [3] I turn to the notion of automaticity. Many philosophers, as well as the overwhelming majority of psychologists, assume that automaticity is an essential feature of habitual acts. However, the term 'automaticity' can be understood in very different ways, so I go on to examine how the variety of automaticity pertaining to habitual acts is best characterized. Habitual acts are commonly taken to be direct responses to the circumstances one is confronted with; the relevant kind of automaticity involves the absence of deliberation. I call the claim that habitual acts necessarily involve automatic processes of this kind(s) the Received View. In section [4] I argue against the Received View. I present certain sorts of contrast case, on the basis of which I argue that automaticity is not an essential feature of habit: habit can explain how an act came about in various ways, some of which do not imply the absence of deliberation. The discussion suggests that acquired automaticity is one outgrowth of habituation that is relevant to

explanation, but not the only one. I conclude that the Received View fails to do justice to the theoretical usefulness of the idea of habit.

2. Habit and Habit Explanation

'Habit' and its etymological progenitors (*habitus*, *hexis*) have a long philosophical history, and have come to express a range of quite different ideas. These range from simple compulsions and additions to full-blown character traits, including all sorts of things in between, such as regularity and consistency of behaviour in certain circumstances, routine everyday activities, and much more besides. It is very doubtful whether one can formulate a conception of habit that is broad enough to encompass all these different ideas, while at the same time rigorous enough so as to be theoretically useful. In pre-theoretical discourse we use 'habit' mostly in connection to repeated everyday activities, which do not ensue from compulsions or physiological addictions, but do not amount to full-blown character traits either. We shall see that this comes closer to fixing on a theoretically useful idea, as long as we impose a further restriction to the effect that habit is a determinant of action: habit *explains* acts, it is not merely regularity in behaviour.

On a quite skeletal characterization of habit found in current philosophical literature, habit is an agential disposition or propensity to do something in certain circumstances, a disposition acquired by having regularly done the same thing in similar circumstances in the past (Brett 1981, 363; Alvarez 2010, 187). This suggests three necessary conditions on the idea of habit: *S* is in the habit of *a*-ing in circumstances *c*, only if (i) *S* is disposed to undertake to *a* when *c* obtain; (ii) *S* has *a*-ed in *c* in a sufficient number of occasions in the past; (iii) the disposition to *a* in *c* is acquired *as a result* of having repeatedly *a*-ed in *c* in the past.¹ The central question to be addressed below is whether habit is subject to some additional condition: automaticity. Notice, however, that I am not looking for a set of sufficient conditions here, i.e. for a definition of habit, partly for the reasons given in the previous paragraph.

Let me briefly explain the import of these conditions. Condition (i) pertains to the

¹ Let me explain the action-theoretic terminology used in these clauses and below. According to the coarse view of action individuation adopted here, actions are temporal particulars (events or processes) that can be variously specified. An action specification corresponds to a property, or to a thing one does in acting. Actions are intentional only under some specification(s). Similarly, an action can be a manifestation of a habit only under some specification(s) (Pollard 2008, 49). The fact that John is in the habit of taking a certain route when coming back from work does not imply that the lower-level specifications of his action on two or more occasions are the same; e.g. he will take a different lane, change gears at different points etc. So 'is habitual' is a predicate of action specifications: the schematic letters 'a' in the clauses above stand for properties or doables. It follows that 'ensues from deliberation' and 'ensues from automaticity' are also predicates of action specifications (on this last point, see Wu 2011; 2013). Moreover, the same thing done (action specification) might ensue from deliberation on one occasion while from the operation of automaticity on another occasion. (This is a standard point in the literature on basic acts; Hornsby 1980, 69 and Enc 2003, 54). So just as 'is basic', 'ensues from automaticity' predicates thing done on occasion (or action-particulars under some specification). So I shall use 'act' as equivalent to 'thing done on occasion', for this allows me to talk simply of *acts* as exhibiting automaticity.

explanatory role of habit. A habit ascription explains an agent's doing something on a certain occasion. We may explain why John is taking a certain route now that he is coming back from work as opposed to some other, by citing his habit of taking that route when he comes back from work. In this respect habit ascriptions have the same role as ascriptions of typical physical dispositions: we may explain why this rubber ball expanded after having been compressed on some occasion by citing a relevant disposition: its elasticity.² Moreover, just as with typical dispositions, the canonical specification of a habit has to mention the circumstances of manifestation. This is because it is the fact that certain circumstances obtain on some occasion that explains why the habit in question has been manifested on that occasion.³ The fact that John is coming back from work now explains why he is taking that route. This is why the circumstances of manifestation are part of the identity conditions of a habit. John's habit of taking the bus when coming back from work is a different habit from the habit of taking the train when coming back from work, because only the latter habit can explain why John is boarding the train now that he is coming back from work. A habit cannot be specified merely as a habit of *a*-ing, for this would fail to explain why the agent *a*-ed on some occasion where *c* obtained, as opposed to some other where *c* did not obtain.

Two points are in order here. First, circumstances of manifestation do not always figure explicitly in habit ascriptions. In many cases the circumstances may be too diverse, or too generic. In such cases the circumstances cannot be specified in a non-trivial manner, and hence are likely to be omitted for pragmatic reasons. Second, the examples above already imply that 'circumstances' here is to be broadly construed: it does not include only the external environmental circumstances that elicit an act, as for instance the habit of picking up the phone when it rings. An agent's objectives, or the activity one is already engaged in, may also constitute the circumstances where one habitually takes certain means: John's objective is to get back home from work, and this constitutes the circumstances where he takes a certain route. In the latter case the habit of *a*-ing in *c*, can be specified equivalently, as a habit of *b*-ing by *a*-ing, since *b*-ing here is the circumstances where one *a*'s. We may call *habitual routines* those cases where the habit elicits at least one means or step. Habitual routines often comprise multiple acts in a sequence; think of the different turns John takes when driving back home via the familiar route, or of making a cup of tea.

The second condition captures another central feature of habit. One cannot be in the habit of *a*-ing in *c* unless she has repeatedly *a*-ed in *c* in the past. As Pollard (2008, 46-47) explains, the claim that one is in the habit of doing something she has never done before

² Notice that just as with other dispositions, a habit is not invariably manifested when the relevant circumstances obtain. So the modal force of a habit ascription '*S* is in the habit of *a*-ing' is not that of a conditional 'if *c* obtain, *S* *a*'s'. The ascription is best paraphrased by a habitual sentence '*S* *a*'s in *c*', as opposed to a universally quantified proposition. Fara (2005) argues that this is true of disposition ascriptions generally.

³ As I mention below, the fact that certain circumstances obtain is (typically) the agent's *reason for acting*.

(or does for the very first time) is not intelligible, because it involves a contradiction. Habit requires past repetition, as matter of conceptual necessity. However, the number of repetitions required to ascribe a habit varies widely depending on the habit in question, and even if we fix on a given habit, the answer is bound to be vague. Notice again that repeated performance does not suffice to ascribe a habit. One needs to have repeatedly done the same thing *in relevantly similar circumstances*. John is not in the habit of going to work by bus no matter how many times he commuted to work by train.

It is the third condition that raises the most controversial questions, however. This condition in effect joins conditions (i) and (ii): it requires that the disposition in question is formed *as a result of* repeatedly acting in the same way in similar circumstances. It is surely a central fact about human nature that by repeatedly acting in a certain way in certain circumstances we acquire the relevant tendency, disposition or propensity. Indeed, it is well known that upon having acquired a habit, this habit can be manifested even when one has decided or resolved *not* to do the thing in question, as in the case of habitual action slips, where an agent acts contrary to one's previously formed intentions.⁴ Moreover, we all know that some habits are hard to break. We all have various 'bad' habits we repudiate and try to get rid of. But these facts about human nature require explanation. It is here that the idea of acquired automaticity becomes highly relevant. On the Received View, the propensity that constitutes a habit is grounded solely on the workings of a certain kind of automaticity: when one is in the habit of *a*-ing in *c*, one has the propensity to *a* in *c* to the extent, and *only* to the extent, that he responds by *a*-ing without any conscious thought or deliberation about what to do. I shall come back to this in what follows.

The basic characteristics of the idea of habit outlined above are also found in psychological research on habitual action. It is true that due to the behaviorist associations of the dispositional idiom in psychology, nowadays psychologists do not generally characterize habits as propensities or tendencies. (One exception is Quelling and Wood [1998, 55] who characterize habits as "tendencies to repeat responses given a stable supporting context"). This stands in sharp contrast with the philosophical tradition, in which the broader idea of habitus is a paradigm case of a dispositional notion. Nevertheless, I think that several characterizations of habit in psychology appeal implicitly to modal notions, or at least that they can be suitably re-phrased in a dispositional idiom without any change of substance. What I want to stress here is that psychologists also characterize a habitual act as a direct response to specific circumstances, and assume that a habit is formed on the basis of repeatedly acting in the same way in the past in these circumstances. For example, Wood and Neal (2007, 845) claim "habits are repeated responses that come to be cued by recurring features of contexts". Gardner, de Bruijn and Lally (2011, 175) formulate the idea of habit as "behavioural patterns learned through

⁴ For influential discussions of action slips in psychology, see Norman (1981) and Reason (1990). For discussion of the philosophical problems raised by slips see Amaya (2013).

context dependent repetition: repeated performance in unvarying settings reinforces context-behaviour associations such that, subsequently, encountering the context is sufficient to automatically cue the habitual response". For Verplanken and Wood (2006, 91) habit is "[a] type of automaticity in responding that develops as people repeat actions in stable circumstances (...) Recurring aspects of performance circumstances come to trigger habitual responses directly without input from people's intentions or decisions to act". Orbell and Verplanken (2010, 374) assert that "[d]ifferent accounts of habit converge on the idea that three aspects of habit are central: a habit is behavior that is frequently repeated, has acquired a high degree of automaticity, and is cued in stable contexts". Gardner (2015, 279) lists eight definitions of habit in current literature. All but one explicitly post that habitual acts are elicited ('triggered', 'cued') by similar circumstances ('stable contexts', 'unvarying settings', or 'recurrent features of the context' or 'stable circumstances'). Notice finally that psychologists have long recognized that 'circumstances' or 'context' is to broadly construed, to account for habitual routines.⁵ Thus despite some terminological differences, these formulations of the idea of habit converge on the conditions outlined above: a habit is formed on the basis of repeated responses to *similar circumstances*, circumstances that *directly* elicit an act and thus explain why the habit has been manifested in some occasion.

In the next section I shall focus on another key element, which figures already of the characterizations above: the notion of automaticity.

3. Habit and Automaticity: The Received View

Until quite recently, the term 'automaticity' did not figure prominently in the philosophical literature on habit. In a classic discussion, Ryle (1949) claims that "[w]hen we describe someone as doing something by pure or blind habit, we mean that he does it automatically and without having to mind what he is doing" (Ryle 1949, 30). But Ryle uses the term rather prosaically (as a "metaphorical title", as he puts it [1949, 95]), and does not pause to consider its possible meaning. More recently, Julia Annas has no much use for this term in her discussion of habit as routine (Annas 2011a; 2011b). But this does not mean that the *idea* of automaticity is not pervasive in these discussions. For as we shall see, automaticity is best understood as an umbrella term for combinations of features such as the absence of deliberation and forethought, diminishing conscious awareness, or diminishing attention. And hardly any philosopher who has written on habit, or at least on habit as understood nowadays, has failed to notice these characteristics of habitual acts. Notice how Ryle alludes to the absence of attention in the quotation above, for instance.

⁵ This is usually formulated as the claim that (the automaticity of) habit guides "routine action sequences". See Bargh and Gollwitzer (1994, 78); Cooper and Shallice (2006); Wood and Neal (2007, 851); Graybiel (2008, 361); Gardner (2014, 280); As Wood and R nger (2016, 292) notice, many psychologists claim that the sequencing of multiple acts is a distinguishing feature of the automaticity of habit (e.g. Dezfouli, Lingawi & Balleine 2014).

Two authors who have more to say on the idea of automaticity in their discussion of habit are Nathan Brett and Bill Pollard. After carefully distinguishing between the claim that an act is *initiated* automatically as a direct response to a cue or circumstances and the claim that automaticity *guides* the performance of the act, Brett (1981) subscribes to the idea that habitual acts are initiated or elicited automatically. And he takes the operation of automaticity to imply that the act is “not the product of a conscious decision”, which is tantamount to saying that it involves the absence of deliberation (Brett 1981, 357; see also 364). The idea of automaticity is given a more prominent place in Bill Pollard’s definition of habitual behavior (Pollard 2008, 55-56; Appendix: Summary of Definitions). Pollard builds automaticity in his definition of habit, only in the sense that habitual acts are *elicited* automatically as a response to the circumstances (Pollard 2008, 61), and the absence of deliberation in his definition of automaticity.⁶

Psychologists present a more united front in taking automaticity to be a defining feature of habit. Indeed, the two ideas came to be so closely connected that, as Wood and Runger (2016, 292) observe that the term ‘automaticity’ is often used interchangeably with ‘habit’ in psychology literature. However, in psychology the term ‘automaticity’ is understood in very different ways. Since Bargh’s (1989; 1994) influential criticisms of earlier monolithic conceptions of automaticity (as for instance Schneider and Shiffrin 1977), the term ‘automaticity’ is not taken to stand for a single concept. On the influential de-compositional approach to automaticity (Moors & De Houwer 2006; 2007), different varieties of automaticity are distinguished by appeal to the different combinations of features that characterize automatic processes. Among the most commonly cited features of automaticity are (i) the absence of deliberation and (its issue) intention; (ii) diminishing attention and the accompanying subjective sense of effort; (iii) diminishing (conscious) awareness; (iv) speed, accuracy and efficacy in overt bodily acts or other kinds of response. The key idea in de-compositional approaches is that different combinations of these features give different *kinds* or determinate concepts of automaticity. These varieties are exhibited in different kinds of automatic behavior.

Which combination or combinations of features pertain to the variety or varieties of automaticity exhibited in habitual acts? A survey of the relevant literature provides no univocal answer to this question. However, this does not matter for present purposes. What matters here is that there is general agreement in the literature that the automaticity of habit involves feature (i): the absence of deliberation and intention. Just as in the case of philosophers such as Brett and Pollard, psychologists take the absence of deliberation

⁶ Pollard’s definition of habit consists of three clauses. A behaviour of *a*-ing in *c* is a habitual action if (a) the agent has repeatedly *a*-ed in *c* in the past; (b) *a*-ing in *c* has become automatic for the agent as a result of this repetition; (c) the agent has rational intervention control over her habitual act. Pollard further defines the idea of automaticity figuring in (b) by two conditions: (i) the absence of conscious deliberation; (ii) the absence of trying (Pollard 2008 Appendix: Summary of Definitions). For a criticism of Pollard’s views on habit see [author b].

to be a necessary (though not sufficient) condition of the automaticity of habit, and automaticity a necessary condition on habitual acts.⁷

This raises the question of what is meant by ‘deliberation’ in these discussions of automaticity. This question is not usually addressed, and when it is, the answers diverge to a smaller or larger extent. It is clear, however, that since deliberation is *contrasted* with automaticity, a rather exacting conception of deliberation is in play. Some philosophers would allow for some variety of more or less unconscious or implicit deliberation, which operates below the agent’s awareness and does not demand resources of attention. They suppose that allowing for such a conception of deliberation is necessary to account for aspects of overt rational behavior that is not preceded by conscious thought. I believe that recourse to such a conception of deliberation testifies an intellectualist prejudice, but this is not the point here. My point is rather that since diminishing (conscious) attention and awareness are par excellence features of automaticity, this alleged variety of deliberation can no longer be *contrasted* with automaticity. Bearing this in mind, I shall take deliberation to be a conscious intentional activity, in which the agent aims to resolve a question.⁸ Deliberation is conscious in that it involves attention and awareness. Moreover, deliberation is an activity one intentionally engages in; it is impossible to deliberate unintentionally. Practical deliberation aims to resolve a practical question: what to do and/or how to do something. The questions addressed in practical deliberation are settled with the formation of intention, that is, the agent’s *decision* about what to do and/or how to do it.

The claim that a habitual act essentially involves automaticity and that this variety of automaticity implies the absence of deliberation is considered to be a consequence of the claim that habitual acts are *directly* ‘activated’, ‘triggered’ or ‘cued’ by the circumstances (broadly construed). The idea is that when an agent is in the habit of *a*-ing in *c*, the perception of *c* elicits an act of *a*-ing without conscious thought or deliberation—which is to say that it elicits this response *automatically*. Indeed, in a widespread experimental research paradigm, the measure of the influence of the automaticity of habit in behavior (“habit strength”) is its potential to elicit acts that conflict with one’s intentions: action slips. Orbell and Verplanken (2010, 378) explain the rationale: given that a habitual act

7 Some examples: Ouellette and Wood (1998, 55); Verplanken and Wood (2006, 91); Graybiel (2008, 361); Neal and Wood (2009, 580); Gardner, Abraham, Lally, and de Bruijn (2012, 1); Gardner (2012, 33); Wood, Labrecque, Lin and Rüniger (2014). For dissent, see Maddux (1997). Of course, this does not mean that the automaticity of habit is specified solely on the basis of this feature. Many authors follow William James (1981, 114) in citing diminishing attention and awareness as central features of the automaticity of habit. As Ouellette and Wood (1998, 55) observe, there is general agreement that the automaticity of habit involves diminishing attention (see also Graybiel 2008, 361; Wood, Labrecque, Lin & Rüniger 2014, 378-79). Other definitions often figure the feature of (diminishing) awareness, rather than attention (van t’Riet, Sijtsma, Dagevos and de Bruijn 2011, 586; Nilsen, Roback, Broström, & Ellström 2010, 1). Nevertheless, the absence of deliberation and intention is a central feature of automaticity if only because, as I explain in the following paragraph, conscious deliberation implies other features, such as attention and awareness, whereas the converse is not the case.

8 I borrow the essentials of this characterization from Owens (2008).

is elicited by contextual cues as opposed to intentions, we can measure the strength of the habit in question by “the likelihood of an individual making an unintended response to a cue previously associated with that response. The stronger the cue contingent automaticity associated with the response, the more likely an accidental, nonintentional response would be made.” Similarly, in their comprehensive review article of the psychology of habit, Wood and R nnger (2016, 292) say that “[t]wo defining features of habit automaticity are (a) activation by recurring context cues and; (b) insensitivity to short-term changes in goals (a.k.a., not goal dependent).” These two features are closely related: it is *because* a habitual act is activated by a context cue that it is insensitive to one’s current goals. And since intentions, the issue of deliberation, are commonly taken to be a kind of goal (Moors & De Houwer 2006, 303, *inter alia*), the claim that habitual acts are not goal-dependent is precisely the claim that they satisfy the feature (i) of automaticity: the absence of deliberation and its issue, intention.

Not all psychologists would agree that the automaticity of habit is not a form of goal-dependent automaticity. An alternative view appeals to the ‘auto-motive’ model of goal-dependent action (Bargh 1990; 1997; Bargh & Gollwitzer 1994), on which agential goals can be activated by being directly elicited by environmental circumstances or other features of the context. Thus Aarts and Dijksterhuis (2000a; 2000b) have argued that, when an agent is in the habit of *a*-ing by taking certain means *b*, forming an intention to *a* automatically elicits the goal to *b* (a habitual routine). But notice that this model in no way implies that the automaticity of habit involves deliberation and intention. On the contrary, the claim that the goals in question are *automatically* activated is a claim to the effect that the process by which these goals are formed does not involve deliberation, attention or (perhaps) even awareness. So it does not issue in intentions proper, but only in a variety of goals which fall short of intentions proper. One of the theoretical incentives for this model is to accommodate the notion that habitual acts can exhibit a form of goal-directed automaticity while not ensuing from conscious deliberation and intentions.⁹

In this section we have seen that many philosophers and almost all psychologists take automaticity to be a defining feature of habitual acts. Habitual acts are direct

⁹ One might wonder how the claim that habitual acts ensue from automaticity squares with the fact that in the explanation and assessment of action we commonly take manifestations of habit to be things intentionally done. Though it is not necessary to argue for this here, I should like to point out that this problem arises only if one assumes the Causal Theory of Action. On this theory, an event is an action only if it is suitably caused by mental states, and it is an intentional action only under the specification figuring in the content of these states. In a nutshell, the problem is that since habitual acts can ensue from the operation of automaticity (here it not does not matter whether they *always* do), and the operation of automaticity is contrasted with the causal operation of mental states, habitual acts can hardly be intentional or indeed manifestations of agency at all. This is why, as several authors have argued (Pollard 2006a; Di Nucci 2011; [author a]), the Causal Theory cannot accommodate the intentional, or even agential, character of habitual acts. However, the problem dissolves on a conception of reasons as facts. When one manifests a habit of *a*-ing in *c*, the fact that *c* obtain (typically, at least) constitutes the agent’s *reasons for a*-ing. The fact that John is getting back home is his reason for taking that route. The fact that the phone rings is one’s reason for reaching for the handset. Habitual acts are elicited by the circumstances that provide one’s reason for acting, and are explained by the fact that these circumstances obtain in the way(s) facts rationally explain an act.

responses to circumstances, which is to say that the relevant variety of automaticity involves the absence of deliberation and intention. Given its wide appeal in contemporary philosophy and empirical psychology, it is fair to call this the *Received View* regarding the relation between habit and deliberation. I should stress, however, that this view has not been predominant throughout of history of philosophy and related disciplines (I shall come back to this in the following section). The Received View stems from a quite recent conceptualization of habit in empirical psychology, some aspects of which I have sketched above. Its influence largely reflects the impact of empirical psychology on current philosophical conceptions of the mind.¹⁰ This observation makes the examination of the credentials of the Received View all the more urgent.

4. Against the Received View

It is not very hard to show that the sorts of acts we call habitual in pre-theoretical discourse may involve conscious thought or deliberation. Surely, to say that Maya is in the habit of going to holidays in Sicily, or of visiting her parents at Christmas, is not to imply that she gives no thought at all about what to do before doing these things. But this fact would seem to be of limited significance in the present context. A defender of the Received View may acknowledge that there is a lax use of 'habit' in pre-theoretical discourse that does not imply the operation of automaticity. But she would suggest that we set this aside because it does not do any explanatory work, in either theoretical or pre-theoretical contexts. The theoretically useful idea of habit is a disposition that *explains* an act: why an agent came to do something on a certain occasion. As Gardner puts it:

Definitions must be judged according to their usefulness for predicting, explaining and changing behaviour. The lay definition of 'habit' as frequent, regular or persistent behaviour is unsatisfactory because it offers no explanatory mechanism for these characteristics (...) (Gardner 2014, 278).

According to Gardner and other proponents of the Received View, the sole explanatory mechanism of habit is *automaticity*:

A psychological operationalisation of habit has emerged, which incorporates an explanatory mechanism: habits are actions that are frequently performed because they are initiated automatically (Gardner 2012, 32).

It is important to distinguish two different claims here. First, an act is habitual in a theoretically interesting sense only if the habit explains an act: why and/or how an agent did something on some occasion. Second, a habit can explain an act only if the act ensues from the operation of the relevant variety of automaticity. It is the second claim that is my target here.¹¹

¹⁰ I would like to thank an anonymous reviewer for pressing this point.

¹¹ Notice that the first claim is also open to dispute. It is not at all evident that a habit explains an

Let us start by noticing an implication of the Received View. Since a habitual act necessarily ensues from automaticity, it is possible that an agent is in the habit of *a*-ing in *c*, that she *a*'s in *c* on some occasion, and yet that the act is not habitual properly speaking, because it does not ensue from the operation of automaticity. For example, suppose John is in the habit of going for a walk before breakfast. But suppose that for some reason John paused to deliberate about whether to do so this morning. Suppose John finally decides to do what he is in the habit of doing, and acts on this intention: he does for a walk. On the Received View, John's act is not really habitual—not *manifestation* of his habit. We may casually say that his going for a walk this morning was habitual, for the pre-theoretical conception of habit allows this. But properly speaking, the act was only in *accidental accord* with his habit. For the act was not directly elicited, 'triggered' or 'cued' directly by the circumstances, but ensued from deliberation.¹² What makes it the case that an act is a manifestation of a habit is the obtaining of a suitable explanatory relation between the habit and the act. Otherwise an act is merely in accidental accord with her habit. And on the Received View the *only* way in which a habit can explain the act is by appeal to the operation of automaticity.

In what follows I shall argue that what the Received View fails to take into account is that there are *various* ways in which a habit can explain an act. The operation of automaticity is only one of them. So an act can be a manifestation of that habit, even when it ensues from conscious thought and deliberation. I shall consider certain sorts of case that serve to illustrate what is wrong with the Received View.

The first sort of case is this. Suppose John plays tennis on Sunday afternoons; he has been doing that for years. His habit is firmly entrenched in life, so he does not usually pause to deliberate. When the afternoon arrives, he just heads for the tennis club. But this Sunday evening John has been invited to a friend's birthday party. So John pauses to deliberate about what to do, weighting the pros and cons. John finally decides (forms an intention) to go and play tennis, and acts on that intention: he plays a game of tennis this Sunday evening as he usually does.

Our question is whether John's habit still plays a role in the explanation of his act. The answer to this question is affirmative. What goes on in the example above is that habit *proposes* or *suggests*—so to speak—the default course of action, or activates the relevant mental representation, but in view of the alternative course of action presented by the invitation, John has to make a choice between these two options. However, habit still plays a role, precisely because it *constrains the options considered in deliberation*. John did not consider all the other things he could do on that sunny Sunday evening. He

act only in terms of explaining how the act came about—the act's occurrence on a certain occasion. Pollard (2006b) argues that habit explanations are a form of *constitutive* explanation, "the sort of explanation we give when we explain something by saying that it is a part of –constitutes – something else" (Pollard 2006b, 240). I set this aside because even if Pollard's claim is correct, it is not straightforwardly incompatible with the claim that the act ensues from automaticity.

12 A proponent of the Received View who subscribes to the Causal Theory of Action, would say that it was the intention that was the triggering cause of the act, as opposed to the circumstances.

likes going the pub, or to visit the local museums, and so on. His habit prevents him from considering all the other good alternatives.¹³ For we may suppose that had John's best option was to go to the museum. Had he paused to consider all his options, he would have acknowledged this and would have acted accordingly. Thus whereas John's habit is an indispensable part of the explanation of what he does, and hence not merely in *accidental accord* with his habit, he still acts on intention formed in deliberation. Generally, when an agent considers refraining from manifesting a habit, this is because he becomes aware of a salient alternative. Typically, the agent will choose between these two options. Considering refraining from manifesting a habit does not typically involve re-considering afresh *all* the available courses of action, or even a suitable sub-set of them. So when the agent finally decides to do what he is in the habit of doing in deliberation, the habit still explains the act.

To see the point more clearly, contrast John's case with the case of Matt. Matt is not in the habit of playing tennis on Sunday evenings. But this Sunday Matt has been invited for a game of tennis by a good friend. Matt has several other options, which he carefully considers in deliberation, but he finally decides (forms an intention) to go and play tennis. Now John and Matt likewise act on an intention formed in deliberation. The key point here is that if we insist that John's act is *not* a manifestation of his habit because it ensues from deliberation, then we have obliterated the difference between the case of John and Matt. But there is an evident difference between these two cases, and even though we might disagree on how exactly to characterize it, it is clear that this difference is due to the fact that only John has the habit in question. In the case of John habit does play *some* role in the explanation of his act, for there were better options open to him, which were occluded by the force of habit. But in the case of James habit does not explain anything, if only because he does not have the habit in question. This suggests that John's going to play tennis on the occasion, while it ensues from deliberation, is not merely in accidental accord with his habit but a manifestation of it.

I am coming now to a different kind of explanatory relation between habit and act, a relation that—again—does not preclude deliberation. A habit can be formed only when one repeatedly finds oneself in a familiar environment. An environment includes other people, places, objects, social events, and much more besides. Some of these will be very

¹³ Verplanken, Aarts and Van Knippenberg claim that the results obtained in their experiments suggest that "habit in daily behaviours may lead to a narrow focus on the habitually chosen option, i.e. habit may limit the decision-maker's consideration set. (...) In other words, habit may be accompanied by less motivation to search out information, regardless of whether this information is unknown, and regardless of its importance." (Verplanken, Aarts and Van Knippenberg 1997, 556). In later research, Danner, Aarts, and de Vries (2007) examine how habit formation involves limiting the consideration of alternative courses of action by inhibiting the retrieval of possible alternatives from memory. Consequently, as a habit is gradually formed, relevant alternatives become more likely to be overlooked in deliberation. So one way to account for the case above is to suppose that while the invitation summons John to consider *one* alternative, the habit is still operative in inhibiting other alternatives, including the best one. But in such a case deliberation will still be required. So although research on habit provides the essentials for making sense of cases such as the ones described above, psychologists who take automaticity to be a defining feature of habit have failed to take such cases into account.

important to an agent's life. So as a result of this familiarity, one tends to develop emotional bonds to other people, places, objects, social processes, and so on, as well as feelings of comfort, security, familiarity, affinity or intimacy with what one is well acquainted with on the basis of previous experience. By contrast, an unknown course of action might be regarded with a sense of reluctance, uncertainty, insecurity and even vulnerability. This is why creatures of habit like us gravitate towards familiar environments and people. The intimate relation between habit and feelings or emotions is a recurring theme in more historically oriented accounts of habit (Carlisle 2006; 2014).¹⁴ And it figures increasingly in the relevant research in psychology.¹⁵

When it is recognized that repeatedly doing something in certain circumstances (which include places and other sorts of environment, objects, people, etc.) promotes the formation of feelings of familiarity and attachment with circumstances, it is not hard to understand how the fact that one is in the habit of doing something can explain an act even though the act ensues from deliberation. For one thing, the emotional bonds forged by habituation may weigh in deliberation in favor of doing what one is in the habit of doing.

The second sort of case serves to illustrate these points. In this case the agent considers many available options in deliberation, but she still forms an intention to do what she is the habit of doing. Suppose Helen is in the habit of staying in some hotel when she visits London. She has done that again and again over the years. However, her friends have by now convinced her that other hotels offer better accommodation and services at this price range. Helen is about to book a room for the next trip. She pauses to take into account the considerations urged by her friends. Yet after pondering about which hotel to book, she decides to go to the one she usually goes. Her reasons are simply that she is

14 On this point it is worth quoting Carlisle in full: "The force of habit suggests a natural affinity with the known, and an aversion to the unknown. It often happens that in a new environment people quickly settle into a certain position. At the beginning of a conference, for example, you find a seat, either accidentally or by design, and then return to it (doggedly) at each new session, as if reassured by its familiarity. You may feel a stab of indignation to find someone in your place – that is, in an uncomfortable chair just like the others in the room. Or you return to the same table in a café, relieved to find it unoccupied. At the start of term, a class of students sit around a table, and without any compulsion or plan will often arrange themselves in the same way for the rest of the term. Like a loyal dog, habit loves what it knows, regardless of its inherent value" (Carlisle 2014, 38).

15 The relation between (repeated) exposure to certain circumstances and emotional bonds has been well documented (see Bornstein 1989 for a review of earlier literature). Recent research has explicitly addressed the more specific relation between habit and feelings of safety, confidence and comfort. One researcher sums up the results obtained as follows: "An examination of the influence of routine behaviour on people's feelings of safety, confidence, and well-being shows, as hypothesized, that these positive emotions increase with routine behaviour. (...) [R]outine enhances feelings of safety, confidence, and well-being in many aspects of everyday life" (Avni-Babad 2011, 223). See also Wood (2017, 10) and references therein. In the same vein, Wood, Quinn and Kashy claim that habitual behavior reduces the negative feelings of stress or loss of control which are associated with deliberating about a new course of action (Wood, Quinn & Kashy 2002, 1283-4; 1294-5). Moreover, Wood (2017, 8-9) presents current research on 'social interaction habits', which are elicited by other people and specific social contexts. These include habits of interacting with people belonging to different social groups. As Wood explains, the emotional substratum involved in the formation and manifestation of such habits has the potential to contribute to our understanding of social discrimination and segregation practices.

accustomed to it, she feels more comfortable or 'at home'. Perhaps she also *feels* safer, even though she has been informed that the other hotels are situated in safer areas. In this case Hannah opts for this hotel on the basis of deliberation. But this does not mean that her habit is not part of a comprehensive explanation of her act. The considerations that account for her decision can only be properly appreciated in the light of her habit.

We can further support this point by contrasting Helen's case with the case of Sophie. Like Helen, Sophie is in the habit of staying in the same hotel when visiting London. Next month Sophie is invited to a conference in London, and is about to book a room. She has been informed that the organizer would propose a hotel to all speakers, and has already decided to book at whatever hotel the organizer would suggest. Luckily for Sophie, it turns out that is the hotel she is in the habit of staying. She is relieved to book at the hotel she usually stays, but this is pure coincidence. Had the organizer proposed a different hotel, she should have booked there.

The acts of Helen and Sophie likewise ensue from deliberation. Yet Helen's case is clearly distinct from Sophie's case. Sophie's overwhelming consideration has always been to stay in the hotel proposed by the organizer. This reason is not related to her habit, so Sophie's act is merely in accidental accord with her habit. This is not the case with Helen, however. So if we insist that Helen's act was *not* a manifestation of her habit because it ensues from deliberation, we would obliterate the differences between the case of Helen and Sophie. We would have failed to take into account that Helen's habit *still* explains why she acted as she did, and hence that her act is habitual, as opposed to Sophie's. This suggests that a comprehensive account of the distinctive kind or kinds of motivation and underlie habitual acts should include the emotional bonds the development of which is part of the relevant kind of habituation. A habit does not explain an act only as a direct automatic response to the circumstances one is faced with. Hence an act may be habitual—a manifestation of a habit—even when it ensues from deliberation.

The case of Helen points to a deeper problem with the Received View. It is a recurrent theme in the history of philosophy that habits are closely related to one's character.¹⁶ When we learn about someone's habits we learn a good deal of the kind of person he is, his affections, preferences and values. And this matters to explanation. It is often the case that a habit explains an act by informing us about the agent's preferences and values, and hence, if somehow obliquely, about her character. But this kind of explanation does not imply the absence of deliberation. Suppose Antony is in the habit of going to the opera on Saturday night, and did so last Saturday. But suppose Antony had other commitments, so he had to deliberate about whether to go or not. Does that mean that mentioning his habit has no explanatory import? When we learn about Antony's habit in this context, we learn about his values and preferences; he loves going to the opera; presumably he enjoys

¹⁶ For a historical survey, see Carlisle (2006). For a recent treatment of this theme see Pollard (2011).

classical music and theatrics generally. But if Antony is that sort of person, this can hardly be irrelevant in explaining why he went to the opera last Saturday. Habit explanations often allow us to understand why an agent did something by pointing to the agent's preferences and values. (Notice that this does not exclude the operation of automaticity; but it does not imply it either). This is why learning about Antony's habit may render his act intelligible even if we know that he paused to deliberate.

I do not mean to suggest that character traits are nothing but habits. Still, the formation of a habit is often explained by a character trait, and conversely, the acquisition of a habit can contribute to the consolidation or attenuation of certain traits. For example, just as Antony's character accounts for his habit, shyness might lead a pupil to repeatedly take a certain seat in the classroom, which thereby develops into a habit, and that habit is henceforth expressive and constitutive of that trait. Conversely, one might set to cultivate healthy habits in order to combat one's unhealthy lifestyle, or try to cultivate habits of moderate behavior in order to ameliorate one's profligate character. The traditional idea of habituation does justice to this close relationship between habits and character. In Aristotelian ethics, to take a prominent example, habituation is a process that shapes, among other things, one's emotional and evaluative propensities. Habituation is the formation of character as a whole, which of course includes (and may start from) the acquisition of certain habits. But the Received View implies a very impoverished conception of habituation, as a process that merely involves the acquisition of automaticity, in the sense of doing something as a direct response to certain circumstances. This makes all the harder to account for the rich interconnections between habits and one's emotional and motivational make-up.

When we think about all the different sorts of things with respect to which we ascribe habits, we can see that we commonly assume this variety of explanatory relations between habits and acts. People would typically pause to think, even if briefly, about where to spend their holidays, about how to dress for an occasion, about which dish to cook, or about which record to play. These are simply not the sorts of things we typically do without *any* thought or deliberation, yet they are among the paradigm sorts of case that call for a habit explanation. I have suggested that habit may explain such acts to the extent that it highlights certain considerations in deliberation while it silences others. Thus by restricting habit explanations to acts that do not ensue from deliberation, the Received View unjustifiably narrows the scope of habit explanation and hence fails to do justice to the theoretical usefulness of the idea of habit.

This takes us back to Gardner's abomination of the 'lay definition' of habit. The above suggest that a more inclusive conception of habitual acts, which does not presuppose the absence of deliberation, has a role to play in the explanation and assessment of action. Perhaps there is a use of 'habit' in pre-theoretical discourse which implies 'no explanatory mechanism', in Gardner's words. But it is deeply erroneous to assume, as Gardner and many others seem to do, that the only mode of habit explanation on offer appeals to

automaticity as a direct response to the circumstances which excludes deliberation. For this implies an impoverished idea of habit, which fails to do justice to its rich explanatory potential in theoretical and pre-theoretical contexts, as well as to its role in the history of thought.

Conclusion

I have started by presenting a rudimentary characterization of the idea of habit, which consists of three necessary conditions. I have suggested that these conditions largely accord with the essential features attributed to habit in current literature. There is an important difference, however: many philosophers, as well as the overwhelming majority of psychologists, also take *automaticity* to be an essential feature of habit. And they suppose that the variety of automaticity involved in habitual acts excludes deliberation, because habitual acts are direct responses to the circumstances. This is the Received View. I have argued against this view, that the idea of a propensity to do something in certain circumstances, a propensity acquired by habituation, does not explain an act only as a direct response to the circumstances—a form of explanation that excludes deliberation. Habit has a richer explanatory potential. I have illustrated different modes of habit explanation by considering certain types of case where habit serves to explain an act while the act ensues from deliberation. And I have suggested that this variety of acquired automaticity appealed to by the Received view is *one* outgrowth of habituation that is relevant to explanation, but not the only one. The view that habit essentially involves the absence of deliberation is bound to illicitly narrow the explanatory scope and hence the theoretical usefulness of the idea of habit.

References

- Aarts H. & Dijksterhuis A. 2000a. "Habits as Knowledge Structures: Automaticity in Goal-directed Behavior." *Journal of Personality and Social Psychology* 78:53-63.
- Aarts H. & Dijksterhuis A. 2000b. "On the Automatic Activation of Goal-Directed Behavior: The Case of Travel Habit." *Journal of Environmental Psychology* 20:75-82.
- Alvarez M. 2010. *Kinds of Reasons: An Essay in the Philosophy of Action*. Oxford: Oxford University Press.
- Amaya S. 2013. "Slips." *Noûs* 47:559-576.
- Annas J. 2011a. *Intelligent Virtue*. New York: Oxford University Press.
- Annas J. 2011b. "Practical Expertise." In J. Bengson & M. A. Moffett (Eds.), *Knowing How: Essays on Knowledge, Mind, and Action*. New York: Oxford University Press (101-12).

- Anscombe G.E.M. 2000/1963. *Intention*. Cambridge, Massachusetts: Harvard University Press.
- Avni-Babad D. 2011. "Routine and Feelings of Safety, Confidence, and Well-being." *British Journal of Psychology* 102:223-244.
- Bargh J. 1997. "The Automaticity of Everyday Life." In R. Wyer (Ed.), *The Automaticity of Everyday Life*. New York: Lawrence Erlbaum (1-62).
- Bargh J. 1990. "Auto-motives: Preconscious Determinants of Social Interaction". In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of Motivation and Cognition: Foundations of Social Behavior* Vol. 2. New York: Guilford (93-130).
- Bargh J. 1994. "The Four Horsemen of Automaticity: Awareness, Intention, Efficiency, and Control in Social Cognition." In R. Wyer & T. Srull (Eds.), *Handbook of Social Cognition*. Hillsdale, New Jersey: Lawrence Erlbaum (1-40).
- Bargh J. & Gollwitzer P. M. 1994. "Environmental Control of Goal-directed Action: Automatic and Strategic Contingencies between Situations and Behavior." In W. Spaulding (Ed.), *Nebraska Symposium on Motivation* Vol. 41. Lincoln: University of Nebraska Press (71-124).
- Bargh J. 1989. "Conditional Automaticity: Varieties of Automatic Influence in Social Perception and Cognition." In J. S. Uleman and J. A. Bargh (Eds.), *Unintended Thought*. New York: The Guilford Press (3-51).
- Bornstein R. F. 1989. "Exposure and Affect: Overview and Meta-analysis of Research, 1968-1987." *Psychological Bulletin* 106:265-289.
- Brett N. 1981. "Human Habits." *Canadian Journal of Philosophy* 11:357-376.
- Carlisle C. 2005. "Creatures of habit: The Problem and the Practice of Liberation". *Continental Philosophy Review* 38:19-39.
- Carlisle C. 2014. *On Habit*. London: Routledge.
- Cooper R. & Shallice T. 2006. "Hierarchical Schemas and Goals in the Control of Sequential Behavior." *Psychological Review* 113:887-916.
- Danner U. N., Aarts H., & de Vries N. K. 2007. "Habit Formation and Multiple Means to Goal Attainment: Repeated Retrieval of Target Means Causes Inhibited Access to Competitors." *Personality and Social Psychology Bulletin* 33:1367-79.
- Dezfouli A., Lingawi N. W., & Balleine B. W. 2014. "Habits as Action Sequences: Hierarchical Action Control and Changes in Outcome Value." *Philosophical Transactions of the Royal Society B* 369:1655.
- Di Nucci E. 2013. *Mindlessness*. Newcastle upon Tyne: Cambridge Scholars Press.
- Enc B. 2003. *How We Act*. New York: Oxford: Clarendon Press.
- Fara M. 2005. "Dispositions and Habituals." *Noûs* 39:43-82.
- Gardner B. 2015. "A Review and Analysis of the Use of 'Habit' in Understanding, Predicting and Influencing Health-related Behavior." *Health Psychology Review* 9:277-95

- Gardner B., Abraham C., Lally P., & de Bruijn G. J. 2012. "Towards Parsimony in Habit Measurement: Testing the Convergent and Predictive Validity of an Automaticity Subscale of the Self-report Habit Index." *International Journal of Behavioral Nutrition and Physical Activity* 9:102.
- Gardner B., de Bruijn G. J., & Lally P. 2011. "A Systematic Review and Meta-analysis of Applications of the Self-report Habit Index to Nutrition and Physical Activity Behaviours". *Annals of Behavioral Medicine* 42:174-187.
- Gardner B. 2012. "Habit as Automaticity, Not Frequency." *European Health Psychologist* 14:32-36.
- Graybiel A. M. 2008. "Habits, Rituals, and the Evaluative Brain." *Annual Review of Neuroscience* 31:358-87.
- Hornsby J. 1980. *Actions*. London: Routledge and Kegan Paul.
- James W. 1981. *The Principles of Psychology*. Harvard University Press.
- Maddux J. E. 1997. "Habit, Health and Happiness." *Journal of Sport & Exercise Psychology* 19:331-46.
- Moors A. & De Houwer J. 2006. "Automaticity: A Conceptual and Theoretical Analysis". *Psychological Bulletin* 132:297-326.
- Moors A. & De Houwer J. 2007. "What is Automaticity? An Analysis of Its Component Features and Their Interrelations." In J. A. Bargh (Ed.), *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes*. New York: Psychology Press (11-50).
- Neal D. T. & Wood W. 2009. "Automaticity In Situ and in the Lab: The Nature of Habit in Daily Life." In E. Morsella, J. A. Bargh, & P. M. Gollwitzer (Eds.), *The Oxford Handbook of Human Action*. New York: Oxford University Press (428-42).
- Nilsen P., Roback K., Broström A., & Ellström P. E. 2012. "Creatures of Habit: Accounting for the Role of Habit in Implementation Research on Clinical Behaviour Change." *Implementation Science* 7:53.
- Norman D. 1981. "Categorization of Action Slips." *Psychological Review* 8:1-15.
- Orbell, S., & Verplanken, B. 2010. "The Automatic Component of Habit in Health Behavior: Habit As Cue-contingent Automaticity." *Health Psychology* 29:374-83.
- Ouellette J. A., & Wood W. 1998. "Habit and Intention in Everyday Life: The Multiple Processes By Which Past Behavior Predicts Future Behavior." *Psychological Bulletin* 124:54-74.
- Owens D. 2008. "Deliberation and the First Person." In A. E. Hatzimoysis (Ed.), *Self-Knowledge*. Oxford: Oxford University Press (261-277).
- Pollard B. 2006a. "Explaining Actions with Habits." *American Philosophical Quarterly* 43:57-69.
- Pollard B. 2006b. "Actions, Habits and Constitution." *Ratio* 19:229-48.
- Pollard B. 2008. *Habits in Action*. Saarbrücken: Vdm Verlag Dr. Mueller.

- Pollard B. 2011. "Identification, Psychology, and Habits." In J. Aguilar, A. Buckareff & K. Frankish (Eds.), *New Waves in Philosophy of Action*. New York: Palgrave Macmillan (81-97).
- Reason J. 1990. *Human Error*. New York: Cambridge University Press.
- Ryle G. 2009. *The Concept of Mind*. London and New York: Routledge 60th Anniversary Edition.
- van t'Riet J., Sijtsema S. J., Dagevos H., & de Bruijn G.-J. 2011. "The Importance of Habits in Eating Behavior. An Overview and Recommendations for future research." *Appetite* 57:585-96.
- Verplanken B. & Wood W. 2006. "Interventions to Break and Create Consumer Habits." *Journal of Public Policy & Marketing* 25:90-103.
- Verplanken B., Aarts H., & Van Knippenberg A. 1997. "Habit, Information Acquisition, and the Process of Making Travel Mode Choices." *European Journal of Social Psychology* 27:539-60.
- Wood W. & Neal D. T. 2007. "A New Look at Habits and the Habit-Goal Interface." *Psychological Review* 114: 843-63.
- Wood W. & Runger D. 2016. "Psychology of Habit." *Annual Review of Psychology* 37:289-314.
- Wood W., Labrecque J., Lin P., & Runger D. 2014. "Habits in Dual-Process Models." In J. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual Process Theories of the Social Mind*. New York: Guilford Press (371-85).
- Wood W., Quinn J. M., & Kashy D. A. 2002. "Habits in Everyday Life: Thought, Emotion, and Action." *Journal of Personality and Social Psychology* 83:1281-97
- Wu W. 2011. "Confronting Many-Many Problems: Attention and Agentive Control". *Nous* 45:50-76.
- Wu W. 2013. "Mental Action and the Threat of Automaticity". In A. Clark, J. Kiverstein, & T. Vierkant (Eds.), *Decomposing the Will*. New York: Oxford University Press (244-61).

Christos Douskos
(Athens)

Deliberation and Automaticity in Habitual Acts

Abstract: Most philosophers and psychologists assume that habitual acts do not ensue from deliberation, but are direct responses to the circumstances: habit essentially involves a variety of automaticity. My objective in this paper is to show that this view is unduly restrictive. A habit can explain an act in various ways. Pointing to the operation of automaticity is only one of them. I draw attention to the fact that acquired automaticity is one outgrowth of habituation that is relevant to explanation, but not the only one. Habituation shapes our emotional and motivational make up in ways that affect deliberation itself. Hence mentioning a habit might be indispensable in explaining an act which nevertheless ensues from deliberation. The view that habitual acts are direct responses to the circumstances implies an impoverished conception of habit, which fails to do justice to its rich explanatory potential in theoretical and pre-theoretical contexts, as well as to its role in the history of philosophy.

Keywords: habit; deliberation; automaticity; action explanation.

Ethics in Progress (ISSN 2084-9257). Vol. 9 (2018). No. 1, Art. #2, pp. 25-43.

Creative Commons BY-SA 3.0

Doi: 10.14746/eip.2018.1.2

Why Can't We Regard Robots As People?

André Schmiljun

ORCID: <https://orcid.org/0000-0002-9415-8495>
(Humboldt-Universität zu Berlin, schmiljun@insystems.de)

1. Introduction

Can a robot¹ be a person²? Although this question seems weird, it is not far-fetched: in spring 2018 at “I/O conference” Google relaunched a software named “Duplex”, a voice system, that sounds like a natural person, being able to arrange appointments on the telephone for its clients like booking a table in a restaurant or a hair cut in a hair salon. “Duplex” deployed human-sounding vocal cues, such as ‘ums’ and ‘ahs’ — to make the conversational experience more comfortable.³ For humans on the other end of the line, the pantomime AI was apparently very realistic.⁴ Due to applications such as this, some philosophers and scientists like Petersen, Sparrow⁵, Veruggio, Abney are already

1 It is important to differentiate at the beginning between the two notions Artificial Intelligence (AI) and robot. Following the Oxford English Dictionary, AI involves the project to develop computer systems that are able to perform tasks, normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages. A robot, on the other hand, is a machine performing certain human functions and movements automatically.

2 Compare (Schmiljun 2017a, 75) in order to read about the complex philosophical debate concerning the term person and how it is differentiated from notions like personality, people or personhood.

3 Source: <https://techcrunch.com/2018/05/10/duplex-shows-google-failing-at-ethical-and-creative-ai-design/>, July 21th 2018.

4 Of course, such an AI-based chat-box is limited to its defined application, due to it can be easily exposed when a spontaneous unexpected question comes up during conversation, for instance, if the waitress on the line suddenly changes the topic.

5 Sparrow also belongs to the optimistic group of robot scientists that believes robots will be regarded as people. In comparison to Petersen, Veruggio and Abney, he argues that robots will become people when they pass the Turing Triage Test, a test that “sets out a necessary and sufficient condition for granting moral standing to artificial intelligence. Machines will be people when we can't let them die without facing the same moral dilemma that we would when thinking about letting a human being die.” (Sparrow 2014, 305-307). Further he adds:

“(…) [M]oral standing of machines cannot be divorced from the question of the proper conditions of application of the only concepts that we possess that might allow us to recognize ‘machine persons’. Any conclusion that we wish to draw about whether or not machines might be persons or what

convinced that it is only a matter of time until the first robot will be developed that can be regarded as a person, probably also inspired by science fictional movies like *Star Trek, I Robot* (2004) or *Ex Machina* (2014).

Initially, their strategy is to show that being a person means to have required certain organizational patterns, emergent properties, like cognitive skills, (self)-consciousness, awareness or moral competences and that being a person is not a question of the particular material that happens to constitute humans (organic biology) (Veruggio & Abney 2014, 353; Petersen 2014, 284). I will demonstrate that this argument implies more philosophical problems than it ought to solve. At the start, I will briefly outline three assumptions of their argument, with the aim to deliver a current example of the first theory of person to be discussed in this paper, defining *a person as a set of (sufficient and necessary) attributes*. Afterwards, I will develop a second common theory of person that defines *a person as a class of all human beings*. I will argue in the second chapter that both theories are no useful candidates in robot ethics, as well as the first and second assumption, and that we need a new approach which I finally reveal in the last chapter.

First assumption: Petersen, Veruggio and Abney seem to favour a variety of *non-reductive physicalism*⁶ that compares human beings to a sort of modular building system of micro and macro properties. Though the macro properties are nomological connected with micro properties, they cannot be derived from knowledge of micro properties (Beckermann 2001, 221). To regard someone as a person, is nothing we can derive from his or her certain physiology or biology. "Being regarded as person" is a macro property that emerges or occurs, once we have merged the right (carbon based or artificial) components or material in a system. "Considerations suggest that biological humanity is not morally crucial if robots could attain first-person self-consciousness and deliberative agency, the usual requirements for moral responsibility and the hallmark for moral personhood" (Veruggio & Abney 2014, 354).

Second assumption: Further, they seem to be convinced that having once fully understood all biological processes of the human body, we can substitute them consequently with artificial body parts. Even human mind with its cognitive abilities and intelligence can be replaced by artificial algorithms. What appears here, is obviously an idea of *materialism*. It is an idea that suggests that our body (but even our world and universe eventually, too) is a result of smaller material particles. It is an idea that

would be required for them to become persons must be drawn from this fact, rather than from claims about empirical human psychology" (cf.). Although Sparrow's theory is more differentiated and develops an interesting argument, he also accepts the categorical identification of machines becoming persons.

⁶ The term „non-reductive physicalism" has been used in literature many times, although it's not really clear what it means. As Beckermann puts it, the term is, if understood literally, a *contradictio in adjecto*. It implies that between mental and physical properties there must be a relationship of dependency, but that mental states cannot be reduced to their physical states (Beckermann 2001, 204). It must be mentioned that all three authors don't explicitly mention that they are physicalists or non-reductive physicalists. I derive the first assumption recurring to their recently published articles.

is strongly influenced by *positivism and scientism*, thinking the world itself exists as a nomological connected whole. Everything that exists has a material origin. In the 18th century, the French philosopher Julien Offray de La Mettrie for example believed that human body and mind work like machines [L'homme machine 1748].

Der Mensch ist eine Maschine (...). Also kann man nur praktisch, oder durch einen Versuch der Zergliederung der Seele, nach Art der Aufklärung über die körperlichen Organe, ich will nicht sagen mit Sicherheit die Natur des Menschen enträtseln, aber doch wenigsten den möglichst höchstens Grad von Wahrscheinlichkeit über diesen Gegenstand erreichen (L'homme machine 1748, 12).

In fact, this group of robot scientists seems to restore a modern variety of La Mettrie's point of view. As Veruggio and Abney mention: "If we could gradually replace all of our higher brain biological functions with mechanical, robotic replacements, until we had a completely robotic brain, with no interruption in first-person self-consciousness, why would the result not qualify as a moral person, even if no longer completely a biological human?" (Veruggio & Abney 2014, 353) To continue, they believe that there is no relevant or at least mentionable difference between biological and artificial organs, concerning moral status of individuals: "And, if so, why would biological arms, or legs or digestive tract be morally crucial? Surely those humans with artificial arms, legs, and so forth, are full moral persons" (Veruggio & Abney 2014, 353-354).⁷

Third assumption: Following all three authors, a person is a set of certain attributes or properties. Abney and Veruggio, for example, believe that one necessary attribute of a person is a "deliberative system capable of agency" (Veruggio & Abney 2014, 47). A person is a being that has morality. To put it concisely, someone we regard as a person, is someone who has moral agency and is able to follow rules among other beings in society. Equipped with what Abney calls a deliberative system, a person is able to structure alternative possible futures as mental representations and can choose its actions based on which representations he or she wishes to become experienced reality (Abney 2014, 47). Following him, even emotions are not important for moral agency. As he argues, psychopaths and sociopaths don't lose their personhood and are still legally and morally responsible for their crimes although they have dysfunctional or missing emotional effect. He says: "But is the ancestral emotional system needed as well? What of hypothetical creatures that could rationally deliberate, yet lack emotions? Would they have morality? In other words – could (emotionless) robots be moral persons? Yes, they could" (Abney

⁷ Another futuristic experiment has been developed by Kevin Warwick. He argues that we can take a human brain and put it into robot body. To prove his theory, he used a neural cortex of rat fetus applying enzymes to disconnect the neurons from each other, then creating a modular closed-loop system between a 'physical' mobile robot and a cultured neuronal network, using the multielectrode array method. As he says, this method allows bidirectional communication between the culture and the robot (Warwick 2014, 319-20). Due to his first successes with the rat neurons, he concludes that we someday have the tools and technologies to develop a human brain in a robot. Again, this argument is inspired by the idea that each and every part of us as human beings can be replaced. In his case, the only thing left of a human being would be the brain. All other parts will be substituted.

2014, 47). Abneys understanding of moral competence is close to Georg Lind's theory. In his opinion, "moral competence is the ability to resolve problems and conflicts on the basis of inner moral principles through deliberation and discussion instead of violence and deceit" (Lind 2016, 13). Developing a robot with such an ability, would be a huge advantage. In many everyday situations, moral dilemma can arise. One popular example concerns the behaviour of autonomous vehicles that are the "first robots to be integrated into our lives at scale" (Lin 2017, 1), and that will likely become a standard transport solution for us in the next years. On the streets, these cars can be involved in conflicting moments, for instance, if they have to decide whether to damage people walking on the pathway or the ones sitting in the car? Lin, therefor, rightly points out that "accidents (...) will continue to happen as matter of physics, (...) even it has the best and most perfect sensors, software, and hardware the future has to offer, there'll still be the risk of crash" (Lin 2017, 2). We can never preclude that a robot car will be surprised by an unseen child who suddenly appears from behind a parked truck.

In another paper Veruggio and Abney add one more property to "their set" of requirements for personhood. They admit that a robot must attain first-person self-consciousness, too. For sure, it is not enough to construct a robot only capable of making decisions. It is important that the robot has a kind of "self, awareness" (Veruggio & Abney 2014, 354) and is able to understand the meaning of its words, in fact it must be guaranteed that the robot doesn't simulate speaking. Wittgenstein proves with his *Sprachspiele* (language games) that language requires agents who learn the meaning of words while interacting with one another. Otherwise we might have a scenario like Jean Searle illustrated it with his "Chinese room" (1984). It seems like the robot is speaking to us, although it is not understanding what its words and sentences mean. Following Petersen, being a person is a matter of having required higher "complicated organizational patterns" that allow us to regard a being as person. These patterns are a result of evolution and can occur not only in human beings, but also in any other similar intelligent subject. Thus, Petersen is convinced that even extra-terrestrial beings could fulfil specifications of personhood and might have everything it takes to be "ethically valuable" (Petersen 2014, 285).

Above-mentioned attempts belong to one – rather classical - theory that *understands a person as a set of (sufficient and necessary) attributes*. Through history of philosophy, yet, different definitions were brought up in the light of this concept.⁸ A second attempt is to *define persons as a class of all human beings*. Robert Spaemann has made this prescriptive idea prominent. He argues that each and every person belongs to a kind of community of persons ("Personengemeinschaft"), from the day they are born. Being part of it, its members receive inalienable and indispensable personal rights. He points

⁸ Again compare (Schmiljun 2017, 75-76) where I summarize briefly some psychological, philosophical concepts of "person". A recommendable extent introduction into the notion has been published by Sturma (1997).

out: “Personenrechte werden nicht verliehen und nicht zuerkannt, sondern von jedem mit gleichem Recht in Anspruch genommen“ (Spaemann 2006, 263). Further, the status of a person depends on biological affiliation to humankind. Spaemann tries to merge semantical meaning of person with the one of human beings. If a human being exists, it automatically can be regarded as person, without exception in regard of beginning and end of life. According to Spaemann, being a person is not a quality of humans. Humans are persons. “Das Sein der Person ist das Leben eines Menschen” (Spaemann 2006, 264).

In the following chapter I will demonstrate that both theories reveal definitional difficulties which doesn't make them suitable options for robot ethics. The same applies to a position of *non-reductive physicalism* and *materialism* which Veruggio, Abney and Petersen use as assumptions for their argument.

2. Why Can't Robots Be People?⁹

I believe current attempts to understand the mind by analogy with man-made computers that can perform superbly some the same external tasks as conscious beings will be recognized as a gigantic waste of time

(Nagel 1986, 16).

One challenge concerning the first theory of person is that its varied definitions never seemed to be completed, with indefinite lists of properties, which fail to give precise parameters of what a person should be. While Locke for example believes that a person is a thinking, comprehending being that has reason, reflection and can consider itself as itself (Locke 1894, 27, 9), Sturma notes that persons live their life self-determinedly, making moral decisions and following individual plans, ideas and beliefs (Sturma 1997, 348). The same uncertainty can be detected among Abney, Veruggio and Petersen who claim different qualities they think to be decisive to constitute personhood. Of course, this definitional vagueness does not really harm scientists' belief that building Artificial People will be someday possible. It is not unlike – they could reply – that robots with “strong AI” in the sense of Wallach or Searle could fulfil some of the (Wallach & Allen 2009, 57) expected qualities like a “deliberative system capable of agency” (Veruggio & Abney 2014, 47).

Philosophical difficulties only appear if we transfer their notion to beings who lack these qualities due to physical or mental diseases, deformities or stage of life. To illustrate

⁹ A prominent opponent of the theory of APs is Selmer Bringsjord who already argues in his book, published in 1992, that Artificial Intelligence might bring up many things but not a robotic person. Although robots will become maybe smarter and autonomous or probably even pass difficult versions of Turing Test, the “Person building project” will fail. As he puts it: “Roughly, *very* roughly AI will down the road give us robots (or androids) whose behaviour is dazzling, but will not give us robotic *persons*” (Bringsjord 1992, 2). Nowadays, Ewa Nowak points out that the notion “person” is too metaphysical and should rather be substituted by less conflicting terms like “agents” (Nowak 2017, 117).

my point, I will provide a short *Gedankenexperiment*: Let's assume being a person meant to be capable of deliberation and acting as fully-blown moral agent. Further, let's assume robot engineers and scientist were able to build a machine with all proper qualities, so that we could regard these new creatures as persons, thus we are confronted with some questions: what about human beings who lose their ability to deliberate rationally, for example after an accident? Will they lose their personhood, too? And what about babies? Evidence suggests that children develop agency gradually through years and are not equipped with a mature sense of morality comparable to adults when they are born. To conclude, such suggested definition – *treating a person as set of properties* – would exclude a significant group of human beings, but artificial beings, too, from personhood, the moment they lack one of these mentioned properties.

Nearly the same result appears when we take the second theory of persons into account. Spaemann's strong definition excludes any other being from personhood, even if its sharing resembling qualities like (self)-consciousness, awareness or moral competences, language. As long as this being is not human, it is impossible to be part of "Personengemeinschaft". Again, we can imagine that the same robot engineers and scientists from our first "Gedankenexperiment" had become even more advanced and had created an artificial being which exactly resembles human beings in their cognitive, psychological and behaviouristic appearance. We could even consider Petersen's scenario of "Person-o-Matic machine" that is able to produce any conceivable person "out of plastic and metal" to just about any specification. To continue, the machine can also create a person out of biomolecules, by synthesizing sequenced human-like DNA from amino acids which might be placed in homegrown cellular container, "allowing the result to gestate in an artificial uterus" (Petersen 2014, 284-285). Whereas Petersen admits that there would not be any perceivable difference between carbon-based and organic-based person, according to Spaemann, these new creatures are *per definitionem* excluded from personhood. Artificial beings, built by "Person-o-Matic machine", will become in our scenario moral agents without being regarded as persons. We would receive a two-class system of two types of species claiming the same legal rights for them. Spaemann explains his restrictive idea as follows:

Personenrechte sind Menschenrechte. Und wenn sich andere natürliche Arten im Universum finden sollten, die lebendig sind, eine empfindende Innerlichkeit besitzen und deren erwachsene Exemplare häufig über Rationalität und Selbstbewusstsein verfügen, dann müssten wir nicht nur diese, sondern alle Exemplare dieser Art ebenfalls als Personen anerkennen, also zum Beispiel möglicherweise alle Delphine (Spaemann 2006, 264).

Self-consciousness and rationality don't qualify beings to be a person. Spaemann contradicts any theoretical approach that predicates personhood on a set of conditions, because it leads to arbitrariness in regard to other species that act on a comparable human-intelligence level. If we accepted self-consciousness and rationality as *sufficient attributes for personhood*, then dolphins might be persons, too. Neither first, nor latter

theory of person provide a helpful definition in order to avoid philosophical conflicts in ethics.

Finally, to favor a position of *non-reductive physicalism* (first assumption) and *materialism* (second assumption) according to possible Artificial People, requires some metaphysical consistency. On one hand, *non-reductive physicalism* claims that we are able to understand **all** microstructures of an object or subject (whatever they are) on the grounds of scientific empirical methods and that everything existing in the world is physical (and a part of a monistic system) bound by “trans-ordinal laws” (Broad 1925, 77-78)¹⁰.

Secondly, although we fully might know all about microstructures of a system, the appearance of its macro properties like for example mental states cannot be derived from bottom-level of the system (Beckermann 2001, 221) which demands a “property dualism” (Kallestrup 2006, 459). It follows, modern robot scientists are driven by an optimistic perspective asserting that we someday gain all necessary information about human biology and may understand all physical laws so that, in a next step, we are prepared to build a machine which appears to develop same organizational patterns “(...) that [let] robot [become] a person” (Petersen 2014, 284).

But if we are not able to explain emergent properties from their microstructures, what makes us sure that we have the right constituents that will get us the desired results like a robot regarding itself as a person or a robot having mental states? And how can mental states cause physical actions and how will they be realized? Such upward explanations are limited to a certain point. Let’s say, we had all knowledge to merge components to a human-like robot, and further add, this robot was intelligent, had self-consciousness and wanted to be regarded as person, *non-reductive physicalism* does not tell us how to reduce subjectivity, emotions and thoughts or the status of being a person to a physicalistic theory.

It even allows us to conclude that macro properties are epiphenomenal and can be multiply realized, a position brought up by Jaegwon Kim¹¹. According to him, *non-reductive physicalism* is faced with two major problems: — the downward causation and causal exclusion arguments. Kim shows that mental states become effective for actions and must not only be realized by physical states. He argues that both upward and same-

10 There already many varieties of emergent theory in literature. It is C. D. Broad who provides a very clear explanation of emergent properties in his outstanding work *The Mind and Its Place in Nature* (1925): “Put in abstract terms the emergent theory asserts that there are certain wholes, composed (say) of constituents A, B, and C in a relation R to each other; that all wholes composed of constituents of the same kind as A, B, and C in relations of the same kind as R have certain characteristic properties; that A, B, and C are capable of occurring in other kinds of complex where the relation is not of the same kind as R; and that the characteristic properties of the whole R(A, B, C) cannot, even in theory, be deduced from the most complete knowledge of the properties of A, B, and C in isolation or in other wholes which are not of the form R(A, B, C). The mechanistic theory rejects the last clause of this assertion” (Broad 1925, 61).

11 Thanks to Jaegwon Kim, a differentiated perspective towards emergence theory was brought up into discussion including his popular theory of supervenience.

level causation entails downward causation (Kim 1999, 22-23)¹². For example, if I burn my hand and remove it suddenly, I do it, because I have pain, precisely, I am in a mental state of pain. On the other hand, I can describe the same situation recurring to a biological story without mentioning mental states. I remove my hand because some signals have been sent to my brain, delivering an information of alarm which again release signals to my arm muscle and let me remove my hand¹³. The question now is which explanation is providing initial cause of my withdrawal?

Kim's second argument of causal exclusion replies to this question by arguing for the possibility of mental causation, implying the dilemma for a *non-reductive physicalist* to choose between either overdetermination or an epiphenomenalism (Schlosser 2006, 139). Causal exclusion involves the idea that no normal behavior or event can have more than one sufficient cause. A behavior cannot have as its cause, both a physical and a mental event, without leading to a kind of overdetermination. A *non-reductive physicalist* now can accept this fact and can give up the position of causal closure or he considers an epiphenomenalism, thus the view that mental states are caused by physical events in the brain, but have no effect upon any physical events (Beckermann 2001, 47).

Taking this strong objection into account, *non-reductive physicalism* is a serious philosophical problem for robot science. If we assume that "being a person" with all its related "what is it like to be"- characteristics¹⁴ and its "complicated organizational patterns" – is a higher-level property of a system (not depending on a certain nature, human carbon based or artificial or even extra-terrestrial DNA), then this theory still needs to explain how property dualism can be integrated into a monistic concept of

12 Kim explains downward and upward causation more in detail which - at this point - is helpful to consider the definitional difference between both terms: "The very idea of downward causation involves vertical directionality – an 'upward' direction and a 'downward' direction. This in turn suggests an ordered hierarchy of domains that gives meaning to talk of something being located at a 'higher' or 'lower' or 'the same position in relation to another item on this hierarchy. As is familiar to everyone, positions on such a hierarchy are usually called 'levels', or sometimes 'orders' (Kim 1999, 19). Same level causation, eventually, involves causal relations between two properties. For example, one thought causes another thought (Kim 1999, 22).

13 The correct logical argument behind this example is given here: "Suppose that a property M, at a certain level L, causes another property M+, at level L + 1. Assume that M+ emerges, or results, from a property M* at level L (M* therefore is on the same level as M). Now we immediately see a tension in this situation when we ask: 'What is responsible for this occurrence of M+? What explains M+'s instantiation on this occasion?' For in this picture there initially are two competing answers: First, M+ is there because, ex hypothesis, M caused it; second, M+ is there because its emergence base M* has been realized. Given its emergence base M*, M+ must of necessity be instantiated, no matter what conditions preceded it; M* alone suffices to guarantee M+'s occurrence on this occasion, and without M*, or an appropriate alternative base, M+ could not have occurred. This apparently puts M's claim to have caused M+ in jeopardy. I believe that the only coherent description of the situation that respects M's causal claim is this: M causes M+ by causing its base condition M*. But M's causation of M* is an instance of same- level causation. This shows that upward causation entails same-level causation; that is, upward causation is possible only if same-level causation is possible" (Kim 1999, 23).

14 I quote Nagel's very prominent question here as a kind of short version of the problem he discusses in his book "What is it like to be a bat" where he outlines the difficulties to reduce the first-person view of a being into physical notions (Nagel 1974, 438). Elsewhere, Nagel asserts: "Eventually, I believe, current attempts to understand the mind by analogy of with man-made computers that can superbly some of the same external tasks as conscious beings will be recognized as a giant waste of time" (Nagel 1986, 16).

world, without risking a case of epiphenomenalism or overdetermination of mental and physical events. Maybe a strategy to solve this vicious circle is to follow Kim's own *theory of supervenience*¹⁵. He argues that higher-level properties must be reducible to lower-level properties and we can even question the difference between these two levels (Kim 1999, 47). If he was right, then mental and physical states are simply two different properties of the same entity. Thus, any change of our mental states includes a change in our physical states. But this approach again brings with it a host of new questions – some of them comparable to the identity theory of mind and brain (Stärk 2013, 32) – which cannot be outlined and answered here, due to it would extend the scope of this paper.

To return to the problem of *non-reductive physicalism*, it appears that – inspired by this theory – somehow strange ideas come up like for example Petersen's scenario of a "Person-on-Matic", a futuristic machine producing Artificial People on artificial DNA. Although creative and detailed as it is, Petersen leaves out the important question of realisation. He develops a metaphysical argument about the possibility of a precise and full reconstruction of human beings, including also the whole process of evolution.

Perhaps, after you feed [Person-on-Matic] with complete information about your DNA makeup and DNA makeup of a willing partner, the Person-on-Matic uses a combination of this information to construct from scratch a viable zygote that matures in an artificial uterus, much later producing an infant, exactly as might have been produced by the more traditional route. (...) After all, it results in people just like the people we now create by traditional means (Petersen 2014, 286).

What robot scientists like Petersen encourage here is their belief in *materialism*. Every elementary particle and atom in the universe without exception is made of material (Gabriel 2013, 16). Given this premise, they think our scientific methods will allow us to detect and analyse every microstructure and substitute it with an artificial one. Nagel argues that this view is "based ultimately on a form of idealism" (Nagel 1986, 26). Materialists believe that the only access of understanding the reality is provided by restrictive objectivity and that we can cross out subjectivity in our research. But as Nagel puts it, "we will not know exactly how scrambled eggs taste to a cockroach even if we develop a detailed phenomenology of the cockroach sense of taste" (Nagel 1986, 25). One result derived from above mentioned premise is that even mental states have to be material. Thus, emotions, thoughts and words are material entities. But do they have the same ontological status as cars, atoms or plants (Schmiljun 2017b). Above, if

¹⁵ Kim introduces this interesting theory of dualist monism named "supervenience" in his in 1993 published book "Supervenience and Mind". Kim differs two versions of supervenience: a weak and strong supervenience. "Weak supervenience (...) requires that *within* any possible world there are not be two things agreeing in B but diverging in A, and this condition is met in each of these cases" (Kim 1993, 60). To illustrate this point: For any being or object x1 and x2 (in the real world) the following applies: if x1 and x2 have the same physical properties, then they share the same mental properties. Strong supervenience is defined like that: For any being or object x1 and x2 in any possible world w1 and w2, the following applies: If x1 in w1 has the same physical properties as x2 in world w2, then x1 in w1 has the same mental properties as x2 in w2 (Beckermann 2001, 208).

really everything was material, then even the verity of theory of materialism must be a configuration of elementary particles and atoms – an objection developed by Gabriel. But not every thought is automatically true, only because it is manifested in our brain. Otherwise probably all of our thoughts, constituted in our brain, might be categorically true. Materialists need to show an explanation how such a materialistic fundament of verity or realisation can be understood, in order to justify their theory. Is verity a result of elementary particles? (Schmiljun 2017b, 133-134).

As illustrated, the argument of Artificial People is based on three conflicting assumptions (first: the idea that it is possible to define persons either as a set of properties or class of human beings and apply it to robots, second: non-reductive physicalism, third: materialism). Thus, without taking these theoretical doubts into consideration, rather weird descriptions of a possible future follow. Petersen asserts: "We should treat APs well, whether organic or inorganic, not because they could be mistaken for people, but because they *are* people (Petersen 2014, 294) In the next chapter I will draw a new approach to the debate on the ground of Christine Korsgaard that avoids metaphysical ballast.

3. Why Should Robots Rather Be Regarded As Ends in Themselves?

On the ground of Kant's transcendental philosophy, Korsgaard develops her argument why we should treat animals not as "mere means" and "instruments" for human purpose, but rather as ends in themselves.¹⁶ I assert that we can apply her argument to robots, too, once these robots are capable of pursuing their own goods. Before I will explain my point, I try to reconstruct Korsgaard's essential ideas.

The first significant assumption of her argument is that animals have agency, understood as the competence to act autonomously and efficaciously, just like any other human being¹⁷. As Korsgaard puts it:

When an animal acts, he is determined by his form, by his instincts, to produce a change in the world, guided by his conception or representation of the world. But an animal's form is what gives him his identity, what makes him the animal he is (...) Action is self-determination, and, to that extent, it is autonomous that the question of its efficacy can come up. If one thing causes another, there is no room for success or failure. But if an animal determines herself to be the cause of something, and yet does not bring that thing about, then she has failed. Autonomy and efficacy are the properties of agents – all agents, not just human agents (Korsgaard 2009, 106-107).

¹⁶ Kant asserts: „Beings the existence of which rests not on our will but on nature, if they are beings without reason, have only a relative worth, as means, and are therefore called things, whereas rational beings are called persons because their nature already marks them out as an end in itself, that is, as something that may not be used merely (...)” (*Groundwork for the Metaphysics of Moral* 4: 428).

¹⁷ Her definition matches also with the explanation, given by Oxford Dictionary, where it said that agency is an action or intervention producing a particular effect. Nevertheless, we have to differ this definition to moral agency which somehow seems to be more complex.

Equipped with agency, Korsgaard further concludes – contradicting Kant – in her second assumptions that animals are *ends in themselves*. Kant believes that we are allowed to kill animals. He advises to kill the animal quickly and without pain, and it shouldn't be for the sake of sport (Korsgaard 2018, 99). The term “ends in itself”, in the sense of Kant, addresses something that is not a mean to any purpose and provides itself a motive to pursue it without any further reason. Kant states that only human beings, persons, in the first place, can be considered as ends in themselves due to their ability to make rational choices, “in general every rational being, exists as end in itself, not merely as means to the discretionary use of this or that will, but in all its actions, those directed toward itself as well as those directed toward other rational beings, it must always at the same time be considered as an end” (Kant, Groundwork 4:428). Rational choices are the result of an assessment, based on reasons and deliberation. They deliver decisions about what should be done (Korsgaard 2012, 9). Above, rational choices have the character of a law, in the sense of an objective or principle (universal law), a value and good absolutely “to which every rational being must then be responsive” (Korsgaard 2012, 8). Whereas Kant presupposes rationality as a criterion for an end in itself¹⁸, Korsgaard suggests a second interpretation of ends in themselves. According to her, the term “end in itself” can have an *active and a passive sense*. Kant refers only to the *active sense* that something is an end in itself if it is a law-maker (capable of rational, and therefore moral choice). But human beings, as she correctly argues, are not “merely rational beings” (Korsgaard 2012, 11). She criticises that the idea that rational choice involves a presupposition that we are ends in ourselves is not the same as the idea that rational choice involves a presupposition that rational beings are ends in themselves. Of course, human beings can be non-rational, too (Korsgaard 2012, 11). Not every decision and not every choice can be traced back to rationality¹⁹. Therefore, she concludes a *passive sense* that something is an end in itself if it pursues its own good and interests (even if it is not capable of rational or moral choice). As she puts it, many things that we take to be good for us are not good for us merely insofar as we are autonomous rational beings. “Food, sex, comfort, freedom from pain and fear” affect us insofar as we are animate beings (Korsgaard 2012, 13).

Given that, Korsgaard believes animals can be ends in themselves at least in the second sense, as they are beings with interests and have (first assumption) autonomy. Animals are able to like or dislike things, be happy or suffer. They have a certain point of view that let things become either “good for” them or “bad for” them. For example, if I stop feeding my cat it will be a matter of time and she will starve. To cease feeding my cat is something that is “bad for” her. On the other hand, if I organize a scratching post,

¹⁸ Kant believes that the original nature of human beings is rational. Therefore, he concludes: “(...) rational nature exists as end in itself. The human being necessarily represents his own existence in this way” (Kant, Groundwork 4:429).

¹⁹ It rather seems like our brain operates on two different methods of making decisions. One method grounds on rational choice, the other method grounds on intuitivism and emotions. Kahnemann developed this theory in his outstanding book named *Schnelles Denken, langsames Denken* (2011).

I know this will be “good for” her, due to it helping her to take care of her claws. But acting on such reasons, I accept my cat as an autonomous being with certain interests. I value her as a being that is able to pursue her own good²⁰. As a result of this, Korsgaard is convinced that human beings owe duties to other human beings as well as to animals, in contrast to Kant who says: “As far as reason alone can judge, a human being has duties only to human being (himself and others), since his duty to any subject is moral constraint by that subject’s will” (Kant, MM 6, 442). To understand Korsgaard’s disagreement with Kant, it is helpful to consider her understanding of rational and instinctive beings. To her, *rationality* is a normative power “grounded in a certain form of self-consciousness” (Korsgaard 2018, 39) and can be seen as a distinctive characteristic of human beings²¹. Animals on the other hand are driven mostly by an *instinctive* behaviour. But Korsgaard doesn’t interpret the term as a mere reaction or movement that is wholly automatic and simply caused, not intentional. Her idea of being instinctive is connected with the ability to learn from experiences, and to solve problems by taking thought. She argues that being instinctive is compatible with being intelligent. Thus, the difference between rationality and intelligence lies in the fact that the former looks outward at the world, questions about the connections and relations like causal relations, special relations, temporal and social relations. Rationality means to look inward, “at the workings of our own minds” (Korsgaard 2018, 40). It means to be able to ask and reply normative and evaluative questions about the connections and relations to be found in the world. With this digression it becomes clear why Korsgaard highlights the value on animals as ends in themselves. As she puts it elsewhere, an animal has a kind of identity, an animal is not just only a substance. It has consciousness that gives the animal a point of view and let it become an agent that is “a subject, (...) and someone” (Korsgaard 2009, 129). To continue: “If she [the animal] is also fairly intelligent, you can interact with her, play with her, get annoyed at her, or adore her. Even if she is not very intelligent you can sympathize with her and enter into her concerns, or be hostile to her as the enemy” (Korsgaard 2009, 129).

²⁰ Korsgaard suggests that her argument could even cover plants considered as “living organisms, functioning so as to survive and reproduce” (Korsgaard 2012, 13). For example, weeds will grow better with rain, which let us assume that rain is something which is good for them. Her idea suggests that there is a dependency between the sense of “good for” and the evaluative attitudes of a being for whom things can be good or bad. She defines “evaluative attitudes” as psychological states like desires, pains, pleasures, fear, loves, hates, ambitions, projects and principles which are experienced by sensate beings (Korsgaard 2012, 13). A weakness of her argument - at least in the quoted text - is the point that Korsgaard doesn’t seem to differentiate between the autonomy of animals and other living organisms like plants. But of course, plants don’t have any brain which constitutes consciousness. A tree’s sensation of environment is obviously different to a sensation of a bat. Based on Aristotle, in “Fellow Creatures” Korsgaard provides a definition of animals in comparison to living things like plants. A living thing, Aristotle claimed, is mostly characterized by reproduction and nutrition, and are self-maintaining entities. An animal, in turn, is a living thing capable of perception and voluntary motions. “Animals maintain themselves in part by forming representations or conceptions of their environment in accordance with those representations” (Korsgaard 2004, 82-83).

²¹ She agrees with the philosophical tradition which can be traced backed to John Locke, David Hume and Francis Hutcheson: rationality is what make human beings become people (Korsgaard 2018, 42).

Dretske argues similar in his philosophical investigation on animal minds. Instead of the term “instinctive”, Dretske rather says that animals are minimal rational (Dretske 2005, 213-215). Animals are rational but not in the sense of human beings. They are rational due to they follow reasons that concern their behaviour. The difference between them and us is only that they don’t evaluate their reasons for their actions if they are good or bad. But like us, animals are equipped with “creature consciousness“, in fact, with the ability to be “conscious of things—of objects (the bug in my soup), events (the commotion in the hall), properties (the color of his tie), and facts (that someone is following me)” (Dretske 2001, 23).

Let me summarize Korsgaard’s premises and conclusions as follows: According to her, animals have agency, understood as the competence to act autonomously and efficaciously, and are beings with interest, for whom things can be good or bad. Further, she differentiates between two versions or senses of “end in itself”:

Ends in themselves are beings insofar as he or she pursues an end only if he or she thinks it is good absolutely, capable of willing his or her principles as universal law (*active sense*).

Ends in themselves are beings with interests, for whom things can be good or bad (*passive sense*).

Based on this, she concludes that animals can be understood as ends in themselves in the second sense as beings with interests, while human beings respectively people cover both senses.

As mentioned, I argue that we can apply Korsgaard’s argument to robots, too. But this requires at least another – but very crucial – assumption, namely, that robots were capable of autonomy, following their interests and good like animals someday. To put it in other words, my assumption involves a presupposition of agency. In this case, we can follow that robots might be considered as ends in themselves²². As ends in themselves

²² I admit that my argument is not as unproblematic as it might seem. This premise as well requires some philosophical assumptions, apart from the question whether robots will ever reach a level of intelligence und autonomy comparable to animals. I will avoid the notion “consciousness” here at this point, due to there is a very controversial debate whether animals are conscious or not. Some scientists like for example Searle (2005, 132), Davidson (2005, 117-118) or Dennett (2005, 402) are convinced that some animals have consciousness and even rationality. My premise assumes that robots will be able to be aware of their aims, interests, guided by their conception or representation of the world. Above, this argument involves the assumption that robots are capable of perception and voluntary motions and can act on instincts and emotions like animals. It might be objected that I probably make the same mistake as Abney, Veruggio or Petersen, assuming that designing a robot with intelligence of an animal is not unrealistic. The difference between them and this approach is that I don’t think we can build a “human sapiens 2.0” or “Artificial People”. Eric Dietrich for example believes that “moral environment of modern Earth wrought by humans, together with what current science tells us of morality (...) *morally* requires us to build our own replacement and the exist stage left” (Dietrich 2018, 531). I don’t share this euphoric forecast of possible Artificial People, due it involves some doubtful and problematic metaphysical assumption, as shown. But to sum, the weakness of my assumption is that it doesn’t precisely define sufficient and necessary conditions, to be fulfilled by a robot, so that the argument runs without “animals”. Without a concrete restriction, I believe this argument would work for any other imaginable creature that somehow meets the nature of animals, too.

we would need to regard robots' autonomy and their certain points of view. We cannot simply treat them as means for our purposes, just like Korsgaard recommends not to treat animals as mere means for our purposes. It will be a matter of fact that these robots are beings with interests, for whom things can be good or bad. Surely, this implies that I need to respect them as non-human and in-organic moral agents, even if they are not capable of rationality like us and can't ground their actions on deliberation. In her recent paper, Nowak already discusses such an idea, saying that an

artificial intelligent device [has] no biological instincts and natural ends, however, analogously to the animal, they [may someday] produce representations of the world and are provided with some laws and ends whose repeated applications, combined with learning process, may give them identity, and even some selfhood as an individual agent (Nowak 2017, 171).

I agree with her that the term *agent* is less metaphysical or spiritual than terms such as "person" or "subject"²³. Besides, Scheutz and Schermerhorn demonstrate in their work that moral agents seldomly make only rational decisions. Time, knowledge and many other resource limitations provide rather "affective evaluations", meaning fast, low-cost mechanisms for "estimating the value of an object, event, or situation for an agent", instead of complex and more computationally intensive *cognitive evaluations*" (Scheutz & Schermerhorn 2009, 74). Both scientists, of course, follow current psychological studies published by Kahneman, Wakker, and Sarin (1997) or Blaney (1986). Thus, Scheutz and Schermerhorn focus in their different experiments on the possibility to integrate affect mechanisms in social robots, in order to let them cope better with "intrinsic resource limitations of the real world" and help them to become "sensitive to human affects", when there are demanded to interact with humans [Scheutz and Schermerhorn 2009, 74]. Together with Kramer, Brick, Anderson and Dingler, they have developed a platform named DIARC – a "distributed integrated affect cognition and reflection" architecture – that should help robots to interact with humans naturally [Scheutz, Schermerhorn Kramer, Brick, Anderson and Dingler 2006, 1]. DIARC integrates cognitive abilities like natural language understanding and complex action planning and sequencing "with lower level activities (such as multi-modal perceptual processing, feature detection and tracking, and navigation and behavior coordination" (Scheutz and Schermerhorn 2009, 74-75). Having social robots acting on the ground of DIARC, it will foster their acceptance in our society. Additionally, it will help humans to regard robots as moral agents and ends in themselves.

²³ Following this interpretation, we certainly avoid mistakes to exaggerate the status of robots in our society like in the case of the social robot "Sophia" that has been produced by Hanson Robotics. Sophia has become famous through several interviews. She can move her head and neck and can gesture with her hands. She can be seen as the next evolution of a chatbot that is able to understand and mimic conversation (Hambling 2018, 198). Her example shows what happens if the abilities of robots are too much overestimated. It may lead to wrong assumptions and confusing consequences. In her case, it resulted into the doubtful political decision of the government of Saudi-Arabia to acknowledge her with the citizenship in October 2017.

4. Conclusion

Responding to the question “Will robots in the coming age (...) be persons?”, Selmer Bringsjord puts it clear in a short reply: “Robots will get very flashy, but (...) they’ll never be people” (Bringsjord 1992, 6). As shown, the modern strategy of robot scientists like Abney, Veruggio and Petersen has to tackle three major metaphysical assumptions which are rather problematic, then a profound solution for robot ethics. Firstly, we have to refrain from the idea that we can develop a precise definition of persons that can be applied to robots or used as a mean to differentiate between artificial and human beings. Secondly and thirdly, non-reductive physicalism and materialism provoke strong philosophical problems like overdetermination, epiphenomenalism or ontological coherence and coherence of variety. Thus, Korsgaard’s argument, applied to robots, given the fact that robots might have autonomy and could follow their interests someday, is maybe a reasonable attempt to encounter robots on a more neutral level. The advantage of this attempt lies obviously in the fact that robots don’t need to fulfil a complicated list of psychological, cognitive or behaviouristic conditions to be regarded as human-like Artificial People. If we regard them as ends in themselves and beings with interests, we need to take care for them the same way we need to take care for other living beings.

References

- Abney K., Veruggio G. 2014. “Roboethics: The Applied Ethics for a New Science.” In Lin P., Leith A. & Bekey G. A. (Eds.), *Robot Ethics. The Ethical and Social Implications of Robotics*. Cambridge: MIT University Press (347-64).
- Abney K. 2014. “Robotics, Ethical Theory, and Metaethics: A Guide for the Perplexed.” In Lin P., Leith A. & Bekey G. A. (Eds.), *Robot Ethics. The Ethical and Social Implications of Robotics*. Cambridge: MIT University Press (35-52).
- Beckermann, A. 2001. *Analytische Einführung in die Philosophie des Geistes*. Berlin: Walter de Gruyter.
- Bringsjord S. 1992. *What Robots Can and Can't Be?* Luxemburg: Springer Science+Business Media S.A.
- Broad C. D. 1925. *The Mind and Its Place in Nature*. New York: Harcourt, Brace & Company, Inc.
- Davidson D. 2005. “Rationale Lebewesen.” In Wild M. & Perler D. (Eds.), *Der Geist der Tiere*. Frankfurt am Main: Suhrkamp (117-31).
- Dennett D. C. 2005. „Das Bewusstsein der Tiere: Was ist wichtig und warum?“ In Wild M. & Perler D. (Eds.), *Der Geist der Tiere*. Frankfurt am Main: Suhrkamp (389-407).
- Dietrich E. 2018. “Homo Sapiens 2.0. Building the Better Robots of Our Nature.” In Anderson M. & Anderson S. L. (Eds.), *Machine Ethics*. Cambridge: University Press (531-37).

- Dretske F. 2005. „Minimale Rationalität.“ In Wild M. & Perler D. (Eds.), *Der Geist der Tiere*. Frankfurt am Main: Suhrkamp (213-222).
- Dretske F. 2001. „Animal Minds.“ *Philosophic Exchange* 31(1):21-33.
- Gabriel M. *Warum es die Welt nicht gibt?* Berlin: Ullstein.
- Hambling D. 2018. *We: Robots. The Robots That Already Rule the World*. London: Quarto Publishing.
- Kahneman D. 2011. *Schnelles Denken, Langsames Denken*. München: Pantheon.
- Kallestrup J. 2006. “The Causal Exclusion Argument.” *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 131(2):459-85.
- Kant I. 1785. *Groundwork for the Metaphysics of Moral*. Ed. and trans. by A. W. Wood. New Haven and London: Yale University Press.
- Kim J. 1993. *Supervenience and Mind*. Cambridge: Cambridge University Press.
- Kim J. 1999. “Making Sense of Emergence.” *Philosophical Studies* 95(1/2):3-36.
- Korsgaard C. M. 2012. “A Kantian Case for Animal Rights”. In Michel M., Kühne D., & Hänni J. (Eds.), *Animal Laws – Tier und Recht. Developments and Perspectives in the 21st Century*. Zürich/ St. Gallen: DIKE (1-28).
- Korsgaard C. M. 2009. *Self-Constitution. Agency. Identity and Integrity*. Oxford and New York: Oxford University Press.
- Korsgaard C. M. 2004. “Fellow Creatures: Kantian Ethics and Our Duties to Animals.” *Tanner Lectures on Human Values* 24:77-110.
- Korsgaard C. M. 2018. *Fellow Creatures: Our Obligations To the Other Animals*. Oxford: University Press.
- Lind G. 2016. *How to Teach Morality?* Berlin: Logos Verlag.
- Lin P., Leith A., & Bekey G. A. 2014. *Robot Ethics. The Ethical and Social Implications of Robotics*. Cambridge: MIT University Press.
- Lin P., Jenkis R., & Abney K. 2017 (Eds.). *Robot Ethics 2.0. From Autonomous Cars to Artificial Intelligence*. Oxford: Oxford University Press.
- Locke J. 1894. *An Essay Concerning Human Understanding*. Ed. by A. Campbell Fraser. 2 vols. Oxford: Clarendon Press.
- La Mettrie J. O 1774. *Der Mensch eine Maschine* (org. L’homme Machine). Berlin: Holzinger.
- Nagel T. 1974. “What Is It Like to Be a Bat?” *The Philosophical Review* 83(4):435-50.
- Nagel T. 1986. *The View from Nowhere*. Oxford – New York – Toronto: Oxford University Press.
- Nowak E. 2017. “Can Human and Artificial Agents Share an Autonomy, Categorical Imperative-based Ethics and ‘Moral’ Selfhood?” *Filozofia Publiczna I Edukacja Demokratyczna* 6(2):169-208. E-access: <https://pressto.amu.edu.pl/index.php/fped/article/view/13198/12903> , <https://doi.org/10.14746/fped.2017.6.2.20>

- Petersen S. 2014. "Designing People to Serve", in Lin P., Leith A., & Bekey G. A. (Eds.), *Robot Ethics. The Ethical and Social Implications of Robotics*. Cambridge: MIT University Press (283-98).
- Scheutz M. & Schermerorn P. 2009. „Affective Goal and Task Selection for Social Robots.“ In *Handbook of Research and Synthetic Emotions and Sociable Robotics: New Applications in Affective Computing and Artificial Intelligence*. Hershey: IGI Global (74-87).
- Schermerhorn P., Kramer J., Brick T., Anderson D., Dingler A., & Scheutz M. 2006. „Diarc: A Testbed for Natural Human-Robot Interactions.“ *Proceedings of AAAI 2006 Robot Workshop*.
- Searle J. R. 1984. *Minds, Brains and Science*. Cambridge, MA: Harvard University Press.
- Searle J. R. 2005. „Der Geist der Tiere.“ In Wild M. & Perler D. (Eds.), *Der Geist der Tiere*. Frankfurt am Main: Suhrkamp (132-52).
- Schlosser M. E. 2006. "Causal Exclusion and Overdetermination." In Di Nucci E. & McHugh J. (Eds.), *Content, Consciousness and Perception*. Cambridge: Cambridge Scholars Press (139-55).
- Schmiljun A. 2017a. "Robot Morality. Bertram F. Malle's Concept of Moral Competence." *Ethics in Progress* 8(2):69-79.
- Schmiljun A. 2017b. „Symbolische Formen und Sinnfelder. Probleme und Unterschiede eines gemeinsamen Projekts.“ In Hamada Y., Favuzzi P., Klattenhoff T. & Nordsieck V. (Eds.), *Symbol und Leben. Grundlinien einer Philosophie der Kultur und Gesellschaft*. Berlin: Logos (129-44).
- Spaemann R. 2006. *Personen. Versuche über den Unterschied zwischen ‚etwas‘ und ‚jemand‘*. Stuttgart: Klett-Cotta.
- Stärk J.-P. 2013. *Das Leib-Seele-Problem, die Hirnforschung und die exzentrische Positionalität*. Hamburg: Diplomica.
- Sturma D. 1997. *Philosophie der Person. Die Selbstverhältnisse von Subjektivität und Moralität*. Paderborn – München – Wien – Zürich: Mentis.
- Wallach A. & Allen C. 2009. *Moral Machines: Teaching Robots Right from Wrong*. New York: Oxford University Press.

André Schmiljun
(Berlin)

Why Can't We Regard Robots As People?

Abstract: With the development of autonomous robots, one day probably capable of speaking, thinking and learning, self-reflecting, sharing emotions, in fact, with the raise of robots becoming artificial moral agents (AMAs) robot scientists like Abney, Veruggio and Petersen are already optimistic that sooner or later we need to call those robots “people” or rather “Artificial People” (AP). The paper rejects this forecast, due to its argument based on three metaphysical conflicting assumptions. Firstly, it is the idea that it is possible to precisely define persons and apply the definition to robots or use it to differentiate human beings from robots. Further, the argument of APs favors a position of non-reductive physicalism (second assumption) and materialism (third assumption), finally producing weird convictions about future robotics. Therefore, I will suggest to follow Christine Korsgaard’s defence of animals as ends in themselves with moral standing. I will show that her argument can be transmitted to robots, too, at least to robots which are capable of pursuing their own good (even if they are not rational). Korsgaard’s interpretation of Kant delivers an option that allows us to leave out complicated metaphysical notions like “person” or “subject” in the debate, without denying robots’ status as agents.

Keywords: artificial people; moral agency; non-reductive physicalism; materialism; ends in themselves; animals.

Ethics in Progress (ISSN 2084-9257). Vol. 9 (2018). No. 1, Art. #3, pp. 44-61.

Creative Commons BY-SA 3.0

Doi: 10.14746/eip.2018.1.3

***Underclass* i kultura ubóstwa w badaniach nad rozwojem kompetencji moralnych i społecznych w środowiskach ludzi ubogich**

Karolina Kowalewska

(Spark Academy w Poznaniu, kowalewska.km@gmail.com)

Wstęp

Liczba perspektyw poznawczych proponujących różne ujęcia zjawiska ubóstwa jest imponująca, i to zarówno na gruncie filozofii społecznej, jak i socjologii, nauk ekonomicznych oraz politycznych. Perspektywy te uwzględniają elementy wyróżniane przez badaczy jako relewantne z uwagi na możliwość zdefiniowania pojęcia ubóstwa. Badacze i teoretycy rozważają bowiem odnośne pojęcie, mając na względzie takie aspekty zjawiska ubóstwa jak jego skutki, przyczyny, cechy charakterystyczne, skalę występowania oraz to, czy ma ono indywidualny, zbiorowy, bądź grupowy (obejmujący wzajemne powiązania, np. międzypokoleniowe powielanie wzorców życia i zachowania) charakter. Wśród badaczy nie ma zgody co do jednego ścisłego wyznacznika – bądź zestawu wyznaczników – które można by uznać za konstytutywne dla ubóstwa, ani też co do tego, jakie znamiona tego zjawiska uznać można za jego cechy istotowe, dystynktywne, odróżniające je od innych zjawisk. I tak też mamy do czynienia z teoriami w ujęciu absolutnym, relatywnym, wielowymiarowym, próbami ujęcia ubóstwa jako brak zdolności do funkcjonowania (Sen), łączeniem zjawiska ubóstwa z marginalizacją i wkluczeniem. Prezentowane są różne nurty, jak kulturowy wiążący ubóstwo ze specyficzną subkulturą, często przejawami patologii i eksponujący mechanizm dziedziczenia nędzy; nurt liberalny, w którym źródło ubóstwa upatrywane jest w nieskrępowanych działaniach mechanizmów rynkowych czy nurt strukturalny, koncentrujący się głównie na problemach stratyfikacji społecznej, jako główny czynnik prowadzący do ubóstwa wskazując zajmowanie niższych warstw społecznych, które blokuje i zamyka możliwości rozwoju.

Przy tak licznych sposobach konceptualizacji problemu ubóstwa, różnicach w metodologii, wynikach badawczych, definicjach i sposobie określaniu jego granic (mających wpływ na charakter i liczebność jednostek, które obejmuje kategoria ubogich),

podejmując badania zjawiska nieuchronny jest wybór ujęcia, które pozwoli zdefiniować grupę bądź kategorię osób, będącą podstawą dla planowanych analiz. W zależności od tego, czy za główną przyczynę ubóstwa przyjmujemy czynniki subiektywne (alkoholizm, niewystarczający wysiłek własny, uzależnienie od zasiłków, zachowania patologiczne itp.), czynniki losowe (np. skutki choroby, wypadków, rozwodów, klęsk żywiołowych), czy też czynniki strukturalne związane z sytuacją gospodarczą, społeczną i polityczną (brak ofert pracy, niskie dochody, silna polaryzacja dochodów, wysokie ceny, etc.), wraz ze zmianą tych kryteriów zmieni się wydatnie charakterystyka i wielkość grup kwalifikowanych jako ubogie.

Ubóstwo i związana z nim marginalizacja separują. Separacja ma zaś przeważanie charakter funkcjonalny (upośledzający funkcjonowanie człowieka), jednakże, jak pokazują badania przeprowadzone na polskich obszarach popegeerowskich, może również chodzić o separację przestrzenną, prowincjonalizację i gettoizację, co bezpośrednio przyczynia się do powstawania swoistych środowisk, ze specyficznymi kulturami, różniącymi się względem kultury głównego nurtu. Życie w środowisku wyizolowanym ogranicza możliwości rozwojowe (np. rozwój kompetencji moralno-społecznych wymagających zróżnicowanego treningu społecznego), sprzyja reprodukcji zachowań i czynników utrwalających tę sytuację, co z kolei wzmacnia mechanizm międzypokoleniowego przekazu i dziedziczenia zarówno ubóstwa, jak i wykluczenia społecznego.

W tej pracy podejmujemy próbę zanalizowania kategorii *underclass* oraz *kultury ubóstwa*, jako specyficznych paradygmatów w badaniach nad ubóstwem. Chcemy ukazać je, jako teorie, które są pomocne w analizowaniu zależności między specyfiką warunków, w jakich żyją i rozwijają się ludzie, a wpływem na posiadanie, bądź nieposiadanie stosownych kompetencji moralnych i społecznych warunkujących skuteczne funkcjonowanie we współczesnym demokratycznym społeczeństwie opartym na wiedzy. Chodzi zatem o takie teorie, które najlepiej uwypuklałyby *specyfikę stanu długotrwałego ubóstwa*, jak również swoistą *kulturową odmiennosc* reprodukowaną pokoleniowo. Teoriami, które skupiają uwagę badawczą na tych aspektach są właśnie: *teoria kultury ubóstwa* i *teoria underclass*.

Oba pojęcia „skażone są” pejoratywnym wydźwiękiem, który nadały im dyskusje polityczne w wielu krajach, marginalizując tym samym ich rolę w obszarze badań naukowych. Celem tej pracy jest wskazanie na wartościowe, z punktu widzenia badacza ubóstwa, elementy kategorii *underclass* oraz *kultury ubóstwa*. Przede wszystkim zważywszy na fakt, że rozszerzają one konceptualny zakres pojęcia biedy głównie w zakresie uwypuklenia pozadochodowego aspektu ubóstwa, dzięki czemu możliwe jest zarysowanie rzeczywistej sytuacji i szans jednostki, ze szczególnym uwzględnieniem szerokiego spektrum deprivacji, zwłaszcza w kontekście pozamaterialnego, psychospołecznego wymiaru zjawiska, uwzględniającego chociażby brak stosownych kompetencji kluczowych dla skutecznego funkcjonowania w sferze ekonomicznej, społecznej czy publicznej. Takie podejście pozwala spojrzeć na problem ubóstwa jako

wyływający nie tylko z uwarunkowań strukturalnych czy indywidualnych predyspozycji, postaw lub działań jednostki, ale także jako na zjawisko zakorzenione w kontekście, środowisku życiowym i kulturowym. To zaś daje możliwość ukazania ubóstwo jako swoistego rodzaju środowisko życia jednostek, które poprzez pewne specyficzne elementy oddziałuje na ich rozwój i sposób funkcjonowania, tym samym na możliwości zdobycia przez nie stosownych kompetencji społecznych czy kompetencji moralnych warunkujących sukces życiowy jednostek pojmowany jako wyzwolenie się z kręgu ubóstwa.

Wybór koncepcji dla badań zjawiska ubóstwa jest istotny również z praktycznego punktu widzenia, przekłada się to bezpośrednio na sposób tworzenia programów przeciwdziałania zjawisku. Wybierając koncepcje podkreślające znaczenie czynników z zakresu deprivacji potrzeb pozaekonomicznych, uzyskujemy wyniki badań, które pozwalają tworzyć programy pomocowe o szerokim spektrum oddziaływania. Nie koncentrujemy się wyłącznie na poprawie materialnego aspektu życia jednostki. Zachodzi tu znacząca reorientacja ukierunkowania zainteresowań poznawczych, działań interwencyjno-prewencyjnych oraz naprawczych. W ramach odnotowanej tutaj tendencji programy przeciwdziałania zjawisku ubóstwa mają przyczyniać się do poprawy jakości zdrowia, poziomu wykształcenia (w tym – wyposażenia w kompetencje kluczowe), a także do podnoszenia efektywności działań organizacji pomocowych i społecznych.

Zarys zagadnienia

Po raz pierwszy pojęcia *underclass* (podklasa) użył w latach 60-tych XX w. szwedzki socjo-ekonomista Gunnar Myrdal (1963). Posłużyło ono do opisu sytuacji społeczno-kulturowej mniejszości etnicznych i rasowych zamieszkujących getta na obrzeżach wielkich amerykańskich miast w czasie długotrwałego bezrobocia i szerzącej się biedy, w okresie transformacji społeczeństwa amerykańskiego. Do *underclass* Myrdal zaliczył jednostki charakteryzujące się **deprivacją warunków materialnych** (spowodowanych głównie długotrwałym bezrobociem, które z kolei wynikało ze zmian strukturalnych i technologicznych na rynku pracy, jakim nie byli w stanie sprostać ludzie z klas niższych) oraz **zmarginalizowane** w obszarze stosunków społecznych. Marginalizacja ta dotyczyła głównie niskiego poziomu edukacji¹ i dyskryminacji rasowej². Pojęcie to zyskało na popularności wśród amerykańskich badaczy społecznych³ i już w roku 1962

1 Powiązania między długotrwałym ubóstwem a niskim poziomem wykształcenia wykazali w badaniach m. in. R. Erol i I. Sawhill (1988, 316-325; por. A. Buckingham 1999; Domański 2002).

2 W przypadku Polski dyskryminacja rasowa nie jest zjawiskiem istotnym statystycznie dla zjawiska *underclass*, ze względu na niski udział mniejszości etnicznych w całej populacji kraju. Według danych z Narodowego Spisu Powszechnego Ludności i Mieszkań 2011 roku (GUS 2015), w Polsce udział mniejszości etnicznych wśród dorosłej ludności wyniósł zaledwie 3,81%. Wskaźniki te mogą jednak ulec zmianie w nadchodzących dekadach, zależnie od polskiej polityki migracyjnej.

3 Prawdziwą międzynarodową popularność pojęcie *underclass* zyskało po publikacji serii artykułów zamieszczonych w magazynie „The Times”. Znaczącą rolę odegrały wówczas artykuły

Michael Harrington wydał wpływową książkę pt. *The Other America. Poverty in United States*, w której charakteryzował *underclass* jako grupę składającą się z osób starszych, niewykwalifikowanych pracowników, samotnych matek etc. skazanych na wykluczenie bez względu na to, jak silny byłby wzrost gospodarczy, a to z racji algorytmu dystrybucji dochodów (Harrington 1962, 14-18). Po ukazaniu się publikacji Harringtona koncepcja zaczęła się silnie rozwijać, wpływały coraz to nowsze ujęcia *underclass*. Klasycy badań nad kategorią podklasy charakteryzowali ją głównie poprzez wyróżnienie szeregu cech, jakimi odznaczać się mają członkowie tej grupy.

Harrington do deprivacji materialnej i marginalizacji wskazanej przez Myrdala, jako czynnik warunkujący pojawienie się *underclass* dołożył aspekt **dziedziczenia** biedy. Dziedziczenie to odbywa się nie tylko na poziomie materialnym, ale przede wszystkim jest to dziedziczenie pozycji, strategii życiowych, stylu życia, kultury i swoistego etosu, co wykazał w swych badaniach Greg Duncan (Duncan 1997, 1998). „Tak skonfigurowane środowisko uczy, że nie warto się kształcić, że wystarczy utrzymywać się z dorywczych prac i zasiłków z pomocy społecznej. Dzieciom przekazuje się postawę bierności, izolacji od wpływu kultury dominującej” (Zabielska 2007, 73).

Ch. Murray w aspekcie dziedziczenia biedy zwrócił uwagę na rolę **uzależnienia od pomocy społecznej**. Zasiłki nazywał Murray „mylącymi zachętami”, oceniając je jako demoralizujące, gdyż stają się one źródłem dochodów niepowiązanych z pracą, co zostało przez członków *underclass* przyjęte jako wygodna strategia życiowa. Znaczy to, że zasiłek nie jest pobierany z konieczności, lecz z wyboru. Styl życia osoby bezrobotnej utrzymywanej przez państwo staje się stylem życia przekazywanym następnym pokoleniom. Takim strategiom życiowym towarzyszy załamanie się tradycyjnej moralności i zastąpienie jej przez kontrkulturę deprecjonującą i odrzucającą społeczną wartość ludzkiej pracy (Murray 2001). „Uzależnienie od pomocy społecznej, wywołuje silną potrzebę wsparcia z zewnątrz i powoduje niesamodzielność jednostek z niej korzystających, niszcząc tym samym inicjatywę do poszukiwania alternatywnego sposobu pozyskiwania środków do życia. W dłuższej perspektywie przyczynia się do stopniowej demoralizacji i całkowitego odrzucenia wartości pracy, co w efekcie może przerodzić się w tworzenie kulturowego wymiaru biedy” (Karwacki 2006, 91-92). Uzależnienie od pomocy społecznej wynika m.in. z faktu, że niskie kwalifikacje zawodowe biednych nie pozwalają na znalezienie

takich badaczy biedy jak W. J. Wilson, R. Aponte, L. Liebermann i K. Auletta. Dzięki wprowadzeniu owego pojęcia do świadomości publicznej zaczęto się nim częściej posługiwać w badaniach i dyskusjach nad problemami społecznymi w USA. Sytuacja osób z ubogich dzielnic wielkich miast zyskała na znaczeniu, a zjawisko ubóstwa otrzymało nowe oblicze. Obok klasycznych ujęć i wyznaczników biedy takich jak bezdomność, inwalidztwo i choroba, pojawiły się nowe zjawiska, tj. chroniczne bezrobocie, zorganizowana przestępczość, uzależnienia i strukturalna izolacja w postaci miejskich gett zamieszkałych głównie przez mniejszości etniczne i rasowe. Problem *underclass* przestał być dla polityków w USA problemem marginalnym, gdyż rozrastająca się *underclass*, zdominowana przez mniejszości etniczne i rasowe, nie wynikała jedynie z czynnej dyskryminacji, o czym świadczyło powstanie „czarnej” klasy średniej, ale kształtowała się w oparciu o czynniki ekonomiczne i szereg elementów kształtujących się jako efekt deprivacji majątkowej i marginalizacji społecznej tych ludzi. Podjęto próby znalezienia rozwiązania problemu podklasy, co zaowocowało wnikliwą analizą zjawiska i wyłonieniem się jego dwóch różnych ujęć, tj. kulturowego i strukturalnego.

stałego zatrudnienia, bądź też zatrudnienia satysfakcjonującego pod kątem samooceny, prestiżu, a przede wszystkim zarobków, które mogłoby konkurować z wysokością otrzymywanych zasiłków. Łatwiej jest dostosowywać swój sposób funkcjonowania do wymogów programów pomocowych poprzez np. większą liczbę nieślubnych dzieci i monopaternizm (pomoc częściej udzielana jest matkom samotnie wychowującym dzieci), niż podjąć wysiłek zabiegania o pracę zarobkową i utrzymania jej. Taka strategia życiowa wpływa jednak na dobór wartości, norm i postaw, jakie przekazywane są dzieciom, szczególnie w odniesieniu do modelu rodziny, znaczenia i wartości pracy oraz poszanowania własności.

Ralph Dahrendorf, sięgający m.in. do badań Wilsona, Nathana i Auletty, w „katalogu” cech członków *underclass* również wymienia brak klasyfikacji zawodowych, słabą więź ze środowiskami ludzi pracujących oraz bezrobocie. Poza czynnikami związanymi z pracą zarobkową, wskazuje na fakt, że wielu członków *underclass* należy do mniejszości etnicznych, żyje w niekompletnych rodzinach oraz wykazuje skłonności do zachowań aberracyjnych, czyli nadużywania alkoholu, uzależnienia od narkotyków czy działań przestępczych, jak również zależności od opieki społecznej, najczęściej trwałej, kilkupokoleniowej (Dahrendorf 1993, 237). Jako istotny czynnik warunkujący kwalifikację danej grupy ubogich do podklasy Dahrendorf wymienia izolację społeczną⁴ (ubodzy są wykluczani z istotnych obszarów życia społecznego) oraz ograniczenie kontaktów społecznych do środowiska lokalnego (Rainwater 1967, 2006).

Dahrendorf był sceptyczny wobec pojęcia **izolacji przestrzennej** jako wyróżnika podklasy, jednak część badaczy, zwłaszcza amerykańskich, włącza do analizy podklasy ten właśnie wymiar, wyodrębniając dwa główne kryteria: (1) kumulację ubóstwa (bliskość przestrzenną, wspólne zamieszkiwanie osób w podobnej sytuacji) oraz (2) izolację, tj. przestrzenną segregację tych środowisk. W przypadku klasycznego ujęcia *underclass* mowa jest głównie o wielkomiejskich gettach, charakterystycznych dla państw zachodnich, z wysokim odsetkiem imigrantów sprawiającą, że gettoizacja kojarzona jest zazwyczaj z czynnikami rasowo-etnicznymi. W krajach homogenicznych pod względem rasowym i etnicznym segregacja przestrzenna ubogich wynika głównie z przyczyn strukturalno-systemowych. W Polsce są to np. środowiska popegeerowskie.

Mimo iż nie ma zgody co do kwestii przestrzennej izolacji *underclass*, to już **pauperyzacja środowiska** życia osób trwale ubogich jest wskazywana przez większość badaczy, jako cecha istotna. To właśnie życie w spauperyzowanym środowisku w

⁴ Dahrendorf w oparciu o szereg badań empirycznych prowadzonych w USA i krajach europejskich poddaje w wątpliwość kategorię izolacji przestrzennej jako klasowego wyróżnika *underclass*. Zgadza się on jednak z R. Nathanem, który z kolei zwraca uwagę na to, iż swoista silna bariera w strefie uprawnień, która skutecznie izolują grupę trwale ubogich od reszty społeczeństwa, wystarczy do tego, by mówić o nich jako o klasie społecznej. Istniejące formalnie uprawnienia często obarczone są formalnymi ograniczeniami w dostępie dla osób trwale ubogich. Jednak ich prawdziwa niedostępność wynika głównie z form materialnej realizacji owych uprawnień przez organy i instytucje państwa, ale również z powodu braku możliwości (np. brak zdolności, wiedzy) do skorzystania z nich przez samych zainteresowanych, np. zagwarantowany jest powszechny dostęp do szkół, ale uczniowie opuszczają je bez podstawowych kompetencji takich jak zdolność czytania i pisanie.

dużym stopniu odpowiada za izolację (tak społeczną, jak i autoizolację) i wykształcenie się specyficznego stylu życia oraz strategii życiowych, nieobserwowanych w innych środowiskach. Izolacyjny charakter *underclass* prowadzi, zdaniem Dahrendorfa, do „jawnej niezdolności lub niechęci wielu ludzi do przyswajania sobie wartości otaczającego ich społeczeństwa” (Dahrendorf 1993, 243). Skutkuje to ograniczeniem kontaktu z przedstawicielami innych grup, co jeszcze silniej izoluje osoby ubogie i pogłębia ich odrębność. Członków *underclass* w większości nie interesują bieżące sprawy społeczne czy polityczne, są wyalienowani z obiegu informacji i dyskursów społecznych, wyrzuceni poza główny, uporządkowany strukturalnie i uregulowany nurt życia społecznego.

Deprywacja materialna i marginalizacja wzajemnie na siebie oddziałują. Jak zauważają William Julius Wilson (1990) i David Smith (1992), marginalizacja w dziedzinie kultury, postaw i stylu życia jest wynikiem ekskluzji, zepchnięcia tej grupy poza obwód spraw, którymi żyje większość społeczeństwa. Wilson dowodzi, że ci członkowie grup mniejszościowych, którzy zamieszkują tereny gett, znajdują się w stanie głęboko strukturalnie uwarunkowanej izolacji społecznej. Charakterystyczny dla tej grupy system zachowań jest wyrazem procesu adaptacji do warunków silnie ograniczających możliwości tych ludzi względem reszty społeczeństwa. Oddziałuje na nich „efekt koncentracji”, który sprawia, że granice między członkami podklasy a resztą społeczeństwa coraz silniej się usztywniają (Wilson 1989). Owa izolacja jako wyróżnik dla trwale ubogich, podkreślana również przez innych badaczy *underclass*, jest jednym z kluczowych elementów ograniczających możliwości rozwojowe w obszarze kompetencji, zwłaszcza kompetencji społecznych, ale również wysokorozwiniętych kompetencji moralnych, których rozwój determinowany jest poprzez różnorodność, ilość oraz jakość interakcji społecznych oraz tzw. treningów społecznych (Argyle 2002; Makowska 2012; Kohlberg 1964, 1976).

Joan Gordon, wykorzystując kategorię Mardala do swoich badań nad jakością życia nowojorskich rodzin, w pracy *The Poor of Harlem: Social Functioning in the Underclass* (1965) jako pierwsza zwróciła uwagę na **normatywne deficyty** wśród badanych, które uznawała za kolejną cechę wyróżniającą *underclass*. Silniejkwestię deficytów normatywnych podkreślali zwolennicy nurtu kulturowego. Sporo uwagi kwestii deficytów normatywnych poświęcił w późniejszym okresie również William J. Wilson (zwolennik podejścia syntetycznego do *underclass*). Opierając się na licznych badaniach przeprowadzonych na terenie ubogich dzielnic, Wilson wskazał na zwiększoną częstotliwość występowania w nich takich zjawisk jak: niezaliczanie klas w szkole, ciężce nastolatków, urodzenia pozamałżeńskie, brak starań o znalezienie pracy, zachowania dewiacyjne, odejście od norm preferowanych przez resztę społeczeństwa, dewastacja budynków (Wilson 1987). Jego zdaniem, kumulacja tych zachowań w danym środowisku sprawia, że wśród osób trwale ubogich zaczynają dominować normy postrzegane przez resztę społeczeństwa jako dewiacyjne, pojawia się też zjawisko przekształcania mainstreamowych norm społecznych we własne poprzez dostosowanie ich do wewnętrznych warunków panujących w ubogim środowisku. W efekcie stają się one w opinii reszty społeczeństwa

alternatywnymi i zgoła dewiacyjnymi wariantami norm z głównego nurtu. To zaś sprawia, że grupy spoza ubogiego obszaru zaczynają marginalizować to środowisko, np. poprzez dyskryminowanie osób z danego obszaru przy rekrutacjach zawodowych, poprzez wygłaszanie negatywnych opinii o mieszkańcach danej dzielnicy, unikanie tych miejsc itp. (Wilson 1987). W taki sposób obszar zamieszkiwany przez dużą liczbę osób ubogich staje się homogeniczny i zamknięty jednocześnie. To zaś zaczyna oddziaływać na zachowanie i postawy prezentowane przez jej mieszkańców, na tworzone w tym obszarze normy czy preferowany styl życia. Pojawia się tym samym znaczna liczba osób przejawiających zachowania i reguły praktyczne odbiegające od prezentowanych przez większość społeczeństwa (Wilson 1991).

Środowisko długotrwałej biedy ma szczególnie negatywny wpływ na dzieci, których procesy rozwojowe i socjalizacja przebiegają w wyizolowanych i obarczonych wieloma deficytami warunkach. Zdaniem Wilsona, pauperyzacja i izolacja sprawiają, że alternatywne względem głównego nurtu społeczne wzory zachowań stają się podstawą procesu socjalizacji i stanowią fundament tworzenia własnego „podkulturowego” świata. Nie są one oczywiście przez członków grup trwale ubogich postrzegane jako dewiacyjne, dlatego rozpowszechnia się je i przekazuje następnym pokoleniom jako naturalne wzory zachowań. Wilson określa je jako normy i wartości powstające w wyniku wielokrotnych transmisji racjonalnych postaw adaptacyjnych, wymuszanych poprzez konkretne sytuacje powstające w wyniku określonych zmian strukturalnych⁵. Im są wyraźniejsze i dłużej przekazywane, tym silniej wpływają na izolowanie się osób ubogich od reszty społeczności. Izolacja ta jest wzmocniana poprzez kolejne symptomatyczne dla środowiska wzory zachowań, jak np. ten opisany przez Wilsona w *The Truly Disadvantaged*. Powiada on, że kobiety z ubogich środowisk mają tendencję do wybierania partnerów z niskim wykształceniem, dotkniętych bezrobociem, należących do grupy wysokiego ryzyka, bądź zajmujących się nielegalnym zdobywaniem pieniędzy. Najczęściej nie wchodzi z nimi w formalne związki, lecz rodzą im dzieci. Posiadanie partnera, który nie jest w stanie zaspokoić potrzeb kobiety i jej dziecka, sprawia, że związek ten nie poprawia ich sytuacji. Odwrotnie, wzmocnia ubóstwo, ogranicza możliwości działania, popycha w stronę aplikacji norm i wzorów zachowań alternatywnych wobec reszty społeczeństwa, a także wzmocnia uzależnienie od pomocy z zewnątrz, np. od systemu pomocy społecznej. Czynniki, które ugruntowują negatywne położenie matki, wpływają również na niskie szanse życiowe jej dziecka, przez co bieda przekazywana jest z pokolenia na pokolenie (Wilson 1987, 57-93).

Od początku funkcjonowania pojęcia *underclass*, brak jest pełnej zgodności co do faktycznych form występowania klasy oraz rozpoznania głównej determinanty tworzenia się tej specyficznej grupy osób ubogich. W miarę tego, jak dyskusje nad zjawiskiem *underclass* nabierały coraz szerszego wymiaru, nastąpił rozłam na dwie odrębne struktury argumentacyjne. Przyczyn powstawania podklasy poszukiwano w ramach

⁵ Szerzej na ten temat w rozdziale „Podejście strukturalne do *underclass*”.

dwóch różnych nurtów: strukturalnego (akcentującego uszeregowanie ludzi ubogich na najniższym szczeblu drabiny klasowej) oraz kulturowego (wskazującego głównie na różnice we wzorach zachowań między członkami *underclass* a resztą społeczeństwa). Na bazie krytyki obu nurtów powstało także podejście syntetyczne, łączące elementy obu nurtów. Poniżej omówię pokrótce każde z nich.

Podejście strukturalne do *underclass*

Podejście strukturalne koncentruje uwagę na strukturalnych uwarunkowaniach *underclass*, które są odpowiedzialne za ekonomiczną deprywację członków tej grupy. Na arenie politycznej reprezentowane głównie przez intelektualną lewicę (partie socjaldemokratyczne w Europie i partię demokratyczną w USA). W ramach tego ujęcia jednostki traktowane są jako ofiary takich zjawisk jak wyzysk, nierówności, wykluczenie i władza. Henryk Domański w książce *Ubóstwo w społeczeństwach postkomunistycznych* wskazuje dwa podstawowe wyróżniki podejścia strukturalnego:

1. odwoływanie się do zestawu makrostrukturalnych czynników związanych z funkcjonowaniem gospodarki rynkowej;
2. wskazanie na relację między uwarunkowaniami strukturalnymi, będącymi immanentnym atrybutem współczesnego kapitalizmu, a zachowaniami i postawami zaliczanymi do syndromu „kultury ubóstwa” (Domański 2002, 106-112).

Wyjaśnienia strukturalne na znaczeniu zyskały na przełomie lat 80-tych i 90-tych XX wieku, kiedy to społeczeństwa kapitalistyczne weszły w fazę postindustrialną, tj. „etap charakteryzujący się m.in. bezzatrudnieniowym niskim wzrostem gospodarczym, kryzysem *welfare state* oraz nową strukturyzacją społeczeństw, powodowaną rozszerzającymi się obszarami bezrobocia i biedy” (Grotowska-Leder 2004, 238). Domański wyróżnia cztery podstawowe strukturalne przyczyny trwałego ubóstwa. Jako pierwszą wskazuje zmienność koniunktury gospodarki kapitalistycznej oraz segmentację rynku, czyli wyodrębnienie się peryferyjnych i kluczowych sektorów na rynku pracy, w poszczególnych branżach oraz przedsiębiorstwach. Oba czynniki wpłynęły na sytuację potencjalnych kandydatów do członkostwa podklasy. Ich niski poziom wykształcenia, brak kompetencji zawodowych i społecznych, niewykształcony odpowiednio etos pracy (długotrwałe bezrobocie rodziców) spowodowały, że znalezienie stałej pracy, nawet niskopłatnej, stało się problematyczne.

Kolejna przyczyna leży w przestrzennych uwarunkowaniach procesów wyłączenia jednostek z rynku pracy (zamykanie zakładów pracy w obrębie jakiegoś terytorium), w rezultacie czego tworzą się tzw. „kieszenie ubóstwa”, których wymownym przykładem w Polsce są tereny popegeerowskie. Obszary te zamieszkuje duża grupa osób wyłączonych z rynku pracy na skutek działania mechanizmów zewnętrznych, często w wyniku likwidacji dużych zakładów pracy.

Jako ostatni czynnik determinujący trwałą biedę (zwłaszcza w Polsce) Domański wskazuje przesunięcie wytwórczości z przemysłu do usług. Zmiana ta wywarła kolosalny wpływ na strukturę zawodową, objawiającą się stopniowym zanikaniem kategorii zawodowych charakteryzujących się niską złożonością wykonywania zadań, związanych z pracą fizyczną, co wpływa na marginalizację, a w dużej mierze także na zanik zawodów niewymagających szczególnych kwalifikacji, które tradycyjnie wykonywali reprezentanci klas niższych np. górnictwo na Śląsku (Domański 2002, 107-110).

To właśnie brak realnych dochodów jest, zdaniem Dahrendorfa, podstawową barierą, która na równi z dyskryminacją społeczną i kulturową różnicuje społeczeństwa. Realne dochody są podstawą rzeczywistego korzystania z praw obywatelskich, a te, wraz ze źródłami dochodów, osoby bezrobotne wydają się tracić bezpowrotnie. Wykształcenie i dochody z pracy to dwa rodzaje zasobów uznawane za podstawowe dla życiowego powodzenia: „kwalifikacje wyznaczają bowiem szanse na rynku pracy, a zarobki są *biletem wstępu do świata zasobów*” (Grotowska-Leder 2004, 237). Ograniczony dostęp członków *underclass* do tych przestrzeni życia społecznego i gospodarczego sprawia, że ich sytuacja ulega utrwaleniu, przez co, jak twierdzi Giddens, zaczynają oni wytwarzać „kulturę ubóstwa”, która jest formą adaptacji do zmiennych warunków egzystencji i pełnienia nowej roli społecznej. „Wpływy społeczne stwarzające podklasy mają charakter strukturalny zanim staną się kulturowe, ale kiedy już działają, mogą spowodować głęboką demoralizację kulturową” (Giddens 2001, 163).

Część badaczy uważa jednak, że to element kulturowy stanowi czynnik sprawczy zjawiska *underclass*. Badacze ci zaliczają się do tzw. nurtu kulturowego *underclass*.

Znamiona *underclass* w ujęciu kulturowym

Perspektywa i podejście kulturowe, zwane także genetycznym, behawioralnym lub patologicznym rozwinęło się w końcu lat 70-tych XX wieku. Głównymi przedstawicielami nurtu są badacze biedy o konserwatywnych poglądach, tacy jak Charles Murray, Mickey Kaus, Lawrence Mead, Ken Auletta, w Polsce zaś Elżbieta Tarkowska i Hanna Palska. W początkowym okresie badacze z tego kręgu wprost nawiązywali do kulturowych ujęć biedy opierających się na występowaniu różnic we wzorach zachowań pomiędzy członkami *underclass* a resztą społeczeństwa. Różnic tych doszukiwano się głównie w uznawanych przez członków podklasy wartościach, przejawianych zachowaniach, postawach oraz mechanizmach ich kształtowania się. Kulturę *underclass* ujmowano w kategoriach zachowań niewłaściwych, nazywanych przez Murraya *godnymi pożałowania* (Murray 2001). Badacze nurtu kulturowego jednoznacznie wskazywali na demoralizację tej grupy wynikającą przede wszystkim z deficytów w sferze norm i wartości, wytwarzanych i utrwalanych poprzez specyficzne warunki życia, niemające odpowiednika w życiu innych środowisk, a także przez nadmierną i niewłaściwie organizowaną pomoc społeczną funkcjonującą w państwach rozwiniętych, która wspomaga rozwój syndromu uzależnienia

od pomocy, funkcjonowania bez zobowiązań, braku ambicji i poczucia odpowiedzialności. Ludzie trwale ubodzy przystosowują się do warunków życia poprzez kreowanie stylu, postaw i strategii adaptacyjnych umożliwiających przeżycie, ale skutkujących takimi zjawiskami jak dezintegracja struktury rodzinnej, izolacja od reszty społeczeństwa, brak zainteresowania sprawami istotnymi dla reszty społeczeństwa, załamanie się tradycyjnej moralności, zwłaszcza w obszarze wartości ludzkiej pracy czy funkcjonowanie na pograniczu prawa (Domański 2002, 100-106).

Do *underclass* zaliczano osoby z grup niezasymilowanych z resztą społeczeństwa. Auletta w jednej z pierwszych monografii poświęconych *underclass* (1982) wyłonił cztery podstawowe grupy społeczne, które tworzą podklasę, tj. 1) ubogich, wieloletnich beneficjentów pomocy społecznej, których deprywacja materialna była wynikiem ich pasywnego podejścia do życia, lenistwa i braku aspiracji życiowych; 2) przestępców, w przeważającej części członków gangów i handlarzy narkotyków; 3) przemytników i handlarzy ulokowanych szarej strefie⁶; 4) nałogowych alkoholików, narkomanów, włóczęgów, bezdomnych, osoby starsze i chore (głównie psychicznie), których nie stać na leczenie i zaspokojenie podstawowych potrzeb życiowych (Auletta 1982). Klasyfikacja Auletty poddana została krytyce, szczególnie dlatego, że bezwarunkowo przypisał dane kategorie ludzi, jako przedstawicieli podklasy, co poczytywano za nieuprawnione⁷, jednakże pewne jego sugestie co do występowania charakterystycznych dla tych grup osób schematów behawioralnych zostały uznane za podstawowe wyznaczniki przynależności- zaczęto koncentrować uwagę na swoistych cechach i schematach działania. Można tu wskazać przede wszystkim całkowite uzależnienie potrzebujących od materialnej pomocy państwa, dużą liczbę związków o charakterze konkubinatu oraz pozamałżeńskich urodzeń, brak aktywności zawodowej zwłaszcza wśród mężczyzn oraz wysoki odsetek przestępczości. Wszystkie z wymienionych przez Auletę czynników, a zwłaszcza uzależnienie od pomocy i brak aktywności zawodowej, sprzyjają „rozpadowi rodziny nuklearnej” i powodują „socjalizację do kontrkultury, która charakteryzuje się niską wartością ludzkiej pracy i sprzyja tworzeniu się trwałych form socjalnej zależności od państwa oraz cechuje się wysokim stopniem kryminogenności” (Murray 2001, 27).

To właśnie dziedziczenie postaw i norm przez członków podklasy jest jedną z kluczowych kategorii podejścia kulturowego do *underclass*. „Jest to kluczowy punkt, albowiem zakłada się, że instytucjonalizacja, czyli utrwalenie się (...) orientacji i postaw, powoduje, że zaczynają one kształtować bieg życia jednostek w postaci samodzielnego czynnika” (Murray 2001, 103). Przedmiotem transmisji międzypokoleniowej są specyficzne normy i wartości. Karwacki podkreśla, że „mamy tu do czynienia z ukształtowaną odmienną tożsamością osoby biednej od „nie-biednej” (socjalizacja w

⁶ Te dwie kategorie również spotkały się z silną krytyką, gdyż osoby ze świata przestępczego rzadko dotyka problem ubóstwa, a już tym bardziej ubóstwa chronicznego.

⁷ Najczęściej krytycznie wskazywano na grupę handlarzy narkotyków czy przemytników, których sposób funkcjonowania trudno jest porównać np. z osobami trwale bezrobotnymi i uzależnionymi od pomocy społecznej.

warunkach/do „kultury biedy”) i poprawa warunków życia jednostek (wzrost zasobów materialnych) nie spowoduje zmian specyfiki ich świata wewnętrznego (wartości, normy, cele życiowe)” (Murray 2001, 92-93). Podejście kulturowe jest zatem powiązane z determinizmem kulturowym. Utrwalanie orientacji i postaw życiowych przez mechanizmy socjalizacji i reprodukcji kulturowej sprawia, że zaczynają one stanowić zewnętrzny względem jednostki czynnik determinujący jej życie w danym środowisku. Przystosowanie do życia w biedzie nabiera z czasem charakteru osobowościowego, przez co życie poza obszarem biedy zdaje się być tak wielkim wyzwaniem, że nawet zmiana uwarunkowań zewnętrznych, np. w postaci poprawy koniunktury gospodarczej, nie stanowi szansy na wydobycie się ze stanu ubóstwa.

Podsumowując, można stwierdzić, że podejście kulturowe do *underclass* opiera się na takich oto gruntownych założeniach:

Po pierwsze, życie w nędzy stawia ludzi w szczególnej sytuacji, niemającej odpowiednika w życiu innych środowisk. Po drugie, osoby funkcjonujące w tych specyficznych warunkach przystosowują się do trudnych warunków, wytwarzając strategie, styl i kategorię adaptacji, a także tworząc zamknięte systemy norm i wartości. Po trzecie, osoby trwale ubogie charakteryzują się niskimi kwalifikacjami zawodowymi, co obniża ich możliwości podjęcia pracy. Po czwarte, są one uzależnione od pomocy społecznej i traktują zasiłki jako formę zarobkowania. Po piąte, rezultatem tej dobrowolnej adaptacji jest formowanie się odrębnego systemu orientacji i postaw, składających się na kulturę, a w gruncie rzeczy na subkulturę ubóstwa (Domańska 2002, 102).

Podejście syntetyczne do *underclass*

Czołowym przedstawicielem nurtu syntetycznego w badaniach nad *underclass* jest William Julius Wilson. W latach 80-tych badał on powstawanie *underclass* w USA w obszarach wielkomiejskich, gdzie na skutek zmian strukturalnych, takich jak wycofywanie się przedsiębiorstw ze zdegradowanych dzielnic dużych miast, zaczęła tworzyć się podklasa. Pozostająca bez pracy, coraz silniej doświadczana skutkami deprivacji materialnej, systematycznie spychana na margines społeczny grupa zaczęła przejawiać coraz wyraźniejsze symptomy odrębności kulturowej, od negacji wartości uznawanych przez większość społeczeństwa poczynając. Wilson zaobserwował wpływ marginalizujących zmian strukturalnych na strukturę psychospołeczną grupy najsilniej dotkniętej tymi zmianami.

Dla Wilsona „kultura ubóstwa” jest odpowiedzią na strukturalne przymusy społeczne, a także generowane przez nie (bezpośrednio lub pośrednio) ograniczenie możliwości (Wilson 1989). W rezultacie można stwierdzić, że status kulturowy członków badanych wykluczonych społeczności jest, z jednej strony, zależny od towarów, usług, umiejętności i możliwości, którymi ludzie dysponują, ale z drugiej strony, towary, usługi umiejętności i możliwości także nie są wolne od wpływu czynników kulturowych (np. rasowych), które

zmniejszają szanse życiowe trwale ubogich społeczności. „W analizie podklasy równie istotne są predyspozycje kulturowe, wytworzone specyficzne dla pewnych zbiorowości normy, wartości, ale przede wszystkim postawy adaptacyjne do zmian strukturalnych w społeczeństwie” (Karwacki 2006, 98). Jest to rozwiązanie zrozumiałe, gdyż w świecie życia społecznego trudno jest ograniczać się do rozpatrywania wpływu jedynie wybranej grupy czynników na los jednostki (Wilson 1990). Dodatkowo należy uwzględnić specyfikę konkretnego społeczeństwa, gdyż w zależności od warunków typowych dla danej grupy, powstają charakterystyczne dla niej i symptomatyczne kryteria ubóstwa, co z kolei wpływa na kształtowanie się swoistych dla tego społeczeństwa grup *underclass* (Wilson jako zwolennik ujęcia relatywnego w badaniach nad ubóstwem akcentował rolę elementów specyficznych środowiska w formowaniu się podklasy i zależnych od środowiska różnic w jej funkcjonowaniu).

Krytyka pojęcia *underclass*

Underclass jako kategoria użyteczna w rozważaniach nad problemem biedy została następnie zaadaptowana w krajach europejskich i po dzień dzisiejszy stosowana jest w analizach tego zjawiska. Jednakże na Starym Kontynencie określenie *underclass* nabrało z czasem wydźwięku silnie pejoratywnego i emocjonalnego. Wpływ na to mieli głównie politycy, którzy poprzez błędne rozumienie czynników wyróżniających członków *underclass* rozpropagowali w mediach obraz grupy składającej się wyłącznie z przestępców, narkomanów i rozwiązłych kobiet w pełni odpowiadających za swoją ciężką sytuację materialną. Pojawiła się tendencja do dualistycznego zróżnicowania i demarkacji, na mocy której zarysowano ostrą linię podziału między „nami” (normalnymi) a „nimi” (zdeprawowanymi). Ludzie zmarginalizowani zaczęli być traktowani jako problem poza głównym nurtem społeczeństwa, a jego rozwiązaniem miało być po prostu „przeciąganie” wykluczonych na „naszą” stronę z pominięciem troski o przyczyny zjawiska wykluczenia, a także z pominięciem zagadnień związanych z coraz to nowymi rodzajami nierówności społecznych, które generują wykluczenie w coraz to szerszych obszarach życia. Cała strategia traktowania w kategoriach *Innego* wpływa na tzw. „strategię symbolicznego wykluczenia”, dzięki której łatwiej jest oskarżać wykluczonych o spowodowanie sytuacji, w jakiej się znaleźli. W tym celu wielu działaczy sceny politycznej wykorzystywało w sposób instrumentalny pojęcia *underclass* i *kultura ubóstwa*, zwracając uwagę głównie na te elementy charakterystyki grupy, które wskazywałyby, że jednostki same, poprzez swoje postawy, działania, zaniechania i zaniedbania, winne są sytuacji, w jakiej się znalazły. Ludzi z *underclass* opisywano jako wyalienowaną, zgettoizowaną, spatologizowaną, pełną przemocy grupę, która wegetuje na dnie społeczeństwa. Takie podejście miało wskazywać na odmienność ludzi dotkniętych biedą, aby w ten sposób pozbawić resztę społeczeństwa poczucia odpowiedzialności za ich los. Demonizowanie obu pojęć i traktowanie ich jak społecznych etykiet i stygmatów doprowadziło w wielu krajach, zwłaszcza w USA, do

zmiany polityki wobec osób dotkniętych ubóstwem na surowszą i mniej efektywną. Retoryka tych, którzy starali się przekonać opinię publiczną o zagrożeniu, jakie rzekomo stanowią członkowie *underclass*, sprawiła, że społeczeństwo odwróciło się plecami do ludzi potrzebujących pomocy, obarczając ich winą i odpowiedzialnością za sytuację, w jakiej się znaleźli, a nawet obarczając ich odpowiedzialnością za inne niekorzystne zjawiska społeczne, takie jak przestępczość czy recesja. Pojęcia zaczęto z czasem używać w potocznej retoryce pogardy i dyskryminacji wobec ludzi zaliczanych do niższych klas społecznych (Bauman 2012, 169-178). Obecnie w oficjalnym, poprawnym politycznie dyskursie społecznym pojęcia *underclass* i *kultura ubóstwa* zostały zakwestionowane jako stygmatyzujące i pejoratywne. Jednakże, obserwując debaty polityczne w sprawie biedy, zwłaszcza amerykańską i brytyjską, nietrudno zauważyć, że są one nadal obecne w świadomości społecznej, choć niekoniecznie otwarcie werbalizowane. Debata nad kategorią *underclass* i *kulturą ubóstwa* przeniosła swój punkt ciężkości na płaszczyznę badań społecznych, gdzie, odchodząc od ewaluatywnego i normatywnego ciężaru tych pojęć na gruncie etyki społecznej i teorii sprawiedliwości, stosuje się je głównie w zakresie deskryptywnym, starając się dowieść ich skuteczności w badaniach nad ubóstwem.

W Wielkiej Brytanii prace badawcze nad zjawiskiem podklasy trwały od końca lat 80-tych. Do tradycyjnego rozróżnienia na trzy klasy: wyższą, średnią i niższą Charles Murray i Lydia Morris dodali czwartą kategorię, właśnie *underclass*. Murray w roku 1989 przeniósł pojęcie podklasy na grunt brytyjski, obserwując analogiczne zjawiska w społeczeństwie brytyjskim i amerykańskim. Jednakże badania Murraya i Morris przynosiły rozbieżne rezultaty, nie dając podstaw do formułowania jednoznacznych wniosków. Stąd np. John Westergaard i Lester I. Morris (Morris 1994; Westergaard 1992) podważają zasadność tezy o istnieniu *underclass* w Wielkiej Brytanii. Zdaniem Murraya nie można jednak podważyć tego, że mamy do czynienia z realną groźbą ukształtowania się społecznej podklasy, mimo że aktualnie obserwujemy dopiero wzrost liczebności i koncentrację osób o niepewnym statusie zawodowym i szansach życiowych.

Badania nad kategorią *underclass* prowadzono także we Francji, gdzie również zwrócono szczególną uwagę na fakt koncentracji ludzi ubogich, wykluczonych, głównie mniejszości etnicznych na obrzeżach miast (pisał o tym Oscar Lewis, por. 1963; 1974). Nadal jednak obszary te nie są uważane za klasyczne kultury getta i nazwano je *underclass* raczej z rozpędu i dla teoretycznego uproszczenia.

Christopher Whelan prowadził badania nad strukturą społeczeństwa w Irlandii pod kątem krystalizacji, agregacji i wyróżnienia *underclass*. Założył, że wyznacznikami *underclass* są: długotrwałe bezrobocie, przynależność kulturowa (głównie mniejszości) i rasowa, a także segregacja przestrzenna. Opierając się na wynikach swoich badań, wskazywał, że, po pierwsze, istnieje duża grupa ludzi, którzy po stracie pracy nie wyróżniają się marginalizacją przestrzenną, a co za tym idzie, tworzeniem gett ubogich. Po drugie, badacz nie znalazł dowodów na międzypokoleniowe dziedziczenie biedy. Podobna sytuacja miała miejsce w Holandii. Nie zidentyfikowano tam jednoznacznie

istnienia tego typu klasy, zauważono jednak, że długotrwałe przebywanie na bezrobociu osób, które nie mają szansy powrotu na rynek pracy, stwarza groźbę powstania podklasy.

Jak wynika z powyższych relacji, badania nad kategorią *underclass* w różnych krajach dawały odmienne rezultaty. W wielu nie potwierdzono obecności grupy charakteryzującej się całym zestawem cech przypisywanych podklasie, jednakże wskazywano na obecność specyficznej grupy osób trwale ubogich. Dyskusja nad istnieniem podklasy jako realnego tworu w wielu krajach nie dała jednoznacznych odpowiedzi.

Krytycy *underclass* podkreślają przede wszystkim nieoczywiste założenie o odrzucaniu przez członków podklasy dominującego w danym społeczeństwie systemu wartości, zwłaszcza w odniesieniu do wartości klasy średniej. Badania Williama Ryana, Leonarda Goodwina oraz Stephena Millera i in. (Goodwin 1972; William 1971; Miller 1971) dotyczące wyboru deklarowanego systemu wartości oraz stosunku do pracy wskazały, że nie istnieje zasadnicza różnica między hierarchią wartości deklarowaną przez członków *underclass* a innymi kategoriami społecznymi, nie zaobserwowano też negatywnego stosunku podklasy do pracy. Miałoby to dowodzić bezzasadności argumentów o znaczeniu moralnym. Zwolennicy teorii kulturowej, a ponadto badacze moralności w nurcie teorii poznawczo-rozwojowych wskazują jednak na fakt, że czym innym są deklaracje ludzi (tu: ubogich), czym innym zróżnicowana wśród nich umiejętność konsekwentnej realizacji przyjmowanych i deklarowanych wartości.

Bardziej umiarkowani zwolennicy *underclass* odrzucają argumenty i oskarżenia mówiące o złej woli, świadomej walce i manifestacji wrogości *underclass* wobec klasy średniej i jej moralności oraz stylu życia. Charakterystyczny dla *underclass* styl życia, strategie i wybory mają raczej podłoże adaptacyjne (racjonalne przystosowanie do specyficznych warunków) i nie są efektem świadomego wyboru alternatywnej moralności czy zgoła amoralności.

Kolejny zarzut kierowany pod adresem zwolenników kulturowo uzasadnianej *underclass* dotyczy tendencji do formułowania nieuprawnionych uogólnień (zwłaszcza pomijających lokalne różnice w obrębie struktury i kultury społeczeństw), uproszczonych interpretacji danych i arbitralności niektórych stwierdzeń i tez.

Zdaniem Domańskiego, jednego z krytyków ujęcia kulturowego, pozytywnie należy ocenić sam fakt identyfikacji i diagnozy syndromów zachowań zaliczanych do kultury ubóstwa, których istnienia nie można zignorować. „Zachowania te pozwalają umiejscowić ludzi ubogich w strukturze społecznej, ze względu na specyficzne atrybuty tej kategorii, jakimi są, przykładowo, niskie aspiracje edukacyjne, skrócone dzieciństwo, fakt dezorganizacji rodzin, niska samoocena, poczucie bezsilności, marginalizacja, fatalizm i świadomość wyłączenia ze świata <równych>” (Domański 2002, 106). Ulf Hannerz wskazał zaś na okoliczność, że krytyka podejścia kulturowego wynika głównie z tego, że nie dokonano w niej wyraźnego rozróżnienia (albo błędnie je zinterpretowano) między przyczynami a symptomami (przejawami), między zewnętrznymi przyczynami (głównie strukturalnymi) a wytworami, analizując i objaśniając kulturowe strategie radzenia sobie z

obiektywnymi skutkami ubóstwa. Błędna interpretacja i niejasne rozróżnienia sprawiają, że pojęcie „kultury ubóstwa” nie jest stosowane jako „strategia radzenia sobie”, lecz jako „sposób na życie”, świadomie wybrany przez jednostki czy nawet całą społeczność. Dlatego Hannerz podkreśla znaczenie łączenia podejścia strukturalnego i kulturowego w badaniach nad kulturą ubóstwa, czego wyrazem jest tzw. podejście syntetyczne (Hannerz 1969).

Sami badacze posługujący się kategorią *underclass*, jak np. W. Wilson zdają sobie sprawę z niedostatków teorii podklasowej, a zwłaszcza z problemów definicyjnych (Jencks 1989) i metodologicznych, z niejasności terminologicznych wynikających ze złożoności badań, negatywnej „prasy” ze strony polityków i negatywnych emocji społecznych wobec ubogich i wykluczonych (niektórzy naukowcy pod presją opinii publicznej wycofali się z badań nad *podklasą*). Mimo iż określenie *underclass* jest niedoskonałe, to stanowi użyteczne narzędzie badań nad ubóstwem. Sugeruje ono bowiem, „że wystąpiły poważne zmiany w dzielnicach śródmiejskich gett oraz, że grupy, które tam pozostają są kolektywnie różne od tych, które zamieszkiwały te tereny wcześniej, a unikanie użycia tego terminu niefortunnie może utajnić znaczenie tych głębokich zmian jakie dokonały się w strukturze klasy i zachowaniach społecznych śródmiejskich gett” (Jencks 1989, 41). Wilson najbardziej zabiega o uznanie jakościowej odmienności *underclass*, co jest istotne w świetle walki z nowym rodzajem ubóstwa.

Brak spójności w wynikach badań nad istnieniem *podklasy* może być spowodowany m.in. brakiem zgodności terminologicznej nie tylko w obszarze pojęcie *underclass*, ale samej klasy. Dotychczas nie udało się wypracować jednoznacznego modelu, który pozwoliłby rzetelnie opracować klasowy obraz struktury społecznej. Nie ma jasnych granic między klasami, ich liczba w zależności od przyjętego paradygmatu jest zmienna, również kryteria określające przynależność klasową są zróżnicowane i płynne. Podstawę większości analiz nad stratyfikacją społeczną stanowią teorie klasy wypracowane jeszcze przez Karola Marksa (klasa w teorii konfliktu), Maxa Webera (ujęcie wielowymiarowe klasy) i Erika Olina Wrighta (klasy w teorii eksploracyjnej).

Dlatego zagadnienie klasy społecznej jest istotne z punktu widzenia dyskusji nad zjawiskiem ubóstwa. Gdy przyjmiemy założenie, że specyficzna grupa ubogich tworzy coś na kształt klasy społecznej, wówczas możemy rozważać reproduktywność pewnych wzorów zachowań, postaw, wartości, kultury, zwłaszcza w obszarze języka, które obok czynników strukturalnych będą oddziaływały na rozwój jednostki i nabywane przez nią (bądź nie) kompetencje społeczne. Należy zatem sprawdzić, czy w oparciu o najbardziej wpływowe teorie i analizy zjawiska klasy społecznej, można uznać ubogich za grupę ludzi wyróżniającą się spośród reszty społeczeństwa charakterystycznym (zależności od przyjętych kryteriów) zestawem cech.

Charakterystyka *underclass* jako klasy społecznej

Underclass można rozpatrywać jako klasę społeczną (dosłownie podklasę) w różny sposób, w zależności od przyjętej przez nas definicji klasy. Gdy przyjmujemy definicję klasy za Karolem Marksem, pojmującym ją jako grupę ludzi o podobnym stosunku do środków produkcji, wówczas ta kategoria osób (głównie z uwagi na długotrwałe pozostawanie poza rynkiem pracy i nieuczestniczenie w życiu publicznym) nie może być zakwalifikowana jako klasa społeczna. Przyczyna leży w niemożności usytuowania jej względem pozostałych klas, gdyż członkowie tej grupy są wyłączeni z posiadania, używania czy jakiegokolwiek innego udziału w środkach produkcji (Kołakowski 1988). Podobna sytuacja zachodzi, gdy zastosujemy neomarksowską koncepcję klasy według Wright (1997). Długotrwałe pozostawanie poza rynkiem pracy członków *underclass* sprawia, że nie podlegają oni kryteriom operacjonalizacji klasy w tym ujęciu (do tych kryteriów zalicza się zaś poziom wykształcenia, tytuł zawodowy i stopień autonomii wykonywanej pracy). Ze względu na pozostawanie biednych poza obrębem procesów produkcji, a co za tym idzie, brak kwalifikacji i potencjału ekonomicznego badacze tacy, jak m. in. Wright, czy Henryk Domański nie włączają osób trwale ubogich w strukturę klasową. *Underclass* jest kategorią, która nie mieści się w schemacie klasy niższej ani też nie przystaje do formuły marginesu ubóstwa.

Jednakże problem klasyfikacji osób pozostających poza systemem produkcji i wyłączonych z relacji do nich nie dotyczy jedynie ubogich i bezrobotnych. Obydwa modele klasy społecznej trudno jest odnieść do jednostek, które są ekonomicznie bierne takich jak emeryci, renciści, studenci, gospodynie domowe czy bezrobotni. Podejmuje się próby kwalifikacji tych kategorii osób do klas np. na podstawie uprzednio wykonywanego zawodu w przypadku emerytów i bezrobotnych, na podstawie kierunku studiów i prognozowanej profesji w przypadku studentów czy na podstawie zawodu wykonywanego przez męża dla ustalenia pozycji gospodyń domowych. Są to jednak starania niedające dobrych rezultatów, głównie ze względu na trudność z ujęciem wielowymiarowych czynników określających przynależność klasową w prognozowaniu (studenci) czy nawiązywaniu do poprzedniego zawodu (emeryci, bezrobotni) (Marshall 1996).

Gdy odwołamy się do klasy w ujęciu Webera, rozumianej jako grupa osób o podobnej pozycji rynkowej, ale też pokrewnym statusie społecznym (wyrażającym się głównie poprzez styl życia) oraz podobnych szansach życiowych (Weber 2002), wówczas określenie *podklasa* będzie się odnosić do najniższej pozycji, jaką zajmują osoby do niej zaliczane, jako że ich szanse życiowe są na relatywnie najniższym poziomie społecznym. W weberowskim modelu stratyfikacji społecznej dodatkowym czynnikiem wyróżniającym tę grupę spośród innych osób o niskich szansach życiowych, zdeprawowanych ekonomicznie, będzie styl życia i orientacje jednostek (implikowane przez „honor” czyli prestiż społeczny) oraz dziedziczenie i utrwalanie owego stylu przez kolejne pokolenia. Zatem ujęcie weberowskie wydaje się bardziej adekwatne od marksowskiego

i neomarksowskiego o tyle, o ile pozwala dowieść odrębności i swoistości *underclass*.

Współczesne ujęcia stratyfikacji klasowej wykraczają poza ujęcia oparte na warunkach materialnych i zatrudnieniu, obejmując czynniki kulturowe takie, jak styl życia i wzorce konsumpcyjne. Pierre Bourdieu albo Mike Savage (Bourdieu 2005; Savage 2013) przekonują, że to właśnie prezentowany styl życia, posiadany przez jednostki *kapitał kulturowy, symboliczny i ekonomiczny* są wyróżnikami klas społecznych współczesnych społeczeństw. Bourdieu postrzega klasę społeczną jako zjawisko wykraczające daleko poza konotacje ekonomiczne czy zawodowe. O przynależności do klasy społecznej decydować ma nie tyle posiadanie czy brak środków produkcji albo konkretny zawód, lecz cały zestaw społecznych cech wpływających na miejsce jednostki w strukturze społecznej. Klasa charakteryzuje się tym, że w jej skład wchodzi jednostki o podobnych praktykach życiowych, podobnym spojrzeniu na świat – posługując się językiem Bourdieu – jednostki wyróżniające się analogicznym *habitusem*. *Habitus* opisuje Bourdieu w następujący sposób:

„Struktury konstytuowane dla pewnego szczególnego typu środowiska (...) wytwarzają *habitus*, system trwałych *dyspozycji*, struktury ustrukturuwane predysponowane do działania jako struktury strukturujące, tzn. jako zasady generowania i strukturywania praktyk i wyobrażeń, które, nie będąc bynajmniej efektem poddania się reguło, mogą być obiektywnie ‚regulowane’ i ‚regularne’; nie wiążąc się ze świadomym dążeniem do określonych celów i wyraźnym opanowaniem działań niezbędnych do ich osiągnięcia, są obiektywnie dostosowane do celów, którym służą; a przy tym wszystkim, nie będąc wynikiem organizującego działania dyrygenta, mogą być kolektywnie orkiestrowane” (Bourdieu 2007, 192-193).

Habitus jest więc wytworem warunków materialnych, społecznych i historycznych, jakim podlega jednostka w danej grupie (klasie). Warunki te wpływają na rezultat oddziaływań socjalizacyjnych, na interioryzowane przez całe życie normy i wartości oraz wytwarzane dyspozycje, oceny, modele działania jednostki, które są zgodne z utrwalonymi i charakterystycznymi dla danego środowiska schematami bądź wzorami zachowań (Bourdieu 2005). W tym ujęciu charakterystyczne cechy *underclass*, tak jej wymiar strukturalny jak i kulturowy, upoważniają do uznania jej za swoistą klasę, grupę o statusie najniższym w drabinie społecznej, która poprzez specyficzne warunki funkcjonowania (*habitat* – zasoby środowiskowe i infrastruktura społeczna) wytwarza *habitus* odmienny względem dominującego czy tradycyjnego.

Teoria strukturacji według Bourdieu, skupiająca się na różnicach w posiadanym kapitale kulturowym, społecznym i symbolicznym (nie zaś tylko ekonomicznym), wskazująca na sposób życia jako efekt zestroju warunków nań wpływających, a także na kryterium przynależności grupowej, pozwala wykazać kulturową swoistość *underclass*, a także rozpatrywać sposób życia jej członków jako odmienny w stosunku do sposobu funkcjonowania innych grup społecznych.

Nurt krytyczny wobec *underclass* nie zgadza się jednak z tezą, że ludzie ubodzy, dotknięci szeregiem niekorzystnych zmian społecznych w krajach kapitalistycznych,

tworzą odrębną klasowo grupę. Krytycy powołują się przy tym na argument wewnętrznej spójności klasy społecznej, która ich zdaniem nie występuje w grupach uznawanych za *podklasowe*. Wskazują na tymczasową przynależność wielu ubogich do *podklasy*, co również narusza zasadę międzypokoleniowej i wewnątrzspołnotowej przynależności klasowej. Nie negują jednak istnienia grupy ludzi wyodrębniającej się w strukturze społecznej wskutek długotrwałego pozostawania bez pracy wraz z towarzyszącą temu deprivacją materialną i uzależnieniem od pomocy społecznej. Nie są jednak skłonni uznać, że grupa ta wypracowała odrębne względem głównego nurtu podstawy kulturowej tożsamość (Lee 1994; Gallie 1988, 1994).

Swoista odrębność trwale ubogich spowodowała, że kategoria *podklasy* była stale rozważana i wykorzystywana w wielu badaniach, głównie jako teoretycznie użyteczna definicja i podstawa do analizy złożonych procesów zachodzących w strukturze nowoczesnych społeczeństw kapitalistycznych oraz antropologicznych, socjologicznych i psychologicznych badań nad kulturowym i behawioralnym aspektem życia w ubóstwie. Wielu badaczy początkowo sceptycznych wobec pojęcia *podklasy* potwierdzało w wynikach swoich badań zasadność tych rozważań, nie tylko teoretyczną, ale też praktyczną. Za przykład posłużyć mogą tu dociekania Alana Buckinghama, który podjął rozważania nad zasadnością używania pojęcia *podklasy* w stosunku do społeczeństwa brytyjskiego. Wyniki jego badań wykazały różnorodność materialnej formy funkcjonowania *underclass* w różnych krajach; podklasa brytyjska nie odpowiadała dokładnie podklasie amerykańskiej, jednakże, zdaniem Buckinghama, nieizomorficzny charakter, tak w strukturze, jak i formach funkcjonowania grup ubogich w różnych krajach, nie powinien być pretekstem do odrzucenia owego użytecznego skądinąd pojęcia i uznania jego obecności w dyskursie fachowym za bezzasadną czy nieuprawnioną. Istotne dla Buckinghama było wykazanie, że wśród analizowanych grup obejmujących osoby trwale ubogie zaobserwować można charakterystyczne, odmienne względem klasy robotniczej, wzorce zakładania rodzin, zaangażowania w pracę, lojalności politycznej etc. Łącząc to z dowodami na międzypokoleniową ciągłość członkostwa oraz z powszechnym brakiem umiejętności zawodowych i obniżonym poziomem zdolności poznawczych, badacz wykazał, że zasadnie jest mówić o zestawie cech odróżniających tę grupę ubogich od innych grup lokowanych w dolnych warstwach struktury społecznej (Buckingham 1999).

Dyskusja nad faktycznym istnieniem *podklasy*, jej cechami i formami funkcjonowania, pozostaje nadal żywa wśród badaczy struktury społecznej. Zdaniem Duncana Gallie, w każdym społeczeństwie znajduje się odrębna grupa ludzi, która ponosi największy ciężar przemian ekonomicznych, cierpi z powodu długotrwałej marginalizacji na rynku pracy oraz doświadcza większej deprivacji niż niska klasa robotnicza. Grupa ta jest na tyle liczna, że można ją traktować jako odrębną klasę, choć sam autor powątpiewa w istnienie świadomości klasowej tej grupy (Gallie 1994).

Również w Polsce toczy się spór między krytykami *underclass* w polskim społeczeństwie a tymi, którzy dostrzegają pewne jej warianty chociażby w postaci

środków popeeagerowskich (Niedzielski & Kisiel 2001; Domański 2002; Poławski 1977; Tarkowska 1999; Wódz 1994). Sporu dotychczas nie rozstrzygnięto, a obserwując jego przebieg, można rokować, że nie nastąpi to szybko. Nie wyklucza to jednak stosowania koncepcji *underclass* jako użytecznej i funkcjonalnej kategorii heurystycznej dla różnego rodzaju badań nad zjawiskiem długotrwałego ubóstwa. Pozwala ona bowiem odróżnić jednostki borykające się z problemem deprivacji materialnej okresowo, od tych, które funkcjonują w specyficznych warunkach wykluczenia, wyróżniających się charakterystycznym systemem zachowań, uwikłanych w długotrwałą, międzypokoleniową biedę. Wielu badaczy problemu ubóstwa, jak chociażby Domański, podkreśla, że bardzo trudno jest badać zjawisko ubóstwa z powodu braku przejrzystych kryteriów wyboru grupy badawczej. Trudno jest wytyczyć granice i ramy, w których dałoby się określić, czy ktoś jest czy nie jest biedny. Stąd też tak ważną sprawą jest wybór teorii, która poprzez swoistą operacjonalizację pojęcia wyodrębnia rzeczywistą grupę osób, podlegających badaniu.

By dopełnić klaryfikację pojęciową, a także jeszcze dokładniej scharakteryzować specyficzne środowisko życia, jakim są enklawy długotrwałej biedy, przyjrzymy się jeszcze pojęciu „kultury ubóstwa”, które wykorzystywane jest przez badaczy *underclass*, jednak faktycznie stanowi odrębną teorię i orientację badawczą.

Koncepcja „kultury ubóstwa” Oscara Lewisa

Oscar Lewis, amerykański antropolog, badał portorykańską społeczność Nowego Jorku i ubogą ludność Ameryki Łacińskiej pod kątem kulturowych podstaw długotrwałego ubóstwa. Rezultaty tych badań zaprezentował po raz pierwszy w 1958 r. w trakcie International Congress of Americanists w San José, wprowadzając nową koncepcję *kultury ubóstwa* (*culture of poverty*). Według jego koncepcji „kultura ubóstwa jest to pewien sposób życia szczególnie niezmienny i trwały, przechodzący z pokolenia na pokolenie w poszczególnych rodzinach. Kultura ubóstwa ma dla tych, którzy należą do jej kręgu, właściwe, przypisane sposoby postępowania i wyróżniające społeczne i psychologiczne skutki. Jest dynamicznym czynnikiem, który bierze udział w szerszej narodowej kulturze, a sam staje się jakąś subkulturą” (Lewis 1973, 25).

Lewis wyraźnie zaznacza w swoich pracach, że subkultura ubogich tworzy się jedynie w specyficznych warunkach i nie dotyczy wszystkich osób żyjących w ubóstwie (czyli ma charakter wybiórczy). Istotna dla tworzenia się tego zjawiska jest specyfika biedy doświadczanej przez ludzi. Bieda krótkoterminowa, incydentalna, czy tzw. ubożenie nie stanowią pożywki dla rozwoju cech właściwych subkulturze ubóstwa. Lewis wskazuje na przykład na członków klasy średniej, którzy, doświadczając epizodów ubóstwa, a nawet żyjąc przez pewien czas w slumsach, za sprawą nabytych wcześniej wartości, przyswojonych norm, wyrobionej umiejętności ich przestrzegania i utrwalonego stylu życia klasy średniej, nie tworzą *kultury ubóstwa*. Podobnie inne grupy, odznaczające się poczuciem przynależności grupowej (np. religijnej i kulturowej wśród Żydów) nie

wykazują skłonności do generowania symptomatycznych dla *kultury ubóstwa* norm i postaw (Lewis 1963, 7-9).

Ustalając warunki, pod jakimi zachodzi tworzenie się tej formy kultury, Lewis wskazywał tak na czynniki strukturalne, jak na ideologiczne, psychologiczne oraz na podłoże kulturowe, z jakiego wywodzą się członkowie ubogiej społeczności. Czynniki strukturalne, na które wskazuje badacz, są podobne do tych, o których była mowa przy wstępnej charakterystyce *podklasy*: bezrobocie, izolacja społeczna, gettyzacja, niskie szanse na zdobycie zatrudnienia, niski poziom edukacji itd. Deprywacja finansowa w każdym środowisku degeneruje warunki życia i generuje podobne problemy, a tym samym, wymusza podobne systemy zachowań. Dlatego też jest to zdaniem Lewisa koncepcja o uniwersalnym zasięgu, jednak jedynie w odniesieniu do pewnego typu ubóstwa, występującego głównie w krajach rozwiniętych, nie zaś tego, z jaki możemy spotkać w krajach rozwijających się, związanego głównie z biedą na poziomie egzystencji biologicznej. *Kultura ubóstwa* tworzy się w społeczeństwach spełniających określone, sprzyjające temu zjawisku, warunki. Karwacki za Lewisem formułuje i wymienia wobec tego następujące warunki:

1. „przede wszystkim istnienie gospodarki towarowo-pieniężnej oraz pracy zarobkowej i produkcji dla zysku;
2. stale wysoki procent bezrobotnych i dorywczość zatrudnienia robotników niewykwalifikowanych;
3. niskie zarobki;
4. brak możliwości tworzenia organizacji socjalnych, politycznych i gospodarczych wśród najbiedniejszych w wyniku niemocy ludzi pod naciskiem rządowym (...);
5. istnienie w klasie panującej szeregu wartości, które kładą nacisk na pomnażanie majątku, awans społeczny i dopatrują się w kiepskiej sytuacji materialnej osobistej nieudolności i niższości” (Karwacki 2006, 99).

Na liście Karwackiego kontrowersje budzi warunek czwarty mówiący o niemożliwości tworzenia organizacji w środowiskach osób ubogich. Lewis zwracał uwagę na to, że brak udziału osób ubogich w organizacjach społecznych czy wszelkich przejawach życia publicznego jest jednym z wyróżników subkultury nędzy, jednak był on wynikiem głównie postaw i przekonań dominujących wśród jej członków, czyli czynników psychologicznych, a nie strukturalnych czy formalnych. Są to jednak ważne czynniki, istotne dla wytworzenia się *kultury ubóstwa*. Do listy Karwackiego należałoby też dodać strukturę zatrudnienia i wysokość zarobków oraz wykształcenie dorosłych przedstawicieli grupy. Podobnie zatem jak w ujęciu syntetycznym teorii *underclass* u Wilsona, *kultura ubóstwa* wytwarza się poprzez połączenie strukturalnej podstawy społecznej oraz wytwarzanych i co najważniejsze, przekazywanych pokoleniowo czynników kulturowych wpływających na postawy, styl życia i działania jej członków. „Kultura ubóstwa jest kategorią (...) determinującą sposób zachowań, hierarchie wartości

ludzi ubogich (...) określającą sposób socjalizacji kolejnych pokoleń. Kultura ubóstwa, gdy już istnieje, ma duży wpływ na dzieci” (Kraczla 1993, 9). Ma więc swój początek niejako w strukturalnych warunkach, jednak to czynnikiem psychologicznym i ideologicznym, które predysponują do rozwoju *kultury ubóstwa*, Lewis poświęcił szczególną uwagę w swojej koncepcji. Wyodrębnił on cztery poziomy analizy dla kultury ubóstwa, wyróżniając na każdym z nich charakterystyczne dla tego stanu cechy:

Po pierwsze, **poziom szerszego społeczeństwa**: *kultura nędzy* odznacza się brakiem czynnego uczestnictwa w dużych instytucjach (np. brak przynależności do partii politycznych, stowarzyszeń). Wynika to m.in. z braku środków finansowych, czasami dyskryminacji, lęku, podejrzliwości bądź apatii. Brak związków z instytucjami szerszego społeczeństwa przekłada się na brak uczestnictwa w życiu społeczności. Samo uczestnictwo w jakiejś organizacji nie jest równoznaczne z tym, że dana osoba nie przynależy do „kultury biedy”. Dodatkowo na tym poziomie „kultura biedy” cechuje się trwałym bezrobociem, dorywczością pracy, częstym zastawianiem rzeczy osobistych, pożyczaniem pieniędzy, nabywaniem głównie ubrań i rzeczy używanych. Charakterystyczna jest też krytyka zasadniczych instytucji społecznych, wyrażana np. niechęcią do policji, dygnitarzy i rządu.

Po drugie, **poziom lokalnej społeczności** – główne problemy to zmaganie się z kiepskimi warunkami mieszkaniowymi, ciasnotą, powszechna niechęć do obcych, zamykanie się w kręgu rodziny, niechęć do sąsiadów i otoczenia.

Po trzecie, **poziom rodziny** – na tym poziomie główne cechy charakteryzujące *kulturę ubóstwa* to brak dzieciństwa jako specjalnego etapu w życiu objętego szczególną ochroną, powszechny konkubinat⁸, tendencja do tworzenia rodzin matriarchalnych.

Po czwarte, **poziom psychologiczny** (indywidualnej jednostki) – główne cechy takich osób to poczucie bycia wykluczonym i zepchniętym na margines społeczeństwa, bezradność, skłonność do uzależnień, niska samoocena. „Rezygnacja, fatalizm i duża pobłażliwość w stosunku do licznych patologii to codzienność trwale ubogich” (Lewis 1973).

Kultura ubóstwa jest więc adaptacyjną (systemem przystosowania do szczególnych warunków życia) subkulturą i stylem życia charakterystycznym dla ludzi żyjących w długotrwałej, wielopokoleniowej biedzie, najczęściej zamieszkujących na przedmieściach wielkich miast (slumsy) lub zacofanych regionach wiejskich (np. byłe PGRy). Jest to zjawisko o pewnym stopniu uniwersalności, o odrębności którego decydują specyficzne stosunki społeczne, wzory zachowań, sposób życia, normy i wartości przekazywane z pokolenia na pokolenie w procesie socjalizacji. Jednostki wychowane w *kulturze ubóstwa* wyróżniają się pewnym zestawem cech, wśród których można wymienić wrogi stosunek

⁸ Ta cecha staje się coraz mniej charakterystyczna dla tych środowisk, co wiąże się ze zmianami społecznymi i podejściem do instytucji małżeństwa, jednakże nadal w tych kręgach konkubinat jest powszechny. Często wynika to z kalkulacji ekonomicznej rodziców, gdyż samotna matka/ojciec jest w stanie uzyskać wyższe świadczenia z opieki społecznej, niż osoba będąca w związku małżeńskim.

do otaczającego świata, wycofanie, orientację na terażniejszość, niski poziom aspiracji (przekładający się bezpośrednio np. na brak zaangażowania w pracę jako główne źródło utrzymania), poczucie rezygnacji i niższości, brak poczucia sprawstwa i wiary w siebie, gotowość do tolerowania zachowań, które klasa średnia uznaje za dewiacyjne oraz fatalizm.

Moralność a ubóstwo

Aby doprecyzować kontekst zjawiska ubóstwa, zaanonsowany w tytule chciałabym teraz skupić uwagę na dwóch obszarach funkcjonowania osób trwale ubogich, a mianowicie sferze moralności oraz szeroko rozumianej partycypacji. Interesują mnie tutaj w szczególności procesy przekazu, nabywania i internalizacji norm, wartości, postaw, wzorów funkcjonowania moralnego i społecznego, które również mają wpływ na wykształcenie i jakość indywidualnych kompetencji moralnych i społecznych.

Analizując pracę czołowych propagatorów koncepcji *kultury ubóstwa* (Lewis, Glazer, Moynihan, Moynihan, Ball, Miller), wyróżnić można zestaw pewnych podstawowych założeń dotyczących sfery aksjologicznej członków tej subkultury. Przytoczmy je za L. Richardem Della Fave:

- istnienie zbioru wartości charakterystycznego dla ubogich;
- wartości te są wynikiem bezpośredniego doświadczenia życia w ubóstwie;
- codzienne działania ubogich znajdują się pod silnym wpływem tych wartości;
- wiele z tych działań postrzeganych jest przez resztę społeczeństwa jako dewiacyjne, zwłaszcza z punktu widzenia klasy średniej;
- wiele z tych zachowań jest dysfunkcjonalna dla ubogich, gdyż działania przez nich podejmowane utrudniają, bądź też uniemożliwiają im podniesienie się i wyjście z ubóstwa;
- w dużej mierze wartości te przekazywane są z pokolenia na pokolenie;
- ze względu na samonapędzający się charakter tych wartości, ubodzy nie są w stanie wykorzystywać szans i możliwości na awans społeczny, nawet w sytuacji, gdy możliwości te się rozszerzają, głównie za sprawą działań instytucji społecznych (Della Fave 1974, 609-610).

Ryzykownym jest jednak wysuwać na podstawie powyższych założeń tezę o uniwersalnym systemie aksjologicznym w kulturze ubóstwa. Krytycy powołują się na wyniki wielu badań (William 1971; Rainwater 1970; Schneiderman 1964; Johnson 1971), które wykazywały, że wyznawane przez członków *kultury ubóstwa* wartości nie różnią się zbytnio od tych, jakie wyznają osoby z innych warstw społecznych. Twierdzą nawet, że ubodzy świadomie wybierają wartości przypisywane klasie średniej. Uważają je mianowicie za cenne, głównie z powodu preferowania ich przez ogół społeczeństwa. Zdaniem oponentów Lewisa dane te zaprzeczają jego teorii. Co prawda, krytycy są skłonni przyznać, że występuje szereg obszarów, w których zachowanie biednych

różni się zdecydowanie od zachowania członków klasy średniej, jednak różnice te nie są spowodowane odmiennością pod względem preferowanych wartości, ale raczej są wynikiem wysiłków podejmowanych przez biednych dla opracowania wzorców zachowań, które umożliwiają im sprostanie wymogom sytuacji ubóstwa.

Współczesny nurt badań nad wartościami deklarowanymi przez różne grupy społeczne pozwala obronić tezę o charakterystycznym dla *kultury ubóstwa* systemie aksjologicznym. Badania te rzeczywiście nie wskazują na istotne zróżnicowanie pod względem wartości deklarowanych w badaniach. Wartości deklarowane przez członków klasy niższej są zbliżone do zestawu wartości szczególnie cenionych przez klasę robotniczą, a nawet do wartości klasy średniej. Jedną z przyczyn takiego stanu jest propagowanie wartości klasy średniej jako dominującego, preferowanego przez ogół społeczeństwa i propagowanego przez instytucje (np. przez szkołę) systemu aksjologicznego. Członkowie klas niższych zdają sobie sprawę ze społecznej przewagi pewnych wartości nad innymi, dlatego też często deklarują je jako swoje. Uzyskujemy wówczas sytuację, w której osoby przejawiające dewiacyjne zachowania, np. przestępcze, w trakcie badań deklarują sprawiedliwość jako ważną dla siebie wartość. Problemem tym zajmuje się m.in. niemiecki badacz Kay Hemmerling (2009, 2013). Prowadził on wieloletnie badania nad systemem wartości osób osadzonych w więzieniach. Wykazał w nich, że deklarowane przez więźniów wartości nie są odmienne od tych, jakie zgłaszają ludzie, którzy nie popełnili przestępstw. Do podobnych wniosków doszli także inni badacze, np. Donald R. Taft (1950, 181-182), Gresham M. Sykes i David Matza (1957, 664-670). Problem nie leży zatem po stronie tego, jakie wartości deklarują badane osoby, ale odnosi się do ich kompetencji w realizowaniu i pojmowaniu danych wartości. Tu właśnie pojawia się trudność, jeżeli chodzi o osoby wychowujące się w kulturze trwałego ubóstwa. Na poziomie deklaracji wartości, grupa ta nie wyróżnia się zbyt, jednak ich świadomość i realizacja wartości są „obciążone” specyfiką środowiska życia, co skutkuje odmiennym stylem życia, wzorami zachowań, stosunkami społecznymi, poglądami oraz normami.

Teoria wartości rozszerzonych według Hymana Rodmana

Interesującą próbę wyjaśnienia tego zjawiska podjął Hyman Rodman w swojej teorii *wartości rozszerzonych* (*value stretch*) (Rodman 1963). Próbował on pogodzić dwie kontrastowe teorie dotyczące charakteru wyznawanych w społeczeństwach wartości, z których jedna reprezentowana np. przez T. Parsonsa (1969) i R. K. Mertona (2002) głosiła, że ogólny system wartości jest wspólny dla całego społeczeństwa, druga zaś, której zwolennikami byli np. Herbert H. Hyman i Allison Davis (por. Hyman 1953; Davis 1946), zakłada, że każda klasa społeczna posiada odrębny system wartości, zatem wartości są zróżnicowane ze względu na stratyfikację społeczną i uzależnione od warunków życia. Rodman założył, że w rzeczywistości obie teorie da się pogodzić, gdyż każda z nich ma swoje odbicie w rzeczywistości. Podstawowe wartości są jego zdaniem podzielane przez

członków wszystkich klas, także przez członków klasy niższej. Jednocześnie, jako wynik reakcji na swoiste, ograniczające możliwości warunki funkcjonowania osób ubogich, wyróżnia się w tej grupie specyficzne wartości, charakterystyczne tylko dla niej. Rodman założył zatem, że obie teorie są względem siebie komplementarne.

Główna teza koncepcji *value stretch* brzmi: osoby z klasy niższej bez rezygnacji z ogólnych wartości (głównie reprezentowanych przez wartości klasy średniej) rozwijają własny, alternatywny ich zestaw, który zawiera szereg wartości niższych, łatwiejszych do realizacji. Efektem tego jest posiadanie przez członków klasy niższej szerszego zestawu wartości niż posiada pozostała część społeczeństwa (Rodman 1963, 209). Koncepcja Rodmana opiera się na następujących założeniach:

- Ogólne wartości społeczne są podzielane przez wszystkich członków społeczeństwa, wszystkie klasy społeczne.
- Członkowie klasy niższej posiadają wartości prezentujące niższy poziom osiągnięć niż ogólne wartości społeczne.
- Członkowie klasy niższej mają szerszy zakres wartości niż członkowie wyższych klas.
- Członkowie klasy niższej poprzez ograniczenia, jakie stwarzają warunki, w których żyją, często nie są zdolni realizować wartości z poziomu klasy średniej.
- Członkowie klasy niższej mają świadomość, że ograniczenia te dotyczą również innych osób z ich środowiska.
- Ludzie mają tendencję do faworyzowania tych rzeczy i celów, które są dla nich możliwe do osiągnięcia. Równają w dół, obniżają swoje cele, gdy przynosi to rezultat w postaci poczucia osiągnięcia sukcesu, realizacji celów, poczucia spełnienia, bądź redukcji dyskomfortu psychicznego⁹.
- Członkowie klasy niższej z czasem zaczynają szczególnie pozytywnie oceniać te osiągnięcia, które są dla nich dostępne, nawet jeżeli są one lokowane (przez resztę społeczeństwa) nisko w ogólnej hierarchii wartości.
- Ze względu na powyższy mechanizm, członkowie klasy niższej mają tendencję do rozszerzania zakresu wyznawanych wartości (Della Fave 1974, 612-613).

W związku z powyższymi założeniami Rodman przyjął, że w klasie niższej koegzystują wartości szerokiego społeczeństwa (głównie wartości klasy średniej) oraz wartości charakterystyczne dla klasy niższej. Ważny jest tutaj nie tylko sam dodatkowy ich zestaw, ale również ich hierarchia, to jak układają się względem siebie i jak na siebie wpływają. Stąd też nazwa teorii, tj. „poszerzenie wartości”¹⁰, jako że zestaw wartości w

⁹ Sam Rodman zauważa jednak, że nie dotyczy to wszystkich osób z kultury biedy. Złożoność ludzkich zachowań i reakcji jest tak wielka, że nie można w tym przypadku traktować takiego zachowania jako uniwersalnego dla wszystkich członków *underclass*, jednak Rodman twierdzi, że jest to jedna z najczęstszych reakcji adaptacyjnych członków podklasy na brak zdolności do realizacji ogólnych wartości społecznych.

¹⁰ Wartości rozumiane są tu jako system, a nie pojedyncza wartość. Rozszerzeniu ulega więc zbiór wartości, choć pod dyskusję poddać można również hipotezę, o rozszerzaniu znaczenia poszczególnych wartości, szczególnie w przypadku wartości abstrakcyjnych, jednakże takie

klasie niższej będzie szerszy niż pozostałych klasach, właśnie ze względu na dodatkowy, specyficzny dla tej klasy segment wartości, których nie dzielają pozostałe klasy.

Rodman opisał również proces, w wyniku którego dochodzi do poszerzania zestawu wartości w klasie niższej. Zrekonstruował go w oparciu o analizę wyników badań i obserwacji członków klas niższych z różnych krajów. Zauważył, że osoby z klasy niższej wchodzi z sobą w częstsze relacje niż z przedstawicielami innych klas¹¹. Mają podobne problemy, funkcjonują w tych samych warunkach, które w sposób istotny ograniczają ich możliwości, przez co bardzo trudno jest im osiągnąć kulturalne cele reszty społeczeństwa przy pomocy uznanych i dopuszczalnych środków. Potrzeba realizacji owych kulturalnych celów jest silna, ponieważ w oparciu o nie społeczeństwo dokonuje oceny jednostek, nagradza bądź karze za sposób ich realizacji. Potrzeba ta jest tak silna, że członkowie klasy niższej pozbawieni dostępu do „legalnych” środków służących do realizacji celów, sięgają po inne, alternatywne, które najczęściej są negatywnie oceniane przez społeczeństwo. Inną reakcją na niemożność zrealizowania kulturowych celów ogółu społeczeństwa jest negacja tych celów. Gdy członkowie klasy niższej spotykają się z negacją, karą ze strony reszty społeczeństwa, wywołuje to niechęć ubogich do tej części społeczności, rodzi frustracje, poczucie niższości, dyskomfort psychiczny. Sytuacje takie sprawiają, że tradycyjne wartości stają się źródłem ich negatywnych odczuć i mimo że z nich nie rezygnują, to coraz częściej zaczynają tolerować, ostatecznie zaś – aprobować, a nawet pozytywnie oceniać odchylenia od takiego sposobu wartościowania. Tak właśnie, w wyniku działań adaptacyjnych zaczynają wytwarzać takie wartości, które są łatwiejsze do osiągnięcia, nie powodują frustracji i poczucia niezdolności do osiągania celów. Dochodzą zatem cele, których realizacja wymaga zasobów i kompetencji, jakie są w posiadaniu członków *underclass*, dlatego też stają się one coraz bardziej popularne, akceptowane przez coraz większą liczbę osób z tego środowiska. Z czasem stają się charakterystyczne dla danej grupy ubogich (Rodman 1963).

Rodman w swojej teorii przyjmuje również dodatkowe założenie, że członkowie tych grup wzmacniają swoją pozytywną ocenę niższych wartości poprzez wchodzenie w interakcje z innymi, którzy są w podobnej sytuacji, przez co są postrzegani jako pozytywna grupa odniesienia. Zatem, w przypadku członków klasy niższej, mamy do czynienia z tzw. równaniem w dół, co pokazują również badania prowadzone w polskich środowiskach popegeerowskich (Niedzielski 2001; Karwacki 2006; Tarkowska 2000; Korzeniowska 2002). Niejednokrotnie badani przyznają, że z punktu widzenia integracji społecznej lepiej jest nie posiadać zbyt wiele, a ci, którzy osiągają jakiś sukces, bądź też czasowo doświadczają poprawy warunków życia, są wykluczani lub poddawani innym negatywnym sankcjom przez resztę grupy. Sprawia to, że nie opłaca się mieć „trochę więcej”¹² niż inni, ponieważ

podejście do problemu nie występuje w teorii Rodmana.

11 Jest to efekt izolacji i autoizolacji.

12 Mowa tu o nieznacznej poprawie warunków materialnych, które nie są w stanie trwale ani znacząco zmienić sytuacji jednostki, nie pozwalając jej wyrwać się z kręgu ubóstwa.

faktycznie nie poprawia to warunków funkcjonowania, ale paradoksalnie je pogarsza.

Rodman poparł swoją teorię poszerzonych wartości o wyniki licznych badań empirycznych, między innymi Raymonda T. Smitha i Lloyda Braithwaite'a, w których to sami badacze zauważali zjawisko poszerzania zbioru wartości, bądź jak to nazywali *współwystępowania dwóch systemów moralnych* – Smith (1956, 149), czy „dualizm przywiązania do wartości” (Braithwaite 1957, 534). Pojawienie się nowych, specyficznych dla środowiska ubogich, wartości nie powodowało w badanych społecznościach zanikania wartości ogólnych. Zatem nowe wartości, jako wynik adaptacji do warunków, są dodatkowym elementem, który wyróżnia środowisko osób ubogich. Przeto, gdy przyjmiemy założenia i argumentacje Rodmana, wówczas wyniki badań wskazujących na deklarowanie przez ubogich zbliżonego do reszty społeczeństwa systemu wartości, które mają być dowodem na niesłuszność teorii kultury ubóstwa stworzonej przez Lewisa, nie dają się utrzymać.

Spółeczność zamknięta i nieuczestnicząca

Zdaniem Braithwaite'a ludzie w kulturze ubóstwa mają silne poczucie marginalizacji, bezradności, zależności oraz braku przywiązania. Poczucie odrębności jest u nich tak silne, że nie ufają instytucją publicznym i nie postrzegają ich, jako służące ich interesom i potrzebom. Przekonanie to wzmacnia wzrost poczucia bezsilności oraz niższości i osobistej niegodności (Braithwaite). Wśród członków *kultury ubóstwa* mała jest świadomość i poczucie przynależności historycznej. Są to ludzie na tyle zmarginalizowani, także w sferze tożsamości, że nie odczuwają więzi z żadną grupą kulturową, ideologiczną czy historyczną. Skupieni są na własnych problemach i to głównie problemach dnia codziennego, na swoim wąskim przestrzennym środowisku (głównie najbliższym sąsiedztwie), na własnym sposobie życia. Brak identyfikacji z innymi może być tak silny, że w większości przypadków osoby z kręgu kultury biedy nie będą się utożsamiać nawet z osobami znajdującymi się w podobnym położeniu. Jest to więc czynnik charakterystyczny i zarazem niezbędny dla powstania *kultury ubóstwa*. Lewis zaznacza nawet, że w sytuacji, w której osoba uboga staje się członkiem organizacji np. związków zawodowych, albo gdy przyjmuje internacjonalistyczną perspektywę na świat wówczas przestaje być częścią kultury ubóstwa, mimo iż finansowo nadal określa się ją jako ubogą (Lewis 1969).

Badacz podkreślił też szczególnie dziedziczność tych cech, czyli sposób życia przechodzący w poszczególnych rodzinach z pokolenia na pokolenie. „Kultura ubóstwa jest nie tylko dostosowaniem się do zestawu obiektywnych warunków szerszego społeczeństwa. Gdy zaistnieje ma tendencję do utrzymywania się z pokolenia na pokolenie, poprzez jej wpływ na dzieci. W okresie, kiedy dzieci slumsów są w wieku sześciu lub siedmiu lat, zazwyczaj absorbują podstawowe wartości oraz postawy z ich subkultury i nie są psychicznie nastawione na pełne wykorzystanie zmieniających się warunków i wzrastających możliwości, które mogą wystąpić w czasie ich życia” (Lewis

1968, 188). *Kultura ubóstwa* jest wobec tego przyswajana w procesie socjalizacji, kiedy to jednostki uczą się pewnych wzorców myślowych, wzorów interakcji społecznych i strategii poznawczych, przyswajają specyficzne normy i wartości, które stają się trwałym elementem ich osobowości, determinują ich tożsamość i zachowania niezależnie od zmieniających się warunków otoczenia. Nie znaczy to, że członkowie subkultury ubóstwa nie ewoluują czy nie podlegają zmianom. Zwraca się jednak uwagę na fakt, że elementy *kultury ubóstwa* są na tyle trwałe i na tyle charakterystyczne, że odpowiadają w sposób dominujący za sytuację życiową osób z jej kręgu i nawet częściowa zmiana statusu materialnego nie gwarantuje, że członek tej kultury przestanie nim być. Dodatkowo instytucje społeczne, takie jak szkoły czy przedsiębiorstwa i firmy, działają w oparciu o system norm wytworzony przez klasę średnią, co stanowi dodatkowe utrudnienie dla osób, które w kulturze klasy średniej nie wyrastają. Widoczne jest to szczególnie w przypadku instytucji edukacyjnych, gdzie etos i kultura klasy średniej są dominujące tak w praktyce działań, jak i w treściach programu nauczania. Normy klasy średniej obowiązują w sposób nieoficjalny, ale nie uczy się ich tych, którzy są spoza kręgu tej klasy, co sprawia, że uczniowie ubodzy mają problemy adaptacyjne, wychowawcze i edukacyjne, wynikające z braku stosownych kompetencji społecznych i moralnych¹³. Widać to szczególnie wyraźnie w przypadku etosu pracy, który w kręgach biedy mocno odbiega od etosu cenionego przez członków klasy średniej. Wykazano, że „podstawowe cechy, którymi mogą się charakteryzować osoby pracujące, które doceniają jak pracowitość, uczciwość, utrzymanie dobrego stanu zdrowia, sprawność, na zasadzie przekazu mogą poprawić życiowe szanse dzieci, niezależnie od dochodów ich rodziców. Dzieci pochodzące z takich rodzin radzą sobie w życiu bez względu na wysokość dochodów ich rodziców, ponieważ dziedziczą pewne wzory” (Mayer 1997, 2-3). Badania w tym temacie prowadziła Susan Mayer, dla której ten międzypokoleniowy przekaz odgrywa znaczącą rolę w kształtowaniu się wzorców zachowań, wartości i norm socjalizowanych przez dzieci. Przekaz ten jest o wiele ważniejszy niż ekonomiczna sytuacja rodziców, co sprawia, że niekoniecznie biedne położenie rodziców musi determinować ubóstwo u ich dzieci, ważniejsze jest to, co rodzice są w stanie przekazać swoim dzieciom w procesie wychowawczym. Analizy losów jednostek lub pewnych grup kulturowych pokazały, że czynniki kulturowe dominują w tym zakresie nad materialnymi (Mayer 1997).

Konsekwencje kulturowego ujęcia ubóstwa

Opierając się na swojej tezie, Lewis opracowywał projekty pomocy osobom ubogim, które zakładały głównie pracę nad wzrostem kompetencji społecznych, edukacją i psychiką osób z kręgu kultury ubóstwa, a nie skupianie się wyłącznie na poprawie warunków zewnętrznych, zwłaszcza ekonomicznych. Z podejściem tym zgadza się wielu

¹³ Problemowi temu poświęcona jest książka Piotra Kostyła (2008), a także badania Zbigniewa Kwiecińskiego.

praktyków, pracowników socjalnych czy terapeutów pracujących z ubogimi, jak np. Linda J. Pearson, która w swojej pracy *Understanding the Culture of Poverty* wyraża taką oto opinię: „Ubóstwo rzadko związane jest z brakiem inteligencji czy zdolności. Jednakże wiele osób tkwi w pokoleniowej biedzie, ponieważ nie są świadome strategii behawioralnych, które otworzyłyby im drzwi do innych możliwości i/lub nie mają nikogo, kto mógłby im pomóc nauczyć się reguł pochodzących z klasy średniej, aby mogli korzystać z zasobów” (Pearson 2003, 6). Praca socjalna w myśl tego podejścia miałaby polegać głównie na owym nauczaniu, a nie przekazywaniu środków finansowych czy materialnych, jak ma to głównie miejsce w większości systemów wsparcia społecznego.

Lewis nigdy nie wypracował jednolitej i spójnej teorii *kultury ubóstwa*. Krytycy Lewisa zarzucali mu, że zbyt skupił się na czynnikach kulturowych i psychologicznych, które obciążone były trudnościami w pozyskiwaniu danych. Szczególnie ostrą krytykę rzetelności empirycznych dowodów teorii Lewisa przedstawił Charles Valentine w pracy z 1968 r. pt. *Culture and Poverty: Critique and Counter-Proposals*. Podobnie jak przytoczeni powyżej krytycy teorii, Valentine zwracał uwagę na fakt, że przy opisie *kultury ubóstwa* brak jest wyraźnego rozróżnienia między zachowaniami reprezentatywnymi dla ubogich a ich wartościami¹⁴. Zarzuty Valentine’a, choć w obliczu dzisiejszej wiedzy nie tak trafne jak to się początkowo wydawało, silnie wpłynęły na postrzeganie i interpretację koncepcji Lewisa, zwłaszcza w zakresie rzekomego „obwiniania ofiar biedy za ich ubóstwo” i oceny ubogich jako dewiacyjnej subkultury. Valentine w wyniku tych interpretacji zaliczył pracę Lewisa do tzw. „pejoratywnej tradycji” badań nad ubóstwem, reprezentowanej między innymi przez pracę *The Negro Family in the United States* Frazier (1966) i *The Negro Family: The Case for National Action* Moynihana (1965).

Pomimo, iż Lewis nigdy nie powiedział, że biedni sami są winni swego ubóstwa, wręcz przeciwnie, zadawał pytanie o potrzeby akceptacji odmienności kulturowej względem tej specyficznej subkultury, to niezależnie od jego podejścia wokół pojęcia „kultury biedy” pojawiła się taka właśnie interpretacja. Lewis podkreślał, że badacze problemu, tak samo jak członkowie klasy średniej, których styl życia i moralność są w większości społeczeństw mainstreamowe, wyjątkowo krytycznie podchodzą do postaw życiowych, wartości czy norm wytwarzanych i preferowanych w ramach *kultury ubóstwa*. Sposób funkcjonowania jej członków oceniany jest w kategoriach aberracji, z przypisaniem winy i intencjonalności działań prowadzących do trudnej, niegodnej sytuacji życiowej. Ocena ta, zdaniem Lewisa, nie jest w pełni uprawniona. Nie było jego intencją chwalenie postaw osób ubogich, jednak w swoich pracach wskazywał na te elementy, które można oceniać w pewnych kontekstach pozytywnie, chcąc odmitologizować długotrwałą biedę postrzeganą jako jedynie źródło patologii, przestępczości i lenistwa. Wskazywał na przykład na element spontaniczności, wynikający głównie z braku myślenia o przyszłości i życia dniem codziennym u osób z kultury ubóstwa, którego brak w dominującej klasie

¹⁴ Valentine używał pojęcia „wartość” w znaczeniu możliwie szerokim, dotyczącym także norm, celów i aspiracji ludzkich.

średniej. Zaznaczał, że zbyt często negatywna ocena osób ubogich wynika ze skupiania się na przestępczych, patologicznych zachowaniach części z nich, które przypisywane są całości społeczności. Zbyt rzadko, jak pisał Lewis, mówi się zaś o ich trudach i cierpieniu (Lewis 1963), nie analizuje się przyczyn pewnych zjawisk, nie próbuje się znaleźć wyjaśnienia, a do oceny zachowań przykłada się miarę z zupełnie odmiennych (ze względu na warunki funkcjonowania) środowisk życia. Elżbieta Tarkowska w trakcie swoich badań nad problemem trwałego ubóstwa zauważyła, że

„sposób rozumienia kultury w badaniach biedy, w kontynuacjach i nawiązaniach do koncepcji Oscara Lewisa nie ma wiele wspólnego z jego oryginalną szeroką antropologiczną koncepcją, obejmującą różne wymiary życia jednostek i rodzin, które tworzą ich specyficzny sposób życia (*way of life*). Kultura w badaniach biedy często bywa sprowadzana do jednostkowych postaw i poglądów, które można dowolnie zmieniać i przekształcać (...) Kultura ujmowana szeroko, antropologicznie, jako składnik zewnętrznej wobec jednostki sytuacji, a nie jako wąska sfera wyboru i odpowiedzialności jednostek – nie mieści się w tym ujęciu” (Tarkowska 2000, 24).

Chaos terminologiczny związany z pojęciem kultury generuje również inne problemy, jak chociażby różnice w jakości i celu badań nad *kulturą ubóstwa*. Można mówić o dwóch dominujących podejściach, które doskonale ilustruje przekonanie charakterystyczne dla tej kultury, że mianowicie przyszłość nie może być w żaden sposób kontrolowana, które jedni badacze traktują jako jedną z przyczyn kultury ubóstwa, inni zaś jako jej cechę (Roach & Gursslin 1965). Mamy więc dwa warianty rozumienia kultury: *kultura jako opis*, gdzie czołową rolę odgrywa opis wspólnego dla członków grupy sposób życia, bądź *kultura jako przyczyna*, gdzie wspólny dla grupy wzór funkcjonowania przekazywany międzypokoleniowo jest postrzegany jako czynnik tworzący zjawisko. Większość badaczy posługujących się pojęciem *kultura ubóstwa* używa go właśnie w tym drugim znaczeniu. *Kultura ubóstwa jako przyczyna*, by można było uznać ją za wiarygodną, wymaga zaistnienia dwóch warunków: po pierwsze, należy udowodnić, że kultura ta jest transmitowana międzypokoleniowo, po drugie, biedni w tej kulturze muszą mieć odmienny od reszty społeczeństwa system normatywny. Jednak jak dowodzą Jack L. Roach i Orville R. Gursslin w pracy pt. *An Evaluation of the Concept „Culture of Poverty”*, niełatwo jest to wykazać. Ich zdaniem elementy, na które najczęściej wskazują badacze *kultury ubóstwa* jako na charakterystyczne dla tej subkultury, jak cechy związane ze statusem materialnym, bliskością przestrzenną czy własnościami społeczno-psychologicznymi jednostki, nie są wystarczające do faktycznego stworzenia i utrzymania oryginalnej kultury. Zdaniem autorów krytyczne znaczenie dla istnienia subkultury ma charakter społecznych relacji wewnątrz grupy oraz ich psychologia społeczna. Wskazują oni wręcz na fakt, że takie cechy społeczno-psychologiczne jak ograniczenia poznawcze, ograniczone umiejętności naśladowania, apatia, i ogólny stan „psychicznego wyczerpania” (Roach & Gursslin 1967, 387). wręcz uniemożliwiają tworzenie kultury jako takiej, gdyż ograniczają społeczne interakcje niezbędne do jej tworzenia i utrzymania. „Normy kulturowe, z definicji są normami grupowymi, które

muszą być dodatkowo podtrzymywane przez transmisję za pomocą systemu skutecznej społecznej interakcji, obejmującego sieć jednostek rodzinnych oraz inne podstawowe grupy” (Roach & Gursslin 1967, 388). Ubodzy zaś nie tworzą instytucji, organizacji, nie formują relacji, które w sposób świadomy miałyby przekazywać normy i wychowywać do *kultury ubóstwa*. Zdaniem autorów, jak i innych krytyków, ani jakość, ani ilość relacji podejmowanych przez większość ubogich opisywanych w badaniach empirycznych, nie kwalifikuje ich do tego typu interakcji społecznych, które są w stanie stworzyć i utrzymać kulturę. Nie wyklucza to jednak możliwości powstania pewnych charakterystycznych cech grupowych związanych ze sposobem życia, jak specyficzne preferencje w ubiorze czy wzory konsumpcyjne (Roach & Gursslin, 387-389).

Roach i Gursslin nie odrzucili jednak całkowicie koncepcji *kultury ubóstwa*, wskazali raczej, że możliwe jest zaobserwowanie jej wśród wyizolowanych grup ubogich, których deprivacja materialna sięga kilku pokoleń. Są jednak sceptyczni wobec uznania uniwersalnej kultury dla wszelkich grup trwale ubogich. Ich zdaniem, bardziej stosowne jest mówienie o kilku rodzajach *kultur ubóstwa*, o odmiennych genezach i systemach przekazu międzypokoleniowego. Krytyka przez nich przedstawiona nie neguje istnienia rzeczonoego zjawiska, jednak wskazuje na jego małą użyteczność w badaniach nad ubóstwem, ze względu na niewielką (proporcjonalnie do liczby osób, które można uznać za ubogie) grupę, która spełnia warunki, w jakich takie zjawisko może się wytworzyć. Z perspektywy pragmatyzmu prac nad ubóstwem rekomendują oni inne podejścia, na przykład „wyjaśnienie sytuacyjne” (*situational explanation*) Louisa Kriesberga (1963), które wyjaśnień dla specyficznych zachowań osób ubogich doszukuje się w strukturalnych i społecznych warunkach funkcjonowania, np. niskich zasobach finansowych, bezrobociu itd. Są to czynniki o większym stopniu uniwersalności, przez co teoria dotyczy większej ilości grup ubogich, nie jest też obciążona dodatkowymi problemami metodologicznymi, z jakimi boryka się „wyjaśnienie kulturowe”, np. pojęcie subkultury.

Dodatkowy problem pojawia się przy używaniu pojęć *underclass* i *culture of poverty*. Brak ich bardziej usystematyzowanych definicji i konsekwentnego stosowania. Bardzo często stosuje się je zamiennie, nie wprowadzając między nimi rozróżnień, co, moim zdaniem, jest błędnym podejściem. Gdybyśmy mieli ustalić relację obu pojęć, to teoria *underclass* byłaby nadrzędna względem *kultury ubóstwa*. Gdy posłużymy się chociażby definicją *underclass* zaproponowaną przez Mary Daly, zauważymy, że *kultura ubóstwa* jest jedną z cech wyróżniających podklasę. Daly postrzega *underclass* jako kategorię społeczną podlegającą marginalizacji społecznej i wykluczeniu przez dłuższy czas, doświadczającą deprivacji w różnych sferach życia, skupioną w zubożałych sąsiedztwach i charakteryzującą się podobieństwem kulturowym, które możemy określić jako *kulturę ubóstwa* (Daly 2001). Ta ostatnia jest więc jedną z cech charakterystycznych podklasy, cechą istotną, gdyż trudno jest sobie wyobrazić *underclass* bez odmiennych względem głównego nurtu elementów kulturowych. Podobnie w przypadku samej *kultury ubóstwa*, powinna być ona rozpatrywana głównie w kontekście *podklasy* i warunków, które jej

wykształcenie wymaga.

Podsumowując można powiedzieć, że *kultura ubóstwa* związana jest z transferem międzypokoleniowym norm, wartości, wzorów zachowań i postaw życiowych. Są one wytworem (najczęściej wynikiem adaptacji do sytuacji) ludzi żyjących w specyficznych warunkach strukturalnych, społecznych i psychicznych. Gdy zostanie wytworzona, stanowi dodatkowy czynnik obciążający dla osób z kręgu biedy, gdyż w wyniku jej odmienności i nieakceptowania przez pozostałą część społeczeństwa (przeważanie oceniana jest jako kultura dewiacyjna, a osoby z kręgu kultury ubóstwa obwiniane są za swoje trudne położenie) realizacja ogólnych (dominujących w całym społeczeństwie) celów kulturowych, realizacja ról społecznych czy wartości, jest dla osób ubogich utrudniona. Warunki strukturalne w połączeniu z *kulturą ubóstwa* ograniczają możliwości zmiany sytuacji życiowej, a procesy przekazu międzypokoleniowego dodatkowo ją utrwalają. Również mechanizm wykluczenia i izolacji, w następstwie których interakcje z osobami spoza kręgu biedy są ograniczone, przyczyniają się do utrwalania tejże kultury. Im bardziej zamknięta społeczność, tym silniej utrwała elementy swojej kultury, które w wyniku swoistej izolacji kontrastują wobec kultury głównego nurtu.

Autorka starała się przeprowadzić powyższą analizę na tyle wnikliwie, aby unaocznić, że *underclass* to pojęcie określające stan wyjątkowy tj. nader szczególny rodzaj doświadczenia ubóstwa, który najwyraźniej też uwidacznia wpływ ubóstwa na zdolność do wykształcenia stosownych kompetencji moralnych i społecznych warunkujących skuteczne funkcjonowanie we współczesnych społeczeństwach demokratycznych. Zygmunt Bauman postrzega podklasę jako „piętno totalnego wykluczenia ze wszystkich miejsc i sytuacji, w których tworzone są, negocjowane, przetwarzane i uchylane wszelkie inne ludzkie tożsamości oraz tytuły do uznania. Stan totalnego wykluczenia w wyniku strącenia do rzędu ‘podklasy’ oznacza bycie odartym ze wszystkich wytwarzanych w wymiarze społecznym i społecznie akceptowanych oznak i wyróżników, które podnoszą nagie życie biologiczne do rangi bytu społecznego, a tłum do rangi wspólnoty. Podklasa nie jest li tylko brakiem wspólnoty – to zupełna *niemożność* wspólnoty. Ostatecznie oznacza to także niemożność człowieczeństwa – tylko bowiem za pomocą sieci wspólnoty, dysponujących i respektowanych tożsamości można dostąpić człowieczeństwa. Jak blisko dwa i pół tysiąca lat temu wskazał Arystoteles, nie sposób być człowiekiem – bądź też, będąc człowiekiem, nie sposób przeżyć – poza polis; jedynie bogowie i zwierzęta potrafią żyć poza polis, dodał. Sokrates zapewne to zdanie podzielał, gdyż nie będąc ani bogiem, ani zwierzęciem, wolał czarę cykuty niż porzucenie na zawsze Aten” (Bauman 2012, 171).

Oczywiście włączenie trwale ubogich osób do pełnej partycypacji społecznej wymaga szeregu czynników. Pośród nich istotnym, a często niedocenianym elementem są stosowne kompetencje społeczne i moralne, pozwalające jednostkom na prawidłowe funkcjonowanie w różnych kontekstach, różnych grupach społecznych, z zachowaniem ich wymogów, co pozwala na realizację celów w ramach tych grup, a także umożliwia włączenie jednostki w jej ramy. Bez stosownych umiejętności społecznych ubodzy nie

mają możliwości stać się pełnoprawnymi członkami społeczeństwa. Dlatego też w pracy tej podkreśliliśmy aspekt relacji ubóstwa i możliwości kształtowania kompetencji społecznych i moralnych.

Literatura

- Argyle M. 2002. *Psychologia stosunków międzyludzkich*, przeł. W. Domachowski.
Warszawa: PWN,
- Auletta K. 1982. *The Underclass*. New York: Random House.
- Bauman Z. 2012. *Straty Uboczne. Nierówności społeczne w epoce globalizacji*. Kraków:
WUJ.
- Bourdieu P. 2005. „Struktury, habitus, praktyki”, w: P. Sztompka & M. Kucia (red.),
Socjologia. Lektury. Kraków: Wydawnictwo Znak (546-58).
- Bourdieu P. 2007. *Szkic teorii praktyki poprzedzony trzema studiami na temat etnologii
Kabylów*, przeł. W. Krokier. Kęty: Antyk.
- Braithwaite L. 1957. “Sociology and Demographic Research in the British Caribbean.”
Social and Economic Studies 6(4):523-71.
- Braithwaite L. 1960. “Social Stratification And Cultural Pluralism.” *Annals of the New
York Academy of Sciences* 83:816-836.
- Buckingham A. 1999. “Is there an Underclass in Britain?” *The British Journal of
Sociology* 50 (1):49-75.
- Dahrendorf R. 1993. *Nowoczesny konflikt społeczny*. Warszawa: Czytelnik.
- Daly M. 2001. *Care Work: The Quest for Security*. Geneva: International Labour Office.
- Davis A. 1946. “The Motivation of the Underprivileged Worker”, w: V. F. Whyte (Ed.),
Industry and Society. New York: McGraw-Hill.
- Della Fave L. R. 1974. “The Culture of Poverty Revisited: A Strategy for Research.” *Social
Problems* 21:609-621.
- Domańska L. 2002. *Deprywacja potrzeb byłych pracowników państwowych
przedsiębiorstw gospodarki rolnej*. Toruń: Uniwersytet Mikołaja Kopernika.
- Domański H. 2002. *Ubóstwo w społeczeństwach pokomunistycznych*. Warszawa: Instytut
Spraw Publicznych.
- Duncan G., & Brooks-Gunn J. (Eds.) 1997. *Consequences of Growing Up Poor*. New York:
Russel Sage Foundation.
- Duncan G. & Brooks-Gunn J. 1998. “How Much Does Childhood Poverty Affect the Life
Chances of Children?” *American Sociological Review* 30:491-498.
- Erol R. & Sawhill I. 1988. „Defining and Mesuring Underclass.” *Journal of Policy, Analysis
and Managment* 7:316-325.

- Gallie D. 1988, "Employment, Unemployment and Social Stratification", w: D. Gallie (Ed.), *Employment in Britain*. Oxford: Blackwell (465-92).
- Gallie D. 1994. "Are the Unemployed an Underclass? Some Evidence from the Social Change in Economic Life Initiative." *Sociology* 28:737-57.
- Giddens A. 2001. *Poza lewicą i prawicą. Przyszłość polityki radykalnej*. Poznań: Zysk i S-ka.
- Goodwin L. 1972. *Do the Poor Want to Work? A Social- Psychological Study of Work Orientations*. Washington D.C.: Brooking Institution.
- Grotowska-Leder J. 2004. „Underclass – kategoria teoretyczna i przedmiot działań praktycznych”, w: K. Wódz & K. Piątek (red.), *Socjologia i polityka społeczna a aktualne problemy pracy socjalnej. Dylematy teorii i praktyki społecznej*. Toruń: Wydawnictwo Edukacyjne Akapit (233-251).
- Gudaszewski G. & GUS 2015. *Struktura narodowo-etniczna, językowa i wyznaniowa ludności Polski*. Warszawa: Departament Badań Demograficznych i Rynku Pracy. Źródło: https://stat.gov.pl/files/gfx/portalinformacyjny/pl/defaultaktualnosc/5670/22/1/1/struktura_narodowo-etniczna.pdf, dostęp dn. 12.09.2018.
- Hannerz U. 1969. *Soulside: Inquiries into Ghetto Culture and Community*. Stockholm: Almqvist & Wiksell.
- Harrington M. 1962. *The Other America. Poverty in the United States*, New York: Macmillan.
- Hemmerling K. & Scharlipp M. 2013. „The Keys of Cognition: Insights into Moral and Democratic Education in Prison”, w: E. Nowak, D. E. Schrader, & B. Zizek (Eds.), *Educating Competencies for Democracy*. Frankfurt/Main – Bern – Bruxelles – New York: Peter Lang Publishing (287-304).
- Hemmerling K., Scharlipp M., & Lind G. 2009. "Dyskusja Konstancka jako metoda edukacyjna w grupach wysokiego ryzyka", w: K. M. Cern, P. W. Juchacz, & E. Nowak (red.), *Etyka życia publicznego*. Poznań: CoOpera, Wydawnictwo Naukowe IF UAM.
- Hymati H. H. 1953. "The Value Systems of Different Classes: A Social Psychological Contribution to the Analysis of Stratification", w: R. Bendix & S. M. Lipset (Eds.), *Class, Status and Power*. Glencoe: Free Press (426-42).
- Jencks Ch. 1989. "What Is the Underclass – And Is It Growing?" *Focus* 12(1):14-26.
- Johnson N. J. & Sanday R. 1971. "Subcultural Variations in An Urban Poor Population." *American Anthropologist* 73:128-43.
- Karwacki A. 2006. *Błędne Koło. Reprodukacja kultury podklasy społecznej*, Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.
- Kohlberg L. 1964. "Development of Moral Character and Moral Ideology", w: M. L. Hoffman & L. W. Hoffman (Eds.), *Review of Child Development Research*, Vol. I, New York: Russel Sage Foundation.

- Kohlberg L. 1976. "Moral Stages and Moralization: The Cognitive-Developmental Approach", w: T. Lickona (Ed.), *Moral Development and Behavior: Theory, Research, and Social Issues*. New York: Holt.
- Kołąkowski L. 1988. *Główne nurty marksizmu. Powstanie – rozwój – rozkład*. Londyn: Aneks.
- Korzeniowska K. & Tarkowska E. (red.). 2002. *Lata tłuste, lata chude... Spojrzenie na biedę w społecznościach lokalnych*. Warszawa: IFiS PAN.
- Kostyło P. 2008. *Wykluczenie jako problem filozofii edukacji*. Kraków: Oficyna Wydawnicza Impuls.
- Kraczla R. 1993. "Badania nad ubóstwem. Przegląd głównych nurtów i teorii", w: K. Wódz (red.), *W kręgu ubóstwa. Próba analizy psychospołecznych aspektów zjawiska*. Katowice: Wydawnictwo UŚ.
- Kriesberg L. 1963. "The Relationship between Socio-Economic Rank and Behavior." *Social Problems* 10(4):334-53.
- Lee P. 1994. "Housing and Spatial Deprivation: Relocating the Underclass and the New Urban Poor." *Urban Studies* 31:1191-209.
- Lewis O. 1973. *Sanchez i jego dzieci. Autobiografia rodziny meksykańskiej*, przeł. A. Frybesowa. Warszawa: PIW.
- Lewis O. 1969. "The Culture of Poverty", w: D. P. Moynihan (Ed.), *On Understanding Poverty: Perspectives from the Social Sciences*. New York: Basic Books.
- Lewis O. 1963. "The Culture of Poverty." *Society* 35:7-9.
- Marshall G., Roberts S., & Burgoyne C. 1996. "Social Class and Underclass in Britain and the USA." *The British Journal of Sociology* 47(1):22-44.
- Martowska K. 2012. *Psychologiczne uwarunkowania kompetencji społecznych*, Wydawnictwo Liberi Libri. Źródło: <http://liberilibri.pl/ksiazka/psychologiczne-uwarunkowania-kompetencji-spo%C5%82ecznych-0>, dostęp dn. 05.04.2014.
- Mayer S. 1997. *What Money Can't Buy. Family Income and Children's Life Chances*. Cambridge, Mass.: Harvard Univ. Press.
- Merton R. K. 2002. *Teoria socjologiczna i struktura społeczna*, przeł. E. Morawska, J. Wertenstein-Żuławski. Warszawa: PWN.
- Miller S. M., Reissman D., & Seagull A. 1971. "Poverty and Self-Indulgence: A Critique of the Non-Deferred Gratification Pattern", w: L. Ferman *et al.* (Eds.), *Poverty in America*. Ann Arbor: University of Chicago Press.
- Morris L. 1994. *Dangerous Classes: The Underclass and Social Citizenship*. London: Routledge.
- Murray Ch. 2001. *Bez korzeni*, przeł. P. Kwiatkowski. Poznań: Zysk i S-ka.
- Niedzielski E. & Kisiel R. (red.) 2001. *Środowisko popegeerowskie – diagnoza stanu*. Olsztyn: Uniwersytet Warmińsko-Mazurski – AWRSP.
- Parsons T. 1969. *Struktura społeczna a osobowość*, przeł. M. Tabin. Warszawa: Państwowe Wydawnictwo Ekonomiczne.

- Pearson L. J. 2003. "Understanding the Culture of Poverty." *The Nurse Practitioner* 28(4):6.
- Poławski P. 1977. „Kulturowe znaczenie underclass. Spór o sens terminu a polska transformacja”, w: K. W. Frieske (red.), *Ofiary sukcesu. Zjawiska marginalizacji społecznej w Polsce*. Warszawa: IS UW.
- Rainwater L. 1970. *Behind Ghetto Walls*. Chicago: Aldine.
- Rainwater L. 2006. *Behind Ghetto Walls: Black Family Life in a Federal Slum*. New York: Transaction Publishers.
- Rainwater L. 1967, *The Problem of Lower-Class Culture and Poverty-War Strategy*. Washington, D.C.: Washington University Press
- Roach J. L. & Gursslin O. R. 1967. "An Evaluation of the Concept ,Culture of Poverty’" *Social Forces* 45(3):383-92.
- Roach J. L. & O. R. Gursslin O. R. 1965. "The Lower Class, Status Frustration and Social Disorganization." *Social Forces* 43(4):501-10.
- Rodman H. 1963. "The Lower-Class Value Stretch." *Social Forces* 42(2):205-15.
- Savage M., Devine F., Cunningham N., Li Y., Taylor M., Hjellbrekke J., Le Roux B., Friedman S. & Miles A. 2013. "A New Model of Social Class: Findings from the Great British Class Survey Experiment." *Sociology* 47(2):219-50.
- Schneiderman L. 1964. "Value Orientation Preferences of Chronic Relief Recipients." *Social Work* 9:13-18.
- Smith D. J. 1992. *Understanding the Underclass*. London: Policy Studies Institute.
- Smith R. T. 1956. *The Negro Family in British Guiana*. London: Routledge & Kegan Paul.
- Sykes G. M. & Matza D. 1957. "Techniques of Neutralization: A Theory of Delinquency." *American Sociological Review* 22:664-70.
- Taft D. R. 1950. *Criminology*. Revised edition. New York: Macmillan.
- Tarkowska E. 1999. „In Search of an Underclass in Poland." *Polish Sociological Review* 125:3-16.
- Tarkowska E. 2000. *Zrozumieć biednego. O dawnej i obecnej biedzie w Polsce*. Warszawa: Typografia.
- Weber M. 2002. *Gospodarka i społeczeństwo*, przeł. D. Lachowska. Warszawa: PWN.
- Westergaard J. 1992. "About and Beyond the `Underclass': Some Notes on Influences of Social Climate on British Sociology Today." *Sociology* 26(4):575-87.
- William R. 1971. *Blaming the Victim*. New York: Pantheon Books.
- Wilson W. J. 1989. "Social Research and the Underclass Debate." *Bulletin of the American Academy of Arts and Sciences* 43(2):30-44.
- Wilson W. J. 1989. "The Underclass: Issues, Perspectives, and Public Policy." *Annals of the American Academy of Political and Social Science* 501:182-92.
- Wilson W. J. 1987. *The Truly Disadvantaged*. Chicago: Chicago University Press.
- Wilson W. J. 1990, *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

- Wilson W. J. 1991. "Studying Inner-City Social Dislocations: The Challenge of Public Agenda Research." *American Sociological Review* 56:1-14.
- Wódz K. 1994. „Underclass w starych dzielnicach przemysłowych miast Górnego Śląska”, w: I. Machaj & J. Styk (red.), *Stare i nowe struktury społeczne w Polsce*, t. I. Lublin: UMCS.
- Wright E. O. 1997. *Class Counts: Comparative Studies in Class Analysis*, Cambridge: Cambridge University Press.
- Zabielska J. M. 2007. *Ubóstwo a procesy marginalizacji społecznej*. Lublin: Wydawnictwo KUL.

Karolina Kowalewska
(Poznań)

*Underclass, Culture of Poverty and the Research
in Sociomoral Competence among the Poor*

Abstract: In this paper, I consider whether a theoretically defined concept of underclass can be helpful in social scientific discourse. I analyze, review and categorize the concept of a underclass and I made some suggestions for a use of the concept application in the study of social, psychological and moral consequences of poverty. It begins with a discussion of the continuing and pervasive appeal of structural explanations as the root cause of poverty. Then are also examined cultural theories, which by contrast find explanation for poverty in various elements of culture (ex. cultural traits of poor are thought to be different from those of the rest society). A third, a synthetic review of the urban underclass concept is explored. The article also describes the concept of a „culture of poverty” (taking into account the context of the subculture) and „lower class value stretch” as influential concepts in analyses of contemporary poverty and in explanations of the behavior of the poor.

Keywords: underclass; culture of poverty; morality; value stretch; contemporary society.

Ethics in Progress (ISSN 2084-9257). Vol. 9 (2018). No. 1, Art. #4, pp. 62-98.

Creative Commons BY-SA 3.0

Doi: 10.14746/eip.2018.1.4

Chinese Adolescents' Conceptions of Teacher Authority and Their Relations to Rule Violations in School

Jianjin Liu

(Guangdong University of Foreign Studies, Guangzhou, liujianjinljj@126.com)

1. Introduction

Teacher authority refers to teachers' power or influence on students which makes them to believe or obey in the process of education and teaching (Zhang & Li 2003). It is manifested as teachers' control and management of students and students' dependence and obedience to teachers (Tian 1998). It not only plays a direct role in education management, but also has a potential impact on the effect of other teachers' educational practices (Yu 2005). Thus, teacher authority is a description of the relationship between teachers and students. Clifton, R. A. and Roberts, L. W. put forward four types of teacher authority: legal authority, traditional authority, charismatic authority and professional authority. The former two originated from the social and educational system, the latter from the teacher's personal factors, and the teacher's general authority was the resultant force of the interaction of the four levels of authority.

After entering the modern society, the lofty status of teachers has been suffering unprecedented challenges. Some scholars have asserted that, part of teacher authority, namely the authority comes from knowledge and institution, is suffering a disastrous decline (Luo 2010). Teacher authority in the Internet age is faced with the danger of being thoroughly dispelled. Some people even claim that teacher authority has become out of date and should be abandoned. But more people believe teacher authority is very necessary for ensuring students' rights, realizing students' freedom and promoting their development in all aspects. Just as Deng & Wang (2007) put it, "in education, teacher authority is indispensable. Therefore, the essence of the problem is not the necessity of teacher authority, but what kind of teacher authority we need." Therefore, how to reconstruct or transform teacher authority has become a hot topic in education. Zhang Liangcai and Li Runzhou (2003) thought that in modern education it is becoming more

and more urgent to transform teacher authority from external dependent authority (traditional, legal authority) into internal generating authority (charismatic, professional authority). Cao Jing and Ming Qinghua (2012) believed that the transformation of teacher authority needed the protection from external system and the improvement of teachers' inner quality, which was reflected in the communication between teachers and students in educational practice.

However, as mentioned above, teacher authority is a concept of relationship, involving the authority of the subject (teachers) and authority of the receptor (students). Therefore, how to reshape teacher authority is not a one-way issue, i.e. it's not just concerning how to make teachers become more authoritative, but also related to how students understand and perceive the authority of teachers, that is, students' cognition of teacher authority matters a lot in the process.

Social Cognitive Domain Theory (hereinafter referred to as Domain Theory) has conducted detailed and in-depth studies of the cognition of teacher authority. Domain Theory is one of the latest development of moral psychology in the west. The tenet of the theory is: the diversity of social interactions results in the heterogeneity of social rules or events, then forms the heterogeneity in social cognitive domains. At present, moral domain, conventional domain and personal domain have been widely recognized and studied. Moral domain focuses on the direct impact of intrinsic consequence on others, and its core concepts are human welfare, fairness and rights. Generally speaking, the consequences of moral violations are the most serious and deserve more severe punishment. Social conventions are the rules of behavior consistency that coordinate social interaction. They are subject to specific social situations and are arbitrary behaviors relative to social situations. Conventional rules mainly involve the normal operation of social organizations and the maintenance of social order. Violations of conventions do not cause losses to the rights and interests of others, so they are not as serious as moral violations (Turiel 1993). Personal domain focuses mainly on actors themselves. The consequences of personal behavior only affect the actor him/herself, and will not have adverse effects on others and social organizations (Nucci 1981). Prudential domain belongs to personal domain, which mainly refers to the potential or actual harm to the actor himself (Tisak & Turiel 1984). There are interactions and interplay between those domains.

Domain Theory researchers found that children and adolescents perceived teacher authority in different ways in different social domains. Smetana and Bitz (1996) found that students recognized teacher authority in moral domain most, followed by conventional and prudential domain; but to a great degree they didn't recognize and accept teacher authority in contextually conventional and personal domain. Yariv (2009) also found that students may sometimes deny teacher authority, and further discussed the boundaries of teacher authority from the perspective of students' understanding and acceptance, and drew the conclusion that there were differences in the acceptance and recognition of teacher authority in different social domains. Pace and Hemmings (2007) emphasized

that legitimate teacher authority was value laden and culturally sensitive. In different sociocultural and instructional contexts, different forms of teacher authority might be perceived differently and have differential effects. Researchers have shown that students in different cultures value different authority attributes (Kim 1998; Laupa & Tse 2005).

Do Chinese adolescents perceive teacher authority differently in moral, conventional, contextually conventional, personal and prudential domain? Is there any correlation between authority cognition in various domains and disciplinary actions in these domains? These are the question that this study tries to answer.

2. Method

2.1 Subjects

The sample consisted 297 adolescents from grade 7, 9 and 11, with 131 males and 166 females, 142 from big cities and 155 from countryside in China. The mean ages were 13.06 years ($SD=.62$), 15.00 years ($SD=.59$), and 17.18 years ($SD=.59$), respectively. 47.8% of the sample came from big cities in China, with more advanced economy and culture; 52.2% of them came from countryside in mainland China, with more traditional culture and lifestyles. The majority of the sample lived in intact, two-parent families (69%); 8.6% lived in single-parent families; 8.9% lived with grand-parents; 3% of them lived in step-parent families; the others lived alone or with other relatives.

2.2 Stimuli

To generate stimuli that were relevant across schools and grades, the author interviewed students and teachers from the seven participating schools about typical rule violations in their schools. The stimuli for this study were drawn from those interviews and were chosen to be relevant to all three grades included in this study. The stimuli were 19 hypothetical transgressions, 4 moral, conventional, personal, contextually conventional items, and three prudential ones. The items were presented in Table 1. Each of the contextually conventional events can be seen as an act that is typically treated as personal in other contexts but was regulated in the school in this study (Smetana 1996). The definitions of morality, conventionality, personal choice and prudence and the assignment of stimuli to conceptual domains were based on previous research (Nucci 1981; Turiel 1983; Smetana 1981; 1988; 1996; Tisak 1993).

2.3 Student questionnaires

2.3.1 Background information.

At the beginning of the questionnaire, students were required to offer their age, gender, grade, parents' marital status, living area (countryside or big city). Students also indicated their average academic performance with choice of "excellent", "average", and

“poor”. Finally, students were asked “how much you like school” on a five-point scale ranging from “dislike very much” (1) to “like very much” (5).

Domain	Stimulus Items
Moral	Cheating in an exam Stealing money from another student Laughing at other students Fighting with other students
Conventional	Being late for school Chitchatting in a class Jumping the queue at lunch Going to school with sleepers
Personal	Keeping short hair Wearing earrings and necklace Sitting next to friends at lunch How to spend pocket money
Contextually conventional	Leaving classroom without permission to go to lavatory Passing a note to friends in class Using cell phones at classroom
Prudential	Smoking cigarettes at school Drinking alcohol at school Surfing the Internet in a bar

Table 1: Stimulus Items for the School Authority Questionnaire

2.3.2 Act and rule evaluations.

Then participants were presented a list of events and asked whether to indicate “whether it is OK or not OK for teachers to make a rule” about each event, assessing the *legitimacy of teacher authority*. The score of “1” indicates legitimacy assigning to teachers; the score of “0” means nonlegitimacy. Next, participants were asked to evaluate how good or bad it would be to make rules for each issue on a 5-point scale ranging from 1 (very bad) to 5 (very good), assessing their *ratings of rules*. Participants then were asked to judge whether or not the acts were “up to the individual to decide”, assessing judgments of *personal jurisdiction*. Responses indicating personal jurisdiction were assigned a score of “1”, and responses rejecting personal jurisdiction were assigned “0”. Then participants were ask to rate the acts on a 3-point scale ranging from 1 (not bad) to 3 (very bad), assessing their *ratings of act severity*.

2.3.3 Alternative authority sources.

Subjects were asked to check off the person or institutions (the self, teachers, parents, school principals, and the law) that should be able to make rules about each issue or have the *authority to govern the behavior*. Participants were instructed to choose as many people or authorities as apply. The score of “1” means the use of the category; “0” means the category was not used.

2.3.4 Ratings of the school context.

The next two questions concerned adolescents’ perceptions of their school context. For each issue, they rated the extent to which issues were governed by rules in their schools or classes on a scale ranging from 1 (no rules or expectations) to 3 (strict rules or expectations), assessing the *existence of rules*. Then they rated the extent that teachers let them decide each issue on a scale ranging from 1 (no input) to 5 (a great deal of input), assessing *opportunities for input into decision making*.

2.3.5 Student ratings of misconduct.

Finally, the participants were asked to indicate their engagement in each of the behavior, over the current school year, on a scale ranging from 1 (never) to 5 (very frequently), assessing *school misconduct*.

2.4 Teacher ratings of misconduct

Class teachers were asked to rate participants' behaviors over the current school year on a 19-item questionnaire that paralleled students' ratings of their behaviors, on a scale ranging from 1 (never) to 5 (very frequently).

2.5. Procedure

Researchers visited junior middle schools and senior middle schools, both in countryside and big cities, to invite students to participate in the study. Students were offered a gift for participating in the research. The sample was randomly selected from students, with the permission of their parents. Then students from the same school were asked to finish the questionnaire at the same period of time (about 30 minutes). The researcher read the questionnaire instructions to the participants.

Class teachers were chosen to rate students' behavior, because they were more familiar with students' performance at school. Teachers involved in the research were offered a U-disk as rewards. A total of 14 teachers participated in the teacher ratings.

2.6. Analyses

Domain scores were obtained by averaging responses for the items in each domain for each judgment. Then mean responses for each judgment were analyzed using separate 3 (grade) \times 2 (gender) \times 2 (area) \times 5 (domain) repeated-measures analyses of variance (ANOVAs) with grade, gender and area as between subject variables and domain as the repeated measure. Multiple regressions on misconduct were performed to examine the relative contributions of different variables on misconduct.

3. Results

3.1 Evaluations of acts and rules

Separate ANOVAs were performed on judgments of legitimacy and personal jurisdiction and ratings of act severity and rules. These analyses showed main effects for domains for all four judgments, $F_s(4,1136)=310.96, 357.38, 407.31, 357.36$, $p_s<.0001$, respectively. As can be seen in Table 2, moral issues were judged to be more legitimately subject to teacher's authority ($M=.95$) than all other issues, and prudential and conventional issues were judged to be more legitimately subject to teachers' authority than were contextually conventional and personal issues. Conversely, personal issues were regarded as under personal jurisdiction ($M=.76$) more than any other issues, and contextually conventional issues ($M=.54$) were seen as under personal jurisdiction more than were prudential, conventional, or moral issues.

GRADE				
	7	9	11	M
Legitimacy				
Moral items	92.8	97.0	95.9	95.2
Conventional	89.2	83.0	80.1	84.4
Personal	53.4	35.3	27.9	39.6
Contextual	78.4	67.0	63.0	69.9
Prudential	91.6	85.6	82.3	86.9
Personal jurisdiction				
Moral items	29.3	44.7	56.3	42.2
Conventional	35.1	50.2	65.8	49.1
Personal	61.1	81.9	87.3	76.2
Contextual	38.9	53.5	72.8	53.6
Prudential	31.1	51.4	65.0	48.0
Rule evaluation^a				
Moral items	4.37	4.30	4.09	4.27
Conventional	4.15	3.83	3.70	3.91
Personal	3.34	2.71	2.54	2.71
Contextual	3.83	3.46	3.27	3.54
Prudential	4.00	4.10	3.85	4.14
Act evaluation^b				
Moral items	2.72	2.70	2.53	2.67
Conventional	2.36	2.06	2.03	2.16
Personal	1.87	1.54	1.49	1.64
Contextual	2.29	2.04	1.86	2.08
Prudential	2.36	2.18	2.07	2.21

Table 2: Mean Judgments (in %) of Authority Legitimacy and Personal Jurisdiction and Mean Ratings of Acts and Rules (by grade)

^a On a scale from 1 (very unreasonable) to 5 (very reasonable).

^b On a scale from 1 (not severe) to 3 (very severe).

Likewise, moral and prudential rules were rated more positively and moral and prudential violations were rated more negatively than all other rules and acts or rule violations (see Table 2 for means). Conventional rules were rated more positively and conventional violations were rated more negatively than contextually conventional and personal rules or rule violations, respectively. Rules concerning contextually conventional issues were rated more positively than rules concerning personal issues,

and violations of contextually conventional rules were rated more negatively than personal violations.

Area			
	Countryside	Big city	M
Legitimacy			
Moral items	95.3	95.1	95.2
Conventional	88.0	80.1	84.3
Personal	46.9	31.3	39.5
Contextual	80.0	58.7	70.0
Prudential	90.4	83.0	86.9
Personal jurisdiction			
Moral items	39.2	45.6	42.2
Conventional	41.4	57.7	49.1
Personal	67.5	85.8	76.2
Contextual	38.6	70.3	53.6
Prudential	39.8	57.1	48.0
Rule evaluation ^a			
Moral items	4.33	4.20	4.27
Conventional	3.99	3.80	3.91
Personal	3.05	2.35	2.71
Contextual	3.87	3.16	3.54
Prudential	4.26	4.00	4.14
Act evaluation ^b			
Moral items	2.71	2.62	2.67
Conventional	2.26	2.03	2.16
Personal	1.78	1.47	1.64
Contextual	2.25	1.88	2.08
Prudential	2.25	2.16	2.21

Table 3: Mean Judgments (in %) of Authority Legitimacy and Personal Jurisdiction and Mean Ratings of Acts and Rules (by area)

^a On a scale from 1 (very unreasonable) to 5 (very reasonable).

^b On a scale from 1 (not severe) to 3 (very severe).

The ANOVAs also revealed significant main effects for grades for judgments of legitimacy and personal jurisdiction and ratings of rules and act severity, $F(2, 287) = 9.23, 24.79, 15.76, 20.61, p < .0001$, respectively. Seventh graders viewed all rules as more legitimately subject to teachers' authority than did older students, whereas older students viewed all acts as under personal jurisdiction more than did seventh graders (see Table 2 for means). Seventh graders also evaluated rules more positively than did ninth and

eleventh graders, and acts more negatively than did older students. Finally, ninth graders rated acts more negatively than did eleventh graders.

There were significant main effects for area in legitimacy, personal jurisdiction, and ratings of rule and act severity, $F(1, 287) = 21.71, 16.46, 26.53, 34.26, p < .0001$, respectively. Students came from countryside believed teachers had more authority over conventional, personal, contextually conventional, and prudential rules, $F(1, 301) = 6.80, 26.11, 40.00, 8.59, p < .05, .0001, .0001, .01$, respectively (see Table 3 for means), but there was no significant difference on moral rules, and the majority (over 95%) of students confessed teachers' authority over moral issues. Conversely, students came from countryside endorse less personal jurisdiction on conventional, contextually conventional, personal, and prudential issues, $F(1, 301) = 12.63, 23.73, 57.09, 12.80, p < .0001$, respectively. However, no significant difference was found between moral issues (see Table 3 for means). The ANOVAs also revealed significant difference across all domains concerning rule evaluation, $F(1, 301) = 4.26, 5.28, 47.72, 51.39, 9.34, p < .05, .05, .0001, .0001, .01$, respectively. Finally, students came from countryside rated act violations in moral, conventional, personal, and contextual conventional domains more severely than did students came from big cities, $F(1, 301) = 5.92, 18.64, 40.80, 46.61; p < .05, .0001, .0001, .0001$, respectively. But there was no significant difference between prudential violations (see Table 3 for means).

3.2 Alternative authority sources

As can be seen in Table 4, the majority of students regarded teachers and principals as having authority over moral and conventional issues, but most of them rejected their authority over personal issues. About half of the adolescents endorsed authority to law, parents, and themselves over moral issues. Also about half of them believed themselves had authority over conventional issues. The majority of students regarded themselves as the authority over personal issues. It's also interesting to notice that almost half (46%) of the students endorsed authority to parents over personal issues. Teachers also had the most authority over contextual conventional and prudential issues (about 71% each), while students themselves and parents were endorsed more authority in personal and prudential issues (about half) than other sources of authority except teachers.

3.3 Ratings of school context

Students reported that, on average, acts were governed by rules ($M = 2.21$), and that they only had a moderate amount of input in decision making ($M = 2.39$). Separate ANOVAs revealed significant main effects for domain in ratings of rule existence and opportunities for decision making, $F(4, 1108) = 115.81, 60.90$, respectively; $p < .0001$. Students reported they had more moral and contextually conventional rules than other types of rules; and they had more conventional and prudential rules than personal rules (See Table 5 for means).

Contextual					
	Moral	Conventional	Personal	Convention	Prudential
Self	47.0	49.6	73.7	54.2	49.7
Parents	47.7	24.4	46.4	39.4	58.7
Teachers	66.5	73.0	37.8	71.2	71.7
Headmaster	57.2	59.2	26.3	47.0	48.9
Law	53.3	10.5	7.6	11.8	50.8

Table 4: Mean Endorsement (in %) of Alternative Authority Sources

There were main effects for areas and grades in ratings of rules and input into decision making, $F(1, 282) = 48.50, 13.47, p < .0001$; $F(1, 282) = 27.92, 6.25, p < .0001, .01$, respectively. On average, students from countryside reported more rules in schools and more input in rule decision making, $MD = .22, .57$, respectively. Seventh graders reported more rules in schools and more input in rule decision making than did ninth and eleventh graders.

There was significant difference between boys and girls in ratings of rules, $F(1, 282) = 4.89, p < .05$. Boys reported more rules of all domains than did girls, $MD = .16$. But there was no significant gender difference in input into decision making.

Grade				
	7	9	11	M
Existence of rules ^a				
Moral items	2.55	2.31	2.16	2.35
Conventional	2.46	2.24	1.99	2.25
Personal	1.95	1.86	1.61	1.83
Contextual	2.43	2.45	2.11	2.35
Prudential	2.51	2.13	2.04	2.24
Opportunities ^b				
Moral items	2.93	2.38	2.44	2.58
Conventional	2.91	2.38	2.54	2.60
Personal	2.22	1.86	1.68	1.93
Contextual	2.72	2.17	2.25	2.38
Prudential	2.88	2.24	2.22	2.45

Table 5: Ratings of Rule Existence and Opportunities for Decision Making

^a On a scale from 1 (no firm rules) to 3 (firm rules).

^b On a scale from 1 (no input) to 5 (a great deal of input).

3.4 Self-reports of misconduct

On average, students reported that they rarely misbehaved ($M = 1.70$). An ANOVA performed on self-reports of misconduct revealed a significant main effect for domain, $F(4, 1092) = 403.18, p < .0001$. Adolescents reported more conventional misconducts than

any other type of misconduct except personal behavior. They reported fewer prudential misconducts than any other types of misconducts and fewer moral violations than conventional and contextually conventional violations (See means in Table 6).

There were significant domain × grade, $F(4, 1092)=9.25, p<.0001$, domain × area, $F(4, 1092)=13.29, p<.0001$ and domain × gender interaction, $F(4, 1092)=9.82, p<.0001$. On average, eleventh graders reported more misconducts than did ninth graders, $MD=1.04$; and ninth graders reported more misbehaviors than did seventh graders, $MD=1.45$. Students came from countryside generally reported more misconducts than did students from big cities. Finally, boys reported more misconducts of moral and prudential issues, while girls reported more misbehaviors of conventional, personal, contextually conventional issues.

Grade				
	7	9	11	M
Teacher ratings ^a				
Moral items	1.20(.37)	1.94(.72)	1.34(.38)	1.46(.60)
Conventional	1.38(.39)	1.98(.59)	1.59(.63)	1.62(.58)
Personal	2.26(.60)	1.81(.40)	3.09(.66)	2.36(.74)
Contextual	1.28(.26)	1.53(.41)	1.87(.91)	1.51(.60)
Prudential	1.16(.30)	1.45(.31)	1.74(.73)	1.40(.51)
Student ratings ^a				
Moral items	1.43(.42)	1.51(.46)	1.61(.49)	1.51(.46)
Conventional	1.76(.58)	2.13(.62)	2.18(.76)	2.02(.67)
Personal	2.43(.78)	3.06(.73)	3.55(.83)	2.97(.88)
Contextual	1.44(.50)	1.63(.54)	2.07(.74)	1.68(.63)
Prudential	1.32(.57)	1.50(.63)	1.46(.61)	1.43(.58)

Table 6: Mean Teacher and Student Ratings Misconduct

^a On a scale from 1 (never) to 5 (very frequently).

3.5 Correlations between Student and Teacher Ratings of Conduct

Mean teacher ratings for conduct in different domains also presented in Table 6. For all students, mean teacher' and students' ratings of students conventional, personal behavior were moderately but significantly related ($r_s=.31, .30, p_s<.001$), whereas their ratings of moral, contextually conventional, and prudential behavior were weakly but significantly related ($r_s=.12, .19, .13, p_s<.05$).

Correlated t tests indicated that across domains, students consistently rated themselves as engaging in more misconduct than did their teachers.

3.6 Multiple factor regressions on misconduct

In the next set of analyses, the relative contributions of adolescents' ratings of rules (as assessed by act severity, rule evaluation, and legitimacy judgments), their evaluations of the school context (as assessed by ratings of rule existence and opportunities for decision making),

background characteristics (gender, age, and parents' marital status, scored as married, intact families vs. all else), and orientations toward school (self-reports of academic performance and liking for school) on school-related misconduct were examined using multiple regression analyses. To alleviate some of the problem associated with the use of self-report data, and because teacher and self-report ratings were significantly correlated, composite teacher and student ratings of misconduct were created by obtaining mean teacher and student ratings of overall misconducts and of misconducts in each domain.

independent/coefficient of R square F standardized				
dependent factor determination				
	change	Regression Coefficient		
	(R ²)	(ΔR)		
Misbehavior (overall) R²=0.412				
age	0.245	0.245	80.422	0.407
authority legitimacy	0.399	0.155	81.216	-0.370
like school	0.412	0.012	57.393	0.117
Moral misbehavior R²=0.169				
rule existence	0.075	0.075	22.417	-0.204
gender	0.115	0.039	12.264	0.215
rule evaluation	0.152	0.037	12.062	-0.200
like school	0.169	0.017	5.169	0.317
Conventional misbehavior R²=0.238				
age	0.095	6.095	29.759	0.265
authority legitimacy	0.161	0.066	22.219	-0.150
gender	0.195	0.034	11.790	0.199
like school	0.213	0.018	6.274	0.116
decision opportunity	0.224	0.012	4.226	0.140
rule evaluation	0.238	0.014	4.987	-0.161
Personal misbehavior R²=0.505				
authority legitimacy	0.288	0.288	112.776	-0.292
age	0.409	0.121	57.146	0.398
area	0.482	0.072	38.661	-0.253
rule existence	0.498	0.017	9.157	-0.135
gender	0.505	0.007	3.933	-0.086
Contextually conventional R²=0.314				
age	0.207	0.207	75.666	0.387
authority legitimacy	0.257	0.	19.643	-0.256
decision opportunity	0.284	0.026	10.527	0.174
rule existence	0.299	0.016	6.474	-0.175
area	0.314	0.015	6.267	0.142
Prudential misbehavior R²=0.254				
rule evaluation	0.125	0.125	41.639	-0.317
age	0.192	0.067	23.995	0.291
gender	0.254	0.062	24.019	0.252

Table 7: Analysis of Multiple Factor Regression on All Domains

As presented in Table 7, teachers' and students' reports of overall misconducts were significantly predicted by student's age, their endorsement of authority legitimacy, and their attitudes towards school (whether they like it or not). Misconducts were reported to be on increase with age in all domains except moral issues. More misconducts also were reported among students who endorsed less authority legitimacy on teachers. Finally, the more they liked the school, the fewer rule violations students conducted. These three variables accounted for 41.2% of the variance in misconduct.

Moral rule violations were significantly predicted by students' ratings of rule existence, evaluations of rules, whether they like school, and gender (as noted previously, boys engaged in more moral violations than girls). Together these variables accounted for 16.9% of the variance in reported moral misconducts.

Teacher- and student-reported misconducts in conventional domain were significantly predicted by students' endorsement on teacher authority, opportunity in rule decision making, and evaluations of rules. The less or lower those factors were, the more rule violations in conventional domain. Meanwhile, when growing older, students reported more conventional rule violations; boys and those who felt less liking to school reported more transgressions in conventional domain. Totally, these six variables accounted for 23.8% of the variance in teacher- and student-rated conventional misconducts.

More teacher- and student-reported misconducts in personal domain were significantly predicted by authority endorsement (being less), ratings of rule existence (being lower), age (being older), and living areas (living in big cities). These variables together accounted for 50.5% of the variance in personal transgressions.

As in other domains, reported contextually conventional misbehaviors increased with age; when students had less endorsement on teachers' authority, lower ratings of rule existence and opportunity in decision making, they engaged in more transgressions in contextually conventional domain; students lived in big cities engaged in more contextually conventional misconducts than students from countryside. These variables totally accounted for 31.4% of the variance in student - and teacher -reported misbehaviors in contextually conventional domain.

Prudential misconducts were significantly predicted by rule evaluation, age and gender. When students held a more negative attitude towards the reasonability of prudential rules, they were more likely to conduct misbehaviors; prudential misbehaviors also increased with students' age; female students reported more transgressions prudential domain than male students did. The three variables accounted for 25.4% of the total variance of misconducts in prudential issues.

4. Discussion

4.1 Evaluations of acts and rules

As anticipated, adolescents' conceptions of school and teacher authority were differentiated according to conceptual domains. Most adolescents viewed moral, conventional, and prudential issues as legitimately regulated by teachers and school administrators (above 80%), but less legitimacy was endorsed to contextually conventional issues (less than 70%), and personal issues gained the least legitimacy of teacher authority (less than 40%). Teachers had the highest authority to make rules on moral issues, followed by conventional and prudential issues; making rules was considered the most unreasonable and unacceptable in personal domain.

These results were consistent with former similar researches and may have important implications for moral and disciplinary education. For example, Killen et al. (1995) found that children and adolescents' acceptance or rejection to principles and rules made by adults in moral and conventional domain depended on their evaluation and judgment of these rules, demands or expectations. Even if they believed that they were reasonable, children and adolescents would actively evaluate the way adults pass these rules to see if these methods were appropriate with domains. When the adult's interpretations were domain-appropriate, adolescents were more likely to accept these rules and interpretations; otherwise they were more possible to reject them. Therefore, it's suggested that teachers keep the concept of domains in mind when react to students' transgressions and violations. Specifically, when a teacher deals with a students' moral violation, he/she should clearly states that their behaviors may harm others, infringe the rights of others, or may result in unfair consequences. In this way, rules and regulations are more likely to be recognized by students, thereby to terminate their violations in then situations and reduce violatons in the future; when a student commits a conventional violation, the teacher should focus on the importance of this conventional rule in maintaining class order and school functioning, rather than simply ordering or forbidding them; for a prudential violation, the teacher should focus on the possible harm and serious consequences to themselves; as for violations in contextually conventional and personal domains, teachers need to be particularly cautious when intervening because students are less receptive to teachers' authority in these two domains. What's more, if the teacher's approach is inconsistent with students' conceptual domains, they are more possible to transgress or resist.

There were main effects of grades in the legitimacy of teachers' authority, personal jurisdiction, rule evaluation and act assessment. Adolescents' personal jurisdiction increased with age in all domains, especially in prudential and contextually conventional domain. This is consistent with previous studies. Smetana (1988) found that the legitimacy of parental authority in the pre-youth (grades 5 and 6) and puberty (grades 7-12) decreased, while individual autonomy increased, especially in the personal and overlapping domains. Coupled

with the dramatic changes in their physical aspects, adolescents' desire and ability on their own body also enhance, and prefer to engage in some risky activities. Therefore, their personal jurisdiction in prudential domain increases dramatically. But parents often prefer to control those activities, which in turn leads to parent-child conflicts. So teachers should be aware of the adolescents' demand for autonomy and, within as far as possible, give them the right to make their own decisions, and avoid all-round control. For prudential events, teachers should make students understand the impulse of their own physical changes, and use real examples to explain the possible consequences of risky activities, rather than blindly force or control them; otherwise it will only lead to strong resistance from adolescents. In addition to the moral domain, adolescents from rural areas and cities had significant differences in the legitimacy of teacher authority, personal jurisdiction, rule evaluation and act assessment. In general, rural adolescents were more likely to recognize teachers' autonomy in conventional, contextually conventional, prudential and personal domains than urban adolescents. This supports some former researches. For example, Geogras et al. (1989) found that even in the same culture, urban subjects were more inclined to individualism, and traditional rural subjects were more inclined to collectivism. In other words, urban adolescents were more autonomous, and could insist on their own inner ideas to action. This may be related to urban students' relatively more extensive knowledge and experience, which is related to the family environment. In general, urban parents have higher culture level, and may adopt more democratic ways in child-raising. That's to say, urban families may pay more attention to cultivate the child's sense of autonomy and critical thinking, so urban adolescents' obedience to authority is more rational or less blind. While rural adolescents are more inclined to obey all the instructions from authorities, including parents and teachers, and tend to suppress their own personal feelings and thoughts. This suggests that students from different cultural backgrounds may have different cognitions and understandings to teacher authority. Teachers should be aware of those differences. On the other hand, it also shows that students' low acceptance to teacher authority is not necessarily a bad thing, because it may suggest that their critical thinking is developing.

4.2 The predicting factors of students' rule violation at school

While students' self-reported rule violations may be fewer than actual ones, it is interesting to note that in all domains, teachers reported fewer rule violations than students did. There are two possible reasons: First, because some disciplinary behaviors, such as smoking, drinking, etc. mostly occurred when teachers were absent, so they had no idea about this; another possible reason is that teachers were more inclined to provide answers according with social expectations. Teachers may regard students' behaviors as indications of their own educational performance; or because teachers often communicate with students, so lead to the so-called halo effect. However, there was a significant correlation between the teacher's assessment of rule violations and the

student's self-reported disciplinary behaviors. So it's reasonable to take the average of the two as a measure of student misconducts, and to reduce the shortcoming of self-report data to some extent. In the future study, it will be more scientific and comprehensive if some relatively objective observation data can be added.

In general, age was a powerful factor in predicting students' rule violations, which explained 24.5% of violations to school discipline. Except moral domain, rule violations in other domains increased with age, especially in contextually conventional domain ($R^2 = 0.207$), personal domain ($R^2 = 0.121$) and the conventional domain ($R^2 = 0.095$). This shows that adolescents' concept was in a period of rapid development, especially in personal and conventional domain. Therefore, teachers should take into account the characteristics of adolescent's psychological development when implementing school discipline, and fully consider their developmental requirements for personal jurisdiction. This is in line with the thoughts of developmental discipline. Watson(2008) proposed a developmental perspective in education—a sense of what the developmental tasks are for children of different ages and the appropriate role of adults in assisting the child's development. That's to say, children should be seen as being in the process of development and naturally want to build their understanding of their world, but they are still developing the competencies needed to succeed. From this perspective, student misbehaviors are mistakes, and mistakes are opportunities for learning. Meanwhile teachers can enhance children's moral understanding by involving them in developing class rules and norms and in the responsibility for seeing that the rules and norms function well. This is also a good way to satisfy their requirements for jurisdiction.

The legitimacy of teacher authority can also effectively predict adolescents' rule violations at school, which can explain 15.5% of the total transgressions. The legitimacy of teacher authority was an effective predictor of violations in conventional, personal and contextually conventional domain, especially in the personal domain, which can explain 28.8% of the transgressions; but it cannot effectively predict violations in moral and prudential domain. This is consistent with the characteristics of domains advocated by domain researchers, in which moral and prudential actions are judged by their inherent consequences (moral transgressions are judged by the harm to others; while prudential transgressions are determined by the harm to oneself). So the assessment of the legitimacy of the teacher's authority did not fundamentally affect the student's judgment of these acts, and naturally cannot effectively predict violations in these two domains. In contrast, the criteria for judging transgressions in conventional, personal, and contextually conventional domain are mainly external, rather than the consequences of the act itself. Therefore, whether students recognize the legitimacy of teacher authority can largely affect students' behaviors. When students highly accept the legitimacy of teacher authority, their violations in these domains will be fewer; on the contrary, when students have a lower recognition of the legitimacy of teacher authority, they may conduct more

transgressions. Therefore, the legitimacy of teacher authority in these domains need to be recognized by students before it can effectively influence the behavior of students. How to obtain teacher authority on those issues is the thing schools and teachers should focus on.

Whether like school or not was also an effective predictor of adolescent delinquent behaviors, especially in moral and conventional domain. The more students liked school, the less disciplinary behavior occurred, and vice versa. This is consistent with previous studies. Such as Smetana & Bitz (1996) found that the more positive the evaluation of school life, and the higher degree of participation in school activities, the less discipline violations occurred among students. Arsenio and others (1999) believed that emotion played an important role in moral development: when children and adolescents were in a warm and fair environment, they would experience the world as a beautiful place, thus promoting their moral development. When children help others, moral behaviors will bring them a positive and happy mood, so as to believe more that the world is warm and beautiful. On the contrary, if children and adolescents suffer from deception and peer exclusion for a long time, they will deviate from the normal moral and reciprocal behavior patterns and show aggression against others. This is also the reason why domain theory researchers attach great importance to emotions in moral development. Dewey (1975) also argued that the mode of social life and the nature of the school community were far more important factors in students' moral growth than direct moral instruction. So teachers must pay attention to students' emotional needs when carrying out school discipline and try hard to build a positive relationship with students. Only when students form positive emotions towards the school, they are more likely to comply with the rules and regulations of the schools.

In addition, gender also was a valid factor in predicting disciplinary behaviors in moral, conventional, personal and prudential domain. Boys were more likely to violate discipline in moral, conventional and prudential domain, which may be related to the patterns of behavior that have been analyzed above. Female students had more violations in personal domain than male students. This may relate to female students' enthusiasm in personal issues, such as hairstyles, jewelry wearing, etc.

References

- Smetana J. G. & Asquith P. 1994. "Adolescents' and Parents' Conceptions of Parental Authority and Personal Autonomy." *Child Development* 65(4):1147-62.
- Smetana J. G. & Bitz B. 1996. „Adolescents' Conceptions of Teachers' Authority and Their Relations to Rule Violations in School." *Child Development* 67(3):1153-72.

- Smetana J. G. 1998. „Concept of Self and Social Convention: Adolescents' and Parents' Reasoning about Hypothetical and Actual Conflicts,” in M. Gunnar & W. A. Collins” (Eds.), *Minnesota Symposium on Child Psychology* 21:79-122.
- Tisak M. S. & Turiel E. 1984. “Children's Conceptions of Moral and Prudential Rules.” *Child Development* 55(3):1030-9.
- Turiel E. 1983. *The Development of Social Knowledge: Morality and Convention*. Melbourne: Cambridge University Press.
- Watson M. 2008. “Developmental Discipline and Moral Rducation,” in L. P. Nucci & D. Narvaez (Eds.), *Handbook of Moral and Character Education*. New York: Routledge (175-203).
- Yu Zh. 2005. „Teacher Authority in the Eyes of Rights and Power.” *Journal of Jiangsu University* (Higher Education Study Edition) 2:57-60.
- Zhang L. & Li R. 2003. “On the Modern Transition of Teachers' Authority.” *Educational Research* 11:69-74.

Acknowledgments

This article is granted by School of English and Education,
Guangdong University of Foreign Studies.

Jianjin Liu
(Guangzhou)

Chinese Adolescents' Conceptions of Teacher's Authority and Their Relations to Rule Violations in School

Abstract: Based on the Social Cognitive Domain Theory, the paper explored the adolescents' conceptions of teacher authority in different domains and their relations to rule violations in school. The main results are: 1) Adolescents viewed moral, conventional, and prudential issues as legitimately subject to teachers' authority and personal issues as under personal jurisdiction, but they were equivocal about contextually conventional issues. 2) Seventh graders judged all acts as more legitimately subject to teachers' authority, all rule violations as more negative than did older students. 3) Compared with adolescents from big cities, adolescents from rural area viewed moral, conventional, contextually conventional, and personal issues as more legitimately subject to teacher authority, and endorsed less personal jurisdiction over those issues; but there were no significant differences in moral domain. 4) Male subjects reported more violations in conventional and prudential domain. 5) Adolescents' older age, less endorsement of legitimacy of teacher authority, and greater dislike for school predicted more teacher- and self-reported misconducts. Implications for moral education from these results were also discussed.

Keywords: Adolescents; Social Cognitive Domain Theory; conceptions of teacher authority; rule violations in school.

Ethics in Progress (ISSN 2084-9257). Vol. 9 (2018). No. 1, Art. #5, pp. 99-117.

Creative Commons BY-SA 3.0

Doi: 10.14746/eip.2018.1.5

An Empirical Study on Chinese and American College Students' Moral Value Recognition

Zhu Hailong

(Guangdong University of Foreign Studies, longhua_88623@126.com)

Introduction

College students are the heirs apparent to the will of the state and the traditional culture. Their identification of moral values not only affects strategic considerations of cultural safety, but also directly affects moral education. Given that China has a deep traditional moral culture and the United States has advanced cultural productivity, this study evaluated moral values among college students in both countries. Specifically, it distributed a survey on moral value recognition to 900 undergraduate students at 12 universities in China and the United States. The study's purpose is to discuss the current role of college students' moral values in Chinese and American identities, verify the possibility of generalizing traditional Chinese moral values, and provide a theoretical basis and data support for delivering moral education to college students.

This study tests whether American college students are willing to accept the moral values also contained in traditional Chinese culture, using experimental measurement and theoretical analysis. This analysis interprets the similarities and differences between Chinese and American college students' cognition of moral values, and explores the possibility of traditional Chinese moral values' generalizability.

Methods

Participants

This study selected Chinese (n=743, 354 males and 389 females) and American (n=157, 80 males and 77 females) undergraduate students as samples. The effective Chinese sample rate is 92.9%, and the Chinese participants come from Beijing Normal University, Xinjiang Financial University, Changsha Normal University, Three Gor-

ges University, Southwest University, Sun Yat-sen University, South China University of Science and Engineering, and Guangdong University of Foreign Studies. The effective American sample rate is 87.2%, and the American participants come from Illinois State University, Fordham University, Boston University, and The Ohio State University.

Measures

The study uses a self-designed questionnaire called the Survey of College Students' Chinese Moral Value Recognition. Because moral values consist of multidimensional psychological structures, a pre-test and factor analysis included interviews with moral education experts, teachers, and college students. Based on the pre-test results, the tested moral values were divided into six dimensions: honesty, civility & self-discipline, benevolence, unity & helping others, esteem, and filial piety. The questionnaire conforms to the core values of traditional Chinese moral culture and contains 31 items. Its Cronbach's alpha coefficient is 0.899 and split-half reliability is 0.867. The internal factors of the questionnaire are distinct and it has excellent structural validity. A 5-point Likert scoring method was used and tested by taking a unit. The SPSS 16.0 software was used to conduct statistical processing.

Results. Moral Value Comparison between Chinese and American College Students

Table 1 presents descriptive statistics on Chinese college students' recognition of the tested moral value dimensions. The scores followed this ranking, from high to low: honesty, civility & self-discipline, filial piety, unity & helping others, benevolence, and esteem.

	N	Minimum	Maximum	Mean	Std. Deviation
Total	743	2.71	4.90	4.055	0.408
Honesty	743	2.40	5.00	4.266	0.455
Civility & Self-Discipline	743	2.00	5.00	4.192	0.538
Benevolence	743	2.13	5.00	3.917	0.601
Unity & Helping Others	743	2.33	5.00	4.044	0.514
Filial Piety	743	1.60	5.00	4.131	0.516
Esteem	743	2.50	5.00	3.884	0.489
Valid	743				

Table 1: Chinese College Students' Moral Values Recognition

Table 2 shows descriptive statistics from the American students, who followed this ranking: honesty, civility & self-discipline, filial piety, and benevolence.

	N	Minimum	Maximum	Mean	Std. Deviation
Total	157	2.91	4.74	4.037	0.382
Honesty	157	3.00	5.00	4.245	0.434
Civility & Self-Discipline	157	2.33	5.00	4.089	0.604
Benevolence	157	2.13	5.00	4.008	0.537
Unity & Helping Others	157	2.17	5.00	3.910	0.615
Filial Piety	157	2.80	5.00	4.077	0.508
Esteem	157	2.75	5.00	3.940	0.475
Valid N	157				

Table 2: American College Students' Moral Values Recognition

These statistics show some differences in moral priorities between the two groups, as summarized in Tables 3 and 4.

Moral value dimension	Honesty	Civility and Self-Discipline	Filial Piety	Unity and Helping Others	Benevolence	Esteem
Chinese College Students	1	2	3	4	5	6
American College Students	1	2	3	6	4	5

Table 3: Overall Ranking of Chinese and American College Students' Moral Values Recognition

	Country	N	Mean	Std. Deviation	Std. Error Mean
Total	U.S.	157	4.038	0.382	0.031
	China	743	4.055	0.408	0.015
Honesty	U.S.	157	4.246	0.435	0.035
	China	743	4.267	0.456	0.017
Civility & Self-Discipline	U.S.	157	4.090	0.604	0.048
	China	743	4.193	0.538	0.020
Unity & Helping Others	U.S.	157	4.011	0.538	0.043
	China	743	3.917	0.602	0.022
Filial Piety	U.S.	157	3.910	0.616	0.019
	China	743	4.044	0.514	0.049
Esteem	U.S.	157	4.078	0.508	0.041
	China	743	4.131	0.517	0.019
	U.S.	157	3.940	0.475	0.038
	China	743	3.884	0.489	0.018

Table 4: Differences in Chinese and American College Students' Moral Values Recognition

As shown in Table 4, the average scores of Chinese college students are higher than those of American students in honesty, civility & self-discipline, unity & helping others, and filial piety, but lower than American students in benevolence and esteem. An independent samples T-test showed that there are significant differences between China and USA in the dimensions of civility and self-discipline and unity and helping others, as shown in Table 5:

	T	Sig (2-tailed)
Total	-0.517	0.606
Honesty	-0.542	0.588
Civility & Self-Discipline	-1.973	0.050*
Benevolence	1.936	0.054
Unity & Helping Others	-2.564	0.011*
Filial Piety	-1.199	0.232
Esteem	1.317	0.189

Table 5: Independent Samples Test

* P <0.05 significant; ** P <0.01 significant; *** P <0.001 significant

To verify the accuracy of the existing conclusion in greater depth, in accordance with the idea of same-sample measurement, we selected data from 157 of the Chinese students at random, and compared it to data from the 157 American students. Further statistical analysis showed although the specific numbers changed, the Chinese college students had higher scores than the American students overall, but they scored lower than the American students in benevolence and esteem. Thus, our earlier finding that Chinese students scored higher in the other four dimensions did not change substantially.

Data Analysis

From the statistical data, we find no significant cognitive differences between Chinese and American college students in the four moral value dimensions of honesty, benevolence, filial piety, and esteem. This verifies the generality of these traditional moral values originating in Chinese culture. However, at the same time there were significant differences in civility & self-discipline and unity & helping others.

Discussion. Similarities Between Chinese and American Students' Moral Values

The independent samples T-test showed a P value greater than 0.05 for both groups' views of honesty, benevolence, filial piety, and esteem, indicating that Chinese and American college students have no significant differences in these four dimensions of moral values. This similarity proves the generality of these traditional Chinese moral values. In fact, in 1989, a number of Western scholars (including 75 Nobel laureates) solemnly declared, "If humans want to survive in the 21st century and avoid confusion, it is necessary to look back 2500 years to the moral wisdom of Confucius" (Wang Dianqing 1999). From the results of this experiment, it is not difficult to find that the impact of moral values originating from traditional culture is extensive and far-reaching for both Chinese and American college students. In addition, the moral values of traditional Chinese culture not only have strong Chinese characteristics but also intersect with traditional American moral values. It is the common accomplishment of human moral values to meet the basic needs of individual moral development. Even though knowledge, benevolence, and courage belong to the Chinese "three kinds of character," while wisdom, justice, boldness, and temperance are attached to the ancient Greek "four main virtues," they undoubtedly all reflect the intersection of Chinese and western moral values. This similarity is a powerful counterblow to the alienation caused by value relativism within multicultural perspectives.

Chinese and American college students' recognition of traditional moral values benefits from the process of virtue ethics culture regression and the two countries' strengthening of moral education. China began to break up the inefficiency of college students' moral education with the idea of educating people with culture. This not only shaped the cultural atmosphere of upholding ethics, esteeming virtue, and doing good things, but also took traditional culture as the background for core socialist values,

cultivating college students' cultural awareness, cultural self-confidence, and cultural self-improvement. Lots of Confucian groups have likewise appeared in American giving American college students closer contact with the enduring charm of Chinese traditional moral culture. As for the US, several rounds of innovation since the 1920s have finally established traditional moral values education, giving American college students significant common ground with Chinese traditional moral values. It is easy for both groups to agree on certain moral values. At the same time, the unique function of Chinese cultural and moral wisdom in solving ecological imbalance, moral decadence, and interpersonal conflict enhances the confidence of Chinese college students in Chinese traditional culture. Thus, it is possible for the two countries' college students to develop value recognition.

Both groups of students recognized honesty, civility & self-discipline, and filial piety in the same order of priority, demonstrating that both cultures share these aspirations. It also proved that Chinese traditional culture has generality in dissolving cultural conflicts. Regarding the specific rankings, honesty is ranked highest due to the rapid development of economic globalization not bringing social and moral growth. Instead, the negative effects of value relativism, individualism, and materialism in society have become more and more serious. Selfishness, cheating, and treachery are common, which is sure to produce a lasting shock to college students. Thus, we can understand college students' desires for an authentic and honest social environment.

Civility and self-discipline is in second place, not only indicating an ardent yearning for civilization, but also reflecting their own educational attainments. In recent years, college students' immoral behaviors have gone beyond disrespect, skipping classes, and cheating: they turn to immoral values, behaving outrageously and even criminally. So, what is the standard of right or wrong? Where is the moral bottom line? (Spring, 1994). These problems which should not be problems are more prominent due to the force of value relativism. The need for civility is not only attracting social attention, but also leads to college students' self-examination. The ideas of putting one's mind in a proper and well-ordered condition, getting true ideas, acquiring knowledge, and studying systematically promote ethical behavior through the moral cultivation of see no evil, hear no evil, and speak no evil. This educational philosophy has a profound effect on students who suffered from value relativism.

Filial piety, the core of Chinese traditional morality is in third place. Chinese and American college students identified it as an important moral value, jointly proving that filial piety plays an irreplaceable role in maintaining human relationships and the process of showing affection. Although China and the United States have different political systems, living customs, cultural traditions, and expressions of filial piety, the Protestant ethic's understanding of the value of filial piety produces the biggest intersection with the concept of filial piety in Confucian moral philosophy, leading to this consensus.

Differences Between Chinese and American Students' Moral Values

As shown in Table 5, the P values for the dimensions of civility & self-discipline and unity & helping others are 0.050 and 0.011, respectively, indicating significant differences ($P > 0.05$) between the groups. Table 4 shows that the scores of the Chinese college students in these dimensions are 4.193 and 4.044, compared to 4.090 and 3.910 for the American college students, proving that Chinese student evaluate these values more highly than Americans. The source of this difference can be traced to each country's traditional culture and direction of moral education. In China, the traditional culture extends back five thousand years, and although it has fused with overseas cultures several times, it still retains the mainstream culture of Confucian virtue ethics. These ethics not only include moral requirements with the core of benevolence and etiquette, but also promote moral restraint for individual and social collectivity, a civil personality, and the idea that collective interests exceed personal interests. Such standards seek to unify families, govern countries, and ultimately match inner peace with world peace (Zhao Xudong 2003).

Moreover, the impact of Chinese traditional culture has long been infiltrated by outside behavioral constraints to the spiritual world, which benefited the construction of college students' moral values and promoted the formation of value recognition, civility & self-discipline, and unity & helping others through the subtle influence of culture. It is the foothold to deal with value relativism, individualism, materialism, and nihilism within a multicultural background. The positive effects of traditional culture can't be ignored. Starting in the 1980s, China began to pay more attention to the importance of traditional culture's moral influence on college students. In this century, unprecedented attention has been attached to the contemporary meaning of traditional culture, as reflected in the content of moral education and social practice of college students. This may be the reason why Chinese college students got higher scores in the value dimensions of civility & self-discipline and unity & helping others.

Compared with China, the United States is a very young country established on the basis of immigrants' multiculturalism. Democracy and freedom are the foundation of civic value, individualism is the long-standing cultural tradition, and moral values are considered as personal affairs in a 'private space' that should not be taught in school by force. Over time, the result has been a lack of preparation for dazzling cultural trends and fuzzy value judgments among American college students. What is the scale of civilization? What criteria should be followed? What is the ultimate goal of unity? Thus, misconduct would follow without a doubt. At the same time, the direction of American culture and moral education is erratic, directly affecting the construction of college students' moral values.

Because United States universities do not offer specialized, systematic, and explicit moral education, hidden moral education mainly relies on humanities and social work

courses. The direction of culture and moral education is fluctuating. From the character education which advocated traditional moral values before the 1920s to the progressive education in the 1920s and 30s, to value clarification education in the 1950s and 60s, to cognitive morality education in the 1970s and 80s, America finally returned to character education. For nearly a hundred years, moral education was wondering "what kind of moral value should be taught" and "how to get moral value" (Holl, 2003). As a consequence, they blurred the standard of judgment for moral values, lost to the invisible temptation of relativism. The question of why to hold certain moral values affected American college students' scores on these two value dimensions.

Improving Students' Awareness of Moral Values

Chinese college students' score for esteem was 3.884, their lowest of the six dimensions of moral value, while American college students scored lowest on unity and helping others (3.910). This proves that they all have deficiencies in their ethical levels. Esteem is a necessary virtue which should be insisted on during interpersonal communication. Unfortunately, judging from the misconduct among college students in recent years, many lack esteem consciousness. Since most Chinese undergraduates were born after the 1990s, they are often at the core of three generations living under one roof. Being spoiled caused the problems of self-centeredness, selfishness, poor ability to adapt, low levels of patience, and a fragile sense of cooperation that should be improved immediately. They are accustomed to a self-centered way of thinking: many students lack respect for fellow students and teachers, damaging communications and interpersonal relationships.

This bad result has affected college life, which should be happy. More seriously, the tragedies due to a lack of basic respect for life have an upward tendency, distorting the personality and mental health status of college students. The actions of Liu Haiyang, a college student at Tsinghua University who poured acid on several bears at a zoo, the Ma Jiajue homicide case at Yunnan University, Yao Jiaxing's intentional killing at Xian Conservatory of Music, and the poisoning case of a Fudan University postgraduate, without exception, exposed college students to disregard for other people's lives, a trend also evident in the increased campus suicide rates. These tragedies not only ruin young lives but also their families' happiness. Chinese moral education is severely lacking in the cultivation of the consciousness of respect. Now, we should enrich the content of college students' moral education. Beginning with respect for others' personality, then developing a respect for others' habits and cultural customs will achieve deeper esteem education. The goal is to make equal moral consciousness the spiritual core of Chinese college students.

For American college students, "unity and helping others" is their moral weakness. Since this dimension manifests as collective spirit more often, American college students who

are used to individualism and the pursuit of freedom feel uncomfortable with it to a certain extent. In recent years, the news that American college students avoid military service, keep away from NGO, hate governmental organizations, and frequently destroy public property prove their lack of unity consciousness. The lack of collectivist value education in American colleges is another cause of this behavior. The content of Dewey's progressive education theory included many traditional moral core values of the United States, and some of them have a close association with unity and helping others. The ecological association tradition is another active influence, so there are no serious flaws on college students' sense of unity. Various moral education innovations focused on some aspect of moral education methods, but ignored the leading role of core moral values to college students. Thus, when American society entered the fast lane of globalization, relativism gave college students who wanted to abandon the collective spirit a plausible moral reason. Fierce competition in society makes them give up unity and collective interests to achieve their goals. On the other hand, college students who are willing to stay together for collective interests become more and more scarce. It is not difficult to see that the collectivist education for American college students needs to be strengthened by the federal government, and college students also should cultivate the self-awakening of unity and helping others.

References

- Holl R. T. 2003. *Theory and Practice of Moral Education*. Trans. Lu Youquan. Hangzhou: Zhejiang Education.
- Spring J. 1994. *The American School 1642–1993*. New York: McGraw-Hill Companies.
- Wang D. 1999. *Oriental Moral Research (3rd Series)*. Beijing: China Industry and Commerce Publishing House.
- Zhao X. 2003. *Reflection on the Construction of Local Culture*. Beijing: Peking University Press.

Acknowledgements

This article was funded by The China national educational science' Thirteenth 5 Years' management 2018 project "Great civilian and values recognition in new era", No. BEA180114.

Zhu Hailong
(Guangdong)

*An Empirical Study on Chinese and American College Students'
Moral Value Recognition*

Abstract: This study used a self-designed questionnaire to empirically evaluate Chinese and American undergraduate students' recognition of moral values in six dimensions: honesty, civility & self-discipline, benevolence, unity & helping others, esteem, and filial piety. In total, 743 valid samples from 8 Chinese universities and 157 valid samples from 4 American universities were collected for an experimental comparison. Measurement results showed that the differences between Chinese and American college students were not significant in the dimensions of honesty, benevolence, esteem, and filial piety. These values, which originated in traditional Chinese culture, had the same priorities in both groups, confirming a certain degree of universality. However, significant intergroup differences existed in the civility and self-discipline dimension and the unity and helping others dimension. These results highlight the importance of enhancing Chinese college students' moral consciousness, especially with esteem, and of enhancing American students' consciousness of unity and helping others.

Keywords: Empirical study; China; American; college students; moral; values; recognition.

Ethics in Progress (ISSN 2084-9257). Vol. 9 (2018). No. 1, Art. #6, pp. 118-127.

Creative Commons BY-SA 3.0

Doi: 10.14746/eip.2018.1.6

Uwagi o sytuacji prawnej zwierzęcia pokrzywdzonego czynem zabronionym w polskim procesie karnym

Paweł Mazur

ORCID: <https://orcid.org/0000-0002-1011-121X>

(Uniwersytet im. Adama Mickiewicza w Poznaniu, pawelmn@amu.edu.pl)

Wstęp

Zwierzęta przez stulecia były traktowane przez człowieka w sposób skrajnie przedmiotowy. Jest truizmem powiedzieć, że prawa stanowione przez ludzi nie stały na straży ich interesów, polegających na minimalizowaniu czy unikaniu cierpienia. Na domiar złego zwierzęta, mimo oczywistego braku zdolności do zawinienia, pociągano również do odpowiedzialności karnej za przypisane im przestępstwa (Lech-Klementowski 2002, 23).

Zgodnie z panującym przez wieki światopoglądem, rolą zwierząt było służenie człowiekowi, jako władcy przyrody, dzierżącym nad nią niepodzielną władzę. Na stosunek do zwierząt w czasach nowożytnych fatalnie wpłynęła kartezjańska wizja zwierzęcia jako dobrze skonstruowanej maszyny, niezdolnej do odczuwania cierpienia – maszyny, której odgłosy podczas zadawania cierpienia powinny być interpretowane jako „skrzypienie albo trzeszczenie popsutego mechanizmu” (Tyburski 2015, 15; Probuska 2013, 182-183).

Od XIX wieku sytuacja prawna zwierząt ulega pewnej poprawie. Dość wspomnieć, że pod wpływem odkryć nauk przyrodniczych między innymi ewolucjonizmu, czy utilitaryzmu Jeremy'ego Benthama promującego sferę interesów zwierząt zmieniała się optyka prawodawcy. Powoli, acz konsekwentnie torowało sobie drogę przekonanie, że zwierzę to istota zdolna do odczuwania cierpienia i zasługująca z tego powodu na współczucie i ochronę prawną. Zaowocowało to wprowadzeniem szeregu zakazów okrutnego traktowania zwierząt obwarowanych sankcjami karnymi w prawodawstwach takich krajów, jak Szwajcaria, Francja, Włochy, a także Polska (Pietrzykowski 2011, 214).

Standardem prawnym obowiązującym w Polsce jest to, że pewne zachowania człowieka godzące w zwierzęta jak np. bicie, głodzenie czy zabijanie w sposób niehumanitarny są uznawane za społecznie szkodliwe w stopniu większym, niż znikomy.

Stanowią one przestępstwa ścigane przez państwo.

Kluczowym aktem normatywnym, który ma służyć ochronie zwierząt jest ustawa z dnia 21 sierpnia 1997 roku. Prawo o ochronie zwierząt (zwana dalej ustawą). Z jej przepisów można odtworzyć elementy norm sankcjonowanych w postaci zakazanych zachowań, jak przykładowo zakaz nieuzasadnionego lub niehumanitarnego zabijania zwierząt (art. 35 ust. 1 ustawy), zakaz znęcania się nad zwierzętami (art. 35 ust. 2 ustawy). Zakres zastosowania i normowania norm dekodowanych z ustawy dotyczy czynów zabronionych popełnianych na szkodę podtypu zwierząt w postaci kręgowców (art. 2 ust. 2 ustawy)

Samo stypizowanie pewnych zachowań człowieka jako zakazanych pod groźba kary nie przesądza jeszcze o skutecznej ochronie prawnej zwierząt. Równie istotne jest bowiem zapewnienie skuteczności ochrony prawnej udzielanej zwierzętom w procesie karnym, którego przedmiotem jest zarzut popełnienia przestępstwa na szkodę zwierząt. Jest to tym bardziej uzasadnione, że zwierzę jest istotą wrażliwą na cierpienie, lecz nie potrafiąca się rewanżować człowiekowi za wyrządzone zło, czy dochodzić swoich praw. Ija Lazari-Pawłowska wymownie podkreśliła że „kto znęca się nad zwierzęciem nie musi obawiać się, że kiedyś role się odwrócą. Zwierzę nie wystąpi w roli mściciela” (Lazari-Pawłowska 1992, 37).

Zadaniem wrażliwego i światłego prawodawcy faktycznego jest nieustanna rewizja obowiązującego ustawodawstwa, zmierzająca do zapewnienie zwierzętom optymalnej ochrony prawnej. Przyczynę do takiej rewizji może stanowić przedłożony artykuł. Tytułową problematykę zaprezentuje on w trzech częściach.

W pierwszym rzędzie przedstawiona zostanie obecna sytuacja prawna zwierząt jako istot pokrzywdzonych czynem zabronionym w procesie karnym. Następnie podjęta będzie próba rekonstrukcji założeń filozoficznych stojących za normami prawnymi, które wyznaczają pozycję zwierzęcia w postępowaniu karnym. Jest ona nieodzowna dla przedstawienia stosownej krytyki. Na koniec wskazane zostaną racje przemawiające za upodmiotowieniem zwierzęcia w procesie karnym, polegającym na przyznaniu mu statusu pokrzywdzonego jako strony postępowania karnego. Gdyby prawodawca zechciał podjąć tak doniosłą decyzję prawną, wówczas wcieliłby w życie postulat upodmiotowienia zwierząt, który pojawiła się pod koniec XX wieku pod wpływem filozofii Toma Reagana i na gruncie polskim jest przekonująco promowany przez Andrzeja Elżanowskiego i Tomasza Pietrzykowskiego (Elżanowski & Pietrzykowski 2013, 18-27).

W ramach uwag wstępnych wypada dodać, że posłużenie się w tytule artykułu słowem czyn zabroniony zamiast przestępstwo nie jest dziełem przypadku. Normę sankcjonowaną zakazującą określonych zachowań można przekroczyć również czynem niedowolnym, w stanie niepoczytalności (Pohl 2007, 83).

Zasadniczym celem procesu karnego jest realizacja zasady trafnej reakcji karnej. Polega ona na wykryciu i pociągnięciu do odpowiedzialności karnej sprawcy przestępstwa i ochronieniu od tej odpowiedzialności osoby niewinnej (Cieślak 1984, 213). Innym istotnym celem tego procesu jest zabezpieczenie prawnie chronionych interesów pokrzywdzonego (art. 2 § 1 pkt 3 kodeksu postępowania karnego).

Jak zauważyła Małgorzata Żbikowska, rzeczona ochrona powinna być udzielana pokrzywdzonemu na kilku poziomach: po pierwsze, na poziomie stanowienia prawa (aspekt formalnoprocesowy). Treść przepisów ustanawianych przez prawodawcę musi przeto uwzględniać interesy tego uczestnika procesu karnego. Po wtóre, na poziomie wykładni, czyli dekodowania norm prawa z przepisów prawnych; i po trzecie, na poziomie stosowania prawa, czyli praxis (Żbikowska 2015, 108).

Pokrzywdzony jest uznany za podmiot, który ma określony interes prawny wyrażający się w uzyskaniu dlań korzystnego rozstrzygnięcia. Jest z mocy prawa stroną w postępowaniu przygotowawczym (pierwszym stadium procesu karnego) i może występować jako strona w postępowaniu rozpoznawczym. Warto odnotować, że "stroną jest podmiot, który w procesie bądź sam występuje w swoim imieniu, bądź w którego imieniu działa ktoś inny (przedstawiciel) i który ma interes prawny w korzystnym dla siebie rozstrzygnięciu o przedmiocie procesu" (Daszkiewicz 2000, 205).

Status pokrzywdzonego ma charakter obiektywny i jest uzależniony od tego, czy w konkretnej sytuacji dany podmiot jest rzeczywiście pokrzywdzonym. Dla takiego uznania zasadnicze znaczenie ma ustalenie przedmiotu ochrony (Stefański 2017, 697). Jego status kształtuje zdarzenie, które ma miejsce poza procesem karnym (Janusz-Pohl 2017, 234)

Czy prawodawca do grona podmiotów pokrzywdzonych zaliczył również zwierzęta chronione rzeczonymi normami sankcjonowanymi?

Odpowiedzi jednoznacznie negatywnej udziela definicja legalna pokrzywdzonego zawarta w przepisie art. 49 §1 k.p.k. Zgodnie z przepisem art. 49 § 1 k.p.k. pokrzywdzonym jest osoba fizyczna lub prawna której dobro prawne zostało bezpośrednio naruszone lub zagrożone przez przestępstwo. Stosownie do §2 wspomnianego przepisu pokrzywdzonym może być także niemająca osobowości prawnej instytucja państwowa lub samorządowa, inna jednostka organizacyjna której odrębne przepisy przyznają zdolność prawną.

Łatwo dostrzec, że w definiens „pokrzywdzonego” nie znalazły się zwierzęta. Ma to oczywiście negatywne konsekwencje dla sytuacji prawnej tej grupy istot żywych. Ta definicja wiąże bowiem organy procesowe zajmujące się ściganiem czynów zabronionych wymierzonych w zwierzęta i osądzaniem ich sprawców (sąd, prokurator, policja). Definicje legalne są wszakże normami nakazującymi określonym podmiotom, mianowicie interpretatorom aby pewnym wyrazom, czy wyrażeniom nadawali odpowiednie znaczenie (Zieliński 2017, 176-177). Wielowymiarowa dyrektywa uwzględniania prawnie chronionych interesów pokrzywdzonego nie dotyczy zatem zwierząt.

Jaki jest status zwierzęcia pokrzywdzonego i kto występuje jako podmiot w

procesach wszczętych z powodu przestępstw popełnianych przeciwko nim?

W przepisie art. 1 ust 1 ustawy zawarta jest następująca definicja zwierzęcia. Jest to mianowicie: "istota żyjąca, zdolna do odczuwania cierpienia, nie jest rzeczą. W tym samym paragrafie ustawodawca zdecydował, że człowiek jest mu winien poszanowanie ochronę i opiekę. Zwierzęta zostały wyłączone zatem z klasy przedmiotów materialnych (Nazar 2002, 134-135). Takie zdefiniowanie zwierzęcia jest uznaniem przez prawodawcę prawdy o świadomości doznaniowej zwierząt i jej moralnego znaczenia (Łętowska 1998, 113). Jednakowoż efekty tak powściągliwej derefikacji są w istocie mizerne. Zwierzęta wszak nie zyskują podmiotowości, co obniża efektywność ich ochrony prawnej (Elżanowski & Pietrzykowski 2013, 25). Szczególnie uwidacznia się to na gruncie procedury karnej.

Trudno odmówić racji Michałowi Gabrielowi Węglowskiemu twierdzącemu, że zwierzę może być istotą faktycznie pokrzywdzoną przez człowieka (Węglowski 2008, 158). Jednakże w sposób podmiotowy w procesach karnych zainicjowanych z powodu pokrzywdzenia zwierzęcia traktowany jest tylko człowiek.

Tego rodzaju wykładnię przepisów prawa przedstawił Sąd Najwyższy, którego orzeczenia mają istotną wagę dla organów stosujących prawo, zwłaszcza sądów powszechnych. W wyroku z dnia 16 stycznia 2014 roku VKK 370/13 wydanym na kanwie sprawy karnej, w której zarzucono oskarżonemu znęcanie się nad 23 końmi, Sąd Najwyższy wyraził pogląd, że zwierzę nie jest pokrzywdzone, ponieważ nie jest „osobą” w rozumieniu przepisu art. 49 § 1 k.p.k. Zwierzę nie jest też rzeczą, o czym stanowi art. 1 ust. 1 ustawy. Jednakże Najwyższa Instancja Sądowa uznała, że zwierzę hodowlane lub domowe należy postrzegać jako przedmiot własności lub posiadania we wszystkich jego formach (tj. posiadania samoistnego i zależnego, ponieważ w sprawach nieuregulowanych w tej ustawie stosuje się odpowiednio przepisy dotyczące rzeczy, zob. art. 1 ust. 2 ustawy). Sąd Najwyższy wywiódł, że jeśli obowiązujące prawo chroni własność, jak i stan spokojnego posiadania, to pokrzywdzonym czynem polegającym na znęcaniu się nad zwierzęciem jest przede wszystkim jego właściciel lub posiadacz (Wyrok SN z dnia 16 stycznia 2014 roku VKK 370/13 Legalis, także Węglowski 2008, 157-163). Zdaniem Michała Gabriela Węglowskiego za pokrzywdzonego można uznać również człowieka, który uprawdopodobni wpływ czynu zabronionego popełnionego na szkodę zwierzęcia na swoje dobra prawne (Węglowski 2008, 159-160).

Kompetencje do podejmowania czynności procesowych mają nie tylko wyżej wymienieni pokrzywdzeni. W sprawach o przestępstwa określone przepisach art. 35 ust. 1, 1 a lub 2 oraz wykroczenia określone w art. 37, a także w postępowaniu w sprawach nieletnich o czyn karalny określony w art. 35 ust. 1, 1 a lub 2 prawa pokrzywdzonego może wykonywać organizacja społeczna, której statutowym celem jest ochrona zwierząt. Dotyczy to również sytuacji, w których nie ma podmiotu pokrzywdzonego, np. w wypadku znęcania się nad zwierzęciem bezdomnym (Węglowski 2008 161).

Konfiguracja osób występujących jako pokrzywdzony bądź jako organizacja stojąca na straży praw zwierząt jest dość urozmaicona. Jednakże niezależnie od takiego, czy innego

układu procesowego do kategorii podmiotów pokrzywdzonych nie zalicza się zwierząt. Perspektywa prawodawcy w tym wypadku jest zatem na wskroś antropocentryczna.

II.

Jakie założenia filozoficzne kryją się za taką, a nie inną definicją pokrzywdzonego oraz zwierzęcia przedstawioną w rzeczonych przepisach prawa? Każdy porządek prawny oparty jest na pewnych metaprawnych przekonaniach ontologicznych, aksjologicznych i epistemologicznych prawodawców, bez których podejmowane przez nich działania i treść stanowionych norm byłaby nonsensowna” (Pietrzykowski 2017, 176). Ich treść jest konsekwencją przyjmowania przez prawodawcę pewnego wyobrażenia świata, kreującego poziom ontologicznych i aksjologicznych presupozycji konkretnych instytucji prawnych (Pietrzykowski 2015, 248; Sarkowicz 1995, 159-172). Poziom presupozycji. obejmuje ”wszelkie informacje o współczesnym prawodawcy, świecie, społeczeństwie, człowieku, jego poglądach i wartościach” (Sarkowicz 1995, 67). Presupozycje mają wymiar ontologiczny. Dotyczą bowiem sfery faktów o człowieku, społeczeństwie. Presupozycje aksjologiczne obejmują wartości przyświecające prawodawcy (Sarkowicz 1994, 67). Ich odtworzenie jest możliwe poprzez wykładnię rzeczonych przepisów na poziomie presupozycji (Sarkowicz 1995, 159-172).

Presupozycje ontologiczne kryjące się za powyższymi przepisami przedstawiają obraz świata zamieszkałego przez ludzi i zwierzęta. W tym świecie, jedynie człowiek bądź twory mające służyć jego interesom majątkowych np. osoby prawne, są uznane za podmiot praw w procesie karnym. Organy procesowe mają obowiązek dbałości o ich interes procesowy na wymienionych płaszczyznach, to jest na poziomie wykładni przepisów prawa i ich stosowania. Prawodawca faktyczny natomiast taką troskę musi przejawiać na poziomie stanowienia prawa.

Z rzeczonych przepisów wyłania się wizja świata, w którym zakazane jest pod groźbą odpowiedzialności karnej między innymi zabijanie zwierząt kręgowców bez powodu, jak również dręczenie i głodzenie. Nie są one traktowane rzeczowo. Wszelako, są podporządkowane człowiekowi i ich krzywda jest relatywizowana do tego, na ile uderza również w dobra prawne człowieka.

Presupozycje aksjologiczne są oparte, jak wydaje się na humanizmie prawniczym, który – jak zauważa Tomasz Pietrzykowski – jest „filozofią prawa” współczesnego prawodawstwa. Opiera się on na przekonaniu, że człowiek jest szczególnie wyróżnionym rodzajem bytu, wyraźnie różnym od innych istot. Z tego powodu jedynie istotom ludzkim przysługuje status podmiotów prawa. Jedynie ludziom przysługuje status podmiotu z uwagi na ich przyrodzoną godność. Prawo postrzegane jest jako instytucja poświęcona zaspokajaniu potrzeb i pragnień człowieka (Pietrzykowski 2015, 21). Prawo karne procesowe uznaje jedynie interesy człowieka jako godne efektywnej ochrony. W konsekwencji człowiek ma status pokrzywdzonego wówczas, gdy zachowanie godzące w zwierzęta godzi w jego dobra (zwierzę jako własność, zwierzę jako istota, z którą

człowieka łączy szczególna więź emocjonalna.)

Założenia filozoficzne ukryte za przedstawionymi rozwiązaniami prawnymi mają charakter hybrydalny. Z jednej strony są wszakże konsekwencją przyjęcia oczywistych praw nauk przyrodniczych, mianowicie zdolności zwierzęcia do wielowymiarowego przeżywania otaczającej go rzeczywistości między innymi poprzez cierpienie fizyczne, czy równie dojmujące cierpienie psychiczne. Z drugiej strony, mimo tej świadomości, nie postawiono przysłowiowej kropki nad „i”, uznając podmiotowość prawną zwierząt. Nie jest też jasne, co jest w istocie przedmiotem ochrony prawnokarnej na gruncie przepisów karnych ustawy.

Romuald Kmiecik ma rację wskazując, że problematyczne jest to, czy przedmiotem ochrony jest zwierzę jako istota zdolna do odczuwania cierpienia, czy też zwierzę jako „obiekt wykonawczy działania naruszającego wzorce moralne, które zabraniają ludziom zadawanie cierpienia (znęcania się, torturowania) także innym istotom” (Kmiecik 2002, 187). Wydaje się, że z uwagi na dość podrzędną pozycję zwierzęcia i kluczową rolę człowieka w procesie karnym, wizja prawodawcy jest bliższa uznaniu, że szkodliwość praktyk przeciwko zwierzętom wynika raczej z tego, że godzą one raczej w ludzi, jak przyjmował to Immanuel Kant, zamiast stanowić samoistne zło, godzące w istoty wrażliwe i cierpiące.

Tego rodzaju niejasność polegającą na braku wyraźnego rozstrzygnięcia przez prawodawcę, jakie wartości są chronione przez przepis stanowi lukę aksjologiczną *intra legem* (Skuczyński & Zientara 2012, 217).

Obecny status zwierzęcia jako istoty faktycznie pokrzywdzonej w procesie karnym wymaga ponownego przemyślenia. Szczególnie aktualne w tym kontekście jest pytanie postawione przez Jana Białocerkiewicza, czy dereifikacja zwierząt jest ostatnim paradygmatem, czy też otwiera drogę do personalizacji (Białocerkiewicz 2005, 59). Decyzja w tym względzie należy do prawodawcy. Petryfikacja obecnego stanu prawnego oznaczałoby w istocie sztywność poznawczą prawodawcy faktycznego zamkniętego na dynamiczny postęp nauk przyrodniczych, gotowych do weryfikowania wyników badań o zwierzętach w świetle miarodajnych kryteriów intersubiektywnej kontrolowalności. Nadanie statusu podmiotu nie jest tożsame z decyzją arbitralną. Jest to wynik kierowania się przez normodawcę pewnymi racjami (Pietrzykowski 2013, 17). Z kolei wartość każdej jednostki ludzkiej i poszanowanie jej praw są ufundowane na podstawach metafizycznych zakładających przyrodzoną godność człowieka (Pietrzykowski 2013, 249). Jest to współcześnie przyjmowane bez zastrzeżeń – jednak pozostaje dalekie od oczywistości w odniesieniu do zwierząt.

Uznanie zwierząt za podmioty prawa wymaga stosownego uzasadnienia aksjologicznego opartego ocen moralnych, żywionych przez prawodawcę (uzasadnienie aksjologiczne heteronomiczne (Ziembinski & Zieliński 1988, 151). Musi być ono na tyle przekonujące, aby zyskało aprobatę i zostało zinternalizowane przez adresatów norm prawnych.

Podstaw do rewizji ocen moralnych przemawiających za upodmiotowieniem

zwierzęcia dostarczają odkrycia nauk przyrodniczych, szczególnie neurobiologii. Rzeczona nauki odsłaniają prawodawcy faktycznemu prawdy o złożoności życia zwierząt i ich wrażliwości na rozmaite doznania, a także niekiedy zdolnościach do empatii, czy zachowań altruistycznych. Empatia, która winna skłonić do zmiany ocen moralnych opiera się zatem na obserwacji na wiedzy naukowców, a nie na uczuciach i opiniach (Hołówka 2015, 73).

Istnieją racje przemawiające za tym, by aplikować na gruncie procedury karnej koncepcję zwierzęcia jako podmiotu nieosobowego i uznać sferę jego interesów za wymagającą skutecznej ochrony w procesie karnym. Różni się on od przedmiotu tym, że jest zdolny do świadomego odczuwania ma swoje własne subiektywne interesy, w tym, aby jego dobrostan był jak najwyższy. Jego interesy nie muszą być wcale uświadamiane (Pietrzykowski 2015, 253).

Nauki przyrodnicze, a zwłaszcza neurobiologia dostarczają coraz więcej przykładów, że zwierzęta są zdolne do odczuwania bólu i emocji. Takie istoty jak kręgowce, a być może również zwierzęta bezkręgowce, mają świadomość. Jej ewolucja sprawiła, że niektóre pozytywne wartości biologiczne, jak np. pokarm czy schronienie, oraz negatywne wartości biologiczne (np. zagrożenie) zaczęły być postrzegane przez zwierzęta jako dobre i złe. Co więcej, niektóre z nich – przykładowo, delfiny, słonie i szympansy – mają samoświadomość czy umiejętność współodczuwania. Ludzie dzielą z pozostałymi ssakami podstawowe doznania negatywne generowane w tzw. układzie limbicznym, takie, jak uczucie bólu, strachu i niepokoju. Również z ptakami homo sapiens dzieli główne kategorie źródeł przyjemności, jak np. pozaseksualne kontakty społeczne, zachowania apetycyjne i eksploracyjne (własna aktywność) oraz zabawę. W mózgu kręgowców nieustannie następuje konwersja wartości biologicznych w doznania. To one są odpowiedzialne za pragnienia, a zatem i interesy (Elżanowski & Pietrzykowski 2013, 20-22). Z kolei te emotywne preferencje i rozpoznania spotykają się u człowieka z procesami racjonalno-decyzyjnymi zachodzącymi w najmłodszym ewolucyjnie, przedczołowym płacie kory mózgowej.

Świadomości zwierząt nie da się jednak zredukować do czystej sensytywności. Zwierzęta domowe, należące wszak nie tylko do ssaków, mają swoje oczekiwania, doznają zawodu lub radości; wykonują one ponadto operacje poznawcze, wykazują empatię bądź jej brak. Trudno tę mentalną złożoność zwierząt wyższych sprowadzić do instynktów i uczenia się wyłącznie za pośrednictwem odruchów warunkowych (Woleński 2015, 29).

Zdolności te i procesy pozostają niepodważalnym faktem również w świetle tego, że zwierzęta nie mają świadomości własnych praw, w tym takich, które posiadałyby w procesie karnym. Jan Woleński rozprawiając się z tym argumentem słusznie zauważył, że uprawnienia przypisuje się również osobom prawnym. Wśród nich są zaś takie, które nie odnoszą się do ludzi (Woleński 2015).

Wreszcie nie sposób nie dostrzec, że w przepisach kodeksu postępowania karnego występują pewne kategorie podmiotów niezdolnych do podejmowania czynności

procesowych, np. osoby ubezwłasnowolnione całkowicie bądź osoby niepełnoletnie (art. 51 § 1 i 2 k.p.k.). Są one reprezentowane przez przedstawicieli procesowych właśnie z racji niezdolności do wykonywania swoich uprawnień. Ich interes prawny i podmiotowość nie budzą wątpliwości i są postrzegane, respektowane, wreszcie realizowane przez podmioty posiadające zdolność do czynności procesowych.

Za uznaniem zwierząt za podmioty na płaszczyźnie procesu karnego przemawia nadto konieczność skorygowania pewnej niekonsekwencji aksjologicznej prawodawcy. Oto bowiem przepisy procedury karnej chronią interesy ekonomiczne człowieka pozwalając występować w roli pokrzywdzonym różnym twórcom organizacyjnym jak osoby prawne. Jednocześnie zwierzęta podobne do człowieka, jeśli chodzi o sposób przeżywania świata, nie mają tak korzystnej pozycji. Taka hierarchia wartości prawodawcy faktycznego w świetle aktualnej wiedzy o zwierzęciu jawi się jako moralnie wątpliwa.

Postulat upodmiotowienia zwierząt w procesie karnym nie polega naturalnie na uznaniu zwierzęcia za podmiot kompetentny do podejmowania czynności procesowych. Tego typu czynności mogą wszak podejmować jedynie istoty wyposażone w zdolności racjonalne, w tym – w zdolność rozpoznania reguł formalizacji i konwencjonalizacji przesądzających o ważności i prawidłowości czynności procesowych (Janusz-Pohl 2017, 62).

Pokrzywdzone – zwierzę nie stanowiłoby typowej strony procesowej, określonej w nauce procesu karnego jako uczestnik który działa we własnym imieniu, a nadto ma interes prawny w określonym rozstrzygnięciu o procesie (Grzegorzczak 1998, 5-6). Podmiotowość prawna przyznana zwierzęciu mogłaby mieć tylko wymiar bierny. Polega on na posiadaniu prawnie chronionych interesów, które przyoblekają postać uprawnień danego podmiotu prawnego, za których naruszenie inne podmioty ponoszą (Pietrzykowski 2013 25).

Zdolność do czynności procesowych przysługuje wyłącznie ludziom pełnoletnim i nieubezwłasnowolnionym (Janusz-Pohl 2017, 244). Tylko te podmioty mogłyby zostać wyposażone w kompetencję do podejmowania czynności procesowych na korzyść zwierząt podmiotów nieosobowych. Oczywiście pozostaje otwarte pytanie czy takim reprezentantem powinien być każdy pełnoletni i nieubezwłasnowolniony człowiek, który współodczuwa ze skrzywdzonym zwierzęciem, czy też czuje się za nie odpowiedzialny, czy ze względów pragmatycznych służących sprawności postępowania karnego jedynie wybrane podmioty.

Ewa Łętowska nie bez racji podkreślała, że skuteczność regulacji prawnych chroniących zwierzęta zależy od tego, jak sprawnie zostanie zorganizowany aparat, dzięki któremu rzeczono uprawnienia będą wykonywane (Łętowska 1997, 91). Być może zatem odpowiedzialność za wykonywanie uprawnień zwierząt pokrzywdzonych-nieosobowych podmiotów prawa powinien spoczywać na specjalnie do tego powołanej instytucji.

Zakończenie.

Upodmiotowienie nieosobowe stanowiłoby kolejny i śmiały krok w kierunku dalszej dereifikacji zwierząt i zapewnienia im efektywnej realnej ochrony ich praw.

Tego rodzaju zmiana nie stanowi z pewnością spełnienia oczekiwań przedstawicieli radykalnego nurtu praw zwierząt postulującego uznanie wielu praw zwierząt, w tym fundamentalnego prawa do życia. Być może nawet rozczaruje niejednego orędownika praw zwierząt swoją powściągliwością.

Jednakże, jak słusznie zauważyli Tomasz Pietrzykowski i Andrzej Elżanowski, "niemożność natychmiastowej przemiany świata w arkadię zwierząt nie powinna stać na przeszkodzie podejmowaniu działań, dzięki którym może on stać się przynajmniej w nieco mniejszym stopniu ich „wieczną Treblinką” (Elżanowski & Pietrzykowski 2013, 25). Takim działaniem jest właśnie upodmiotowienie zwierzęcia poprzez poszerzenie zakresu definicji pokrzywdzonego w przepisie art 49 kodeksu postępowania karnego.

Zmiana definicji legalnej pokrzywdzonego w postulowanym kierunku wzmocniłaby ochronę prawną zwierząt, choćby poprzez włączenie troski o ich prawnie chronione interesy do wartości, które ma realizować proces karny.

Wydaje się, iż rzeczona zmiana wpłynęłaby także na język prawniczy, a tym samym na praktykę prawniczą użytkowników i kreatorów tego języka.

Do języka prawnego należą wszelkie wyrażenia występujące w jakimkolwiek tekście prawnym. Wprowadzenie aktu prawodawczego zawierającego choćby jedno nowe wyrażenie wzbogaca ten język (Patryas 1997, 108-109). Tworzenie ważnego języka prawniczego opiera się zaś na języku prawnym (Patryas 1997, 114). Języki prawne mają swoje odpowiedniki wchodzące w skład języka prawniczego. Taki składnik Wojciech Patryas określa mianem przedmiotowego odpowiednika swojego wzorca. W skrajnym przypadku prawny wzorzec jest identyczny ze swoim przedmiotowym odpowiednikiem. Z reguły jednak jest tak, że prawny wzorzec jest jedynie fragmentem swojego przedmiotowego odpowiednika. Odpowiednik zawiera z kolei wszystkie reguły swojego prawnego wzorca, a nadto konstytuują go dodatkowe reguły językowe (Patryas 1997, 112). Można zaryzykować stwierdzenie, że zmiana definicji projektującej pokrzywdzonego, polegająca na włączeniu *definiens* zwierząt, wpłynęłaby na korzystne dla zwierząt zmiany języka prawniczego, wzbogacając ten ostatni o nowe pojęcie zwierzęcia pokrzywdzonego w procesie karnym.

Tego rodzaju zmiana dostarczyłaby argument na rzecz uznania, że przedmiotem ochrony prawnokarnej przepisów ustawy o ochronie zwierząt jest – zapożyczając semantykę od Iji Lazari-Pawłowskiej – „autonomiczne dobro samych zwierząt a nie tak czy inaczej pojmowana korzyść ludzi” (Lazari-Pawłowska 1992, 33).

Literatura

- Elżanowski A. & Pietrzykowski T. 2013 „Zwierzęta jako nieosobowe podmioty prawa”.
Forum Prawnicze 1:18-28.
- Daszkiewicz W. 2000. *Prawo karne procesowe. Część ogólna*. Poznań – Bydgoszcz:
Wydawnictwo Brandt.
- Grzegorzczak T. 1998. *Strony, ich procesowi przedstawiciele i inni uczestnicy
postępowania karnego*. Warszawa: Wydawnictwo Naukowe PWN.
- Hołówka J. 2015. „Czy istnieją zwierzęta ze statusem osoby”. *Przegląd Filozoficzny* 2:65-
88.
- Janusz-Pohl B. 2017 *Formalizacja i konwencjonalizacja jako instrumenty analizy
czynności procesowych w prawie polskim*. Poznań: Wydawnictwo Naukowe UAM.
- Kmieciak R. 2002. „Dereifikacja zwierząt czy antropomorfizm prawniczy?”, w: M.
Mozgawa (red.), *Prawna ochrona zwierząt*. Lublin: Wydawnictwo Verba (185-
187).
- Lazari-Pawłowska I. 1992. *Etyka. Pisma wybrane*. Wrocław – Warszawa – Kraków:
Zakład Narodowy im. Ossolińskich.
- Lech-Klementowski M. 2000. „O odpowiedzialności karnej zwierząt i ich udziale w karze
w średniowiecznej i wczesnonowożytnej Europie”, w: M. Mozgawa (red.), *Prawna
ochrona zwierząt* (19-43).
- Łętowska E. 1997. «Dwa cywilnoprawne aspekty praw zwierząt: dereifikacja i personi-
fikacja», w: A. Szpunar (red.), *Dwa cywilnoprawne aspekty praw zwierząt. Dereifik-
acja i personifikacja. Studia z prawa prywatnego. Księga pamiątkowa*. Łódź: Wy-
dawnictwo Naukowe Uniwersytetu Łódzkiego (71-93).
- Nazar M. 2002. „Normatywna dereifikacja zwierząt-aspekty cywilnoprawne”, w: M.
Mozgawa (red.), *Prawna ochrona zwierząt* (129-154).
- Pohl Ł. 2007. *Struktura normy sankcjonowanej w prawie karnym*. Poznań: Wydawnictwo
Naukowe UAM.
- Węglowski M. G. 2008. *Przestępstwa przeciwko humanitarnej ochronie zwierząt*.
Warszawa: Wydawnictwo Dom Organizatora.
- Woleński J. 2015. „Szowinizm gatunkowy, humanitaryzm i animalocentryzm”. *Przegląd
Filozoficzny* 2:25-34.
- Patryas W. 1977. *Definiowanie pojęć prawnych*. Poznań: Wydawnictwo Naukowe UAM.
- Pietrzykowski T. 2011. *Etyczne problemy prawa*. Warszawa: Wydawnictwo Lexis Nexis.
- Pietrzykowski T. 2015. „Problem podmiotowości prawnej zwierząt z perspektywy
filozofii prawa”. *Przegląd Filozoficzny* 2:247-259.
- Pietrzykowski T. 2013. „Podmiotowość prawna-ujęcie teoretyczne”, w: A. Bielska-
Brodziak (red.), *O czym mówią prawnicy mówiąc o podmiotowości*. Katowice:
Wydawnictwo Uniwersytetu Śląskiego (15-31).

- Pietrzykowski T. 2017. *Naturalizm i granice nauk prawnych. Esej z metodologii prawoznawstwa*. Warszawa: Wolters Kluwer.
- Probučka D. 2015. *Prawa Zwierząt*. Kraków: Universitas.
- Skuczyński P. & Zientara A. 2012. "Prawnokarna ochrona zwierząt a filozoficzny i teoretycznoprawny problem wartości i praw podmiotowych", w: T. Gardocka & A. Gruszczyńska (red.), *Status zwierzęcia : zagadnienia filozoficzne i prawne*, Toruń: Wydawnictwo Adam Marszałek (202-227).
- Żbikowska M. 2015. *Zasada lojalności w procesie karnym w odniesieniu do pokrzywdzonego*. Toruń. Wydawnictwo Adam Marszałek.
- Zieliński M. & Ziemiński Z. 1988. *Uzasadnienie twierdzeń ocen i norm w prawoznawstwie*. Warszawa: Wydawnictwo Naukowe UAM.
- Zieliński M. 2017. *Wykładnia prawa. Zasady – reguły – wskazówki*. Warszawa: Wydawnictwo Wolters Kluwer.

Paweł Mazur
(Poznań)

*Remarks About Legal Situation of Animal Injured By Prohibited Act
in Polish Criminal Process*

Abstract: The article is focused on the issues relating to a legal situation of an animal as a victim of a criminal offense. Currently, according to the rules of criminal procedure, an animal, as an entity, cannot be legally recognised a victim, which significantly lowers efficiency with which its interests can be protected. There are strong moral and pragmatic reasons to grant animals the status of a non-personal carrier of legal rights. Legal personality of animals as victims during a criminal trial should be of only passive nature, allowing recognition of legally-protected interests of the animal.

Keywords: animal; non-personal carriers of legal rights; criminal process; interests.

Ethics in Progress (ISSN 2084-9257). Vol. 9 (2018). No. 1, Art. #7, pp. 128-139.

Creative Commons BY-SA 3.0

Doi: 10.14746/eip.2018.1.7

Imperatyw kategoryczny w obronie praw zwierząt. Ekstensjonizm etyczny Christine Korsgaard

Małgorzata Adamska

(Uniwersytet im. Adama Mickiewicza w Poznaniu, adamska.gos@gmail.com)

Wprowadzenie

Niniejszy artykuł podejmuje problematykę praw zwierząt, a zatem i ich podmiotowości moralnej w poglądach amerykańskiej filozofki Christine Korsgaard. Autorka ta ukończyła Uniwersytet w Chicago, obecnie zaś wykłada na Uniwersytecie Harvardzkim. Określana jest mianem bodaj najbardziej znanej uczennicy John Rawlsa i jednej z najwybitniejszych współczesnych kantystek. Jej prace obejmują zagadnienia moralne, przede wszystkim zaś pojęcie „normatywności” (*The Sources of Normativity*). Obecnie podejmuje tematykę podmiotu działającego (zdolności do działania) czy też samokonstytucji w ogóle (*Self-Constitution: Agency, Identity, Integrity*). Jako zdeklarowana obrończyni zwierząt, włączyła ona do swej pracy filozoficznej także kwestię obrony ich praw.

Pierwszy rozdział artykułu stanowi syntetyczną rekonstrukcję poglądów Korsgaard zawartych w jej rozlicznych esejach. Skupię się przede wszystkim na statusie moralnym zwierząt, krótko też wspomnę o ich sytuacji na gruncie filozofii prawa. Wskażę także pokrewne względem założeń Korsgaard topoty filozoficzne. Drugi rozdział wyklada metodologię, jaką zamierzam się posłużyć w analizie poglądów amerykańskiej filozofki. Analiza ta będzie mieć charakter semantyczny. Trzeci rozdział stanowi już odczytanie poglądów amerykańskiej filozofki z użyciem konceptów wyłożonych w rozdziale drugim.

Na wstępie chciałabym zaznaczyć pewną trudność, jaka pojawia w sposobie prowadzenia przez Korsgaard wywodu czy też stosowanej przez nią formie narracji. Filozofka ta jako zdeklarowana kantystka pracuje na tekstach i ideach Immanuela Kanta. Nie jest to jednak praca historyczna, lecz – jak sama podkreśla – czysto filozoficzna (Korsgaard 2014). Taki rodzaj działalności nieuchronnie skutkuje powstawaniem nowych propozycji, nieznanych Kantowi lub obcych jego etyce. Korsgaard, niewątpliwie

świadoma takiego obrotu spraw, posługuje się pomimo to zwrotami sugerującymi, że jej własne propozycje i tezy mają proveniencję kantowską. Trudność polega wobec tego na tym, że niełatwo ustalić, w którym momencie kończy się jej interpretacja myśli Kanta, a zaczyna się twórczy rozwój, rewizja bądź nawet refutacja jego założeń.

Pragnę ponadto podkreślić, że nie analizuję tutaj bezpośrednio filozofii Kanta ani też filozofii związanej z tzw. ruchem wyzwolenia zwierząt. Unaoczniam jedynie sposób, w jaki – celem rozszerzenia podmiotowego zakresu zastosowania praw – przekształcane są pojęcia i rozumowania kluczowe dla etyki Kanta. Przekształceń tych dokonała Christine Korsgaard.

1. Rekonstrukcja poglądów Christine Korsgaard

W niniejszym rozdziale postaram się zrekonstruować poglądy Korsgaard na istotę i zasadność praw zwierząt. Jej filozofia czerpie przede wszystkim z koncepcyjnego bogactwa myśli Kanta, a w drugiej kolejności z dzieł Arystotelesa. Tutaj bowiem należy szukać podstaw idei prawa zwierząt u Korsgaard.

Aby uchwycić istotę i znaczenie zabiegów myślowych, jakich dokonać musiała ta amerykańska filozofka, by przekuć „kantyzm”¹ w skuteczne narzędzie ochrony praw zwierząt, rozpocznę od przypomnienia prac, które powstały zanim zainteresowała się ona tą właśnie problematyką. Unaoczniam to modyfikacje wprowadzone przez autorkę *The Sources of Normativity* do jej wcześniejszych odczytań Kanta (modyfikacje te są inspirowane m.in. Arystotelesowską kategorią duszy wegetatywnej).

1.1 Stanowisko Korsgaard poprzedzające jej koncepcję praw zwierząt. Kantowska Formuła Człowieczeństwa

W 1986 roku Christine Korsgaard opublikowała artykuł *Kant's Formula of Humanity*, w którym dokonała – uważanej do dziś za przełomową – egzegezy formuły celowościowej Kantowskiego imperatywu kategorycznego: „Postępuj tak, byś człowieczeństwa tak w twojej osobie, jako też w osobie każdego innego używał zawsze zarazem jako celu, nigdy tylko jako środka” (Kant 2017, 46). Uznając brak treściowej ekwiwalencji pomiędzy poszczególnymi sformułowaniami imperatywu kategorycznego Korsgaard doszła do wniosku, że człowieczeństwo w znaczeniu racjonalnej (rozumnie-autonomicznej) natury jako wartość autoteliczna (cel sam w sobie) może stanowić o materii prawa moralnego, Formuła Powszechnego Prawa² (uniwersalności) – o jego formie, zaś autonomiczne prawodawstwo (łącznie z legislacją definiowaną przez Kanta jako prawodawstwo zewnętrzne, tworzące ład prawny całej społeczności) reprezentuje kompletny zbiór maksym i całość celów (Korsgaard 1996a, 106). Każda z trzech formuł

1 Na pytanie, czy filozofię tę nadal można klasyfikować jako „kantyzm”, odpowiem w dalszej części pracy.

2 Temu sformułowaniu Korsgaard poświęciła osobny artykuł (Korsgaard 1985, 24-47), dostępny również w tomie *Creating the Kingdom of Ends* (1996a).

ma zatem jej zdaniem oznaczać odmienną cechę charakterystyczną racjonalnych zasad praktycznych³. Formuła Powszechnego Prawa nadaje maksymom formę zasad rozumu praktycznego dzięki testowi kategorycznego imperatywu, przez co jest przydatna w rzeczywistym procesie podejmowania decyzji; z kolei człowieczeństwo traktowane jako cel sam w sobie stanowi jej właściwy materiał (materię) (Korsgaard 1996a, 107). W toku interpretacji Formuły Człowieczeństwa autorka *Creating the Kingdom of Ends* stara się odnaleźć dystynktywną cechę człowieczeństwa, która pozwala mu stanowić materialne wypełnienie prawa moralnego. Inaczej mówiąc, zastanawia się nad tym, jakie jest znaczenie pojęcia „celu samego w sobie”. Jej podejście jest więc najpierw podejściem semantycznym, a dopiero w dalszej kolejności staje się normatywne.

Struktura argumentu z Formuły Człowieczeństwa ma charakter dwupoziomowy. Jak pisze Korsgaard, Kant najpierw postuluje deontycznie bezwarunkową celowość, a następnie dopiero wskazuje, czym ona jest, a raczej *być musi*:

Kant najpierw argumentuje, że musi istnieć bezwarunkowy cel, a później wskazuje, że tym celem musi być człowieczeństwo (Korsgaard 1996a, 109).

Pierwszy krok opiera się na stwierdzeniu, że rozum w swym użytku tak teoretycznym, jak i praktycznym poszukuje tego co bezwarunkowe i nieuwarunkowane (*unconditional*). Niemniej jedynie rozum praktyczny kończy owe poszukiwania w sposób efektywny, ponieważ rozum teoretyczny napotyka antynomie, nie dochodząc do rozwiązania (abstrahując naturalnie od antynomii, które się pojawiają w związku z rozumem praktycznym, chodzi jednak o samo odnalezienie „nieuwarunkowanego”).

Konstrukcja bezwarunkowego celu korzysta z dwóch dysjunkcji w pojęciu dobra⁴, które odzwierciedlają następujące pary:

a) dobro wewnętrzne (*intrinsic good*) – dobro zewnętrzne lub innymi słowy, dobro bezwarunkowe – dobro warunkowe;

b) Cel – środek lub dobro finalne (*telos*) – dobro instrumentalne.

Kryteria demarkacyjne pomiędzy powyższymi dystynkcjami opisuje Korsgaard następująco:

powiedzieć, że coś jest wewnętrznie dobre to nie przez definicję powiedzieć, że jest cenione dla samego siebie: to jest powiedzieć, że ma to coś dobro w sobie. Odnosi się to, jak powiedzieliby niektórzy, do usytuowania lub źródła wartości dobra raczej niż do sposobu, w którym cenimy rzecz (...). Rozróżnieniem pomiędzy rzeczami, które są cenione same dla siebie i ze względu na coś innego jest rozróżnienie pomiędzy celami i środkami lub między finalnymi a instrumentalnymi dobrami (Korsgaard 1996a, 250).

Chodzi więc o dobro źródłowe w takim sensie, że jest ono nienabyte i niezbywalne, przez nikogo niewytworzone, niezależne od jakichkolwiek okoliczności zewnętrznych, a

³ W ten sposób Korsgaard dystansuje się od takiego ujęcia imperatywu kategorycznego, które uznaje, że jest on tożsamy z Formułą Powszechnego Prawa, a pozostałe formuły stanowią racjonalne zasady jako treściowe ekwiwalenty formalnego ujęcia imperatywu.

⁴ W sprawie tej dysjunkcji zob. Korsgaard (1983, 169-195).

ponadto, niemogące stanowić środka do żadnego innego celu, tj. nieinstrumentalizowalne w sensie normatywnym⁵.

Jakkolwiek sposób użycia przez autorkę powyższych terminów sugeruje inspiracje moore'owskie związane z pojęciem „wartości wewnętrznej” (zob. Moore 1919; 1922; 1980), odrzuca ona zdecydowanie stanowisko realizmu moralnego. Uznaje, że źródło wartości leży po stronie podmiotu. Co więcej, zakłada, że Kant był świadom dwóch rozróżnień w obrębie pojęcia dobra, lecz wiązał je z innymi konsekwencjami aniżeli te, które miały później wynikać z teorii proponowanych przez Moore'a i Rossa (Korsgaard 1996a, 256). I tak, wartość wewnętrzna przynależna *dobrej woli* udziela się aktom dobrej woli jako źródłu działań moralnych. Wynika to z Kantowskiej metaetycznej konstatacji otwierającej rozdział I *Uzasadnienia metafizyki moralności*:

Nigdzie w świecie, ani nawet w ogóle poza jego obrębem, niepodobna pomyśleć żadnej rzeczy, którą bez ograniczenia można by uważać za dobrą, oprócz jedynie dobrej woli (Kant 2017, 11).

Oznacza to, że z chwilą rozpoczęcia realizacji aktów dobrej woli, kiedy opuszczają one inteligibilne pole dobrej woli, jej aksjologii i normatywności, tracą one swój bezwarunkowy i koniecznie dobry charakter (gwarantowany przez szereg warunków wewnątrzpodmiotowych i leżący w mocy podmiotu). Są one wówczas narażone na kontyngencję, uwarunkowania i determinizmy świata rzeczywistego, albowiem dobroć i (źródłowo inteligibilna) wolność podmiotu są u Kanta nierozłącznie powiązane ze sobą. Dla przykładu szczęście, które wprawdzie stanowi jeden z celów człowieka, definiowane jest przez warunki zewnętrzne (*extrinsic, conditional*) i przedmiotowe. Jest obarczone przygodnością i zmiennością czynników przedmiotowych i okoliczności rzeczywistych, które mogą zakłócać, wspomagać, czy po prostu zmieniać pierwotny zamysł podmiotu co do „szczęścia”. Jego ostateczny kształt, materia, realizacja itd. w większości nie leżą w mocy podmiotu; nie tylko on jest źródłem własnego szczęścia, często jest nim zwykły przypadek. Tymczasem całkowicie w mocy podmiotu leżała uregulowana przez jedno jedyne prawidło moralne „dobra wola”. Jest ona w pełni zasługą wewnętrznej aktywności podmiotu, podczas, gdy szczęście jest (w nieprzewidywalnym dla nikogo stopniu) zasługą okoliczności zewnętrznych. Ujęcie szczęścia jako celu zdeterminowanego lub co najmniej współdeterminowanego przez czynniki przedmiotowe, cudzą wolę itd. zamiast podmiotowych dodatkowo przemawia za utrzymaniem wskazanych wyżej rozróżnień, zdaje się argumentować Korsgaard.

Przechodząc do kroku drugiego w argumencie z Formuły Człowieczeństwa można dojść do wniosku, że w rozumowaniu Korsgaard występuje założenie o tym, że wartościowanie (jako proces waloryzowania tj. przyznawania wartości) poprzedza

⁵ W sensie technicznym można wyobrazić sobie użycie takiego dobra dla osiągnięcia innych celów (instrumentalizację), jednak w sensie normatywnym stanowi to *nadużycie*. Takich nadużyć ludzkość rzecz jasna dokonuje, instrumentalizując to, co autoteliczne; niniejsza praca pokazuje zaś, dlaczego i jak w sposób normatywny, tj. etycznie i prawnie, nadużycia takie wyeliminować, i w ten sposób uregulować praktyki ludzkie wobec zwierząt.

konstytucję wartości. Wartość powstaje zatem i przysługuje czemuś (niech to będzie akt podjęcia decyzji poprzedzający akt realizacji działania) wskutek prawodawczego aktu podmiotowej woli, który warunkuje wartość, czyli jest aksjotwórczy. Innymi słowy, nie tkwi w przedmiocie (np. przedmiotowych skutkach działania) ani w samej czynności, jak sądzą przedstawiciele realizmu moralnego i jak sądzą niektórzy przedstawiciele filozofii działania. W następstwie tego, że wartość wewnętrzna (bezwarunkowa i bezwzględna) przysługuje pierwotnie wyłącznie dobrej woli, czyli prawodawczo-decyzyjnemu źródłu, rzeczywiste akty działania moralnego i stany przedmiotowe będące ich skutkami mogą nabierać jedynie pochodnej wartości moralnej, która ponadto będzie mniej pełna w porównaniu z „dobrą wolą” jako tym, co jako jedyne jest dobre w pełni. Im dalej w koncentrycznym porządku wartości znajduje się realizacja owej dobroci, tym mniejsza może być wartość, której źródłem jest podmiot działania. Zależy to od skali udziału czynników zakłócających lub wspomagających, występujących poza podmiotową sferą dobrej woli. Wywód ten można przedstawić w sposób tabelaryczny:

Dobro wewnętrzne (dobra wola, człowieczeństwo) jako cel ostateczny	Dobro zewnętrzne jako cel ostateczny (np. szczęście)
Brak (tj. według Korsgaard nie istnieje taka kategoria jak dobro wewnętrzne instrumentalne)	Dobro zewnętrzne instrumentalne (np. pożywienie, dzieło sztuki, zwierzęta)

Tabela 1: Dobro u Kanta (interpretacja Korsgaard) a tradycja analityczna

Niezależnie od tego czy wartość ma charakter bezwarunkowy (wewnętrzny) czy warunkowy (zewnętrzny) przysługuje jej przymiot obiektywności.

Możemy powiedzieć, że rzecz jest dobra obiektywnie albo gdy jest dobra bezwarunkowo albo gdy rzecz jest warunkowo wartościowa i warunki jej dobroci są spełnione (Korsgaard 1996a, 260).

Obiektywność dobra, mimo jego podmiotowego (nieontologicznego) charakteru, wynika z tego, że dobro jest konceptem rozumowym. Ów rozumowość dobra sprawia, że ludzie determinując swoje cele, które uznają za dobre dla siebie (a zatem subiektywnie), poczytują jako dobre absolutnie (dobro zewnętrzne finalne).

Cel sam w sobie (dobro wewnętrzne finalne) w ujęciu Korsgaard stanowi bezwarunkowy warunek (*unconditioned condition*) wszelkiej (innej, „dalszej”) wartości i tożsamy jest z człowieczeństwem, możliwością racjonalnego determinowania celów czy dobrą wolą. Filozofka używa tych określeń zamiennie, jednak, jak zauważa Jerome B. Schneewind, zróżnicowanie używanych przez nią terminów jest nieistotne (Schneewind 1998, 37), o czym będzie mowa w dalszej części tego podrozdziału. Istotne jest zaś, że cel sam w sobie stanowi źródło wartości. Takie rozwiązanie zapewnia spełnienie

warunku konkluzywności uzasadnienia, *ergo* skutkuje zabezpieczeniem argumentacji moralnej przed błędem *regressus ad infinitum*. Toteż określenie „wartość wewnętrzna” (*intrinsic value*) mogłoby zostać zastąpione terminem „wartość autoteliczna”.

Ostatecznie można więc uznać, że drugi krok w argumentacji z Formuły Człowieczeństwa zasadza się na tym, że:

tylko jednej rzeczy – dobrej woli – jest przypisana wartość wewnętrzna (*intrinsic value or inner worth*) (...). Jeśli postrzegamy siebie jako mających moc uzasadniania naszych celów, to argument nakazuje nam postrzegać siebie jako mających wartość wewnętrzną – a także nakazuje traktować innych, którzy lokują wartość w swoich celach na podstawie własnego człowieczeństwa mającego tę samą wartość wewnętrzną (Korsgaard 1996a, 272).

Z powyższego *passusu* wyłania się wyraźnie dlaczego właśnie człowieczeństwo / dobra wola / możliwość racjonalnego determinowania celów musi być owym bezwarunkowym celem. Tylko bowiem istota ludzka posiada status przyznającej wartości (*conferring-value status*). Taki wybór dystynktywnej cechy człowieczeństwa wskazuje, że to w ogóle możność racjonalnego wyboru jakiegokolwiek, a nie tylko moralnego, celu przesądza o wartości człowieczeństwa. Skoro ja – jako istota ludzka – jestem w stanie wybrać swój cel, uznać przedmiot moich interesów za wartościowy, tylko dlatego że jest to przedmiot moich interesów, to samego siebie muszę afirmować jako wartość. Tym samym staje się za celem sam w sobie, co pociąga również za sobą konsekwencję w postaci zyskania obywatelstwa państwa celów. Obywatelstwo to na płaszczyźnie moralnej (abstrahując od politycznego wymiaru tego pojęcia) implikuje, że ludzie – jak pisze Ewa Nowak – „powinni zatem uznawać się nawzajem za osoby ludzkie, umożliwiając sobie wzajemnie realizację szczęścia” (Nowak-Juchacz 2002, 66). Dla Korsgaard zaś – uznawanie samego siebie za członka państwa celów – stanowi warunek związania prawem moralnym.

Człowieczeństwo w takim ujęciu może być stosowane alternatywnie właśnie z możliwością racjonalnego wyboru. Sam Kant w *Uzasadnieniu metafizyki moralności* stosował przecież terminy „człowieczeństwo” i „istota racjonalna” wymiennie. Twierdzenie to jednak nie uzasadnia zamiennego posługiwania się sformułowaniami „człowieczeństwo” i „dobra wola”. Utożsamienie tych pojęć wymaga dodatkowych założeń. W tym celu Korsgaard sięga do *Religii w obrębie samego rozumu*, a konkretnie – do wyróżnionych przez filozofa z Królewca trzech predyspozycji ludzkiej natury, mianowicie do zwierzęcości, człowieczeństwa i osobowości człowieka. Przytaczając odnośny ustęp, w którym mowa o tym, że:

(...) korzeniem pierwszej [predyspozycji, M.A.] nie jest żaden rozum, drugiej wprawdzie rozum praktyczny, ale służący tylko innym pobudkom, natomiast jedynie trzeciej praktyczny dla siebie, tj. rozum bezwarunkowo prawodawczy (Kant 2007, 50)

harwardzka filozofka konstatuje, że „człowieczeństwo totalne i udoskonalone staje się osobowością” (Korsgaard 1996a, 114), zatem jeżeli cenimy człowieczeństwo,

to *a minori ad maius* cenimy również osobowość, która ma wymiar moralny. Przy takiej wykładni uzasadnione zdaje się stwierdzenie Korsgaard, iż Kant za wyróżnik człowieczeństwa uznał determinowanie własnych celów, przedmiotów swoich interesów, a zatem świadomość działania na podstawie racji zarówno na poziomie hipotetycznym, jak i kategorycznym.

1.2 Zmodyfikowana wersja argumentu. Normatywność

W *The Sources of Normativity* Korsgaard przedstawia „nową fantazyjną” wersję argumentu z Formuły Człowieczeństwa, której przytoczenie wydaje się istotne z uwagi na to, że jest to chyba najbardziej znana praca tej autorki. Rozprawa ta poświęcona jest analizie normatywności moralności poprzez odpowiedź na tzw. pytanie normatywne (*normative question*), które w swej najbardziej lakonicznej, najłatwiejszej formie brzmi: „Dlaczego powinienem/powinnam być moralny/a?” (Korsgaard 1996b, 9). Pytanie to przybiera formę koniecznie pierwszoosobową, gdyż powstaje w sytuacji, gdy podmiot mając podjąć działanie na podstawie jakiejś normy zastanawia się czy faktycznie musi to działanie wykonać. Innymi słowy, pytanie normatywne jest pytaniem o to co czyni normę moralną wiążącą. Jednocześnie presuponuje ono, że nasza aktywność moralna ma charakter normatywny. W kwestii samego pojęcia normatywności, warto napomknąć, że „choć często powoływane, szczególnie w filozofii prawa, moralności i języka – rzadko jest precyzyjnie definiowane” (Brożek 2013, 19).

Korsgaard wskazuje na cztery nowożytnie stanowiska, rozmaicie lokujące źródło normatywności: woluntaryzm, realizm, refleksyjne potwierdzenie (*reflective endorsement*) oraz autorytet refleksji (*the authority of reflection*), który nawiązuje do Kantowskiej autonomii woli. Autorką ostatniego z nich jest sama Korsgaard.

Punktem wyjścia dla budowy własnej koncepcji normatywności czyni ona spostrzeżenie o refleksyjnej naturze ludzkiej samoświadomości, która problematyzuje działanie człowieka w kategoriach normatywnych, czego nie sposób zaobserwować u mniej rozwiniętych zwierząt (*lower animals*). „Zwierzęta bowiem są zaangażowane w świadome czynności, nie będąc ich jednak świadome” (Korsgaard 1996b, 93), gdyż działają na podstawie percepcji i pragnień. Człowiek, z uwagi na przynależną mu gatunkowo refleksyjność, uznaje pobudki te za niewystarczające. Istota ludzka, by działać, potrzebuje racji (*reason*). Wyraźnie wyczuwalne są zatem w rozumowaniu Korsgaard inspiracje Kantowskim ujęciem człowieka jako istoty noumenalnej, wolnej, potrafiącej abstrahować od pobudek empirycznych.

Refleksyjność ludzka wymusza na człowieku posiadanie jakiejś koncepcji siebie, tj. praktycznej tożsamości (*practical identity*), która obejmuje wiele przypadkowych składowych – bycie obywatelem konkretnego państwa, członkiem rodziny, czy czyimś przyjacielem. Praktyczna tożsamość jest źródłem bezwarunkowych obowiązków, norm – jako, że „twoje racje [do działania, M.A.] wyrażają twoją tożsamość” (Korsgaard 1996b, 101) – których naruszenie powoduje utratę własnej integralności oraz tożsamości,

i sprawia, że już więcej nie jesteś tym kim byłeś. Swoistą treść praktyczne tożsamości zyskują poprzez identyfikację z innymi, tj. jak pisze Robert Piłat odnosząc się do koncepcji Korsgaard: „tożsamość jest wprawdzie autonomiczną relacją do samego siebie, lecz posiada także przedmiotowy biegun, przygodną podstawę, która owej relacji do samego siebie nadaje treść” (Piłat 2013, 235). Tym przedmiotowym biegunem, tj. nadaniem treści, będzie wobec tego „utożsamienie się z pewnymi ludźmi, grupami lub ideałami” (Piłat 2013, 235).

Takie ujęcie praktycznych tożsamości powoduje, że większość z nich jest całkowicie przygodna, niezależna od nas samych. Ponadto dyktują one różne normy, co prowadzi do konfliktów między nimi. Dlatego autorka *The Sources of Normativity* zastrzega, że istnieje tylko jedna konieczna (nieprzygodna, przynależna każdej istocie ludzkiej) praktyczna tożsamość, a mianowicie człowieczeństwo. Pozostałe praktyczne tożsamości nadbudowują się i wyrastają z tej podstawowej tożsamości bycia człowiekiem, a więc zwierzęciem refleksyjnym, normatywnym, które do tego by działać potrzebuje racji. Czyli, jak hiperbolicznie ujmuje to Richard Rorty, człowieczeństwo rozumiane jako praktyczna tożsamość

stoi, by tak rzec, w cieniu za moją tożsamością rodzica, kochanka, przedsiębiorcy, patrioty, mafiosa, profesora lub nazisty, czekając na jej odsłonięcie przez refleksję (Rorty 2009, 292).

Reasumując, struktura Korsgaardiańskiego argumentu przebiega następująco: człowiek jest istotą refleksyjną, a refleksyjność wymaga tego by działać na podstawie racji. Racje, którymi się posługuje w działaniu odzwierciedlają praktyczne tożsamości jednostki. Lecz to, że jednostka w ogóle może świadomie postępować na podstawie norm wpływających z praktycznej tożsamości wynika z faktu, że jest człowiekiem, a więc zwierzęciem refleksyjnym, co przesądza, iż to jest jego podstawowa tożsamość. Wszelkie racje mają swe źródło w człowieczeństwie, które musi być zatem cenione same dla siebie, gdyż ono umożliwia nam działanie. Jak ujmuje to Korsgaard – „afirmujesz swoją wartość jako posiadacz praktycznej tożsamości, ilekroć działasz na podstawie racji, których jest ona źródłem” (Korsgaard 2011, 23), czyli mówiąc prościej – działając, musimy cenić nasze człowieczeństwo. Wynikające już z samego działania ceniecie samych siebie wymaga cenięcia innych ludzi w ten sam sposób, co skutkuje moralnymi obowiązkami wobec nich. Najbardziej podstawowa praktyczna tożsamość ma zatem wymiar moralny.

Założenie o człowieczeństwie jako celu w samym sobie w *The Sources of Normativity* konkretyzuje się jako racja wszelkich racji. Można zatem powiedzieć, że zarówno na poziomie wartości, jak i racji, Korsgaard postrzega Kanta i samą siebie jako konstruktywistę. Krzysztof Saja takie rozumienie konstruktywizmu zalicza do etycznego racjonalizmu (realizmu racji), który „uznaje, że istnieją obiektywne, poznawalne powody właściwego myślenia i działania” (Saja 2015, 28). W świetle poglądów tego autora etyczny racjonalizm przynależy do realizmu etycznego w szerokim znaczeniu.

1.3 Pomysł Christine Korsgaard na prawa zwierząt

W poprzednim podrozdziale została przedstawiona koncepcja człowieczeństwa, które jako jedyne wśród wartości jest wartościowe autotelicznie. Z uwagi na swą dystynktywną cechę, a mianowicie rozumną, a w konsekwencji prawodawczą naturę obywatelami państwa celów są zatem jedynie istoty ludzkie. Takie charakterystyczne dla etyki Kantowskiej myślenie wyklucza traktowanie zwierząt innych niż ludzie jako celów samych w sobie. Co więcej, Kant wprost wyrażał się o zwierzętach jako jedynie środkach:

To, co istnieje, a czego istnienie zależy nie od naszej woli, ale od przyrody, ma mimo to – o ile jest czymś pozbawionym rozumu – tylko względną wartość, a to jako środek, i zwie się dlatego rzeczą (Kant 2017, 45)

Obowiązków bezpośrednich wobec zwierząt w etyce filozofa z Królewca nie ma, co nie oznacza, że człowiek może w sposób okrutny traktować zwierzęta. Co więcej takie traktowane winno być oceniane negatywnie. Jak pisze Tomasz Pietrzykowski: „Jej źródłem [negatywnej oceny moralnej] nie jest (...) naruszenie obowiązku człowieka wobec zwierzęcia, ale obowiązku właściwego postępowania jakie człowiek jest winny innym ludziom” (Pietrzykowski 2007, 23). Taka wizja ochrony zwierząt jawi się dla amerykańskiej filozofki jako niewystarczająca, wciąż bowiem są one traktowane jedynie instrumentalnie.

1.3.1 Zwierzęta racjonalne i zwierzęta nieracjonalne

Aby wyjaśnić w ogóle czym jest zwierzę sięga Korsgaard do Arystotelesowskiego pojęcia substancji (lub jestestwa), przy czym istoty żywe są szczególnym typem substancji. Substancja stanowi funkcjonalnie zorganizowaną materię, żyjące substancje (jestestwa żywe) mają za swą funkcję zachowanie życia i reprodukcję. Takie określenie charakteryzuje – niezależnie od gatunku – wszystko co żyje.

Tak pojęte zwierzęta realizują swoją rolę utrzymywania życia poprzez sformułowanie koncepcji swojego środowiska (przekonanie) i sterowanie sobą w tym środowisku zgodnie z tą koncepcją (działanie) (Korsgaard 2004, 82). Zadania te mogą być realizowane przy użyciu rozumu bądź instynktu.

W celu wyjaśnienia funkcjonowania zwierząt tak rozumnych, jak i nierozumnych harwardzka filozofka odwołuje się do „Kantowskiej teorii działania” obejmującej pobudki (*incentive*), a więc swoiste bodźce do działania, mentalne reprezentacje obiektów, oraz zasady (*principles*⁶), które determinują co zrobi lub spróbuje zrobić podmiot działający (*agent*) w sytuacji, gdy taka pobudka wystąpi. Pobudki i zasady występują w naturalnych parach, dla każdego podmiotu działającego zasady determinują działanie w stosunku do bodźców, na które jest wystawiony. Według Korsgaard zasadami dla zwierząt są instynkty. Filozofka nie traktuje instynktu wyłącznie jako zestawu wrodzonych zasad – obejmuje

⁶ Korsgaard używa tego terminu zamiennie z maksymą (*maxim*), co sugeruje, że utożsamia to pojęcie semantycznie z pojęciem maksymy u Kanta.

on bowiem także możliwość uczenia się z wcześniejszych doświadczeń zwierzęcia. „Instynktowna reakcja” nie stanowi zatem antonimu względem „reakcji wyuczonej”. Konkludując, „powiedzieć, że zwierzę działa na podstawie instynktu to powiedzieć, że działa na zasadzie ustalonego połączenia pomiędzy szczególnym typem reprezentacji (pobudką) a prymitywnie normatywną odpowiedzią” (Korsgaard 2004, 83-84).

W duchu kantyzmu autorka *Self-Constitution: Agency, Identity, and Integrity* sięga też do dystynkcji między rozumnością a inteligencją. Tę pierwszą definiuje jako szczególną formę samoświadomości (będącą zdolnością normatywną), która pozwala nam być świadomymi podstaw naszego działania i podstawy te wartościować, tę drugą zaś jako „zdolność do tworzenia skojarzeń lub zasad, do powiększania zasobu swoich właściwych reakcji [na pobudki]” (Korsgaard 2004, 85). W konsekwencji tak rozumianego instynktu i inteligencji zachowania zwierząt mają „prymitywnie normatywny” charakter, a inteligencja zawierałaby się w instynkcie. Prymitywna normatywność zachowań instynktownych opiera się na tym, że zachowania te powstają w wyniku dwóch czynników, a więc pobudki i zasady. Brak świadomości podstaw (zasad) swojego działania powoduje, że tak naprawdę zwierzęta nie działają na podstawie norm, jednakże ich działania tworzą normy *sui generis*. Różnica zatem zdaje się zasadzać na tym, że rozumność jest możliwa wtedy i tylko wtedy, gdy istnieje samoświadomość, albo inaczej mówiąc samoświadomość skutkuje rozumnością, która traktowana jest synonimicznie względem zdolności do normatywnego zarządzania sobą (*normative self-goverment*⁷). Jeśli zaś weźmiemy pod uwagę założenie Korsgaard o normatywnym charakterze moralności, to uzyskamy stwierdzenie, że tylko zwierzęta samoświadome są zwierzętami moralnymi. Pozostałe zwierzęta nie posiadają zdolności działania moralnego, nie mogą więc posiadać też moralnych obowiązków. Czy to implikuje jednocześnie, że ludzie – jako istoty moralne – nie mają żadnych obowiązków wobec nich? Takie wnioski wysnuwał sam Kant:

Moralność jest warunkiem, od którego jedynie zależy to, że istota rozumna może być celem samym dla siebie, ponieważ tylko dzięki niej można być prawodawcą w państwie celów (Kant 2017, 51)

Korsgaard uważa, że Kant dochodzi do powyższych wniosków, gdyż stosuje pewną strategię myślową, którą filozofka określa mianem tzw. argumentu z wzajemności (*reciprocity argument*), zakładającą, że tylko podmioty znajdujące się w pewnym rodzaju wzajemnej relacji ze sobą, mogą się nawzajem wiązać prawem. Ową wzajemną relacją jest status współprawodawcy w państwie celów, którego obywatelstwo zyskuje się poprzez moralność (na co wyraźnie wskazuje przytoczony z *Uzasadnienie metafizyki moralności* cytat). Jest to – wywodzi Korsgaard – „duchowa lub hipotetyczna wspólnota ukonstytuowana na relacjach między istotami ludzkim z podzielanym przez nie oddaniem dla pewnej koncepcji ich samych” (Korsgaard 2004, 82). Zatem jeżeli podmiot A posiada

⁷ Zdolność ta wiąże się także z inną zdolnością właściwą jedynie człowiekowi, tj. normatywną konceptualizacją samego siebie (*normative self-conception*) (zob. Korsgaard 2009).

tak jak podmiot B zdolność do działania moralnego czy moralnie prawodawczego, a podmiot C takiej zdolności nie posiada, to na mocy argumentu z wzajemności, podmiot A posiada moralne obowiązki wobec jedynie podmiotu B.

1.3.2 Naturalne dobro. Paradoks życia wegetatywnego

W podrozdziale 1.1 wykazywałam, że w pierwotnej wersji swych poglądów Korsgaard utożsamiała wartość wewnętrzną z dobrą wolą. Moralnie istotną cechą była więc szeroko pojęta rozumność. W świetle koncepcji naturalnego dobra tą cechą jest bycie „organicznym systemem, który liczy się sam dla siebie” (Korsgaard 2004, 103) lub innymi słowy bycie istotą, dla której pewne rzeczy mogą być dobre lub złe w sposób naturalny. Samo przyznanie wartości wewnętrznej jedynie ludziom ocenia Korsgaard jako rodzaj twierdzenia metafizycznego. Wynika ono z zawężenia statusu metafizycznego do istot ludzkich. Wybór kryterium rozumności, przy wykazaniu przez Kanta, że poznanie nie ma dostępu do rzeczy samych w sobie, jest nie tylko nieuprawniony epistemologicznie, lecz także arbitralny, tzn. obciążony błędem szowinizmu gatunkowego, czyli zakładający nieuzasadnione⁸ roszczenie do zwierzchności istot ludzkich nad zwierzętami. Możliwość działania moralnego jest co prawda cechą przynależną jedynie ludziom, ale w żaden sposób nie przesądza o przewadze czy nadrzędności nad innymi zwierzętami. Nadrzędność człowieka nad pozostałymi stworzeniami podbudowana jest u Kanta antropocentrycznie, przy asekuracji Boga, którego poznanie wykracza poza granice rozumu, ale którego rozum ów „postuluje”.

Naturalizując dobro, uczennica Rawlsa odwołuje się do filozofii Arystotelesa. Sięga przede wszystkim do dzieła *O duszy*. Jednocześnie – krytykując tym samym utylitaryzm – uznaje, że dobro jest przywiązane (*tethered*) do podmiotu, a ściślej: do funkcji i dyspozycji, jakie uosabia podmiot. Utylitaryzm – jak argumentuje Korsgaard – uznając za swe naczelne kryterium maksymalizację dobra, odrywa to pojęcie od podmiotu, bowiem teoria ta nie mogłaby w innym przypadku zakładać (możliwości) sumowania dóbr. Takie podejście przypomina podnoszony w literaturze wobec utylitaryzmu zarzut, który mówi o pustości zasady maksymalizacji dobra. Dobro nie stanowi własności pomnażalnej (mimo, że Kant postuluje nieskończony postęp ludzi „dobrej woli” ku *summum bonum*). Saja argumentuje, że „przekonanie to opiera się na tzw. atrybutywnej teorii dobra” (Saja 2015, 173) Wywodzi się ona z tradycji analitycznej i pochodzi od Petera Geacha, który uznaje, że „dobry” i „zły” są zawsze atrybutywnymi, nigdy predykatywnymi przymiotnikami” (Geach 1956, 33) (niezależnie od tego co sugeruje struktura gramatyczna zdania). Dystynkcja atrybutywny – predykatywny opiera się o kryterium na możliwości oderwania od przymiotników ich „znaczenia od tego, o czym orzekają, bez semantycznej starty” (Makowski 2011, 21). Brak możliwości takiego oderwania skutkuje atrybutywnością.

⁸ Albo też uzasadniane teologicznie, co jednak nie wyczerpuje w żaden sposób uzasadnień na gruncie etyki, zwłaszcza naturalistycznej.

Znaczenie terminu „dobry” jest zatem każdorazowo relatywizowane względem tego o czym orzeka (używając przykładów samego Geacha np. dobry włamywacz a dobry ludzki czyn). Argument ten odnosi się zatem do właściwości języka, Korsgaard odwołuje się zaś raczej do tego co powinno być właściwym wyznacznikiem moralnego działania. Jej strategia jawi się *simpliciter* jako deontologiczna. W podobny sposób wypowiada się Tom Regan, który uznając, że konfrontacja utylitarystycznej zasady równego traktowania i zasady użyteczności wymusza przyznanie prymatu jednej z nich. W sytuacji, gdy ów prymat zostanie przyznany tej drugiej, prowadzi to w sposób konieczny „do zgody na preferowanie wybranych, jednostkowych interesów i poświęcanie jednych istot w imię ratowania innych. Tym samym wiodący charakter zasady użyteczności oznaczałby w praktyce instrumentalne traktowanie zarówno zwierząt, jak i ludzi” (Probuska 2015, 87).

Amerykańska filozofka wskazuje, że jej argument nie jest bynajmniej skierowany przeciw każdemu rodzajowi agregacji. Z ideą powiązania dobra z podmiotem zgodna jest taka agregacja, która zakłada „kierunek działania, który przynosi korzyść większej ilości różnych podmiotów, tak długo jak nikt nie jest pokrzywdzony przez ten kierunek działania” (Korsgaard 2011, 6). Używając słów Ronalda Dworkina z *Biorąc prawa poważnie*: „Ogólna korzyść [w sensie utylitarystycznym, M.A.] nie może więc usprawiedliwiać ograniczania praw” (Dworkin 1998, 349). Jakkolwiek pojęcia „dobro” i „prawo” różnią się od siebie, to u Korsgaard ów dobro staje się podstawą dla ukonstytuowania pozycji normatywnej, a więc posiadania praw właśnie.

Mając już na uwadze przywiązanie dobra do podmiotu, powrócić można do Arystotelesa. Ch. Korsgaard jednak w żadnym miejscu nie pisze w jaki sposób używa arystotelesowskich kategorii, ogranicza się jedynie do ukonstytuowania naturalnego pojęcia dobra bazującego na funkcji jestestwa żywego, czyniąc jedynie przypis o tym, że Stagiryta szcążkowo pisze na ten temat w dziele *Metafizyka*, rozwija go zaś w *O duszy*. Traktat *O duszy* czyni z Arystotelesa protoplastę nauki nazywanej dzisiaj psychologią. Jakkolwiek jego metoda w dużej mierze jest empiryczna – to jak zauważa tłumacz Paweł Siwek – oprócz dotarcia do faktów, Stagiryta pragnie odkryć także ich naturę (metafizyczną istotę) i to „wyróżnia metodę przyrodniczą Arystotelesa od metody fizyków i w ogóle przyrodników współczesnych” (Siwek 1988, 18), którzy od tego problemu abstrahują. Wszelkie zatem pojęcia w ramach psychologii Arystotelesowskiej trzeba interpretować metafizycznie, a kluczowym dla niej pojęciem jest „dusza”, „ona bowiem stanowi jakby podstawę istnienia zwierząt” (Arystoteles 1988, 45).

Aby odkryć, czym jest dusza trzeba odwołać się do polisemicznego terminu substancji. Jak pisze Arystoteles:

[Istnieje] pewien szczególny rodzaj bytu, który nazywam substancją, jedną jej odmianą jest materia, która sama przez się nie stanowi rzeczy określonej, drugą jest forma i istota, dzięki której nazywamy [materię] taką a taką rzeczą; jeszcze inną jest to, co złożone z nich obydwu [czyli z materii i formy] (Arystoteles 1988, 73).

W takim układzie dusza jawi się jako substancja w sensie formy (forma substancjalna), która wraz z materią tworzy ciało naturalne. Ze stwierdzenia, że „materia jest możliwością; forma jest aktem” (Arystoteles 1988, 73)

...otrzymujemy zaś definicję wspólną wszystkim rodzajom duszy (...): jest to pierwszy akt ciała naturalnego organicznego (Arystoteles 1988, 74).

Posiadanie duszy stanowi cechę odróżniającą jestestwa żywe od innych jestestw. Na samą duszę zaś składają się pewne władze, a mianowicie wegetatywna, zmysłowa oraz umysłowa (władze te nazywane są po prostu duszami czy formami życia, w niekiedy funkcjami). Jakkolwiek jestestwa żywe wyróżnia posiadanie duszy, nie każde jestestwo posiada wszystkie jej wszystkie rodzaje bądź stopnie rozwojowe o coraz większej złożoności i potencjale. Najbardziej podstawową z „dusz” jest dusza wegetatywna, która przysługuje wszystkim rodzajom jestestw żywych – od roślin przez zwierzęta po ludzi. Zadaniem władzy wegetatywnej jest wobec tego „rodzić i przyswajać pokarm” (Arystoteles 1988, 82).

Arystotelesowskie dusze w sposób organiczny zakładają jedna drugą, tzn. zmysłowa nie może istnieć bez wegetatywnej, a umysłowa bez zmysłowej i wegetatywnej). Zależność ta – jak już zostało wskazane – nie zachodzi w drugą stronę, tzn. funkcje wegetatywne nie wymagają udziału intelektu i zmysłów (czego jednak nie da się powiedzieć o zmysłach i umyśle w świetle dzisiejszego stanu nauki). Bez zagłębiania się w zbędne szczegóły należy podkreślić, że Korsgaard swoje pojęcie naturalnego dobra wiąże przede wszystkim z życiem wegetatywnym i potencjałem, jaki posiada ono u zwierząt. Na pierwszy rzut oka wydaje się, że obejmuje ono również rośliny, jednak życie wegetatywne roślin, zwierząt (a także ludzi) bardzo się różni pod względem jakości i złożoności. Dlatego metafora *człowieka-rośliny* jest w tym kontekście wyłącznie metaforą, wskazującą na ograniczenia organizmu w stanie wegetatywnym, a nie na analogie do wegetacji roślin. Korsgaard jednak precyzuje, że naturalne dobro jest raczej czymś, co nas łączy z innymi zwierzętami, o których powiada się, że są „stworzeniami na tej planecie, które dzielą z nami nasze zaskakujące przeznaczenie – jako z innymi, ale przecież bliźniaczymi istotami, dla których coś może być naturalnie dobre lub złe” (Korsgaard 2011, 23). O roślinach, które z pozoru obejmuje pojęcie naturalnego dobra, Korsgaard powiada:

Co by było, gdybyśmy odkryli, że rośliny mają świadomość? (...) Zwierzęta mają znaczenie same dla siebie w takim sensie, że dostrzegają swe naturalne dobro i dążą do niego. Nie myślimy zaś o roślinach jako dostrzegających swe własne dobro i dążących do niego; nadal jednak, podobnie jak zwierzęta, traktują one „samozachowanie” jako coś koniecznego i w tym sensie wydają się być zorientowane na ich własne dobro (...). Z pewnością nie ma żadnej, wyraźnie widocznej linii demarkacyjnej w naturze, która oddzielałaby rośliny od zwierząt (Korsgaard 2004, 106-107).

W powyższym fragmencie autorka nie ustosunkowuje się jednak do sytuacji moralnej roślin. Jakkolwiek widoczne jest, że przywiązuje dużą wagę do świadomości

zwierząt (co skądinąd ważne, ponieważ wielu filozofów, a także biologów, odrzuca taką konstrukcję „zwierzęcego podmiotu”) i możliwości identyfikowania przez nie własnego. W tym właśnie aspekcie zdaje się nawiązywać do założenia Arystotelesa, który w traktacie *O duszy* twierdzi, że „wszystkie zwierzęta posiadają spośród zmysłów przynajmniej dotyk; jestestwo które ma zmysł doznaje także przyjemności i bólu” (Arystoteles 1988, 80). Tym niemniej, można dostrzec pewną niekoherencję w poglądach Korsgaard (która sama mówi o braku wyraźnej linii demarkacyjnej); otóż biorąc za podstawę dobra naturalne, powinna zasadniczo bronić statusu moralnego roślin i zwierząt. Na tym tle rysuje się pewien paradoks, który określiłabym mianem „paradoksu życia wegetatywnego”. Biorąc pod uwagę, że naturalne dobro odnosi się do bardzo wielu istot, rodzi się pytanie następujące: czy jesteśmy w stanie realizować ukonstytuowane na jego podstawie normy moralne? (i w konsekwencji dalsze pytanie, czy waga powinności może być uzależniona od liczby jej adresatów). Korsgaard stawia w tym kontekście takie oto pytanie: czy „powinien” implikuje „może”? (*does ought to imply can?*) (Korsgaard 2004, 106). Zgodnie z intuicją filozofka odpowiada na to pytanie przecząco, przywołując sytuację, kiedy to wiążące normy moralne są niemożliwe do zrealizowania w praktyce. Natomiast w jej przekonaniu na gruncie filozofii Kanta – „powinność” zakłada „możliwość” w takim znaczeniu, że „możliwe jest, abyśmy byli [jedynie, M.A.] zmotywowani przez to, co powinniśmy zrobić” (Korsgaard 2004, 107), nie zawsze mogąc powinności dopełnić. Takie rozumowanie skutkuje rozdzieleniem obowiązywania (a zwłaszcza bycia zobowiązanym w sensie podmiotowym) od winy i odpowiedzialności. Ale z jakich to powodów niedopełnienie powinności i obowiązków może nie pociągać za sobą winy? Z pewnością nie z takich, że obowiązek zaniedbano lub że wypełniono go tylko wobec tych, do których zdołano dotrzeć. Wszystko zależy od natury warunków ograniczających lub uniemożliwiających wywiązanie się z obowiązku. Jeśli mają one podłoże subiektywne, ale niezawinione przez podmiot (np. fizyczną lub psychiczną niepełnosprawność rozważaną przez Kanta) lub jeśli ich podłoże jest przedmiotowe, związane z obiektywnymi, nieprzekraczalnymi dla podmiotu przeszkodami, wówczas odpowiedzialność i wina sprawcy z tytułu niedopełnienia powinności są mniejsze, jednak te „okoliczności łagodzące” w żadnym razie nie zwalniają z obowiązków ani nie umniejszają zobowiązań. Zdaniem Kanta człowiek winien rozwijać i doskonalić swoje moralne zdolności, sprawności i ambicje; z drugiej strony musi liczyć się z tym, że pewne wymogi moralne go przerosną. Może w tym względzie liczyć na pomoc nadprzyrodzoną, edukację lub samodoskonalenie własnej cnoty moralnej (zob. Hare 1996, zwł. rozdział „Puffing up the capacity”). Dyskutując ten wątek w Kantowskiej etyce, John Hare przyznał, że utylitaryzm jako jedyny wypracował strategię przykrwania wymogów moralnych (obowiązków) do możliwości człowieka, ale to nie on jest w stanie zmniejszyć „uskok” metafizyczny (normatywny) między kantowskim ideałem i rzeczywistością (Hare 1996, zob. rozdział „Kant and the moral demand”). Za jeden ze sposobów na redukcję tego uskoku Hare uważa nie to, że inny ma być dla mnie celem samym w sobie jako „rozumny podmiot” (*rational agency*), lecz

to, że znajduje się ze mną w bliższych relacjach (Hare 1996, zob. rozdział „Reducing the demand”), do których można przecież zaliczyć relacje transgatunkowe.

Należałoby oczekiwać, że naturalizacja dobra, jaką postuluje Korsgaard, powinna sprzyjać redukcji tego uskoju jako *naturalizacja* właśnie, oznaczająca coś w rodzaju redukcji metafizyki. Jest to zabieg demonstrujący, jak coś może być dobre lub złe dla określonego, żywego organizmu i wyjaśniający przy okazji, co to znaczy być żywym organizmem z punktu widzenia jego funkcjonowania. W sposób jednoznaczny wynika z niego, że dla żywego organizmu dobre jest to, co umożliwia mu realizację jego funkcji (lub sprzyja takiej realizacji); z kolei złe jest to, co realizację uniemożliwia (zakłócając ją, ograniczając, wynaturzając, itd.). W argumentacji tej pobrzmiewa echo przekonań Arystotelesa, z których Kazimierz Leśniak wyłowił kluczowe, tj. mówiące o „nierozzerwalnym związku życia psychicznego z podłożem fizycznym” (Leśniak 1989, 69).

Naturalizm Korsgaard w pewnej mierze przypomina poglądy Philippy Foot z okresu *Natural Goodness*. Obydwie autorki odwołują się do tradycji Arystotelesowskiej i rozumowania w kategoriach biologicznych o celowości i wartości życia w rozmaitych, również zaawansowanych ewolucyjnie przejawach. Koncepcja Foot obejmuje podobne do Korsgaardiańskiego założenie, że „naturalne dobro (...), które związane jest tylko z żyjącymi istotami i ich częściami, charakterystyką, i działaniami, jest wewnętrzne...” (Foot 2001, 26-27). Wydaje się ponadto, że Korsgaard zgodziłaby się z tym, że dobro danej istoty jest w dużej mierze wyznaczone jej przynależnością gatunkową, a więc można je oszacować dopiero wówczas, gdy zrelatywizuje się je do jego cech, właściwości, a także *differentia specifica*, które pozwalają mu realizować naturalne cele.

Niezależnie od wszystkiego, Korsgaard wykorzystuje pojęcie „naturalnego dobra”, by „zbudować” swój pomysł na prawa zwierząt. Opiera się on na założeniu, że wszystkie zwierzęta żywią ten sam interes manifestujący się w byciu taką istotą, dla której określone rzeczy lub stany rzeczy mogą być dobre lub złe na sposób naturalny. Pozostając w dalszym ciągu na wysokim poziomie ogólności stwierdza ona następnie, że funkcją żywych stworzeń jest zawsze samozachowanie (przetrwanie) i reprodukcja. Uznaje, że istota żywiąca tak rozumiany interes, przyznaje sobie „pozycję normatywną”. Co ciekawe, u Kanta można wskazać epistemologiczne warunki rozpoznawania celowości bądź interesu natury (*Naturzweck*) lub inaczej celu immanentnego w przyrodzie, przykładowo „w ciałach organicznych (...), kiedy dla wydania sądu o jej wytworach przypisujemy jej nasze pojęcie celu (...) jako cel naturalny. Transcendentalna zasada przygotowała już intelekt do stosowania pojęcia celu (...) do przyrody” (Kant 2004, 47-49), a więc również do rozpoznawania go w zaludniających ją organizmach. Rozpoznanie od poszanowania dzieli zaś dokładnie taka odległość, jaka dzieli *Krytykę czystego rozumu* (epistemologię) od *Metafizyki moralności*.

1.3.3 Pojęcia normatywne a pojęcia naturalne. Prolegomena do praw zwierząt w sensie jurystycznym

Normatywność przynależy zarówno moralności, jak i prawu. W konsekwencji obie te dziedziny charakteryzuje namysł dokonywany przy pomocy pojęć normatywnych (*normative concepts*). W winnych dyscyplinach, przede wszystkim w naukach przyrodniczych, refleksja prowadzona jest za pomocą pojęć naturalnych (*naturalistic concepts*). Pojęcia normatywne są zdaniem Korsgaard w inny sposób związane z rzeczywistością empiryczną niż pojęcia naturalne, gdyż mają implikacje normatywne. Niemniej bazują one w jakiś sposób na pojęciach naturalnych, tj. przedstawicielom rodzaju *homo sapiens sapiens* przypisywane jest na gruncie normatywnym pojęcie osoby⁹. Tym samym – w większości koncepcji – jedynie ludziom przysługuje podmiotowość moralna i prawna, składające się na ich pozycję normatywną (*normative standing*). Pozycja normatywna pozwala wysuwać roszczenia moralne i prawne. Jakże jest jednak uzasadnienie przejścia od pojęcia naturalnego do pojęcia normatywnego? Ten problem nazywa Korsgaard problemem „twierdzenia pomostowego” (*bridge-claim*)¹⁰. „W Kantowskiej koncepcji twierdzeniem stanowiącym pomost między istotami racjonalnymi i osobowością normatywną jest dokładnie to: twierdzenie (*claim*)” (Korsgaard 2013, 8). Rozwijając swoją myśl, harwardzka filozofka powiada, że idzie o twierdzenie w sensie praktycznym – a więc w ścisłym sensie *twierdzenie zawierające roszczenie do ważności*, które ujawnia się w moralnej legislacji, podczas której nadajemy prawo sobie i innym (lub – na gruncie koncepcji powszechnego prawodawstwa – wszystkie jednostki nadają sobie prawa).

Korsgaard uznaje, że z własności ludzkiej natury jaką jest normatywność (konieczność posiadania racji do działania) nie wynika, że tylko człowiek posiada pozycję normatywną, tj. twierdzenie o posiadaniu roszczeń moralnych względem innych osób nie jest podnoszone przez daną jednostkę w imieniu istot racjonalnych, a wszystkich istot, dla których coś może być dobre lub złe w sposób naturalny (naturalne dobro). Reprezentowanie siebie jako celu samego w sobie obejmuje co prawda to, że traktujemy nasze wybory jako prawo, nie dokonujemy ich jednak z uwagi na naszą prawodawczą naturę, a z uwagi na to, że uznajemy, że to a to będzie dla nas dobre, a z racji własnej racjonalności, która zmusza do poszukiwania uzasadnienia, uznajemy, że to, co dobre dla nas, jest też dobre absolutnie (nie mamy bowiem wglądu do świata wartości jako takiego, albo też wychodzimy z założenia o podmiotowym źródle wartości, nawet, gdy rozważamy podmiot w wersji znaturalizowanej, a nie metafizycznej). Coś dobrego w sposób absolutny może mieć charakter powszechnie ważny (a więc i powszechnie wiążący/obowiązujący). Zatem, to, że jesteśmy w stanie podjąć jakiś wybór, działać moralnie wynika z rozumnej autonomii, niemniej „materiałem” dla takiego postępowania, czyli *de facto* sposób, w jaki postępujemy, jest uzasadniony uznaniem czegoś za dobre dla nas – a to z kolei przesądza o podobieństwie łączącym nas z innymi żywymi istotami.

Reprezentacja samych siebie jako celów samych w sobie ujawniająca się zarówno

⁹ Swoją drogą trzeba wspomnieć, że jest to ujęcie współczesne. Trzeba bowiem pamiętać o niewolnikach i ich rzeczowym statusie przełamanym u Kanta.

¹⁰ W analogii do prawa pomostowego (*bridge law*) w matematyce.

w (moralnie) prawodawczym działaniu, jak i uznaniu własnego dobra za dobro absolutne znajduje odzwierciedlenie „w dwóch delikatnie różnych sensach „celu samego w sobie” (Korsgaard 2012, 10), które zakładają jak ujmuje to sama Korsgaard, że „muszę cię szanować jako cel sam w sobie w sensie aktywnym, jeśli postrzegam cię jako istotę zdolną do nadawania mi praw, i ty sam nakładasz na mnie obowiązek, bym szanował twe wybory i pomagał ci w osiągnięciu twych celów. Muszę postrzegać cię jako cel sam w sobie w sensie biernym, jeśli jestem zobowiązany do traktowania twych celów, a przynajmniej rzeczy, które są dla ciebie dobre, jako dobrych absolutnie” (Korsgaard 2012, 10). Poziom celu samego w sobie w sensie biernym wyznacza zatem kategoria naturalnego dobra; z kolei poziom celu samego w sobie w sensie aktywnym pozostaje tradycyjnie kantowski. Tym samym, Korsgaard próbując zachować wagę rozumnej autonomii jako koniecznego („moralnego” i zarazem „transcendentalnego”) warunku możliwości uzyskania statusu prawodawcy, włącza zwierzęta do wspólnoty moralnej, jaką jest królestwo (inaczej: państwo) celów, definiując je jako istoty nieposiadające obowiązków, a jedynie prawa.

Jeżeli zaś idzie o prawo (w sensie jurystycznym), sytuacja ma się podobnie. Korsgaard wskazuje na wywodzącą się z prawa rzymskiego prawną dychotomię (*legal bifurcation*), która dzieli niemal wszystko, co istnieje na osoby (ludzie) i własność (rzeczy, zwierzęta, a do pewnego momentu historycznego także niewolnicy). Podobnie jak w przypadku pojęć normatywnych na gruncie moralności (osoby i rzeczy), tak i na gruncie prawa Kantowskie kategorie normatywne, zachowujące rzymski podział, opierają się na niewłaściwie zidentyfikowanych podstawach.

1.3.4 Zwierzęta a prawo

Prawo w filozofii I. Kanta tyczy się wolności zewnętrznej, a więc wolności jednego podmiotu od przymusu nakładanego na niego przez inny podmiot, co wynika wprost z powszechnego pryncypium prawa:

Zgodne z prawem jest każde działanie, które (lub maksyma tego działania) sprawia, że wolność arbitralnej woli może współistnieć z wolnością wszystkich innych podług powszechnego prawa (Kant 2005, 41).

Jedynie pozornie uprawnienie do stosowania przymusu stoi w konflikcie z wolnością obywatela, gdyż przymus może zostać użyty wyłącznie w sytuacji, gdy obywatel korzysta ze swojej wolności w sposób naruszający wolność innego członka wspólnoty.

Podstawy do budowania moralnego i prawnego statusu w Kantowskiej metafizyce moralności – jak zauważa Korsgaard – są odmienne. Niemniej pomiędzy położeniem moralnym zwierząt jako wyłącznie środków a położeniem prawnym jako wyłącznie przedmiotu prawa własności istnieje korelacja.

Kazus praw zwierząt polega na tym, że jako istoty nierozumne nie uosabiają one wolności w sensie „metafizycznym” przyjętym przez Kanta, co skutkuje m.in. niemożnością zastosowania sankcji przymusu, której użycie wymaga spełnienia określonej przesłanki,

a mianowicie zagrożenia wolności zewnętrznej. Jeżeli ochrona zwierząt ma skupiać się na interesach zwierząt, a nie na ich wolności, to może się wydawać, że w kantowskim systemie nie ma miejsca dla praw zwierząt. Korsgaard jednak uznała, że dokładniejsza egzegeza Kantowskiej filozofii prawa podważa takie przypuszczenia. Sięgając do uzasadnienia istnienia własności, filozofka tłumaczy, co tak naprawdę – w jej opinii – jest podstawą do powstania tego prawa.

Sam Kant pojmował jednak prawo własności jako „tożsame z uprawnieniem do prywatnego użytkowania przedmiotu, który (jako własność nabyta w sposób pierwotny bądź wtórny) staje się mój na mocy tego, że jego posiadaczem źródłowym była zrazu cała ludzkość” (Kant 2005, 83). Własność – podobnie jak u Johna Locke’a (abstrahując od innych różnic, które pojawiają się u obu myślicieli w kontekście umowy społecznej i prawa własności) – powstaje zatem przez swoisty akt zawłaszczenia w stanie natury. Instytucja prawa własności ma zatem genezę przedpaństwową, a samo państwo powstaje z potrzeby jej zabezpieczenia. „Własność jest jednym z podstawowych i niezbędnych składników wolności praktycznej” (Szymaniec 2009, 33).

Jak pisze Ewa Nowak w przedmowie do polskiego tłumaczenia *Metafizyki moralności* „prawo własności nie wypływa z niezapośredniczonego stosunku osoby do rzeczy, lecz właśnie ze stosunków międzyosobowych” (Nowak 2005, xxviii). Korsgaard zupełnie odwraca tę optykę, dokonując rewizji oryginalnych założeń Kanta. Wskazuje ona m.in. na to, że źródłem prawa własności nie są stosunki między osobami, tzn. nie posiada ono genezy społecznej (np. kontraktualistycznej), lecz niezbywalna, biologiczna, żywotna i egzystencjalna konieczność korzystania ziemskich zasobów naturalnych. Konieczność ta wynika stąd, że warunkują one przetrwanie. Sprawia to, że uzasadnienie (podstawa) prawa własności przybiera postać empiryczną¹¹ (acz bez ingerencji w społeczną legitymizację innych praw). W zasadzie, jeśli idzie o rodzaj ludzki, Kant już uczynił użytkowanie naturalnych zasobów Ziemi jednym z podstawowych praw ludzkich. Jego naturalistyczne zręby zarysował w *Projekcie wiecznego pokoju* oraz – skrótowo – w *Metafizyce moralności*, czyniąc to w kontekście *ius cosmopolitanum* i swobodnej „cyrkulacji ludzkiej”. Jak pisze Kant,

Cała ludzkość tworzy już *od zarania* coś na kształt wspólnoty ziemi (*Gemeinschaft des Bodens*) (...) Jest wielce prawdopodobne, że poszczególne ludy będą się ze sobą stykać fizycznie, a nawet *oddziaływać na siebie*, nawiązując trwałe lub przejściowe stosunki wzajemne (Kant 2005, 205).

Gdyby terminy „ludzkość” i „ludy” zastąpić określeniem żywe istoty, tj. wyposażone w zdolność sensomotoryczną zwierzęta, mielibyśmy przed oczyma schemat

¹¹ Aczkolwiek nabywanie praw może być w kantyzmie zapośredniczone intersubiektywnie zarówno na sposób transcendentally-społeczny w sensie epistemologicznym (np. założenie o wolnej i rozumnej woli wszystkich prawodawców) i czysto normatywnym (tytuł prawny do własności, rozporządzania nią, itd.), jak i na sposób transcendentally-społeczny w sensie empirycznym (np. założenie o „wzajemnym” przymusie wpisanym w pojęcie prawa), zob. (Kant 2005)

powszechnego dostępu do naturalnych zasobów. W wypadku człowieka konieczny byłby korektyw uwzględniający gospodarkę rolną i leśną, problematyczną gospodarkę mięsną, przemysł i wszystko, co sprawia, że zasoby częściowo utraciły swój naturalny charakter pod wpływem stosowanej przez człowieka technologii i nienaturalnego stylu życia.

Niemniej, owo przeznaczenie wyrażające się w nieuchronnej potrzebie korzystania z ziemskich zasobów przynależnych do sfery *physis* człowiek dzieli z innymi istotami naturalnymi, czyli zwierzętami¹². Jest to jakiś podstawowy i naturalny rodzaj wolności, w który Korsgaard włączyła również „wolność do używania własnego ciała, aby wypracować sobie rodzaj godnego życia” (Korsgaard 2012, 23). Zasluguje ona na ochronę prawną w nie mniejszym stopniu niż wolność uwarunkowana metafizycznie i racjonalnie. W świetle naturalnej wolności zakwalifikowanie zwierząt do kategorii normatywnej, jaką stanowią własność i rozmaite jej formy (property) okazuje się nieuzasadnione. Korsgaard nie wysnuwa jednak absurdalnych wniosków z tej konkluzji, np. o nadawaniu własności zwierzętom. Podkreśla jedynie, że z tak pojętego uzasadnienia prawa własności wynika ciężący na ludziach obowiązek zapewnienia możliwości godnej egzystencji także innym żywym istotom.

Dalej posunięte „odspołecznienie” bądź deintersubiektywizacja podstaw prawa mogłyby zrujnować Kantowską doktrynę prawa, jednak tego wątku nie mogę tutaj rozważać. Na szczęście różnorodność teorii i filozofii prawa pozwala przyjmować różnorodność podstaw, a od typu ustroju prawnego zależy, które z nich – i czy w ogóle jakieś – będą uznawane za wiążące.

2. Analiza semantyczna. Sens i odniesienie

Rozdział ten ukonstytuuje metodologię, jaką przyjmę w analizie przedstawionych powyżej poglądów Korsgaard. Analiza ta będzie analizą semantyczną – opierać się będzie na zmianie rozumienia pojęcia „celu samego w sobie”. Będzie więc mowa o swoistej semantyce terminu moralnego (czy też rozpatrywaniu go z perspektywy metaetycznej). Przy takiej metodzie trzeba jednak określić o jakiej semantyce mowa.

Przede wszystkim chciałabym wskazać na wywodząca się od Gottloba Fregego¹³ różnicę między „sensem” a „odniesieniem”. Jak zauważa John Lyons, „innymi parami terminów oznaczającymi te same (bądź przynajmniej podobne) przeciwstawienie są: znaczenie i referencja (...), konotacja i denotacja; treść i zakres; intensja i ekstensja” (Lyons 1984, 176).

Nie sposób przecenić wpływu dystynkcji między „sensem” a „odniesieniem” i

¹² Można z tym naturalnie dyskutować, albowiem np. prawo do korzystania z zasobów naturalnych przez nomadów na terenie Afryki okresu prekolonialnego czy Ameryki prekolumbijskiej nie przerodziło się w prawo własności; ponadto, użytkowanie i własność również w świetle prawnym stanowią osobne kategorie.

¹³ Najczęściej jako filozofa, który dokonał tej dystynkcji wskazuje się właśnie Fregego. Niemniej można doszukiwać się jej już wcześniej m.in. u Johna Stuarta Milla. Frege wprowadził także inne istotne innowacje, takie jak „rewolucja składniowa” i zasada składalności sensów.

w ogóle dzieł Fregego na współczesną filozofię języka oraz filozofię analityczną. Dość powiedzieć, że pisma Fregego doczekały się rozmaitych, niekiedy wzajemnie sprzecznych interpretacji, a także wielu twórczych rozwinięć. Znanymi komentatorami Fregego są m.in. Tyler Burge i Michael Dummett.

Pragnę poczynić w tym miejscu także pewne ustalenia terminologiczne. Jak do tej pory doczekaliśmy się dwóch przekładów *Über Sinn und Bedeutung* na język polski. „Bedeutung” przez Bogusława Wolniewicza tłumaczony jest jako „znaczenie”, zaś przez Jerzego Pelca jako „nominat”¹⁴. Idąc za intuicjami tłumaczy na język angielski (*On the sense and reference*) autorka niniejszej pracy, mając na myśli „Bedeutung”, będzie używać odpowiednika „odniesienie”. Podobnie czyni Janusz Maciaszek (Maciaszek 2008). W literaturze można się spotkać także z określeniem „korelat semantyczny” (Besler 2010).

Wątpliwości może budzić jednak kwestia, iż jak pisze Maciaszek, teoria znaczenia nie winna być traktowana jako „metateoria humanistyki” (Maciaszek 2008, 212), bowiem jej zadaniem jest odpowiedzieć na pytanie, „jak możliwe jest rozumienie języka, czyli dzięki czemu język może spełnić swoją funkcję komunikacyjną” (Maciaszek 2008, 213). Jednak, jak konstatuje Jan Woleński, „jeżeli się pewne rozważania nazywa semantycznymi, to trzeba jasno określić, o jakiej semantyce się mówi” (Woleński 1972, 42),

Moim celem jest pokazanie, jak zmiana pewnego założenia filozoficznego dokonana celem objęcia ochroną nowej grupy podmiotów, tzn. podmiotowe rozszerzenie zastosowania praw, implikuje nie tylko zmianę odniesienia pewnych pojęć, ale i zmianę sensu pewnych wyrażań filozoficznych zawierających te pojęcia. Mówiąc prościej, chciałabym unaocznic, w jaki sposób zmiany takiej dokonała Korsgaard. Cel ten zamierzam zrealizować poprzez szkicowe nakreślenie genezy tych kategorii semantycznych (a raczej pewnej kategorii pragmatycznej oraz semantycznej). Badanie zmian semantycznych nie będzie tak trudne, jak w języku potocznym, ponieważ język filozoficzny często niweluje nieostrość języka potocznego poprzez definicje projektujące. W tym przypadku będą to definicje projektujące *sui genesis*, jako że znaczenie analizowanych przeze mnie terminów będzie wypływać z założeń, jakie legły u podstaw koncepcji Korsgaard. W tym miejscu chciałabym również zaznaczyć, że – po pierwsze – nie będę odnosiła się do tego, jak Frege traktuje zdania, gdyż nie będzie mi to potrzebne w dalszej części pracy. Po

¹⁴ Źródła tej konfuzji translatorskiej tłumaczy się faktem, iż „Bedeutung” używane jest w potocznej niemczyźnie jako znaczenie. Z kolei sam Frege w pismach z lat 1892-1895 tłumaczy, że różnica między „znaczeniem” (*sense*) i „odniesieniem” (*meaning*) polega na tym, że nazwie własnej (*proper name*) i pojęciu (*concept-word*) odpowiada desygnowany bądź nazywany przez nie przedmiot, który pod nie „podpada”. Reguły określające relację odpowiedniości wyznaczają z kolei prawdziwość twierdzeń, w których zostały one użyte. Wyrażenia o charakterze pojęć (*concept-word*) posiadają ekstensje będące przedmiotami, a nie pojęciami. Ekstensjonista powie ponadto, że „dane dwa pojęcia znaczą to samo wtedy i tylko wtedy, gdy ich ekstensje koincydują ze sobą”. Frege wymienia również nazwy i znaki jedynie wskazujące na przedmiot (*a sign which only indicates an object*) (Frege 1979, 120 -122).

drugie, zastosowane przeze mnie kategorie nie będą używane tak, jak wymagałby tego ortodoksyjny fregizm (jeśli takowy istnieje).

2.1 Sens i odniesienie

W przedmowie do przetłumaczonych przez siebie *Pism semantycznych* Bogusław Wolniewicz pisze, że „trzonem pism semantycznych Fregego są jego trzy fundamentalne artykuły z lat 1891–1892” (Wolniewicz 2014, xvii). Są to *Funktion und Begriff* (*Funkcja i pojęcie*), *Über Begriff und Gegenstand* (*Pojęcie i przedmiot*) oraz *Über Sinn und Bedeutung* (*Sens i odniesienie*¹⁵). Skoncentruję się – jak już wyżej zostało powiedziane – na „sensie” i „odniesieniu”. Zaznaczyć jednak trzeba, że Frege w *Über Sinn und Bedeutung* wychodzi od nazw własnych, bowiem w przeciwieństwie do Milla przyznaje im – mówiąc językiem tego drugiego – zdolność konotowania (konotatywność).

Swoją refleksję Frege rozpoczyna od pytania o naturę tożsamości, a także od stwierdzenia, że „ $a = a$ ” oraz „ $a = b$ ” są zdaniem o nierównej wartości poznawczej” (Frege 2014, 60). To spostrzeżenie implikuje, że tożsamość nie jest stosunkiem między nazwami przedmiotów czy znakami, lecz – wbrew wcześniejszym poglądom Fregego – między przedmiotami. Przedmioty te stanowią odniesienie nazwy. Tym zaś, co odróżnia „ a ” od „ b ”, jest *sens*. Sens, mówiąc najbardziej lakonicznie, to „sposób dania przedmiotu”. Stanowi on oświetlenie jednego aspektu odniesienia i „chwytą [go, M.A.] każdy kto zna wystarczająco język lub symbolikę” (Frege 2014, 62), do których dana nazwa przynależy. Sens determinuje odniesienie, a odmienne sensory po prostu w różny sposób doprowadzają do odniesienia, pełniąc tym samym doniosłą funkcję poznawczą. Sens ma charakter konieczny, gdyż istnieją wyrażenia, które legitymują się jedynie sensem, a nie posiadają odniesienia (np. Pegaz).

Colin McGinn w kontekście semantyki Fregego pisze:

Jeżeli staramy się analizować wyrażenie takie jak „prezydent Stanów Zjednoczonych”¹⁶, to mamy aż cztery poziomy do przeanalizowania (i) wyrażenie językowe (ii) sens, który oświetla aspekt (iii) aspekt oświetlony przez sens oraz (iv) odniesienie, przedmiot (McGinn 2015, 15).

W dalszej części pracy nie będę odnosić się do nazw własnych, lecz do pojęć. Różnicę między nimi można zilustrować następująco:

„X jest Krzysztofem Kolumbem”

„X jest Włochem”

Struktura gramatyczna obu powyższych zdań wygląda identycznie, niemniej słowo „jest” ma inną wymowę semantyczną. W pierwszym przypadku między „X” a „Krzysztofem

15 Wolniewicz tłumaczy ten tytuł jako „Sens i znaczenie”.

16 Deskrypcja ta traktowana jest równoważnie z nazwą własną. Ponadto Frege rzeczowników poprzedzonych rodzajnikiem określonym nie traktował jako pojęć, co unaocznia znany problem „the horse concept”.

Kolumbem” zachodzi stosunek tożsamości, równoznaczności i pełnej koincydencji (*complete coincidence*) (Frege 1979, 121); w drugim przypadku zachodzi subsumcja, tzn. „X” podpada pod pojęcie „Włocha”. Pojęcia mają zdaniem Fregego naturę predykatów, „predykatywny charakter pojęć jest natomiast szczególnym przypadkiem niezupełności czy nienasyconia, które stanowią istotę funkcji” (Maciejczak 2016, 78) bądź ekstensji, o której wspomniano w jednym z przypisów. „X” jest tutaj przedmiotową ekstensją pojęcia „Włoch”. Na ten temat traktuje rozprawa autorstwa Fregego pt. *Funktion und Begriff*.

Czy jednak wolno powiedzieć, że predykaty – podobnie jak nazwy własne – posiadają sens i odniesienie?

Michael Dummett uznaje, że „tylko filozof stawiający całe pojęcie sensu pod znakiem zapytania mógłby wątpić w to, czy wyrażenia nienasycone, takie jak predykaty, mają sens” (Dummett 1973, 204). Niemniej sam Frege „rzadko kiedy wypowiada się o sensie wyrażen nienasyconych” (Dummett 1973, 206), stąd twierdzenie to wzbudza liczne kontrowersje u komentatorów. Autorka niniejszej pracy przyjęła, że pojęcia mają sens analogiczny do nazw własnych, co potwierdza Frege w *Comments on Sense and Meaning*.

Odniesienie predykatów stanowi wobec tego zbiór przedmiotów podpadających pod dane pojęcie (zob. Besler 2010, 75). U Fregego stosunek podpadania pod pojęcie ma naturę ontologiczną, tymczasem stosunek oznaczania ma naturę semantyczną, co należy skrupulatnie rozróżniać. Chciałabym tutaj podkreślić, że nie jestem rzeczniczką jednej, określonej koncepcji semantycznej. Przywołując (skądinąd dynamiczną) koncepcję Fregego, staram się jedynie zwrócić uwagę na to, że z analizy języka nie da się całkowicie wykluczyć płaszczyzny sensu, np. poprzez zredukowanie go – jak czyni to Donald Davidson – jedynie do ekstensji. Nie zamierzam przy tym rozstrzygać w tym miejscu, czy kategoria intensji jest kategorią natury semantycznej (co prowadziłyby oczywiście do określonych konsekwencji ontologicznych), czy pragmatycznej¹⁷. Będę jedynie analizować zmianę w obrębie tego, co dane terminy znaczą na płaszczyźnie pojęć, oraz tego, co oznaczają na płaszczyźnie podpadających pod nie przedmiotów, a także świadomie odróżniać znaczenie od oznaczania. Terminy będę traktować jako pojęcia. Okaże się przy tym, czy znaczenie może być wyznaczane przez logiczne reguły odniesień między pojęciami: przypuszczalnie tak właśnie będzie.

3. Rewidując imperatyw kategoriowy

Jak wiadomo, Korsgaard dokonała rewizji imperatywu kategoriowego, wywracając jednocześnie do góry nogami podstawy etyki kantowskiej. I tak, od imperatywu kategoriowego w sformułowaniu celowościowym, tj.

¹⁷ Jako pragmatyczną ujmuje ją Woleński uznając jednocześnie, że ma to poważne konsekwencje, bowiem pozwala na ujęcie kategorii znaczenia (sensu) jako znaczenia psychologicznego (Woleński 1972, 44). W kontekście niniejszej pracy dyskusje wokół „statusu znaczenia” mają wagę drugorzędą.

IKc1: „Postępuj tak, byś człowieczeństwa tak w twojej osobie, jako też w osobie każdego innego używał zawsze zarazem jako celu, nigdy tylko jako środka” (Kant 2017, 46)

przechodzi Korsgaard do nowej zasady moralnej, którą można wyrazić następująco:

IKc2: Postępuj tak, byś zwierzęcej natury (organizmu, który ma znaczenie sam dla siebie) tak w twojej osobie, jako też w osobie każdego innego zwierzęcia używał zawsze zarazem jako celu, nigdy tylko jako środka.

Jak dokładnie rozumie ona pojęcia „człowieczeństwa” i „zwierzęcej natury”, pokazane zostało we wcześniejszych rozdziałach. Na tym etapie chciałabym skoncentrować się na opisie zabiegów dokonanych przez nią celem wykazywania obowiązywalności IKc2. IKc2 opiera się na

1. deracjonalizacji Kantowskiego dobra poprzez jego naturalizację (bądź przejściu od pojęcia rozumności do pojęcia racjonalności jako własności umysłu nie tylko ludzkiego);
2. zmianie członu w relacji tożsamości;
3. tym samym zmianie sensu wyrażenia „cel sam w sobie”, które ma naturę pojęciową;
4. zmianę odniesienia (rozszerzenie ekstensji) powyższego wyrażenia pojęciowego.

Aby uniknąć konsekwencji w postaci wykluczenia zwierząt ze wspólnoty moralnej (metafizycznego królestwa celów, którego definicja w kategoriach teleologicznych również ulegnie zmianie wraz ze zmianą wskazaną w punkcie 3), potrzeba skutecznych strategii i narzędzi do obrony ich praw. Powinny one umożliwić zbudowanie takiego systemu moralnego, w którym istoty ludzkie będą traktować zwierzęta jako cele same w sobie, tzn. mieć wobec nich bezpośrednio pewne zobowiązania lub obowiązki moralne. Pojęcie „celu samego w sobie” musi więc zostać rozszerzone. Aby tak się stało, należy koniecznie zlokalizować wartość wewnętrzną w odmiennej aniżeli „rozumność” cesze dystynktywnej, dzielonej przez wszystkie zwierzęta.

Nową cechą dystynktywną pozwalającą wszystkim zwierzętom stanowić materię prawa moralnego, a więc cechą moralnie relewantną, stanowić będzie właśnie naturalne dobro.

3.1 Naturalizacja dobra. Metakomentarz

O tym jak wygląda kategoria „naturalnego dobra” pisałam już w 1.3.2. W tym miejscu pragnę skupić się na konsekwencjach zabiegu naturalizacji.

Tym, co najbardziej uderza w koncepcji Korsgaard jest odmienne ugruntowanie twierdzeń normatywnych. Jak pisze filozofka:

Mówię, że to jest dobre dla mnie i dlatego to jest dobre. To sugeruje adekwatny fakt o mnie – podstawę na której uznaję swoją pozycję normatywną – czyli że jestem typem istoty, dla której albo z której punktu widzenia, rzeczy mogą być dobre lub złe (Korsgaard 2013, 14).

Naturalne dobro jest tym dobrem, na podstawie którego przypisujemy sobie (podstawową) pozycję normatywną, *ergo* posiadanie pozycji normatywnej zależy jedynie od łączącego nas z innymi zwierzętami swoistego statusu biologicznego, a nie od autonomii w rozumieniu kantowskim.

Pietrzykowski zabieg dokonany przez Korsgaard ocenia pozytywnie¹⁸ uznając, że wprowadzenie kategorii naturalnego dobra stanowi „swego rodzaju naturalizację etyki Kantowskiej i otwarcie możliwości uzgodnienia jej ze współczesnym obrazem świata wyłaniającym się z postępów wiedzy przyrodniczej” (Pietrzykowski 2015, 119). Czy taka naturalizacja nie niweczy jednak *metafizyczno-intelektualnego* projektu etyki Kantowskiej i przy okazji nie degraduje misternych założeń i racji, na których Korsgaard oparła swój konstruktywizm aksjologiczny? Wydaje się bowiem, że osadzenie i ugruntowanie wartości w naturze zamiast w poznawczym i rozumnym potencjale umysłu sprawi, że człowieczeństwo będzie można uznać za źródło wartości jedynie w takim sensie, że człowiek będąc istotą mimo wszystko racjonalną jako jedyny będzie w stanie wartości identyfikować, tzn. rozpoznawać je w naturze. A jednak stanowisko Kanta, przywołane tu wcześniej na podstawie *Krytyki władzy sądowniej*, jest pod tym względem bardziej optymistyczne, zakłada bowiem, że umysł dysponuje kategorią celu i wartości, i dlatego właśnie potrafi dostrzegać je w naturze, co nie przeszkadza mu tworzyć wartości poprzez działanie wypływające z „dobrej woli”. Nie jest więc tak, że obiektywność wartości pochodzących z konstrukcji transcendentального podmiotu zostanie zastąpiona i wyparta przez obiektywność wywodzącą się z natury i mającą inne podstawy ontologiczne niż ontologia episteme i aksjologii w ludzkim umyśle. Jest raczej tak, że pierwsza obiektywność znajdzie dopełnienie w drugiej, i że obiektywności te będą „koincydować” ze sobą, by użyć tu określeń Fregego.

Nie pod każdym jednak względem można się zgodzić z powyższym rozumowaniem. Trzeba bowiem pamiętać, że zwierzęcej naturze przysługuje wartość wewnętrzna – a to z powodu jej cechy dystynktywnej, pozwalającej stać się „materiałem” dla prawa moralnego, w zgodzie z wyjściowym twierdzeniem Korsgaard o nieekwiwalencji treściowej trzech formuł imperatywu. Zwierzęta dzięki naturalnemu dobru „awansują” z kategorii dobra zewnętrznie instrumentalnego do kategorii dobra wewnętrznie finalnego. Pozostałe wartości (zarówno zewnętrzne finalne, jak zewnętrzne instrumentalne) na gruncie tak znaturalizowanego dobra mogą nadal pozostać. Ponadto konstrukcja ta – co wydaje się najbardziej istotne – wciąż bazuje na kwalifikacji ludzkich zachowań jako dobrych lub złych w kontekście racji działania podmiotu („motywów” w języku Kanta), a nie

¹⁸ Zapewne wynika to z faktu, iż w swych pracach z zakresu prawoznawstwa zwraca się ku naturalizmowi (zob. Pietrzykowski 2017).

naturalnych skłonności.

Można zatem powiedzieć, że kantyzm Korsgaard ma charakter nieortodoksyjny czy wręcz nawet heretycki. Eufemizując nieco, można go określić mianem wybiórczego. Wybiórcza wydaje się też interpretacja Arystotelesa, bowiem filozofka odwołuje się do kategorii życia wegetatywnego, nie wyciągając jednak wszystkich konsekwencji związanych z jej użyciem.

3.2 Cel sam w sobie

Dla porządku przywołam ponownie dwie wersje imperatywu kategorycznego w formie celowościowej – tę oryginalną oraz tę zrewidowaną:

IKc1: „Postępuj tak, byś człowieczeństwa tak w twojej osobie, jako też w osobie każdego innego używał zawsze zarazem jako celu, nigdy tylko jako środka” (Kant 2017, 46), oraz

IKc2: Postępuj tak, byś zwierzęcą naturę tak w twojej osobie, jako też w osobie każdego innego zwierzęcia używał zawsze zarazem jako celu, nigdy tylko jako środka.

Podstawą traktowania kogoś jako celu nigdy tylko jako środka jest w obu formach „bycie celem samym w sobie”, „podpadanie pod pojęcie celu samego w sobie” (dobro zewnętrzne finalne). Aby orzec ten predykat o osobie (podmiocie), musi ona charakteryzować się określonymi cechami dystynktywnymi. U Kanta, jak i u „wczesnej” Korsgaard cechą taką stanowi (przynajmniej potencjalna¹⁹) rozumność: „istoty rozumne nazywamy *osobami*, ponieważ już ich natura wyróżnia je jako cele same w sobie” (Kant 2017, 45). Tym niemniej, przełom zainicjowany przez Korsgaard na gruncie koncepcji dobra zaskutkowało nową cechą dystynktywną – tj. *byciem istotą, dla której coś może być dobre lub złe w sposób naturalny*. Naturalizacja dobra pociągnęła zaś za sobą zmianę sensu i odniesienia pojęcia „celu samego w sobie”.

W IKc1 „bycie celem samym w sobie” tożsame jest z człowieczeństwem, dobrą wolą czy możliwością racjonalnego determinowania celów (działania bądź co najmniej ochrony, nietykalności, biernego szacunku), oczekiwanymi od istot rozumnych. Wynika to z przyznania człowiekowi jedynej nieredukowalnej wartości z uwagi, na to, że on sam jest źródłem wartości. Zatem stosunek tożsamości kształtuje się następująco:

Bycie celem samym w sobie = bycie wartościowym wewnątrznie = bycie istotą rozumną ($F(x) = G(x) = H1(x)$)

Zarówno $F(x)$, $G(x)$, $H1(x)$ denotują ten sam zbiór (zbiór wszystkich istot rozumnych, tj. każdy podmiot cechujący się rozumnością podpada pod powyższe pojęcia), nie są

¹⁹ Tzn. w toku ontogenezy nie zawsze rozwinięta i zaktualizowana zdolność, uznana za gatunkową cechę dystynktywną (*sapiens*). Kant wymieniał cały szereg potencjalnie rozumnych przedstawicieli gatunku. Prawodawstwo moralne wymaga aktywnej rozumności, jednak podleganie biernej ochronie prawnej tego nie wymaga, co podkreśla Paweł Łuków, wywodząc z tego ustalenia możliwość praw zwierząt (zob. Łuków 2006).

jednak synonimiczne, lecz zastępowalne (wymienialne) *salva veritate* (por. Frege, 1979, 118: "can replace one another *salva veritate*, so too can concept-words, if their extension is the same"). Abstrahuję w tym momencie od argumentu z tzw. przypadków marginalnych (argument from marginal cases, w skrócie AMC), których zasada się na tym, że niezależnie jaką ludzką cechą charakterystyczną wybierzemy jako podstawę dla posiadania praw, to znajdują się ludzie (np. osoby ciężko upośledzone), które będą wykazywać daną cechę w stopniu mniejszym niż część zwierząt (w tym przypadku rozumność). W IKc2 „bycie celem samym w sobie” jest tożsame ze zwierzęcą naturą. Zatem:

„bycie celem samym w sobie = bycie wartościowym wewnątrznie = bycie organizmem, dla którego coś jest dobre lub złe w sposób naturalny ($F(x) = G(x) = H_2(x)$)

W nowym sformułowaniu imperatywu $F(x)$ i $G(x)$ tożsame są z $H(x)$ i denotują zbiór wszystkich istot, dla których coś może być dobre lub złe w sposób naturalny. Oczywiście zbiór wszystkich istot racjonalnych zawiera się w zbiorze istot, dla których coś może być dobre lub złe w sposób naturalny. Niemniej, aby uzyskać poprawny logicznie stosunek tożsamości, Korsgaard musiała zmienić także sens $F(x)$ i $H(x)$, przez co mamy *de facto* do czynienia z $F_1(x)$ i $F_2(x)$ oraz $G_1(x)$ i $G_2(x)$. Ponadto, w kontekście wspomnianego w 1.3.3 rozumienia aktywnego i pasywnego celu samego w sobie, chciałabym wskazać, że należy raczej mówić o celach samych w sobie w sensie aktywnym jako o klasie wyróżnionej spośród celów samych sobie poprzez możliwość rozpoznania samego siebie jako celu samego w sobie aniżeli o zwykłym podziale na cel sam w sobie bierny i cel sam w sobie aktywny. W przekonaniu Korsgaard podział taki byłby bowiem moralnie irrelevantny. Najbardziej elementarne rozumienie „celu samego w sobie” konstituuje się w oparciu o naturalne dobro.

Z meta-perspektywy można zatem przyjąć, że Korsgaard zmodyfikowała zmienną w stosunku tożsamości. Jak zatem modyfikacja ta wpłynęła na pojęcie „celu samego w sobie” i „wartości wewnętrznej”?

Jak wykazywałam w 1.1, „cel sam w sobie” stanowił bezwarunkowy warunek (*unconditioned condition*) wszelkiej (innej, „dalszej”) wartości z uwagi na to, że warunek wszelkiej wartości sam musi być wartościowy²⁰. Warunek wszelkiej wartości to człowiek, bowiem, to on nadaje wartość – gdyby nie istniał, nie istniałaby też żadna wartość. Za takim stwierdzeniem stoi przyznanie człowiekowi – jako jedynej istocie żywej – wartości wewnętrznej ze względu na potencjał rozumnej autonomii, a co za tym idzie, pozycję prawodawczą w państwie celów. Na tym etapie poglądy Korsgaard nawiązywały jeszcze do ujęcia człowieka z perspektywy transcendentnej, jako inteligibilnego *homo noumenon*. Nawiązując do dwóch dystynkcji w pojęciu dobra (a więc do tradycji analitycznej, w której obraca się filozofka), aspektem, który oświetlało pojęcie „celu samego w sobie”

²⁰ Analogicznie twierdzi Kant w *Metafizyce moralności*, że bezprawie nie może być źródłem prawa.

w pierwotnej wersji była nieinstrumentalność wartości ludzkiego człowieczeństwa, zaś pojęcie „wartości wewnętrznej” wskazywało na aspekt ulokowania tej wartości. Pierwszy termin odnosi się zatem do sposobu, drugi do źródła, które miało charakter transcendentálny.

Tymczasem na poziomie ICK2 „cel sam w sobie” przestaje być wartościowy z racji tego, że warunek wszelkiej wartości sam jest wartościowy. Korsgaard uważa, że „wyodrębniając ideę celu samego w sobie jako jedyne go możliwego źródła wartości w świecie faktów” (Korsgaard 2004, 108) Kant miał rację, ale „powodem, dla którego ta wartość w ogóle istnieje, jest to, że istnieją takie istoty, które są istotne same dla siebie” (Korsgaard 2004, 108). Wartościowe wewnętrznie będzie więc w pierwszym rzędzie wszystko to, co żywe. Twierdzenie to nie tylko zbliża Korsgaard do naturalizmu, lecz prowadzi ją także go odrzucenia – jak u Foot – błędu naturalistycznego, bowiem to natura poprzez taki a nie inny sposób „zaprojektowanie organizmów” dyktuje po części normy moralne. Prowadzi ono również do jakiegoś rodzaju realizmu moralnego, od którego filozofka tak bardzo odżegnywała się i czyni to nadal w aktualnych pismach. „Cel sam w sobie” oświetla w ICK2 aspekt odniesienia (zbioru zwierząt), jakim jest ich autoteliczna wartościowość, zaś wartość wewnętrzna – źródłowość natury (biologicznie pojmowanej) w porządku wartości.

Z powyższego wynika, że pierwotny zakres (ekstensja) pojęcia „celu samego w sobie” był zbyt wąski, jak na potrzeby rozszerzenia zasady bezwzględnej celowości i wynikających stąd praw również na zwierzęta. Dzięki naturalizacji pojęcie zmieniło swój sens, a zmiana na poziomie sensu pociągnęła za sobą modyfikację (rozszerzenie) odniesienia.

Konsekwencje swoistego buntu semantycznego wobec transcendentalnego języka, logiki i etyki Kantowskiej w kontekście pojęcia „celu samego w sobie” są takie, że dokonana przez Korsgaard rewizja imperatywu kategorycznego jest de facto rewizją „materialną”, tj. obejmującą materialne podstawy stanowienia prawa moralnego dla istot rozumnych, którymi idealiter są podmioty „królestwa celów” resp. obywatele „państwa celów”. Ich naturalizacja przez relację z innymi żywymi jestestwami skutkowałaby przybraniem przez owo „królestwo” znamion postmetafizycznych. Jego mieszkańcy dążyliby wszak nie tyle do moralnego dobra, ile do dobra przełożonego na walory naturalistyczne, takie, jak szczęście i dobrostan. Warto w tym miejscu przypomnieć, że Kant uznał za każdorazowo „mój” obowiązek dążenie do *cudzego szczęścia* (Kant 2005, 251), ponieważ do własnego szczęścia każdy dąży z natury rzeczy, a więc nie ma potrzeby nakazywać takiej dążności w formie obowiązku moralnego względem samego siebie. Najważniejszym obowiązkiem wobec samego siebie jest zaś osiągnięcie doskonałości moralnej. Ale właśnie ten stan rzeczy, tj. moralne samodoskonalenie poprzez dążność do zapewnienia warunków umożliwiających realizację naturalnych dóbr, dobrostanu i szczęścia zwierzętom, miały szansę na realizację w nowym, postmetafizycznym „królestwie” czy też „państwie” celów.

Kanta często uważa się za prekursora współczesnego pojęcia ludzkiej godności.

Godność ta wyrastająca z własnej przyczynowości – wolności transcendentalnej polega na tym, że człowiek w przeciwieństwie do innych istot kieruje się prawami rozumu, tj. *według przedstawienia praw*, a nie według praw. Rozumna natura stanowiąca kryterium dla orzekania o istocie predykatu, jakim jest bycie celem samym w sobie została zastąpiona u Korsgaard – skrótowo rzecz ujmując – naturalnym dobrem. Wolność u amerykańskiej filozofki schodzi zatem na drugi plan (wciąż bowiem to człowiekowi przysługuje status prawodawcy w państwie celów), a godność ludzka zostaje zrównana z godnością wszystkich jestestw żywych. Brak zatem prymatu rozumu nad instynktem.

3.3 Ekstensjonizm etyczny Korsgaard

Ekstensjonizm etyczny jest strategią etyczną, która polega w ogólności na tym, że pewne istniejące już koncepcje filozoficzne, które postulują ochronę określonych grup podmiotów rozszerza się podmiotowo. Rozszerzenie takie oznacza zmianę podmiotowego zakresu ich zastosowania z węższego na szerszy, a sam ekstensjonizm, jak pisze Dariusz Gzyra, „jest wyrazem centrystycznej strategii określania poszerzania granic wspólnoty moralnej” (Gzyra 2017, 135).

Program ten ma charakter centrystyczny, gdyż wychodzi się tu od jednego rodzaju istot posiadających już status moralny; w większości przypadków punktem wyjścia są istoty ludzkie. Ponadto, Gzyra wyróżnia tzw. schemat ekstensjonizmu, na który składają się: „wartościowy wzorzec”, „narzędzie włączania (analogia)” oraz „włączani (wartościowani na podstawie analogii z wzorcem)” (Gzyra 2017, 136). Ekstensjonizm wykorzystuje się w etyce środowiskowej i w jej ramach przeciwstawia się go progresywizmowi („nowej etyce opartej na holistycznym paradygmacie”) (Ganowicz-Bączyk 2005, 50). Elementem wspólnym dla wszystkich ekstensjonistów, niezależnie od tego, o jakie teorie i jakie podmioty (Engel 2008, 396) chodzi, jest przekonanie, że zawężenie zakresu podmiotowego tych teorii nastąpiło w sposób arbitralny.

W wersji Korsgaard strategia ta polega na zmianie sensu, a tym samym, rozszerzeniu odniesienia kantowskich terminów etycznych. Jest to właśnie konsekwencja opisanej wyżej naturalizacji dobra. Wychodząc od normatywnego podmiotu, filozofka konstatuje, że człowiek jako jedyny ma wprawdzie możliwość posługiwania się konstrukcją praw, ale nie implikuje to wcale tego, że jest on jedynym posiadaczem i podmiotem owych praw. Na gruncie moralnym podstawowym prawem jest prawo do realizacji własnego funkcjonowania jako żywego organizmu; na gruncie jurydycznym zaś jest to prawo do używania zasobów naturalnych. Naturalizacja jest wobec tego procesem dwojakim, zachodzącym na dwóch odrębnych poziomach: 1) na poziomie podstaw prawa moralnego i 2) na poziomie podstaw prawa w rozumieniu jurydycznym. Korsgaard dokonała przejścia od wolności woli i autonomii jako podstawowych wyznaczników statusu moralnego jednostki do interesu jako pojęcia fundującego ten status na gruncie – by tak

rzec – postmetafizycznej, naturalistycznie nachylonej etyki i filozofii prawa.

Z kolei analogią w ekstensjonistycznym schemacie jest żywienie tego samego interesu. Analogia ta powstaje ostatecznie dzięki zastosowaniu wykładni rozszerzającej do Kantowskiego pojęcia „celu samego w sobie”.

3.4 Korsgaard a ruch wyzwolenia zwierząt

Na koniec warto krótko zarysować w jakie bardziej generalne tendencje w ruchu wyzwolenia zwierząt wpisuje się projekt Korsgaard. Ruch ten charakteryzuje opozycja pomiędzy „nurtem poszanowania interesów zwierząt (animal welfare) a zwolennikami koncepcji ochrony praw zwierząt (animal rights)” (Pietrzykowski 2007, 79). Głównymi zaś przedstawicielami tych antagonistycznych nurtów są odpowiednio – Singer oraz Regan, którzy nawiasem mówiąc również stosują strategię ekstensjonizmu etycznego. Pierwszy z nich w stosunku do utilitaryzmu (w nawiązaniu do Benthamowskiej idei uprawnień moralnych jako „piramidalnej bzdury”), drugi do dyskursu praw człowieka. Utilitarystyczny ekstensjonizm wydaje się mniej radykalny z uwagi na odwołanie się do pojęcia dobrostanu, które nie zakłada konieczności przejścia całej ludzkości na wegetarianizm, bowiem dopuszcza zabijanie zwierząt, o ile warunki ich hodowli i uboju charakteryzowałyby się dostateczną humanitarnością. Zaznaczenia wymaga także fakt, że to właśnie na gruncie utilitaryzmu – u Benthama – narodziła się refleksja dotycząca sytuacji moralnej zwierząt.

Oczywiście pomysł Korsgaard zdecydowanie przynależy do obozu związanego z Reganem z uwagi na posługiwanie się pojęciem praw czy też w ogóle stosowania argumentacji, którą z grubsza można określić mianem deontologicznej i prawnonaturalnej opartej – jak powiada z kolei Dorota Probuca – „na idei podstawowych uprawnień naturalnych przynależnych zwierzętom z samej ich natury (*natural basic rights*) np. prawa do życia, do niebycia niczyją własnością, do niebycia traktowanym jak rzecz, czyli w instrumentalny sposób” (Probuca 2015, 70). Co więcej, Regan również uznawał, że „Kantowskie stanowisko dotyczące podmiotowości jest zbyt silne, przynajmniej wtedy, jeżeli wymagane jest dla pełnej moralnej istotności” (Engel 2008, 397) i dokonywał jego *reductio ad absurdum* poprzez odwołanie się do tzw. argumentu z przypadków marginalnych. To doprowadziło go do przyznania podmiotowości moralnej jedynie niektórym zwierzętom, czyli tzw. podmiotom życia (*subject-of-a-life*).

Niemniej strategia Korsgaard odbiega od tej, którą obrał Regan, bowiem odwołuje się do pojęcia interesu, który staje się podstawą dla powstania praw. Podobną strategię w wyjaśnianiu natury praw (uprawnień) przyjmują niektórzy filozofowie prawa w opozycji do tych, którzy opierają się na koncepcji woli. Zgodnie z twierdzeniem Josepha Raza zawartym w eseju *On the Nature of Rights*, poszukującego uniwersalnej definicji praw²¹, prawo istnieje, gdy osoba ma zdolność do posiadania praw i „aspekt jej dobrobytu

21 W założeniu Razowska definicja jest formalna i nie implikuje żadnych konsekwencji normatywnych, przez co ma stanowić narzędzie w rozmaitych koncepcjach, które „rzuca światło

(interesu) jest wystarczającym powodem, by nałożyć na kogoś obowiązek” (Raz 1984, 195). Sam Raz nie przyznaje zdolności do posiadania praw zwierzętom twierdząc, że ich dobrobyt (*well-being*) nie jest aksjologicznym priorytetem (nie jest najwyższej wartości (*ultimate value*), ale strategia Korsgaard zdaje się odwoływać do podobnych założeń. Dąży ona najpierw do nadania zwierzętom podmiotowości moralnej, by następnie uznać funkcje wegetatywne danego organizmu (ich „interes”) jako istotny dla nich samych po to, by uczynić je „źródłowo” wiążącymi w świetle nowo ustanowionych obowiązków, które będą ciążyły na istotach ludzkich. Mimo, że pojęcie interesu w filozofii moralnej kojarzone jest raczej z tradycją utilitarystyczną, tutaj ma ono dodatkowo konotację aksjologiczną powiązaną z dobrem naturalnym, którą Korsgaard zdołała wykazać.

Z racji tego, że pomiędzy statusem moralnym i statusem prawnym zwierząt występuje korelacja, przedstawiciele *Animal Rights Movement* – w przeciwieństwie do swych adwersarzy z obozu *Animal Welfare*²² – negują status zwierząt jako przedmiotów prawa własności. Postulują zarówno zmiany w kulturze, jak i w prawie stanowionym, bowiem postrzeganie zwierząt jako własności uniemożliwia ich skuteczną ochronę. Z podobnego założenia wychodzi Korsgaard uznając, że zwierzęta nie powinny podpadać pod kategorię normatywną, jaką jest własność.

Literatura

- Arystoteles. 1988. *O duszy*, przeł. P. Siwek. Warszawa: Państwowe Wydawnictwo Naukowe.
- Bentham J. 1958. *Wprowadzenie do zasad moralności i prawodawstwa*, przeł. B. Nawroczyński. Warszawa: Wydawnictwo Naukowe PWN.
- Besler G. 2010. *Gottloba Fregego koncepcja analizy filozoficznej*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Brożek B. 2013. „Pojęcie normatywności,” w: A. Brożek, B. Brożek, & J. Stelmach (red.), *Fenomen normatywności*. Kraków: Copernicus Center Press (19-44).
- Dummett M. 1973. *Frege: Philosophy of Language*. New York – Evanston – San Francisco – London: Harper & Row Publishers.
- Dworkin R. 1998. *Biorąc prawa poważnie*. Warszawa: Wydawnictwo Naukowe PWN.

na tradycję dyskursu politycznego lub moralnego w różnych teoriach oferujących rozbieżne poglądy na to, jakie prawa są i dlaczego” (Raz 1984, 195). Teoretycznie za pomocą tej definicji można zatem analizować dowolny system filozoficzny, jakkolwiek charakterystyczne jest w niej odwołanie do „interesu”.

22 Niemniej J. Bentham uznawany za ojca tego nurtu uznawał, że niesłusznie zwierzęta są zakwalifikowane do klasy rzeczy – inne zwierzęta, zdegradowane do klasy rzeczy na skutek zlekceważenia interesów przez nieczułych prawników starej szkoły czy też: Oby nadszedł dzień, w których reszta żywych stworzeń otrzyma prawa, których mogła pozbawić ich tylko ręka tyrana (Bentham 1958, 418-419).

- Engel M. 2008. „Ethical Extensionism,” w: J. Baird Callicott & R. Frodeman (Eds.), *Encyclopedia of Environmental Ethics and Philosophy*, Vol. 1. Detroit: Macmillan Reference USA.
- Foot P. 2001. *Natural Goodness*. Oxford: Clarendon Press.
- Frege G. 1979. “Comments on Sense and Meaning”, w: idem, *Posthumous Writings*. Oxford: Basil Blackwell.
- Frege G. 2014. *Pisma semantyczne*, przeł. B. Wolniewicz, Warszawa: Wydawnictwo Naukowe PWN.
- Ganowicz-Bącznyk A. 2005. „Narodziny i rozwój etyki środowiskowej”. *Studia Ecologiae et Bioethicae* 13(5):39-63.
- Geach P. T. 1956. “Good and Evil”. *Analysis* 17:33-42.
- Gzyra D. 2017. „Problemy ekstensjonizmu etycznego”, w: B. A. Nowak & K. Maciąg (red.), *Filozoficzne rozważania o naturze człowieka*. Lublin: Wydawnictwo Naukowe TYGIEL.
- Hare J. E. 1996. *The Moral Gap: Kantian Ethics, Human Limits, and God’s Assistance*. Oxford: Clarendon Press.
- Kant I. 2004. *Krytyka władzy sądenia*, przeł. J. Gałęcki. Warszawa: Wydawnictwo Naukowe PWN.
- Kant I. 2005. *Metafizyka moralności*, przeł. E. Nowak. Warszawa: Wydawnictwo Naukowe PWN.
- Kant I. 2007. *Religia w obrębie samego rozumu*, przeł. A. Bobko. Kraków: Wydawnictwo Homini.
- Kant I. 2017. *Uzasadnienie metafizyki moralności*, przeł. M. Wartenberg. Kęty: Antyk Marek Derewiecki.
- Korsgaard Ch. M. 1983. “Two Distinctions in Goodness”. *Philosophical Review* 92 (2):169-195.
- Korsgaard Ch. 1985. “Kant’s Formula of Universal Law”. *Pacific Philosophical Quarterly* 66(1-2):24-47.
- Korsgaard Ch. M. 1996a. *Creating the Kingdom of Ends*. Cambridge, Mass.: Cambridge University Press.
- Korsgaard Ch. M. 1996b. *The Sources of Normativity*. Cambridge, Mass.: Cambridge University Press.
- Korsgaard Ch. M. 2004. “Fellow Creatures: Kantian Ethics and Our Duties to Animals”. *Tanner Lectures on Human Values* 25.
- Korsgaard Ch. M. 2009. *Self-Constitution: Agency, Identity, and Integrity*. New York: Oxford University Press.
- Korsgaard Ch. M. 2011. “Valuing Our Humanity”. <http://www.people.fas.harvard.edu/~korsgaard/CMK.Valuing.Our.Humanity.pdf> (dostęp z dn. 20.08.2018).

- Korsgaard Ch. M. 2011. "Interacting with Animals: A Kantian Account", w: T. Beauchamp & R. G. Frey (Eds.), *The Oxford Handbook on Ethics and Animals (Oxford Handbooks Online)*. New York: Oxford University Press.
- Korsgaard Ch. M. 2012. "A Kantian Case for Animal Rights", w: M. Michel, D. Kuhne, & J. Hanni (Eds.), *Animal Law – Tier und Recht: Developments and Perspectives in the 21st Century. Entwicklungen und Perspektiven im 21. Jahrhundert*. Dike Verlag: Zürich – St. Gallen.
- Korsgaard Ch. M. 2013. "Kantian Ethics, Animals, and the Law". *Oxford Journal of Legal Studies* 1-20.
- Korsgaard Ch. M. 2014. *Christine Korsgaard – Ethics & Morality – Extended*. Rozm. przep. N. Swierstra, <https://www.youtube.com/watch?v=jzBLPDt-BI0&t=573s> (dostęp z dn. 25.08.2018).
- Leśniak K. 1989. *Arystoteles*. Warszawa: Wiedza Powszechna.
- Lyons J. 1984. *Semantyka*, t. 1, przeł. A. Weinsberg. Warszawa: Państwowe Wydawnictwo Naukowe.
- Łuków P. 2006. „Kantowskie obowiązki względem przyrody – człowiek a pozostałe zwierzęta”. *Diametros* 9:51-73.
- Maciaszek J. 2008. *Znaczenie, prawda, przekonania: problematyka znaczenia w filozofii języka*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Maciejczak M. 2016. *Filozofia jako krytyka języka. Od Kartezjusza do Dummetta*. Warszawa: Oficyna Wydawnicza Politechniki Warszawskiej.
- Makowski P. 2011. *Po metaetyce. Dobro i powinność w etykach naturalistycznych*. Kraków: Universitas.
- McGinn C. 2015. *The Philosophy of Language: the Classics Explained*. Cambridge, Massachusetts: The MIT Press.
- Moore G. E. 1922. *The Conception of Intrinsic Value*. <http://www.ditext.com/moore/intrinsic.html> (dostęp z dn. 20.08.2018).
- Moore G. E. 1919. *Zasady etyki*, przeł. Cz. Znamierowski. Warszawa: M. Actra.
- Moore G. E. 1980. *Etyka*, przeł. Z. Szawarski. Warszawa: Państwowe Wydawnictwo Naukowe.
- Nowak-Juchacz E. 2002. *Autonomia jako zasada etyczności. Kant, Fichte, Hegel*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.
- Pietrzykowski T. 2015. „Kant, Korsgaard i podmiotowość moralna zwierząt”. *Archiwum Filozofii Społecznej i Politycznej* 2:106-119.
- Pietrzykowski T. 2017. *Naturalizm i granice nauk prawnych. Esej z metodologii prawoznawstwa*. Katowice: Wolters Kluwer.
- Pietrzykowski T. 2007. *Spór o prawa zwierząt*. Katowice: Wydawnictwo Sonia Draga.
- Piłat R. 2013. *Aporie samowiedzy*. Warszawa: Wydawnictwo Instytutu Filozofii i Socjologii PAN.
- Probučka D. 2015. *Prawa zwierząt*. Kraków: Universitas.

- Raz J. 1984. "On the Nature of Rights" *Mind. New Series* 73 (370):194-214.
- Rorty R. 2009. *Filozofia jako polityka kulturalna*, przeł. B. Baran. Warszawa 2009: Czytelnik.
- Saja K. 2015. *Etyka normatywna. Między konsekwencjonalizmem a deontologią*. Kraków: Universitas.
- Schneewind J. B. 1998. "Korsgaard and Unconditional in Morality". *Ethics* 109 (1):36-48.
- Szymaniec P. 2009. „Problematyka własności w filozofii Immanuela Kanta”. <http://www.bibliotekacyfrowa.pl/Content/34388/0003.pdf> (dostęp z dn. 22.08.2018).
- Woleński J. 1972. „Logiczne problemy wykładni prawa”. *Zeszyty Naukowe Uniwersytetu Jagiellońskiego*. Kraków: UJ.

Małgorzata Adamska
(Poznań)

*The Categorical Imperative in Defence of Animal Rights.
Christine Korsgaard's Ethical Extensionism*

Abstract: In this paper, I present Christine Korsgaard's style of building a conception of animal rights protection based on a concept of Kantian provenance, namely "end in itself". Considering that an end in itself in Immanuel Kant's philosophy includes only human beings, the American philosopher needed to modify (extend) the meaning of the abovementioned concept. This study aims at showing this change of the meaning in categories derived from Fregean semantics (sense and reference). Moreover, I will attempt to prove that by broadening the extension of "end in itself" Korsgaard uses a strategy called ethical extensionism and situates the revised (naturalized) categorical imperative *ipso facto* within environmental ethics.

I start with elucidating Korsgaard's views on the meaning of "end in itself" before she took interest in animal rights (set out mainly in *Kant's Formula of Humanity* and *The Sources of Normativity*) and then I synthetically describe her animal rights philosophy with a special regard to the category of "natural good". The second part of the article explains the methodology used in the paper, while the last part is intended to be an interpretation of Korsgaard's thought.

Keywords: animal rights; ethical extensionism; Christine Korsgaard; Immanuel Kant; an end in itself; sense and reference.

Ethics in Progress (ISSN 2084-9257). Vol. 9 (2018). No. 1, Art. #8, pp. 140- 173.

Creative Commons BY-SA 3.0

Doi: 10.14746/eip.2018.1.8

Changes on the Qualities of University Counsellors in China

Lan Wen

(Institute of Psychological and Education, South China Business College,
Guangdong University of Foreign Studies, windlives@163.com)

Peijue Zheng

(School of Interpreting and Translation Studies, Guangdong University of Foreign Studies,
491367347@qq.com)

Shaogang Yang

(Institute of Psychological and Education, South China Business College,
Guangdong University of Foreign Studies, ysgrime@163.com)

1. Literature Review

Since 19th century, student affairs in American universities were separated from academic affairs to provide specific counselling for students' on issues such as psychological health, academic performance, and professional development, and thereby the job of "student affairs officer" emerged (YanJun 2007). Although many scholars have concluded that student affairs officers play an important role on both formal and informal teaching occasions (Moore & Marsh 2007). Gregory, a senior student affairs officer (SSAO) pointed out that, in fact, this job is not usually a career option for young people, and there aren't any relevant undergraduate courses available now to prepare students for this job (Blimling 2002). The earliest systematic study on student affairs in universities traces back to the theory proposed by Chickering in 1969. He put it forward that student affairs officers should be assessed based on 7 vectors related to students' development: Developing competence, Managing emotions, Developing autonomy, Establishing identity, Freeing interpersonal relationships, and Developing integrity (Moore & Upcraft 1990). McClelland proposed 8 qualities for student affairs administrative staff, namely achievement orientation, analytical thinking, developing others, flexibility, impact and influence, information seeking, self-Confidence, and team leadership (McClelland 1998). Spreitzer McCall and Mahoney listed 14 dimensions of work potentials: Sensitive to Cultural Differences, Business Knowledge, Courage, Brings Out the Best in People, Integrity, Insightful, Committed, Takes Risks, Seeks Feedback, Uses Feedback, Is Culturally Adventurous, Seeks Learning Opportunities, Open to Criticism, and Flexibility (Spreitzer McCall & Mahoney, 1997).

There are also specific studies on student affairs officers. Nicole believes student affairs experts must learn to be innovative, observant, make full use of resources at

disposal and bear in mind the whole picture (Eramo 2006). Then Kidder shed light on the four criteria for student affairs staff recruitment: assessment, accountability, service learning, the value of cocurricular education efforts on campus (Kidder 2010). Building upon "Student Talent View" by American Council on Education (1937), Kathleen and Jeanne summarized 8 life aspects of students that student affairs officers are supposed to focus on: intellectual capacity and achievement, emotional make up, physical condition, social relationships, vocational aptitudes and skills, moral and religious values, economic resources, and aesthetic appreciations (Kerr & Hart-Steffes 2012). American universities generally divide student counsellors into 5 categories based on their professional qualities: psychological counsellors, vocational counsellors, learning counsellors, life counsellors and accommodation counsellors (Ping 2006). Leilani Kupo concluded that for scholar-practitioners, multitasking has been a way of existence. It is a frame of mind. It includes navigating the realms of practitioner and scholar as well as having the ability to understand how to translate relevant research and data and incorporate it into daily practice (Leilani Kupo 2014).

It has been about 50 years since Chinese "University Counsellor System" was initiated by President Jiang Nanxiang of Tsinghua University in 1953, and the student affairs administrative staff were originally called "political counsellors". This system is intended to establish an ideological and political leading force among teachers and students to ensure the leadership of The Communist Party of China (CPC) over universities. The research on university counsellors began with research on their "qualities" or "competence". The early "ternary theory" divides counsellors' qualities into three categories: ideological and political qualities, literacy, professional competence (Wang Lifen 2001). At the same time, some scholars divide them into 4 aspects: political qualities, ideological and moral qualities, expertise and competency (Lin Sifeng 2004); or "leadership, coordination ability, literacy and the ability to summarize" (Sun Jiayan 2001). According to Zhang Shuming, counsellor qualities fall into 2 categories: basic qualities and basic abilities, respectively including political qualities, moral qualities, expertise, psychological qualities; and coordination ability, communication and cooperation ability, ability to serve students, ability to psychologically counsel students, scientific research and innovation ability, ability of self-control and addressing complex situations and linguistic proficiency (Zhang Shuming 2010). Gu Qian divides counsellor' professional qualities into 16 dimensions: expression ability, communication ability, adaptability, coordination ability, caring for students, innovation ability, observational acuteness, professionalism, ability to promote students' development, self-retrospection, personal charisma, moral standard, psychological counselling ability, the ability to understand and respect students, being principled, and the ability to engage in students' activities (Gu Qian 2004). Shao Guoping concluded 4 dimensions for occupational competence through modeling: education, management, service and professionalism (Shao Guoping 2013).

Furthermore, some scholars focus on the study of counsellors' competence. Yang

has put forward the most important 5 qualities in terms of counsellors' competence as follows: understanding and respecting students, expression ability, being principled, caring for students, professionalism (Yang Jiping 2004). There is another 7-dimension model: professional competence, student-cultivating ability, moral qualities, professionalism, awareness for the whole picture, influence strategy, caring for students (Huo Xiaodan 2013). Moreover, there is also a two-level model with 8 factors at the first level and 12 factors at the second level. The 8 factors are social awareness, expertise, political and moral qualities, professionalism, personality, adaptability, professional skills, and values (Zou Tao 2017).

Since 2010, researches on qualities of university counsellors have begun to emerge. It is proposed that counsellors' professional qualities composition includes expertise, professional skills, professional ethics and professional affection (Ma Jing 2010). Zhang Jingpeng theorizes that qualities composition includes political qualities, ideological qualities, professional ethics, literacy, psychological qualities, competence (Zhang Jingpeng 2010). Based on the above review, it is realized that most previous studies only theoretically analyzed the qualities composition of counsellors with highly subjective research methods. There are few empirical studies employing research methods such as survey and mathematical analysis, therefore lacking reliability, validity, and item analysis.

2. Methods and Findings

Chinese university faculty members and students are our subjects. The pre-test is based on Likert 7 point scale (1 = disagree very much, 7 = agree very much). SPSS22.0 and AMOS21.0 are applied to data analysis and processing. The correlation coefficient between each question in the questionnaire and the total score will be calculated after the test, and all the questions with correlation coefficient $r < 0.4$ will be eliminated until only 51 questions are left. Exploratory Factor Analysis (EFA) shows that 9 factors were extracted, the cumulative variance contribution rate reaches 82.83% and the model tends to be stable. Then the final version of questionnaire titled „university counsellors' qualities questionnaire" is completed with 54 items in total, among which 3 are lie detection questions. Of 800 pieces circulated, 696 pieces are collected, and 537 pieces are found to be effective after the screening through lie detection questions .

The results show that Cronbach's alpha coefficient and the split-half reliability coefficient are all above 0.9, indicating that the actual value is in good accordance with the true value. The internal alpha coefficient and the split-half reliability coefficient of the 5 factors are all above 0.8. Correlation analysis between the 5 sub-scales and the total scale of the test shows that the correlation coefficients are all above 0.8. The correlation coefficients between different sub-scales range from 0.7 to 0.9, indicating that there's relatively high correlation between these sub-scales, various factors are closely connected, and that the structure validity of the test is high.

Factor	Number of item	Cronbach's α coefficient	Split-half reliability
Moral quality	15	0.95	0.93
Business execution	14	0.96	0.95
Political consciousness	10	0.93	0.90
Guiding strategy	7	0.87	0.85
Interpersonal coordination	5	0.85	0.86
Total test	51	0.98	0.95

Table 1: Test reliability

A 5-factor model is formulated based on the research results, including: moral quality; professional competence; political awareness; instructive ability; interpersonal skills. KMO is 0.979, which is a very suitable figure for factor analysis. Bartlett's test of sphericity is 23516.560, $df = 1275$, $P < 0.001$, the cumulative variance has reached 65.72%, all 51 items are included in the EFA model.

We have conducted confirmatory factor analysis has been performed on the data, showing all fit indexes of this model are sound. $X^2/df < 5$, indicating that the model is acceptable. The closer GFI, AGFI, NFI, CFI, and IFI are to 1, the higher the fitness is, and it is more desirable if RMSEA is close to 0.

	X^2	f	X^2/df	GFI	AGFI	NFI	CFI	IFI	RMSEA
Model Fit	4411.32	214	3.634	0.74	0.71	0.82	0.86	0.86	0.07

Table 2: Model Fit.

Moral quality refers to that counsellors' daily behaviors should follow well-accepted moral standards, being a role model for students so as to help them develop correct values, outlooks on life and the world as a whole. Professional competence is indispensable skills that counsellors have to acquire to handle daily work involving many aspects such as leadership, instruction, observation, management, coordination, communication and practice. Political awareness puts much weight on political and legal knowledge. Counsellors are supposed to be familiar with Party-related paperwork, school regulations to ensure daily work is done in compliance with laws and regulations. Instructive ability mainly has to do with counsellors' ability to guide students' long-term development and to foster the leadership of students. As we step into new eras, the counsellors are expected to master more extracurricular, social and internet student-instructing strategies so that education can keep pace with the times. Interpersonal skills involve the skills of communication and emotional interaction with students. Even if the

counsellors are under immense work pressure, they should be able to adjust themselves and still be generous and forgiving with students.

This 5-factor model reflects the evaluation of faculty members and students in universities on qualities of counsellors and transformation in their expectation for the role counsellors play as time goes by. On the one hand, counsellors are expected to be competent in problem-solving and management; on the other hand, they are also required to serve as an exemplar with their own uprightness and honorable deeds to help students form positive as well as correct outlooks on value, life and the world. In addition, counsellors ought to have a high political awareness, which is characterized by being well-informed of Party paperwork, laws and regulations to ensure work is carried out in their compliance. Counsellors' work runs a whole gamut of management from Party paperwork, student associations, scholarship, student loans and subsidies, suspension or dropping out of school, student registration affairs, graduation and employment, student behaviour monitoring, dormitory management and so on.

3. The Changes

The qualities of counsellors are constantly changing without deviating from intrinsic nature of student affairs. These changes are based on the realistic situations and the current and future developmental needs of student administration.

3.1 Objectivity of times

The qualities of counsellor is characterized by objectivity, and its existence, development and changes are affected by the objective environment. The new situations and demands arising from social progress drive the reform in student administration models and ideas, and in professional qualities composition of counsellors so that they are in line with the ever-changing educational environment. The comprehensive application of new technologies have rapid and irresistible impact on the counsellors' professional qualities and working environment. It even can cause qualitative changes in their mindset. In this case, qualities such as communication skills and abilities to operate new media are included in our questionnaire. A great deal proportion of student affairs have been handled online so that they can be carried out in a low-cost, low-entry-threshold, high-efficiency manner. This is a testimony to the phenomenal development of student administration in this new era.

3.2 Initiative of inner

Despite the objectivity of counsellors' professional qualities, they are also in a way influenced by counsellors' subjectivity, which are internal factors leading to the changes in their qualities. The subjectivity is manifested in counsellors' awareness for their own professional qualities, namely their own ability to evaluate and reflect on their qualities, in the-

ir ability to perceive and respond to changes in physical working environment, namely the ability to judge the compatibility between their own professional qualities and the working environment, the ability to look into the future, to grasp the rules of student administration in the future, the ability to explore and forge an optimal professional qualities composition. Self-awareness, perception and response ability and the ability to live by the natural rules of student administration constitute the subjectivity of counsellors themselves.

3.3 Predictability of practice

The changes in counsellor's professional qualities are positive ones in accordance with the reality and objective rules, which therefore can be recognized, predicted and even created. Only by conducting abundant investigation, keeping updated on the situation, acquiring field experience and carrying out lone-term observation, analysis, comparison can the counsellors form systematic and correct understanding, predict the developmental trend of student administration, make foresighted decisions to guide counsellors' future professional practices and improve the feasibility, scientificity and practicality of their practices. In other words, the predictability of the future comes from the reliable understanding of the reality, the profound mastery of the natural rules of development, and the logical analysis for the future trend.

In China, people lay emphasis on counsellors' ideological and political qualities, instructing skills, and transferable skills to organize, serve students and promote their long-term development and psychological health. The professionalization of the counsellors fundamentally requires this career to be sustainable, stable and widely-received by the society (Zhu Ping 2007). Being well aware of the trend of this career is conducive to the formation of a sound relationship between counsellors and students, measurability for morality-oriented education, and improvement in counsellors' overall qualities.

References

- Blimling G. S. 2002. "Reflections on Career Development Among Student Affairs Leaders." *New Directions for Student Services* 98:27-36.
- Eramo N. P. 2006. „Rethinking Student Affairs Practice." *Journal of Colleges Student Development* 123-5.
- Gu Qian. 2004. *Compilation and Preliminary Application of College Counsellors' Competency Questionnaire*. Shanxi: Shanxi University Editions.
- Huo Xiaodan. 2013. *Quality Standards and Development of College Counsellors – Analysis Based on Competency Model*. Beijing: Peking University Press.

-
- Kerr G. & Hart-Steffes J. S. 2012. „Sustainability, Student Affairs, and Students.” *New Directions for Student Services* 137:7-17.
- Kidder R. 2010. „Administrators Engaging in the Research Process.” *ACPA Developments* 8(1).
- Lan Haitao & Wei Xing. 2017. „The Quality and Ability Structure of College Counsellors and Their Training Methods.” *China Higher Education* 5:42-44.
- Leilani Kupo V. 2014. „Becoming a Scholar-Practitioner in Student Affairs.” *New Directions for Student Services* 147.
- Lin Sifeng. 2004. „Reflections on Strengthening Quality Training of Political Counsellor Team.” *Journal of Jimei University (Philosophy and Social Sciences)* 2.
- McClelland D. C. 1998. „Identifying Competencies with Behavioral-event Interviews.” *Psychology Science* 9(5):331-9.
- Moore E. L. & Marsh R. S. 2007. “College Teaching for Student Affairs Professionals.” *New Directions for Student Services* 117:3-11.
- Moore L. V. M. & Upcraft L. 1990. „Theory in Student Affairs: Evolving Perspectives.” *New Directions for Student Services* 51:3-23.
- Shao Guoping, Miao Delu, & Yang Lin. 2013. „The Structure and Measurement of Occupational Ability in College Counsellors.” *Psychological Research* 6(2):85-89.
- Spreitzer G. M., McCall M. W. Jr., & Mahoney J. D. 1997. „Early Identification of International Executive Potential.” *Journal of Applied Psychology* 82(1):6-29.
- Sun Jiayan. 2001. „Discussion on the Qualities That College Counsellors Should Have.” *Journal of Social Science of Jiamusi University* 4.
- Wang Lifan. 2001. „On the Quality of University Instructors.” *Journal of Taiyuan University of Technology (Social Sciences Edition)* 3:114-6.
- Yan Ping. 2006. „Comparison of The Work of College Instructors in China and The United States – A Brief Discussion on the Development Direction of China’s College Counsellor System.” *Journal of Hunan Mass Media Vocational Technical College* 7:115-6.
- Yang Jiping & Gu Qian. 2014. „A Preliminary Study on the Competency of College Counsellors.” *Journal of Shanxi University (Philosophy and Social Sciences)* 12.
- Zhang Jingpeng. 2010. *Research on Role Orientation and Professional Quality of College Counsellors*. Shenyang: Shenyang Aerospace University.
- Zhang Shuming. 2010. *College Counsellors’ Team Building*. Jinan: Mountain Tai Press.
- Zhang Yanjun. 2007. „Comparing the Recruitment Conditions of College Counsellors between China and the United States.” *Journal of Nantong University (Education and Science)* 23(3):40-43.
- Zhu Ping. 2007. „Interpretation of College Instructor’s Professionalization and Specialization.” *Journal of Anhui Normal University (Humanities and Social Sciences)* 35(2):218-223.

Zou Tao, Ai Hong, & Chen Xiang. 2017. „A Study on The Difference Assessment of Instructor Competency Based on The Empirical Survey of Three Groups of College Students.” *Journal of Southwest University of Nationalities (Humanities and Social Sciences)* 1:222-227.

Acknowledgements

The study was conducted as a part of moral education project (code number: 2017JKDY16), 2017; Guangdong high education Grant of Special Creativity (Humanities and social science); Young Creative talents project (code number: 2017WQNCX208), 2016; and Guangdong Philosophy and social science “13th five-year plan” project (code number: GD16XJY31).

Lan Wen, Peijue Zheng, Shaogang Yang
(Guangdong)

Changes on the Qualities of University Counsellors in China

Abstract: Counsellors are important key members in Chinese universities. In this paper, on the one hand, we compared student affairs workers' occupational characteristics, classification and development of western as well as Chinese university. On the other hand, Chinese university faculty members and students are our subjects. Of 800 pieces circulated, 537 pieces are found to be effective. We verified the reliability and validity of data, then conducted correlation analysis, exploratory factor analysis (EFA), and confirmatory factor analysis (CFA). The test has good reliability and validity. Cronbach's alpha coefficient and the split-half reliability coefficient are all above 0.9, indicating that the actual value is in good accordance with the true value. The internal alpha coefficient and the split-half reliability coefficient of each factor are all above 0.8. $X^2/df < 5$. GFI, AGFI, NFI, CFI, IFI, RMSEA and the other indicators are fitting, and this test meets measurement requirements. A 5-factor model is formulated based on the research results, including: moral quality; professional competence; political awareness; instructive ability; interpersonal skills. This 5-factor model reflects the evaluation of faculty members and students in universities on qualities of counsellors and transformation in their expectation for the role counsellors play as time goes by. So we summarize the qualities of counsellors are constantly changing. These changes have been constantly evolving in time, and their traits are such as: Objectivity of the times, Initiative of inner, Predictability of practice. Being well aware of the trend of this career is conducive to the formation of a sound relationship between counsellors and students, measurability for morality-oriented education, and improvement in counsellors' overall qualities.

Keywords: University counsellor; quality; changes; China.

Ethics in Progress (ISSN 2084-9257). Vol. 9 (2018). No. 1, Art. #9, pp. 174-182.

Creative Commons BY-SA 3.0

Doi: 10.14746/eip.2018.1.9

Über die Philosophie von Barbara Skarga – eine Anregung zur Auseinandersetzung mit ihrem Leben und Werk

Małgorzata Bogaczyk-Vormayr

ORCID: <https://orcid.org/0000-0001-7612-8535>
(UAM Poznań, bogaczyk@amu.edu.pl)

Et te Socraticae docuerunt tollere chartae
Quod fas iusque homini est sidere in corde odium
Totaque corde tuo veteris sentina doloris
Iam dudum exhausta est, mens tibi pura viget.
Juliusz Domański, *Jubilanti*¹

Einleitung: Zur Person von Barbara Skarga

Barbara Skarga (1919–2009) war eine polnische Philosophin und Expertin im Bereich der französischen Philosophie der Neuzeit, aber auch eine ausgesprochene Kennerin der Werke von Bergson oder Lévinas, jedoch tief verankert in der Tradition der deutschen Klassik (Kant, Hegel), gemäß der Begriffsgeschichte von Gadamer verstand sie das Philosophieren als eine nicht historische, sondern problemorientierte Auseinandersetzung mit der Gegenwart und deren Verwurzelung in der Geschichte. Sie war auch eine kritische Kommentatorin zu Heidegger, angezogen jedoch von seiner Genialität.

Barbara Skarga war auch eine Gulag-Gefangene, eine Kämpferin im polnischen Untergrund, in Armia Krajowa. 1944 wurde sie in Vilnius (wo sie lebte und schon vor dem Krieg Philosophie studierte) von einem sowjetischen Gericht als westliche Spionin und Verräterin des Kommunismus verurteilt (Kommunismus als richtige Antwort auf den

¹ Das Zitat stammt aus einem im Jahre 1999 von Juliusz Domański (Altphilologe, Professor an der Universität Warschau) verfassten Jubiläumsgedicht – zum 80. Geburtstag von Barbara Skarga (vgl. Domański 2017, 226).

Faschismus verstanden) und für 15 Jahre in sibirische Gefangenschaft geschickt. Dort war sie jahrelang als Krankenschwester in den Gulag-Lazaretten tätig. 1954 (Amnestie nach dem Tod Stalins) hatte man sie aus dem Gulag in eine sibirische Kolchose geschickt. Als sie Ende 1955 entlassen wurde, inskribierte sie zum zweiten Mal in Philosophie, diesmal in Warszawa, und war motiviert, ihren schon berühmten Philosophiekollegen nachzueifern – allen voran Leszek Kołakowski –, deren Überzeugung für das neugeborene kommunistische Land sie jedoch nicht begreifen konnte.

Barbara Skarga war meine Lehrerin an der Polnischen Akademie der Wissenschaften. Mein dortiger Forschungsschwerpunkt umfasste andere Bereiche als die, welche man in ihren Seminaren behandelte, aber wir fanden uns irgendwie in der Mitte des Weges. Es war nicht schwierig, denn auch innerhalb der Sozialethik und politischen Philosophie – die damals ihre und nicht meine Schwerpunkte waren – blieb Skarga eine Metaphysikerin – ja, im richtigen, inhaltvollen, ernsthaften Sinn dieses Wortes. Sie stellte die Urfrage des Philosophen – nach dem Sein.

Barbara Skarga ist außerhalb von Polen so gut wie unbekannt. Deswegen erlaube ich mir, heute über ihre Person und ihre Schriften zu erzählen.² Ein nächster Schritt werden deutsche Übersetzungen sein. Ich erinnere mich jetzt daran, als ich vor 15 Jahren den ersten Vortrag an der Uni Salzburg besuchte: Prof. Edgar Morscher sprach in der Einleitungsvorlesung über die Thematik des beginnenden Semesters und erklärte seine absolut wichtigste Absicht, uns die Wahrheitstheorie von Alfred Tarski beizubringen. Er sagte, ich zitiere wörtlich: „Zum Glück hatte Tarski seine Theorie auf Englisch verfasst. Denn was wäre mit Polnisch? So eine exotische Sprache! Vielleicht hätten wir es bis heute nicht kennen gelernt!“ Was die Exotik betrifft, wie gesagt, muss ich zustimmen. Aber ist diese Aussage ganz richtig? Die Beschreibung der Welt in der Sprache der Logik findet seine Wege durch geniale Menschen, unabhängig davon, in welcher Sprache diese ihre Gedanken aussprechen, weil diese Beschreibung die mathematischen Objekte, Naturgesetze behandelt. Somit muss diese Beschreibung die Bedingung erfüllen, auch in einer formalen Sprache – fehlerlos – erfasst werden zu können. Anders scheint es mir innerhalb der praktischen Philosophie zu sein. Hier muss vor allem im Diskurs selbst ein Raum (oder: ein Rahmen) geschaffen werden für widersprüchliche Aussagen, für Differenzen – wenn auch nur am Beginn des Diskurses. Somit bin ich im Kern des Philosophierens von Barbara Skarga angelangt, weil ich ihr Werk der Tradition der späten, französischen Begegnungsphilosophie zuschreiben würde und ihr Konzept als „Philosophie der Differenz“ bezeichnen möchte.

1. Rückkehr zu metaphysischen Begriffen

An der Wende vom 20. zum 21. Jahrhundert entstanden zahlreiche Essays von

² Der Text wurde als Vortrag verfasst und auf der V. Tagung für Praktische Philosophie an der Universität Salzburg im Oktober 2017 vorgetragen.

Skarga. Sie hörte damit auf, typische philosophische Sachbücher, Monographien zu veröffentlichen. Nach dem Buch *Identität und Differenz. Metaphysische Essays* (1998) erscheinen *Spur und Anwesenheit* (2002), *Metaphysisches Quintett* (2005), *Metaphysisches Quartett* (2005), *Der Mensch ist kein schönes Tier* (2007) sowie *Metaphysisches Terzett* (2009). *Identität und Differenz* wurde in Polen mit wichtigen Literatur- und Friedenspreisen ausgezeichnet, obwohl dieses Buch durchaus einen sehr intellektuellen, philosophischen Duktus hat. *Identität und Differenz* – ein Titel, der einem deutschen Leser sofort an die Schrift von Heidegger aus dem Jahre 1957 erinnert, Skarga jedoch beginnt dieses Buch mit folgenden Worten:

Das den Lesern präsentierte Buch ist aus einer trotzigen Haltung heraus geboren. In der Regel erwartet man von einem Philosophen und einem Forscher, dass er auf die Probleme der Gegenwart reagiert, die ihm nicht gleichgültig sein sollen. Aber was heute innerhalb der Philosophie vorgeht, stößt mich ab und langweilt mich. Fremd ist mir die analytische Philosophie, detailliert, aber oft leider nur eine Haarspalterei. Es scheint mir, dass sie auch für sich selbst unfruchtbar ist. Sie zwingt tatsächlich zur Präzision, aber nicht zur Reflexion. Andere philosophische Strömungen unterliegen der Soziologisierung oder gar Politisierung. Es handelt sich bei diesen um die Kommunikation, um das kommunikative Handeln, um die normativen Grundlagen unterschiedlicher Sozialtheorien u.Ä. Und diese Strömungen wenden sich stark dem Thema von Identität und Differenz zu, immer aber nur in sozial-kulturellen Zusammenhängen. Es handelt sich also um die europäische Identität, um die Identitäten von Nationen, bestimmten Gruppen, Minderheiten usw. Dieser Thematik widmet sich eine Tagung nach der anderen. Aber die Debatten über diese Begriffe, welche ausschließlich im Bereich ihrer kulturellen und gesellschaftlichen Aktualität verbleiben, übersehen den eigentlichen Sinn dieser Begriffe, oder abstrahieren vielleicht sogar bewusst von einer Sinnfrage. Es sind nämlich metaphysische Begriffe, und mit deren metaphysischer Analyse müsste man die Diskussionen beginnen (Skarga 2009, 5)³.

Die Spaltung, in der heute die philosophischen Disziplinen gesehen werden, war nach Skarga nicht lediglich ein Problem eines Akademiker-Alltags oder eines geschlossenen Diskurses, der in den Fachzeitschriften entsteht und dort ohne Auswirkungen bleibt. Nein, es ist ein Problem, aus dem schreckliche gesellschaftliche Folgen resultieren können. Gerade auf diese Erwartung, dass der Philosoph sich in der Öffentlichkeit äußern soll, soll er – meint Skarga – damit antworten, dass er *das Philosophische* in die Debatte einbringt, dieses als adäquat aufzeigt, daraus eine Antwort, einen Vorschlag schöpft.

Daran dachte ich selbst immer, wenn ich mir in den letzten Jahren die langen, heißen Diskussionen meiner Studierenden zur europäischen Flüchtlingspolitik anhörte. Nachdem sie schon einen modernen, anspruchsvollen Text gelesen hatten, in dem sie mit einer Stellungnahme konfrontiert wurden, die sie noch mehr aufheizte, die ihnen Gegenargumente brachten – während sie bis jetzt bloße Meinungen austauschten und Fernsehdiskussionen zitierten –, biete ich ihnen in einem zweiten Schritt einen klassischen Text an. Also wir bearbeiten zu Beginn beispielsweise ein Buch von Harald Welzer und anschließend, zum Beispiel Kants Schrift *Vom ewigen Frieden*. Beide Autoren bringen erst-

3 Alle Zitate in eigener Übersetzung.

mals eine Ordnung in die laufenden Gespräche. Welzer liefert Daten, die Vermutungen und Klischees ausräumen, und dann zeigt sich bei Kant – als brillante Satire! – die Ähnlichkeit der Problematik und die Aufgabe jedes Moralphilosophen, ein metaphysisches Konzept als Antwort, also als Lösung für die konkrete politische Handlung zu erstellen (wie es zum Beispiel Kant mit der Klärung des Begriffes Fremdheit und des Prinzips der Gastfreundschaft abhandelt). Die aktuellsten Fakten und die klassischen Problemstellungen sollen gemeinsam eine Arbeitsgrundlage bilden. Dies habe ich von Barbara Skarga gelernt. Dies haben von ihr drei Generationen polnischer Philosophen und Sozialethiker gelernt. Mit folgenden Worten hatte sie ihr Freund Juliusz Domański porträtiert:

Sie war schon seit langem eine unbestrittene Grande Dame der polnischen Philosophie, als ich sie so vor fast zwanzig Jahren⁴ in Paris dem Ehepaar Ilsetraud und Pierre Hadot vorgestellt habe. Und als solche, mit diesem Titel geschmückt, blieb sie schon in meinem Bewusstsein. Ich habe natürlich an ihr Werk gedacht, aber nicht nur das. Auch dachte ich an ihr Leben, wie es in der Kriegszeit und unmittelbar nachher war. Und noch mehr dachte ich daran, wie ich selbst ihr Leben in den vielen Jahren unserer Freundschaft beobachtete, besonders seit Beginn der Achtziger, als unsere Bekanntschaft intensiver geworden war. Ich musste an ihre Haltung in den schwierigen, gar dramatischen Zeiten denken, angesichts der Gefahren, welche die Dummheit und Gemeinheit für das öffentliche Leben darstellten. Ich dachte über ihre Entscheidungen nach, die sie zur empathischen und scharfsinnigen Stimme des Gewissens machten – und zwar nicht nur für den polnischen Philosophenkreis, sondern für die gesamte polnische Öffentlichkeit, zumindest in ein paar historisch wichtigen Augenblicken (Domański 2017, 238-9).

Demokratie, Solidarität, Gerechtigkeit, Dialog – all dies waren anfangs *Ideen*. Gleichheit und Anerkennung sind doch keine Naturgegebenheit, sondern ein Konzept, eine ethische Idee – man entscheidet sich, diese zu berücksichtigen und anzuwenden. Natürlich, aus dem Faktum der Asymmetrien und Differenzen, die die menschliche Welt beinhaltet, kann man ganz unterschiedliche Regeln und Verhaltensweisen für sich wählen. Deswegen bemühten sich Ethiker und Philosophen in der Nachkriegszeit, die Differenzen, oder: den Modus der Differenzen, zu nivellieren und das Denken über das Gemeinsame, das alle ähnlich Auszeichnende zu betonen, also: die bei allen Menschen gleichen Bedürfnisse nach Schutz und Gerechtigkeit, nach Aus- und Einwanderung usw. Barbara Skarga meint, dass es zu wenig sei, oder stärker: dass es eine Falle sei, in die wir immer wieder tappen werden – in unserem Denken und dann im sozialpolitischen Handeln –, weil wir die metaphysische Kraft der Differenz aus unserem Horizont verbannt haben.

2. Kultur der Immanenz und Vernichtung der Identität

Wenn man unter „Differenzen“ sofort und dann ausschließlich in einen sozialen Kontext subsumiert, wie zum Beispiel das Ethnische, die Herkunft, dann das Kulturelle und Religiöse (Religion selbstverständlich als Teil der Kultur verstanden, nicht im

⁴ Anmerkung der Autorin: Dieser Text stammt aus dem Jahre 2009.

theologischen Sinne), da bemüht man sich – allerdings in bester, zutiefst ethischer Absicht –, die Differenzen im strahlenden Licht des Gemeinsamen zu entblößen. Ich bin der gleichen Überzeugung wie Barbara Skarga, dass dies ein Fehler ist, weil die Anerkennung via Analogien de facto eine Bestimmung des Eigenen ist. Auch wenn man das Eigene immer weiter fasst (wenn man immer moralischer handelt), wird im Endeffekt das Eigene doch als Abgrenzung zum Anderen bestehen bleiben. Skarga sagt:

Die Kultur ist der Sinn – ein Sinn, den man dem Sein verleiht. Die [nach Lévinas – M.B.-V.] sogenannte Kultur der Immanenz wäre also eine durchaus intellektuelle und gleichfalls strikt anthropozentrische. Für sie bilden nämlich das menschliche Bewusstsein und seine Weltsicht den Orientierungspunkt. Der Kontakt eines Ich mit der Welt hat einen erkenntnismäßigen Charakter, den man als grundlegend und ursprünglich erkennt. Der Mensch will die Welt verstehen. Aber das Verstehen benötigt es, das Unbekannte zum Bekannten zurückzuführen, zu dem, was schon *meines* ist (Skarga 2009, 129).

Man könnte diesen Gedanken weiterführen: Wenn die Differenzen der Kulturen jemandem Sorgen bereiten, findet er sein Hauptargument in der Hauptanalogie zwischen den einzelnen Menschen, in der Zugehörigkeit zur selben Spezies – und dies ist eine biologistische Argumentation. Skarga meint, diese sei nicht nur zu kurz gegriffen, sondern könne leicht entgegengesetzten Zielen dienen – denn wenn man den Menschen lediglich in seiner Biologie auffasst, bringt man die Mechanismen des Reduzierens in Gang. Diese Argumentation muss, so hält sie dagegen, metaphysischer Natur sein: *Jede/Jeder andere ist ein für mich transzendentes Individuum. Die Grenze meiner Handlungen bestimmt diese ursprüngliche, ontologisch nicht reduzierbare Differenz.*

Diese Differenz ist für eine anerkennende Haltung vorrangig gegenüber den zahlreichen Gemeinsamkeiten. Ohne Annahme dieser Anerkennung der Differenz wird jede noch so „menschenfreundliche“ Debatte nur kurzfristige Wirkungen haben, weil sie nämlich keine philosophischen Grundlagen hat.

Ich will jedoch hier betonen: Skarga selbst beginnt ihre Vorträge, Texte oft mit einem Beispiel zur Wahrnehmung eines Subjektes, zu seinen Emotionen, in denen das Gemeinsame – so sagt uns doch die Intuition – das Hauptargument darstellt. Die Philosophin macht aber sofort eine Kehrtwendung: vom anfänglichen Phänomen zur Reflexion. Nicht nur in den metaphysischen Essays, sondern auch in ihren Erinnerungen an den Gulag, mit dem Titel *Nach der Befreiung...1944-1956*, gibt es viele solcher Stellen:

Ja, genau hier, im Großlager, vergeht der Hass, der das Herz eines Gefangenen, während er angesichts des Aufsehers sitzt, erfüllte. Denn womit unterscheidest du dich von mir, Herr Kapitän, der du den Dienst eines Lagerleiters erledigst? Ja, du kannst zehnmal besser essen als ich, dies stimmt, du bist nicht jeden Augenblick dem Tod ausgesetzt, du bewegst dich ohne Aufsicht und darfst mich jederzeit anbrüllen. Aber bist du tatsächlich frei? Du bist hierher, in dieses weite Land geschickt worden, wie ich, ohne nach deiner Meinung gefragt worden zu sein. Man hat dich angewiesen, uns, diese *dochodiagi*⁵, zu beaufsichtigen, man

5 Anmerkung der Autorin: „Dochodiaga“ – Bezeichnung für einen Gulag-Gefangenen, der an der Grenze zum Sterben ist; bekannt u.a. als Selbstbezeichnung von Schalamow, der sich als „lebende Leiche“ beschrieben hat; aus dem Russischen von *дойти* - ankommen, bezeichnet es jemanden,

hat dich instruiert, dass es hier die schlimmsten Verbrecher gibt. Du gehst aber zwischen uns herum und siehst, dass es sich um normale Menschen handelt, vielleicht um bessere als die Offiziere in deinem Regiment, aber diese Gedanken wagst du nicht auszusprechen, ha, nicht einmal zu denken! Du tust also, was sie von dir erwarten, bemüht dich eifrig zu sein, weil du genau weißt, wonach der Widerstand riecht, weil durch deine Hände Formulare gehen, aus denen du erfährst, dass manche Gefangene einmal höhere Offiziere waren, als du es bist - und doch, die sitzen hier, deine ehemaligen Vorsitzenden und Kameraden, sogar deine Freunde oder Verschwägerten. Und du selbst fühlst dich wie ein Kandidat, nach dem alten Spruch, dass ganz Russland sich dreifach teilen lässt: in Menschen, die gerade im Gefängnis sitzen, in solche, die gesessen sind und in solche, die noch sitzen werden. Also du fürchtest dich und diese Angst ist deine Gefangenschaft. Du fürchtest dich, einen Schritt zu machen, aber du hast doch ein Plansoll zu erfüllen. Du hättest gerne eine straffe Ordnung, aber mit Regime wird man hier den Plan nicht fortsetzen. Dazu sitzt auf dir dein treuester Helfer, der Vertraute, der jederzeit bereit ist, auch dich zu denunzieren. Also du spielst einen sehr strengen, aber gleichzeitig weißt du, dass übertriebene Strenge hier im Lager sich nicht auszahlt, dass man den Leiter des benachbarten Lagers vor Kurzem getötet hat. Wenn du ein Stück intelligenter bist, wahrst du nur den Schein und machst die Augen zu für das Geschehen in der Zone (Skarga 1990, 61).

Es ist die provisorische Gerechtigkeit, eine Situationsethik unter vielen komplexen Bedingungen, die das Leben in Gulag bestimmen. Es ist der Verzicht auf das Schlimmste (der Kapitän ist doch kein Mörder...), um sich selbst zu schützen, weil doch alle Menschen das gleiche Schicksal treffen kann. Verständlich, begründet, wir würden sogar sagen: menschlich. Aber doch: Es bleibt nur ein Provisorium der Gerechtigkeit, der Anerkennung in der Beziehung zum Anderen. Denn die Hierarchie in dieser Situation (einer ist Sklave, der andere Herr) bildet doch gleichzeitig ein anderes, ein weiteres Provisorium der Gerechtigkeit. Skarga ist der Meinung, dass die Identität der Macht der Totalität mit Hinblick auf die Differenz widerstehen kann. Sie sagt das in Anlehnung an Lévinas.⁶ Wie man weiß, zeigte Lévinas diese Mechanismen auf, in welchen einen die Sehnsucht nach der eigenen Identität – einer klaren, starken Identität, die Stabilisierung und Selbstsicherheit bedeutet – dazu bringt, diese innerhalb einer Totalität anzusiedeln (man erliegt dabei dem Irrtum, die eigene Identität erst dort zu finden).

Die Anerkennung dieses Faktums, dass die Differenzen auch vorrangig sind, ermöglicht es, die eigene Identität zu pflegen. Im Namen meiner Identität – die u.a. eine Zugehörigkeit zu gewissen Gruppen bedeutet (zum Kultur-, Sprachkreis) – werde ich nicht das ganze Eigene aufgeben. In der Unifizierung wird vielleicht eine Illusion oder sogar ein Rest von Sicherheit geboten, aber keine wahre Identität erlebt, die sich in der Andersheit jedes Einzelnen manifestieren könnte. Wiederum schreibt Skarga in einem politischen Kontext:

Ohne Identität zu sein, bedeutet eine Reduzierung auf ein Objekt – ein Objekt für die anderen, einfach zu manipulieren, ohne Widerstand und nicht selten bereit, sich einer starken Hand zu unterwerfen, wenn nur diese Hand imstande ist, dem

der schon fast am Ziel ist, sprich: tot ist. Stilistisch bildet es eine Vorstellung des langsamen, schmerzhaften Todes sowie eine Assoziation mit dem Tod eines Tieres.

⁶ Das Werk von Emmanuel Lévinas hat Barbara Skarga sehr viel bedeutet; meines Erachtens ist der beste Text von Skarga ihre fesselnde Einleitung zur polnischen Ausgabe von *Totalität und Infinität*.

Objekt irgendeine Form aufzudrücken. Es scheint möglich zu sein, in der starken Faust des Anderen zum eigenen Ich zu finden – ganz gleich ob es die Hand eines Meisters oder eines Tyrannen ist. In uns existiert die Furcht davor, die Identität zu verlieren, eine verständliche Furcht, die nicht nur bei einzelnen Individuen auftritt, sondern analog auch bei Menschengruppen oder gar Gesellschaften (Skarga 2009, 13).

Ich möchte auch diese Gedanken in eine eigene These kleiden:

Der Kern einer Identität ist die Differenz. Die Totalisierung des Seienden geschieht durch das Verlassen – das heißt: durch die Verleugnung – der Differenz.

Dies ist meine eigene ethische Stellungnahme. Ich entferne mich nicht von der Axiologie, vom Konzept der *Geltung von Werten*. Normen existieren doch nicht, sie kommen nur zur Geltung. Man anerkennt gutes Verhalten, richtige Lösungen – aus einer Praktik und sodann aus einer Sitte generiert man eine Regel. Wenn auch die Normen nichts an Geltung verlieren, so werden sie dennoch nicht beachtet und praktiziert. Deswegen schreibt Skarga einem rein axiologischen Konzept einen Mangel zu – dieses sei zwar gut konstruiert, aber nicht wirksam. Dies muss uns an das gutbekannte Konzept von Lévinas erinnern, dass nämlich die Verantwortung der Freiheit vorangeht.

Zum Abschluss: Die Begriffsklärung und das Kritische

Skarga begann ihre Herangehensweise an die brennendsten und aktuellsten Fragen – in der Öffentlichkeit, in der Politik – jeweils mit einer Begriffsklärung, Systematisierung, und bei der Antwort auf die Probleme des Alltags bezog sie sich auf die klassischen philosophischen Texte. Sie gehörte also zu jenen Philosophen, die an die unerschöpfliche Kraft des Anfangs der Philosophie glauben. Sie fand in den klassischen Texten Impulse zur Entlarvung unseres mehr oder weniger publizistischen Geredes. Ich möchte jetzt aus dem Buch *Metaphysisches Quintett* zitieren, einem Plädoyer für die Zusammengehörigkeit von Erfahrung und Reflexion. Ich zitiere aus dem 3. Teil, betitelt mit „Vom Bösen“:

Wenn wir nach dem Bösen fragen, steigen wir sofort in das Territorium der hohen Metaphysik ein, uns ist aber zugleich bewusst, dass dieses immer auf einem konkreten Grund wirkt, hier, unter uns, oder einfach – in uns. Dieser Gedanke verlässt mich nicht – obwohl die Metaphysiker behaupten, dass man es mit Sicherheit nicht wissen kann –, dass der Ursprung des Bösen im menschlichen Sein liegt, dass der Mensch selbst daran schuld ist, obwohl man durch Jahrhunderte hindurch versuchte, diese Last zu verwischen, an etwas oder jemanden abzugeben, irgendwelche erklärende Umstände aufzuzeigen. (...) Aber die ganze Thematik des Bösen scheint – vielleicht auch deswegen, weil sie bis dato nicht geklärt ist – wiederum so unfruchtbar zu sein, zu keinem Ziel zu führen, trotz der langen philosophischen Tradition, diese Frage nach dem Bösen zu stellen. Somit hatte man diese Thematik auch in unserer Nachkriegszeit aufgegriffen und dabei alte, inzwischen in Vergessenheit geratene Theorien in Erinnerung gebracht. [...] Denn irgendwo im tiefsten Grund geht es immer um die Wahl einer Maxime. Die Entscheidung für eine Maxime bildet die Grundlage für den Gebrauch unserer Freiheit. Und wenn wir diesen Kant'schen Gedanken weiterführen, bekommt das radikale Böse mehr an Kraft und Bedrohlichkeit. Also es ist so, dass wir das Böse völlig bewusst, als Folge einer rationalen Entscheidung

ausführen. Wir tun es nicht zufällig, nicht aus Verwirrung, nicht aufgrund unserer Schwächen. Im Gegenteil, aus einem Kalkül heraus, im Übergehen von ethischen Skrupeln, verweigern wir die moralischen Maximen, die uns im Weg stehen. Das radikal Böse entsetzt. Und doch hat uns der sehr schöne Text von Kant keine Lösung gebracht (Skarga 2005, 86).

Die zitierten Sätze bringen mich selbst zu noch einer anderen Reflexion – zur Unterscheidung von der Moral und dem Moralisieren. Es ist diese Tatsache, auf welche Ethiker und Moralphilosophen besonders aufmerksam sein sollen und auf welche wir sehr direkt verweisen müssen: dass sich *die Moral/das Moralische* und *das Moralisieren* grundlegend unterscheiden. Das Moralisieren ist eine Art Macht, vielleicht sogar Gewalt, es ist ein Eingriff, der allerdings keine wahre Wirkung hat, weil er keine moralische Entwicklung bringt. Das Moralisieren bleibt ein abgehobenes Gerede, das jedoch viele problematische Nebenwirkungen mit sich bringen kann. „Das Moralische“ dagegen bedeutet meines Erachtens unter anderem „das Kritische“, die Kritik selbst, welche jedoch synonym mit „Reflexion“ zu verstehen wäre.

Das In-Frage-Stellen ist *die Kritik* selbst. Ein gesellschaftspolitisches Thema steht nie außerhalb der Ethik – und gleichermaßen will die praktische Philosophie das Politische und Historische daran erinnern, dass sie die gleiche, ethische Verwurzelung haben. Für den Philosophen bleiben die sozialen und politischen Konzepte vor allem Ideen, wir werden an ihnen weiterarbeiten, auch wenn sie gerade in einer Gesellschaft erfolgreich umgesetzt werden, und wir werden an ihnen auch dann weiterarbeiten, wenn sie gerade von irgendwelchen Kräften vernichtet werden.

Literatur

- Domański, J. 2017. *Szkice do portretów humanistów XX wieku*. Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.
- Skarga, B. 1990. *Po wyzwoleniu...1944-1956*. Poznań: W drodze.
- Skarga, B. 2005. *Kwintet metafizyczny*. Kraków: Universitas.
- Skarga, B. 2009. *Tożsamość i różnica. Eseje metafizyczne*. Kraków: Znak.

Małgorzata Bogaczyk
(Poznań)

On Philosophy of Barbara Skarga: An Approach to Get to Know Her Life and Work

Abstract: Barbara Skarga (1919–2009) was one of the most important Polish Philosophers of 20th century. She was an expert in classical and contemporary French (e.g. Comte, Bergson, Lévinas) and German Philosophy (e.g. Kant, Hegel, Heidegger). In this paper I present some important biographical facts (participation in organized resistance in Vilnius, interments in Gulags) as well Skarga's philosophical, mostly sociopolitical and ethical, ideas. I called its philosophical concept "philosophy of difference".

Keywords: Barbara Skarga; Polish contemporary philosophy; philosophy of dialogue; metaphysics; difference; identity.

Ethics in Progress (ISSN 2084-9257). Vol. 9 (2018). No. 1, Art. #10, pp. 183-191.

Creative Commons BY-SA 3.0

Doi: 10.14746/eip.2018.1.10