Peter JONKERS
Giovanni SCARAFILE

ISHIKAWA Tomoki
AMANO Emiri
WATANABE Kazuki

Tetsugaku

International Journal of the
Philosophical Association of Japan

Special Theme

World Philosophy

Volume 8, 2024

ISSN 2432-8995

ISSN 2432-8995

Tetsugaku

International Journal of the Philosophical Association of Japan

Volume 8

2024

Special Theme

World Philosophy

The Philosophical Association of Japan

Tetsugaku
International Journal of the Philosophical Association of Japan
Volume 8, Special Theme: World Philosophy
December 2024

Edited and published by the Philosophical Association of Japan
College of Arts
Rikkyo University
3-34-1 Nishi-Ikebukuro, Toshima-ku, Tokyo 171-8501, Japan

ISSN 2432-8995

Table of Contents

Tetsugaku
International Journal of
the Philosophical Association of Japan
Volume 8, 2024

I Special Theme: World Philosophy

Ethical Education in a Multicultural World

Peter JONKERS
Emeritus Professor of Philosophy, Tilburg University

***Abstract**: This paper addresses a pressing question of world philosophy, viz. the impact of living in a multicultural world on ethical education. To explore this question, the introduction gives an overview of the ambivalent attitudes of modern, detraditionalized societies toward their embedded traditions and substantial values, based on the theories of social scientist Ronald Inglehart and philosopher Charles Taylor. The second section, which builds on Aristotle's insights and their interpretation by Martha Nussbaum, discusses a traditional component of ethical education, viz. the virtue of practical wisdom. Sections three and four focus on the challenges of multiculturalism to ethical education. In today's societies, it is difficult to reach a consensus and to find plausible points of orientation in pressing ethical questions. This is why a very influential way of responding to the multicultural situation, viz. to focus on universal but at the same time formal and procedural moral principles, and to leave the contents of substantial values up to the individuals' discretion, has fallen short of expectations. In exploring an alternative approach to this question, I use Paul Ricoeur's views on the value of traditions, viz. appreciating them as culturally embedded mediators between universal moral principles and the contingencies of the life-world. This is the second goal of ethical education. Yet, to avoid the looming deadlock of ethical traditionalism and particularism a reasonable debate is necessary to find out if and how particular ethical traditions have a universal potential, which can be integrated into universal principles. This approach aims at fostering the dialogue among ethical traditions with the help of an enlarged idea of rationality. Thus, the third goal of ethical education is training people in self-reflection about the ethical values of their own traditions, as a first step to understanding the values of other traditions as potential universals.*

Introduction: Setting the scene

These days, we see a growing interest in ethical education. Throughout the European Union and in many other countries, ethics is a part of the curriculum of high schools and universities, either as a separate course or as a point of attention in other courses,

like citizenship, religion, economics, law, etc.[1] Yet, as I will argue in more detail in this paper, ethical education is more than obtaining knowledge about moral theories, like utilitarianism, deontological ethics, or virtue ethics. Students can easily reproduce these theories without relating them to their moral behavior; reversely, people can be exemplars of ethical behavior without being knowledgeable about these theories. Hence, ethical education is not only a cognitive affair but should also include training in several non-cognitive skills and virtues, like ethical reflection, balanced judgment, and appropriate action.[2] An even more important element that needs to be taken into account is the rise of socio-cultural diversity in many societies around the world. This challenges ethical education to include a historical and intercultural dimension in its teaching and practical training.

One of the reasons for the growing attention to ethical education, not only among educators but also among politicians, religious leaders, and opinion formers concerns the ambivalent effects of detraditionalization, the shift from material to postmaterial values, and increasing socio-cultural diversity in many modern societies. Detraditionalization can be defined as the gradual fading of the societal impact of traditional values, moral institutions, and common patterns on people's (ethical) behavior. As the different waves of the World Values Survey show, the influence of (religious) ethical traditions upon modern societies has decreased dramatically since the end of the Second World War, while the importance of secular-rational values has been constantly growing. Secular-rational values "place less emphasis on religion, traditional family values and authority,"[3] and support a procedural, instrumentalist, and individualist approach of ethic, based on the principles of disengaged, impartial reason. Because of its claim to universal validity this procedural ethic and the secular-rational values on which it rests tend to overrule all substantial, traditional and hence culture-specific values.

In rich, postindustrial societies, this evolution runs parallel to another one, viz. the diminishing relevance of material or survival values. These values emphasize economic and physical security and are linked to a relatively ethnocentric outlook and

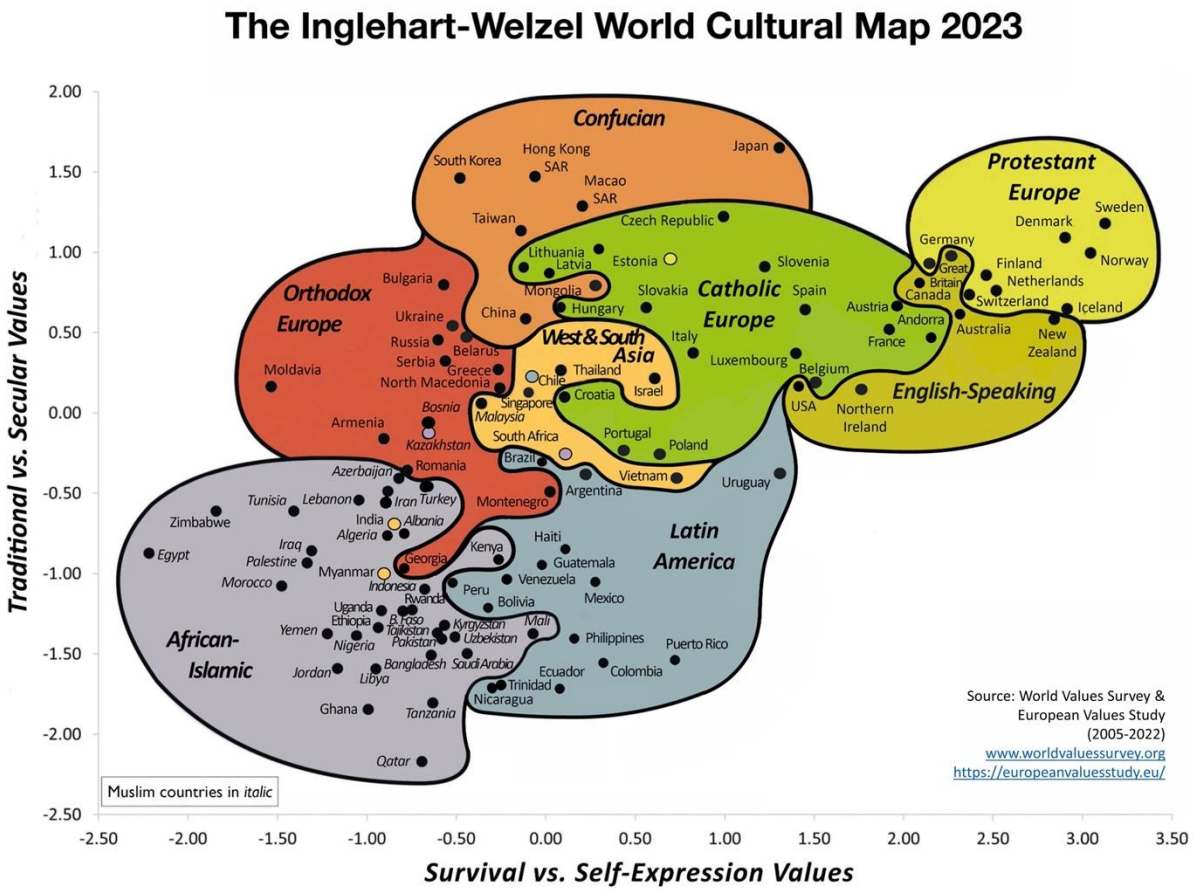
[1] For an overview of the situation in Europe, see Natascha Kienstra, "Ethics Education in a Globalized World: A Community of Ethics Teachers in Europe (COMET)," *Age of Globalization. Studies in Contemporary Global Processes* 33, no. 1 (2020), 109-119.
[2] Natascha Kienstra, *Zorg dragen voor ethische vorming [Caring for Ethics Education]* (Tilburg: Tilburg University Press, 2023), 16.
[3] *Inglehart-Wenzel World Cultural Map – World Values Survey 7 (2023)*.
Source: http://www.worldvaluessurvey.org/

low levels of trust and tolerance.[4] The three decades after World War II were a period of rapid economic growth and a more equal distribution of wealth in these rich societies. Because physical and economic security could be taken for granted during that period, material, survival values were replaced by the postmaterial values of individual self-expression, which emphasize human autonomy and choice. Self-expression values "emancipate people from traditional constraints that are no longer necessary for survival, allowing greater freedom of choice in how to live one's life. For many groups such as women and gays, emancipation from traditional constraints makes a major contribution to life satisfaction and happiness".[5] It goes without saying that the shift toward individual self-expression substantially increased socio-cultural diversity in Western societies, a trend that was further enhanced by the influx of immigrants, who brought their own cultures and values with them. In sum, rich, postindustrial societies, which are mainly Western ones, went through a process of pervasive cultural changes, leading them to embrace secular-rational and self-expression values, to the detriment of traditional and survival values.[6]

[4] Ibid.

[5] Ronald F. Inglehart, *Cultural Evolution. People's Motivations Are Changing and Reshaping the World* (Cambridge: Cambridge University Press, 2018), 39.

[6] The Inglehart-Welzel World Cultural Map of 2023 gives an excellent graphic overview of the current situation regarding the importance of these four groups of values under discussion in different countries across the world. See *Inglehart-Wenzel World Cultural Map – World Values Survey 7 (2023).*



From a philosophical perspective, Charles Taylor interprets this evolution in terms of the rise of the ethics of authenticity. This expression refers to a new and potent ethical ideal, which stimulates each of us to discover and follow our own way of realizing our humanity and to live it out, as against surrendering to conformity with a model imposed on us from outside.[7] This ideal, whose origin goes back to the Romantic movement at the end of the 18th century, has become mainstream in Western societies. However, the spread of this ideal has gone hand in hand with its narrowing down to the principle of free, individual self-determination, thus enhancing the individualizing and liberalizing trends in society. This evolution, in combination with the increase of socio-cultural diversity, explains the procedural character of contemporary ethics; it tends to become restricted to following correct and fair rules and procedures for autonomous decision-making, while leaving it up to the individuals to define the substance of their ethical values. A relatively recent factor,

[7] Charles Taylor, *A Secular Age* (Cambridge, MA: Harvard University Press, 2007), 475.

enhancing this trend, is that the principle of self-determination and the proceduralist approach to ethical issues have been implemented in the legislation of most Western societies, resulting in a host of new laws that allow and facilitate individuals to choose their way of life as long as it does not cause harm to other people. This shows that the state has become loath to define, let alone impose traditional ideas of the common good and substantial values upon its citizens and confines itself to laying down the procedures that need to be respected. Examples of this trend are the legal recognition of a multiplicity of family models and work systems, and the voting of laws that allow people to decide autonomously about beginning- and end-of-life arrangements, gender, sexuality, parenting, etc., under the condition that they follow the correct procedures.

Yet, as Taylor has convincingly argued, the drawback of detraditionalization and the reduction of the ethics of authenticity to promoting the values of individual self-expression and free self-determination is that they reduce people to unsituated, even punctual selves and sap the broader horizons of meaning, against which every ethical decision gets its significance.[8] Authenticity is a genuine moral ideal, but it cannot "stand alone, because it requires a horizon of issues of importance, which help define the *respects* in which self-making is significant. […Hence,] the agent seeking significance in life, trying to define him- or herself meaningfully, has to exist in a horizon of important questions".[9] Therefore, ethical traditions and substantial values, which constitute the frame of reference and meaning of a person's ethical deliberations and actions, keep their relevance even in a detraditionalized society. This shows that "authenticity is not the enemy of demands that emanate from beyond the self; it supposes such demands".[10] Precisely the recognition of the importance of these horizons of meaning distinguishes the moral ideal of authenticity from its flattened mode, viz. free self-determination, and, hence, serves as a philosophical critique of the dominance of a secular-rational and procedural ethics and the value of free self-expression. According to Taylor, the current dominance of these ideas is self-defeating because it ignores the dialogical character of the self and the fact that the self cannot determine on its own what is of value since this depends on the recognition by others.

[8] Charles Taylor, *Sources of the Self: The Making of the Modern Identity* (Cambridge: Cambridge University Press, 1989), 514.
[9] Charles Taylor, *The Ethics of Authenticity* (Cambridge, MA: Harvard University Press, 1991), 39f.
[10] Ibid., 41.

Interestingly, Taylor's philosophical insights concerning the importance of ethical traditions in modern societies as horizons of meaning and orientation and his critique of procedural ethics are confirmed by empirical research into some recent evolutions of the values in post-industrial, high-income societies. In the USA and most European countries, the spread of secular-rational values and postmaterial, self-expression values, which were promoted by the younger, economically more secure, and better-educated strata during the "Golden Sixties", is going into reverse. According to Ronald Inglehart, this backlash is partly caused by growing economic inequality and, hence, insecurity among older and less educated people since the end of the previous century.[11] This has resulted in a return to material values, which goes hand in hand with a rise in xenophobic and ethnocentric attitudes and a greater attachment to family and other traditional values.

A second explaining factor of this reverse trend is that "postmaterialism was its own gravedigger".[12] One of the effects of the growth of self-expression values and the concomitant liberalization and individualization of society has been the emergence of radical cultural changes and their implementation into the legal systems of most rich, post-industrial societies. More than that, sociological research shows that "the new non-economic issues introduced by Postmaterialists overshadowed the classic economic Left-Right economic issues, drawing attention away from redistribution to cultural issues".[13] This has provoked a counterreaction among older and less secure strata who felt threatened by the erosion of familiar values. Apparently, the return to traditional and material values is not only caused by growing economic insecurity but also by socio-cultural insecurity. An additional element that contributed to the rise of socio-cultural insecurity is the large immigration flows from low-income countries with different cultures and religions, causing a feeling of socio-cultural estrangement among the native population. "Rapid cultural change, coupled with large-scale immigration, tends to make older [and less educated] people feel that they are no longer living in the country in which they grew up".[14]

Of course, the above counterreactions against radical cultural changes have been present for decades. What is new is that this trend has become dominant in many Western societies: it receives support from people who do not belong to the older and less educated strata of the population, resulting in a diminishing tolerance toward

[11] Inglehart, *Cultural Evolution*, 191-193
[12] Ibid., 175; see also 188-191.
[13] Ibid., 175; see also 181f; 185f.
[14] Ibid., 188.

cultural minorities and more attention to socio-cultural matters in the programs of mainstream political parties. Unsurprisingly, this evolution has led to rising tensions between the more and less educated, younger people and older ones, city dwellers and people who live in the countryside, native inhabitants and migrants, etc.

Interpreting the above analysis of radical cultural change from a larger perspective explains why the reverse movement to traditional values is not as unexpected as it seems at first sight. Social scientists argue that "cultural change is path-dependent: a society's values are shaped by its entire historical heritage, and not just its level of existential security".[15] This means that cultural change certainly happens but at a very slow and gradual pace. This is because traditional values are deeply embedded in a society, so that they never really disappear but continue to play an important role in the background. Therefore, in times of adverse period effects, such as economic crises or periods of cultural insecurity, people still can and actually do fall back on these traditional values. In sum, despite all the (empirically well corroborated) talk about detraditionalization and the shift toward postmaterial and self-expression values, traditions and traditional values remain important in rich, postindustrial societies as ways to give orientation and motivation to people's ethical behavior.

The upshot of the above overview illustrates the predicament in which many contemporary societies find themselves. Universalist, secular-rational ethics has fallen short of expectations because of its procedural character, which means that it is unable to serve as the only or even most important source of ethical orientation and motivation. At the same time, a simple return to traditional values offers no solution to the above tensions either because it ignores the profound and irreversible societal changes that Western societies have been going through since the end of the Second World War. Therefore, it is no wonder that the contents and goals of ethical education in a multicultural world have become an important issue. This paper aims to contribute to this debate by analyzing three important elements of ethical education and the cognitive and attitudinal virtues underpinning them. The next section discusses the virtue of practical wisdom, thereby building on the insights of Aristotle and the interpretation that American philosopher Martha Nussbaum gives of them. Sections

[15] Ibid., 24; see also 42f. See also Wil Arts, "The European Values Study and Grand Theory: A Fruitful Alliance?", in *Reflections on European Values: Honouring Loek Halman's Contribution to the European Values Study (European Values Series, 2),* eds. Ruud Luijkx, Tim Reeskens, and Inge Sieben (Tilburg: Open Press Tilburg University, 2022), 36: "Cultural traditions create forces to sustain themselves even though the circumstances that gave rise to and reinforced them in the past may now no longer be relevant".

three and four analyze the virtues that are needed to respond to the ethical challenges of today, viz. the gap between procedural ethics and substantial values, and the consequences of rising socio-cultural diversity for peaceful co-existence in a multicultural world. The analyses of these two sections build on the insights from French philosopher Paul Ricoeur.

Practical wisdom

Following Aristotle, the first goal of ethical education is to acquire practical wisdom, defined as the reasoned capacity to act regarding the things that are good or bad for man.[16] Although practical wisdom cannot attain the same degree of exactness and certainty as theoretical wisdom because of the contingency of human life, practical wisdom is a *reasoned* capacity and can therefore not be reduced to a private opinion or practical skill. Rather, practical wisdom is knowledge about what is truly good for all human beings, and its insights serve as a horizon against which the ethical quality of individual decisions and actions can be assessed.

Knowledge about moral principles and obligations is an important goal of practical wisdom for yet another reason. As summaries of the wise judgments of others, these principles and obligations are guidelines in moral development. They serve as rules of thumb, guiding virtuous people tentatively in their approach to the particular, and helping them to pick out its salient features. When there is no time to formulate a concrete decision or to deliberate about the impact of all the features of the case at hand, it is better to follow a good rule of thumb than to make a hasty and inadequate concrete choice. Furthermore, these rules and obligations give constancy and stability in situations in which bias and passion might distort judgment. In sum, knowledge about moral principles and obligations is necessary because we are not always good judges.[17]

The final and most important reason why knowledge about moral principles is crucial for moral education is that this knowledge is essential to grasp the ethical salience of particular situations: “The particular case would be surd and unintelligible without the guiding and sorting power of the universal. […] Nor does particular

[16] Aristotle, *Metaphysics,* I 1, 981b-982b; *Nicomachean Ethics,* VI 3-7, 1139b-1141b; VI 8, 1142a 23-4.

[17] Martha Nussbaum, *The fragility of goodness. Luck and ethics in Greek tragedy and philosophy* (Cambridge: Cambridge University Press, 1986), 304.

judgment have the kind of rootedness and focus required for goodness of character without a core of commitment to a general conception – albeit one that is continually evolving, ready for surprise, and not rigid. There is in effect a two-way illumination between particular and universal".[18] Since all knowledge is essentially a matter of interpreting the particular experiential world with the help of universal concepts, we always need to perform a back-and-forth movement between the universal and the particular, in this case, between the life-world and its broader, conceptual horizons of meaning. Hence, conceptual knowledge is needed to orient and sift out our spontaneous experiences of right and wrong, while moral experiences are crucial to give content to these concepts and to amend and modify them in the light of new ethical challenges. Yet, there is a difference between theoretical and practical knowledge: whereas the laws of nature apply to the material world in a univocal way, the world of ethical praxis is continually evolving in sometimes unpredictable ways, so that it requires a kind of knowledge that is flexible and ready to respond to unexpected situations. To illustrate the flexible character of practical wisdom, Aristotle uses the image of the builders of Lesbos, who measure a round column with a flexible lead rule that "adapts itself to the shape of the stone and is not rigid".[19]

In sum, insight into moral principles and obligations is an important goal of ethical education. This insight can be acquired in various informal and formal, experiential and theoretical ways: The informal and experiential ones, which are by far the most effective, include parental education, life stories about ethical dilemmas and decisions, conversations with relatives and friends, and the public (including online) debate on topical ethical issues. Formal knowledge about the theories that underpin these principles and obligations is acquired through courses and books on ethics, religious teachings, the study of human rights and constitutional essentials, etc.

Stressing the importance of moral knowledge for ethical education is a critique of the idea that self-expression and individual self-determination are the only foundations of ethical deliberations and actions. Knowledge about moral principles and obligations is essential to contain the passions of the heart, to give orientation to ethical choices, and to foster critical self-reflection. This marks the crucial distinction between Immanuel Kant's idea of moral autonomy, according to which the will of every rational being should be a universally legislating will,[20] and the principle of

[18] Ibid. 306.

[19] Aristotle, *Nicomachean Ethics,* V 10, 1137b 29-32.

[20] Immanuel Kant, *Groundwork of the Metaphysics of Morals* (Cambridge: Cambridge University Press, 1998), 39 (*Akademie Ausgabe* 4:431). Of course, it is essential to interpret the idea of autonomy in a non-egological and non-monological way, taking into account the

self-determination, which does not refer to any universal moral principle or obligation and hence lacks the broader perspective that ethical actions require. Furthermore, Kant's categorical imperative highlights the need for an ethical reflection about whether individual maxims for action can become universal moral laws,[21] whereas the value of individual self-expression lacks this reflexive, self-critical potential.

However, moral knowledge is only one aspect of practical wisdom. It also comprises sensitivity to the moral dimensions of particular situations and actions, the ability to apply universal moral principles to concrete situations, and, reversely, the capacity to seize these principles and values in a confrontation with concrete, moral situations. In other words, universal moral principles need to be contextualized in historical and communitarian ways.[22] This aspect of practical wisdom is the result of a (reflexive) praxis, which is why ethical education requires not only cognitive insight but also practical training.

The practical aspect of practical wisdom comes to the fore when we realize that moral principles are plural and incommensurable, and therefore cannot be measured univocally, as if morals were a kind of *technè*. There is no single common notion of the good that practical wisdom only needs to apply to pass a correct ethical judgment in specific situations. Instead, the best human life should be conceived as a life inclusive of a number of different constituents, each being defined apart from each of the others and valued for its own sake; each virtue is defined separately, as something that has value in itself. To put it concretely: "If I should ask of justice and of love whether both are constituent parts of *eudaimonia* […], I surely do not imply […] that we are to hold them up to a single standard, regarding them as productive of some further value. […] Something can be an end in itself and at the same time be a valued constituent in a larger or more inclusive end". To choose a value "for *its own* sake (for the sake of what it itself is) not only does not require, but is actually incompatible with, viewing it as qualitatively commensurable with other valuable items".[23] Different values can be legitimately valuable for their own sake and cannot be weighed up against each other without considering the particular situation.

---

plurality and otherness of persons, which implies accepting a certain degree of heteronomy; see Paul Ricoeur, *Oneself as Another* (Chicago, IL: University of Chicago Press, 1992), 275.

[21] Kant, *Groundwork of the Metaphysics of Morals,* 47 (*AA* 4:441)

[22] Ricoeur, *Oneself as Another,* 274.

[23] Nussbaum, *The fragility of goodness,* 297. For this argument see also Thomas Nagel, "The Fragmentation of Value," in *Mortal Questions* (Cambridge: Cambridge University Press, 2012), 128-141.

Hence, because moral principles are manifold and because the practical field in which they are realized is mutable, indeterminate, and particular or non-repeatable, it is crucial that the person of practical wisdom is a thoroughly human being, that is, someone who does not attempt to stand outside of the conditions of human life but bases her judgment on a long and broad experience of these conditions,[24] and is able to understand their complexities. This is why confining ethical education to teaching universal moral principles and values, like human rights, the principle of reciprocity, etc., is inadequate. Instead, ethical education should also comprise training people in accepting the confrontation of these principles with the particular conditions of human life.

In sum, the first goal of ethical education is to foster practical wisdom, which consists not only of moral knowledge and insight into the multiple and incommensurable nature of moral principles and obligations but also practical training in how to apply these universal principles to particular situations. Because the judgments and decisions of practical wisdom are always situational, a person of practical wisdom needs to realize that they are always temporary and open to improvement.

Revaluing ethical traditions

The second goal of ethical education concerns a revaluation of the role of traditions in contemporary, detraditionalized societies. This introduces a new element in comparison with Aristotle's theory about practical wisdom. Aristotle developed his ideas in a traditional, morally homogeneous society with a small number of well-defined relationships, so that there was a broad consensus about the substantial values constituting the good life. By contrast, today's societies have become irreducibly plural and even fragmented; in the words of Ricoeur, they are societies "after Babel".[25] This implies that reaching a consensus about substantial values and their mutual relationships has become difficult. To deal with this situation, contemporary polities tend to limit themselves to defining rather formal and procedural moral principles, like free self-determination, reciprocity, causing no harm to other people, etc. and are

[24] Nussbaum, *The fragility of goodness,* 290.

[25] Paul Ricoeur, *On Translation* (London: Routledge, 2006), 11. Ricoeur refers to George Steiner, *After Babel: Aspects of Language and Translation* (Oxford, Oxford University Press, 1998).

loath to interfere directly in people's preferences regarding the prioritization of these principles and the ensuing substantial values. This means that modern societies leave the contents of the good life up to the discretion of individual citizens and only try to influence this process indirectly through nudging.[26]

However, this procedural approach to values generates its own problems, thus showing the flipside of the shift from traditional to secular values. The rationality that underpins secular values is not as neutral as it seems but is a heritage of the Enlightenment, dominated by a functional and quantitative outlook on the world. Therefore, functional rationality is inclined to treat traditions and their ideas of the good life as non-functional, non-measurable, and thus irrational. As Jürgen Habermas has argued, because of its dominant position in most modern societies, this type of rationality has colonized the life-world of many people by re-organizing it through more impersonal and strategic exchanges of money and power, within the context of the economy and the modern administrative state and judiciary.[27] The upshot of this development is that, in today's detraditionalized societies, formal principles and secular-rational values appear unable to orient and motivate people's ethical life in the broad sense. This explains why they nowadays feel left to their own devices when it comes to finding ethical orientation in the life-world. This causes increased mental pressure, as the growing numbers of burnouts illustrate,[28] and estrangement from the functioning of the modern state, as the rise of populist political parties in many Western countries shows.[29] The reverse movement from secular to traditional values, as analyzed above, is an expression of society's discomfort with and resistance against the dominance of procedural rationality and its formal moral principles. This shows that the universal but formal principles of justice, fairness, reciprocity, etc., however important they are, cannot be entirely detached from traditional, teleological ideas of the good,[30] thus providing a strong argument for the need to revalue the role of traditions in detraditionalized societies: Traditions are a source of inspiration for a life

[26] For examples, see the Introduction.
[27] Nick Crossley, *Key Concepts in Critical Social Theory* (London: Sage Publications, 2005), 37f. See also: Timo Jütten, "The Colonization Thesis: Habermas on Reification", *International Journal of Philosophical Studies* 19, no. 5 (2011): 704.
[28] Joep de Hart, Pepijn van Houwelingen, Willem Huijnk, *Religie in een pluriforme samenleving. Diversiteit en verandering in beeld. Deel 3: buiten kerk en moskee [Religion in a Pluriform Society. Diversity and Change in Pictures. Part 3: Outside Church and Mosque]* (Den Haag: Sociaal en Cultureel Planbureau, 2022), 147.
[29] Inglehart, *Cultural Evolution*, 176-187.
[30] Ricoeur, *Oneself as Another*, 284.

beyond the domain of functional and quantifiable rationality and for the substantial ethical values embedded in them.

Traditions are also important to modern societies for another reason: they bridge the gap between formal moral principles and the life-world. Traditions can be defined as successful examples of the good life, thus showing people how to live by substantial values. Moreover, since traditions have a narrative character, they are usually better suited than abstract moral principles to make people aware of ethical dilemmas and trade-offs. Finally, because of their embeddedness in the life-world, ethical traditions have a far greater motivational potential than formal principles or theoretical arguments about ethical matters. To avoid misunderstanding, this argument in favor of revaluing traditions should not be understood as a plea for traditionalism. Modern, democratic societies are by definition plural and, therefore, need to be based on a consensus about universal moral principles as a result of reasonable deliberation. This raises the fundamental question about the recognition of particular traditions and their substantial values in a plural, multicultural society, which will be explored in the next section.

In this section, I will confine myself to showing how traditions can bridge the gap between universal, but formal moral principles and people's particular convictions and substantial values in the life-world. To do so, I will summarize Ricoeur's critical discussion with Hans Küng about the pros and cons of a global ethic.[31] Küng states that the principle of reciprocity, also known as the Golden Rule, is the most universally accepted moral rule in all cultures and throughout the times.[32] Because this principle "is not just hypothetical and conditioned, but is categorical, apodictic and unconditioned – utterly practicable in the face of the extremely complex situation in which the individual or groups must often act,"[33] it has tremendous moral authority and reveals the profound unity that underlies the diversity of human

[31] "Entretien Hans Küng – Paul Ricoeur [autour du "Manifeste pour une éthique planétaire" (Ed. Du Cerf) de Hans Küng]," in *Les religions, la violence et la paix. Pour une éthique planétaire* http://www.fondsricoeur.fr/uploads/medias/articles_pr/entretien-hans-kung-paul-ricoeur-v2.pdf (retrieved Feb. 5, 2016).

[32] Hans Küng, *Global Responsibility: In Search of a New World Ethic*, (New York: Crossroad Publications, 1991); "Global Politics and Global Ethics. Status Quo and Perspectives", *Seton Hall Journal of Diplomacy and International Relations* (Winter/Spring 2002), 8-20. For an analysis of the complexities of reciprocity as the principle of a global ethics, see Peter Jonkers, "Can Reciprocity Be the Principle of a Global Ethics?", in *Reciprocity: A Human Value in a Pluralistic World*, eds. Peter Jonkers, Wang Tianen, Astrid Vicas (Washington DC: Council for Research in Values and Philosophy, 2022), 275-294.

[33] Küng, *Global Responsibility*, 59.

experience. This is why Küng proposed reciprocity as the principle of the "Declaration Toward a Global Ethic", which was accepted by the Parliament of World Religions in 1993.[34] In contrast to the idea of self-determination, the principle of reciprocity recognizes that human beings are dependent on and in permanent interaction with each other, thus creating room for commitment to fellow human beings and the world around them.

Despite its obvious merits, a fundamental problem of reciprocity as the most fundamental and universal ethical principle is its disembodied formalism. Küng accepts this drawback to be able to unite all people around this principle and expects them to endorse it. He is of the opinion that the Declaration Toward a Global Ethic rightly does not mention specific ethical issues, such as sexual behavior, contraception, abortion, euthanasia, etc. because there is no consensus either among religions or within each single religion.[35] This is a fair argument, but Ricoeur queries the relevance of this global ethics if the consensus on which it is based does not reach beyond the abstract universality of the Golden Rule and proves to be unable to address concrete, pressing ethical questions.[36]

Hence, it is no surprise that Ricoeur is quite critical of the high hopes that Küng places in reciprocity as the principle of a global ethic. He compares Küng's project with the attempt to create a universal language, viz. Esperanto. This attempt has been unsuccessful because people are profoundly attached to their native language. In a similar vein, because global ethics only rests on the universal principle of reciprocity, it does not recognize the passion that arises from people's deep attachments to their individual life-worlds, including their heterogeneous ethical traditions as an orienting and motivating force for moral praxis.[37] This insight shows that universal moral principles can only become effective if they can be embedded in ethical traditions as concrete exemplifications of the good life, just like the universal capacity to express oneself and communicate linguistically can only be realized through a multiplicity of particular languages. Therefore, universal moral principles cannot overtrump culture-

[34] See *Declaration Toward a Global Ethic,* http://www.parliamentofreligions.org/content/toward-global-ethic-initial-declaration. For an excellent overview of the background and the various stages of Küng's project of a World Ethos see José Casanova, "The Sacralization of the *Humanum:* A Theology for a Global Age", *International Journal of Politics, Culture and Society* 13, no. 1 (1999): 21-40.
[35] "Entretien Hans Küng – Paul Ricoeur."
[36] Ibid.
[37] Ibid.

specific traditions, nor should the relation between them be conceived as antagonistic. Rather, societies need to find a reflexive equilibrium between them.

Hence, the second goal of ethical education consists of making people sensitive (again) to the importance of ethical traditions, especially their role as bridges between formal principles and the concrete life-world. In fact, ethical life in the broad sense consists of intertwining three aspects: first, there are universal moral principles; second, these principles have to be applied to the complexities and contingencies of the life-world; and, third, these life-worlds are embedded in historical and communitarian traditions of the good life, which are, in turn, related to these universal moral principles. Inevitably, there will be conflicts between these three dimensions of ethical life, since they result from the conflicting nature of human existence itself. However, it is impossible and undesirable to extract people's individual ethical actions from their embeddedness in larger traditions or to treat these traditions as if they could be reduced to objects of detached moral reasoning, as the failure of Küng's proposal about reciprocity as the principle of a global ethics shows. As human beings, we are always situated in such a way that we cannot bring our individual moral praxis face-to-face with universal ethical principles, be it reciprocity or another one. Rather, we need traditions as examples of the good life, embedded in the life-world, thus enabling them to mediate between the two other aspects of ethical life.

Responding to the challenges of a multicultural world

The plural, multicultural character of liberal democracies adds yet another new dimension to the Aristotelian idea of ethical life and, hence, points to a third goal of ethical education. As argued in the previous section, the detraditionalization of modern societies needs to be counterbalanced by ethical traditions and their substantial values as means for orientation and a mediation between formal moral principles and the life-world. Yet, typical of a multicultural world is that these traditions are plural, heterogeneous, and thus potentially conflictual.[38] Social scientists have demonstrated that the demographical composition of contemporary Western societies, especially in the big cities, is becoming increasingly diverse, evolving towards an increasing number of minority cultures without a majority culture. The members of each of these cultures are attached to their own traditions and

[38] Peter Jonkers, "How to Respond to Conflicts over Value Pluralism?" *Journal of Nationalism, Memory, & Language Politics* 13, no. 2 (2020): 1-22.

substantial values and claim the right to manifest their attachments in the public sphere and even ask for the political recognition of these values. Hence, it is no option to (politically) impose a limited number of well-established values of the majority culture upon society as a whole. This shows how vital it is that ethical education trains people to respond to the challenges of today's multicultural world and its concomitant ethical pluralism in a constructive, non-exclusivist way. In this section, I further explore Ricoeur's ideas on this matter.

This goal of ethical education should first make people aware of the passion, which arises from their deep attachment to traditions and the substantial values embedded in them. As demonstrated in the first section, the cause of many controversies about these issues in our times is not the persistence of these traditions and values but rather the fact that their crucial contribution to the formation of people's socio-cultural identity has been ignored. Hence, the pluralism of traditions and substantial values is not only a matter of fact but their (political) recognition is also a legitimate claim from communities the world over. Thus, the challenge of our times is not so much finding a minimal consensus about universal moral principles, as Küng argued, but rather becoming sensitive to what distinguishes us from people who belong to another tradition and learning how to prevent this distinction from degenerating into a means of exclusion.

To develop this sensitivity, we need a profound understanding of the – sometimes conflictual – plurality of traditions, because only on that base can we become aware of the substantial ethical values embedded in them. This understanding should primarily take the form of a self-reflection on why and how the values of our own tradition separate us from those of other ones. The crucial question in this respect is: how can I recognize from my own conviction that something vital is not said in my tradition but may be addressed in another one?[39] Such a self-reflection can lead to the conclusion: yes, we can understand that other people endorse their values from a different point of view than ours, i.e., from their embeddedness in a different tradition. Hence the most appropriate way of responding to our multicultural condition is to discover and recognize the inexhaustible richness of ethical life at the heart of particular traditions instead of abstracting from them.

To realize this goal of ethical education, Ricoeur proposes that we should learn to assume the following paradox: "On the one hand, one must maintain the universal claim attached to a few values where the universal and the historical intersect, and on the other hand, one must submit this claim to discussion, not on a formal level, but on

[39] "Entretien Hans Küng – Paul Ricoeur."

the level of the convictions incorporated in concrete forms of life. Nothing can result from this discussion unless every party recognizes that other potential universals are contained in so-called exotic cultures. The path of eventual consensus can emerge only from mutual recognition on the level of acceptability, that is, by admitting a possible truth, admitting proposals of meaning that are at first sight foreign to us".[40]

In this quote, Ricoeur expresses his agreement with the basic idea of Habermas and John Rawls that the problem of cultural and ethical diversity can only be solved by means of public political debates, which are the cornerstone of liberal, deliberative democracies. Yet, Ricoeur is skeptical about their argument that these debates require the translation of these traditions and substantial values into the language of secular reason as the only common ground of liberal democracies.[41] In contrast to Habermas and Rawls, Ricoeur is of the opinion that the translation proviso favors secular moral principles and attributes them a seemingly universal status, overtrumping the values of particular ethical traditions. Moreover, the formal character of the moral principles of secular reason makes them unsuited to fully appreciate the substantial values embedded in these traditions. This also explains why most religious and cultural traditions reject this proviso as an unfair demand. Instead, to create a level playing field between different ethical traditions, including the secular one, Ricoeur thinks it necessary to broaden our conception of rationality in such a way that it is hospitable to the substantial values of all ethical traditions.

The major advantage of this broadened rationality is that it enables people to give a constructive response to the ethical challenges of today's multicultural world: with the help of this enriched idea of rationality ethical education can create an awareness that the substantial values of exotic traditions potentially have a universal significance. According to the traditional view, which goes back to Aristotle, the task of practical wisdom is to apply universal moral principles to contingent situations, resulting in a prudential judgment. Yet, this idea of practical wisdom leaves these principles unaffected and holds them superior to the values of particular traditions. Hence, the possible cultural biases of these principles remain unnoticed; an example of such a bias in Aristotle's ethics and political philosophy is the inferior position that he attributes to slaves, craftsmen, and women in the polis. The innovative character of Ricoeur's proposal lies in recognizing the universal potential of ethical traditions,

[40] Ricoeur, *Oneself as Another,* 289.

[41] See e.g. Jürgen Habermas, "Faith and Knowledge," in *The Future of Human Nature* (Cambridge: Polity Press, 2003), 109; John Rawls, "The Idea of Public Reason Revisited", in *Political Liberalism. Expanded edition* (New York: Columbia University Press, 2005), 462.

which means that they can possibly modify and enrich universal moral principles. In other words, the relation between universal moral principles and particular traditions and their values is not a one-way movement from the former to the latter but also the other way around. Stressing the potential universality of particular ethical traditions makes this relation much more dynamic without yielding to ethical traditionalism, since the question of whether this potential universality is actually realized depends on reasonable deliberation.

The Christian value of mercy is an excellent example of the universal potential of a particular tradition. This value offers a necessary modification and complement to the principle of reciprocity, which Küng saw as the basis of global ethics. Reciprocity is based on the symmetrical relations between free and equal citizens in a well-ordered society, and hence the basic principle of a fair distribution of primary goods.[42] However, the value of mercy, which is embedded in the Christian tradition and is narrated, among others, in the parable of the Good Samaritan, points to the importance of asymmetric relationships in every society, as the examples of the relationships between parents and children, healthy and disabled people, etc. show. Hence, the value of mercy shows the limits of the principle of reciprocity: a society based on reciprocity alone is unforgiving because it fails to take into account that the altruistic attitude of giving something without expecting something in return plays a crucial role in all forms of human interaction.[43] As is common knowledge, this insight has not remained limited to the Christian community of faith but is nowadays recognized as vital for every human society and implemented in the legislation of most states. This clearly shows the universal potential of the value of mercy.[44] This conclusion does not automatically put particular ethical traditions on a par with universal moral principles, but it definitely rebalances the traditional conception of their relationship by adding a historical and intercultural dimension to the idea of universal moral principles. In sum, learning to see insights of particular ethical traditions as potential universals is an essential element of ethical education in a multicultural world, because it staves off the danger of apriori excluding the

[42] John Rawls, *Political Liberalism. Expanded Edition* (New York: Columbia University Press, 2005), 17.

[43] For an analysis of the value of reciprocity and its problems, see Jonkers, "Can Reciprocity be the Principle of a Global Ethics?", 275-294.

[44] Another example of the potential universality of the values of a specific tradition is the idea of human dignity, which was originally a Christian and Western value, but is now universally recognized as the foundation of human rights. See Ricoeur, *Oneself as Another*, 289 and Hans Joas, *The Sacredness of the Person: A New Genealogy of Human Rights* (Washington, DC: Georgetown University Press, 2013).

substantial values of the cultural other under the pretext that they run counter to the universal principles of secular reason. At the same time, this approach also avoids an unqualified upgrading of particular values to universal ones, which would inevitably lead to cultural relativism: whether or not potentially universal values are actually recognized as universal always depends on the outcome of reasonable deliberation.

Conclusion

In this paper, I discussed three goals of ethical education. All of them are focused on teaching and training people to find a reflexive equilibrium between three different dimensions of ethical life, viz. universal moral principles and obligations, ethical traditions as examples of the good life in a specific socio-cultural context, and the indeterminateness and contingencies of ethical praxis. First, people need to familiarize themselves with moral principles and obligations through formal and informal learning. These principles should be applied in a flexible, yet appropriate way to various contexts of action. Hence, to realize ethical praxis, people need training in the virtue of practical wisdom, that is, determining which moral principles are salient in a particular situation and finding ways to apply these principles to contingent situations. However, it has become difficult to find plausible points of ethical orientation in today's multicultural societies. This explains why a very influential way of responding to the multicultural situation, namely to promote a formal, procedural approach to ethical issues has fallen short of expectations. To address this issue, I used Ricoeur's alternative approach to the value of traditions, viz. appreciating them as culturally embedded mediators between universal moral principles and the life-world. This is the second goal of ethical education. Yet, to avoid the looming deadlock of ethical traditionalism and particularism a reasonable debate is necessary to find out if and how the universal potential of particular ethical traditions can be realized. This approach aims at fostering the dialogue among ethical traditions with the help of an enlarged idea of rationality. In today's multicultural society, the third goal of ethical education is, therefore, to train people in self-reflection about the ethical values of their own traditions, as a first step to understanding the values of other traditions as potential universals. Yet most importantly, the analysis of all three goals of ethical

education has shown that its aim is not so much to *inform* the learners about moral theories and insights, but to *(trans)form* their lives.[45]

[45] Pierre Hadot, *La philosophie comme manière de vivre. Entretiens avec Jeannie Carlier et Arnold I. Davidson* (Paris: Albin Michel, 2001), 144.

Out of step?
Reimagining the role of moral philosophers in a changing world

Giovanni SCARAFILE
Associate Professor, University of Pisa

***Abstract**: This paper analyzes the evolving role of moral philosophers in addressing contemporary challenges, arguing for a more dynamic "philosophical posture" that balances theoretical rigor with practical relevance. It examines how philosophers can effectively engage with complex ethical issues while maintaining their core principles, envisioning them as modern argonauts navigating diverse fields from bioethics to artificial intelligence. The discussion critiques the potential reduction of philosophical knowledge to technical expertise and emphasizes the need for actionable insights grounded in community realities. The paper concludes by calling for a renewed philosophical engagement that combines comprehensive understanding with practical wisdom to address our era's existential challenges. Concluding, the paper calls for philosophers to actively engage with societal and ethical issues and reclaim philosophy's foundational vocation. This involves a comprehensive understanding and wisdom to navigate the existential and global challenges of our era.*

1. Philosophy in Step with Time: Reflecting on Method and Adaptation in a Changing World

In this paper, I would like to discuss the possible existence of a "philosophical posture" suitable for the times in which we are living. When I speak of "posture", I am mainly referring to the method used in philosophy. In this regard, some preliminary clarifications are required:

a. Obviously, there are clearly many areas of philosophy (the history of philosophy, moral philosophy, theoretical philosophy—just to name a few). Although it may be too obvious to reiterate this, I do not intend to argue that all these dimensions of doing philosophy should be reduced to a single frame of reference to be adopted in practice;

b. In principle, I mean to refer to the moral philosophy called upon to erect the bridge that connects theories to practices. Now, staying within the metaphor of the bridge, my impression is that not all bridges are identical and that some of them, because of the architecture adopted in construction, are more effective than others. Moreover, it is important to emphasize that while principles and theories are fundamental, moral philosophy, like any other discipline, should never lose sight of human reality in all its complexities. Consequently, the bridges we are talking about should be robust and flexible, able to bear the weight of the toughest ethical issues and adapt to the changing conditions of the real world. Out of metaphor, my discussion is therefore aimed at trying to isolate at least some characteristics of the work that the moral philosopher must do in order to be true to his mission. Regarding the work of the moral philosopher, in my view, this is based on a twofold responsibility: on the one hand, the rigorous analysis of ethical theories, their assumptions and implications; on the other hand, the commitment to communicate the results of these analyses in an accessible and meaningful way. It is only through this balance between intellectual rigor and communicative commitment that the moral philosopher can hope to erect bridges that are not only solid, but also walked concretely by those who are called upon to do so in everyday life;

c. From these two premises, a third follows. Indeed, I do not rule out the possibility of proposing an extension of the posture best suited for the moral philosopher to those militating in other areas of philosophy as well. Such an extension, however, should always be understood as enriching and listening to each other, rather than as a forced alignment or assimilation. Finally, the idea of extending the moral philosopher's approach to other areas of philosophy should not be seen as an attempt to impose a kind of monolithism. On the contrary, it stems from the conviction that interdisciplinary dialogue and respect for the specificities of each field can only enrich philosophical research as a whole, allowing for constructive confrontation and a deeper understanding of the issues being reflected upon.

These three premises can be summarized in a single metaphor that gives my paper its title. It is absolutely necessary to prevent the philosopher from being in the same situation as someone who persists in dancing the twist while house music is playing in the background. This metaphor means that it is imperative that the philosopher, in whatever area of philosophy he or she is operating, adapt to the pace of current times, keeping a cohesive pace with the social, cultural and ethical

dynamics that are shaping our world. That said, this does not mean that the philosopher should necessarily abandon the cardinal principles of his work. On the contrary, the need for adaptation should encourage reconsideration of those principles in light of changed circumstances. This will be somewhat like the dancer who, while remaining true to his skill and mastery, must know how to use his skills in different dance contexts without ever risking finding himself out of step.

2. From Ivory Towers to Grassroots: Philosophy's Adaptive Dance with Contemporary Contexts

Perhaps now more than ever before, the staggering number of crises that modern societies face has increased calls for philosophical intervention. It is not easy to define exactly what is asked of philosophy in these times. In an era defined by disintermediation, where expertise is often dismissed, the common perception of a philosopher's role is unclear. People certainly don't expect philosophers to readily solve the problems before us. Instead, they are probably perceived as unique intellectuals capable of providing a holistic perspective on ongoing debates. This synthesizing ability is what often eludes industry specialists, who are understandably focused on their own areas of expertise. Thus, the philosopher might be seen as a modern argonaut, navigating the boundaries of knowledge, seeking connections, pondering the existence of possibilities, and imagining breakthroughs. Such anticipations of a future to be hastened are not a game for philosophers, quite the contrary. The serious reason why this unique cross-disciplinary expertise is sought by non-philosophers is apparent. In most cases, the issues at stake are severe, and people turn to philosophers when it seems all options have been exhausted. In such scenarios, seeking philosophical insights is not merely an act of faith in others' expertise but rather a last-ditch attempt before surrendering and declaring oneself unable to solve a particular problem.

Under such environmental pressure, the philosopher—and particularly the moral philosopher, which will be the primary focus of this discourse—is required to show great versatility. They might deal with bioethics, ICT ethics, doctor-patient communication, artificial intelligence, or ethics and eco
nomics, to name a few fields.

Given these coordinates, practicing philosophy must balance at least two distinct yet converging demands:

1) The first demand concerns the effectiveness of the philosopher's prescriptions. Their solutions cannot be seen merely as testimonies. The philosopher's guidelines must have prescriptive value; they must embody the reasons that induce listeners not to remain indifferent towards them. Therefore, once a rule has been defined, it must have binding force. Simone Weil aptly noted, "Nothing that is ineffective has value".[1] Such an observation captures the expectations vested in philosophers' work. While they may approach problems from legitimately different perspectives (the notorious scientific-disciplinary sectors!), they cannot refrain from contributing to the needs of the community they operate within;

2) The second demand, arising from the first but of a different order, concerns how the philosopher interacts with contexts where their prescriptions must be implemented. A widespread tendency considers contexts merely as application sites for philosophical theories. This tendency derives from a specific—and now outdated—conception of knowledge, according to a top-down model, which attributed the task of defining knowledge content only to certain institutions or authorities. Both sociology of knowledge and philosophy of science have abundantly signaled how this model is outdated. In a hypothetical map of contemporary knowledge, we should include stakeholders, decentralized network animators, and so-called "totipotent" citizens who, due to their digital devices, feel entitled to opine on any topic. Such grassroots demand for protagonism has been labeled the "demotic turn",[2] referring to the increasing visibility of the "common person" who becomes media content through celebrity culture, reality TV, DIY websites, talk radio, etc. Of course, contexts remain ideal places where philosophical theories are applied, but the transformation of the dynamics within these contexts requires the philosopher—and intellectuals in general—to adopt a more inclusive posture than in the past. These considerations suggest that contexts are not optional and should not be considered as mere receptacles. Interactions within a context, such as those between doctors and patients in a hospital ward, are crucial to define the action rules to be followed there. If these interactions are disregarded, we risk devising decontextualized norms that become impossible to implement, which would undermine the efficacy of the moral norm. To avoid such an outcome, philosophers must adopt a receptive stance towards the demands of the

[1] Weil, Simone. *Quaderni,* Vol. 1. Translated by Saverio Gaeta. Milan: Adelphi, 1994, p. 334.
[2] See Turner, Graeme. 2010. *Ordinary People and the Media: The Demotic Turn*. Theory, Culture & Society. Los Angeles: SAGE.

contexts. Intellectuals must be able to interpret these demands, setting aside any semblance of arrogance, which would be completely unjustified given the changing conditions under which their authority is recognized.

Three questions arise:

1. How can philosophy effectively engage with and contribute to the diverse needs of the modern community without compromising its holistic perspective?
2. Considering the shift from a top-down model of knowledge to a more inclusive one, how can philosophers adapt their approach to ensure their theories remain relevant and implementable?
3. Given the dynamic and decentralized nature of contemporary contexts, how can philosophers maintain their authority and credibility while also embracing a receptive, interactive stance toward these contexts?

3. Philosophy Unbounded: A Call to Reclaim Comprehensiveness in a Fragmented Epoch

Recently, the Italian philosopher Vittorio Possenti in *Considerazioni sulla filosofia italiana contemporanea*[3], suggested that the very versatility of philosophy, recalling the broad spectrum of topics philosophers are called to deal with today, could be the basis for reducing philosophical knowledge to "technical knowledge". This would represent a paradoxical outcome, and therefore, Possenti's thesis should be given due consideration. If confirmed, we would be facing a profound subversion of the tasks traditionally assigned to philosophy.

The broad perspective inherent in philosophy is nothing more than the ability to grasp the connection between various fields of knowledge, or otherwise stated, the unity of knowledge that is today divided among countless disciplines. Mastering this perspective implies the flexibility to break disciplinary boundaries, understanding the reasons that lift our gaze from the specialization that is often considered the only viable option today, thus giving rise to technicism. If, as hypothesized, the

[3] The text *Considerazioni sulla filosofia italiana contemporanea* (Observations on Contemporary Italian Philosophy) is available at the following address: https://shorturl.at/gmqyB.

philosopher's method becomes a "technical method", we are witnessing a short circuit: it's as if philosophy has lost its vocation.

In an attempt to clarify this dynamic, it is useful to listen to what Possenti observes: "Contemporary philosophy [he writes] tends to live in the present and has significantly lowered its horizons, it is content with the everyday, it is reluctant to project, it is rather retrospective. In this sense, it is a time of Lent, especially in the theoretical-metaphysical field. We are witnessing a persistent decline of theoretical thought and an explosion of moral and political philosophy that dominate the field. When a problem of any kind arises, we turn to the expert in ethics". Within such a scenario, Possenti continues, we should ask "whether the philosophy scholar has become an academic expert, a professional holder of technical knowledge".

It is only a matter of observing that a philosophy transformed into technical knowledge represents the subversion of the philosophical enterprise itself, which is called to provide an integrated and profound vision of human reality and the world, to stimulate critical thinking, to interconnect the different fields of knowledge, and to guide ethical and political action, rather than to reduce itself to a mere tool of specialized analysis.

A number of issues flow from this warning. Although I do not have time to develop them, I would at least like to recall them:

1. Since when did philosophy begin to move away from its root, thus losing its holistic view of knowledge and ending up focusing only on specific areas of thought, resulting in its transformation into "technical knowledge"?
2. Is there a way to recover the original vocation of philosophy, so that it does not become a mere technical tool, but returns to its role as a guide for understanding and navigating the world?
3. Could the current perception of philosophy as "technical knowledge" be related to its reluctance to engage with worldviews other than its own, and how might we overcome such self-referentiality? Is Western philosophy willing to question itself and recognize the limits of its own perspective while integrating ideas and concepts from non-Western philosophies?

4. The Solitude of Echoes: The Struggle for Original Thought in Modern Philosophy

In his essay, Possenti refers to the current setting where philosophy operates. He notes the "rethinking of A in the light of B, replacing direct view of reality with theories of theories and cross-references of authors". It's noteworthy how Possenti's analyses were somewhat anticipated in 1973 by Walter Kaufmann, a German scholar teaching at Princeton. Indeed, analyzing the operative modes of his contemporaries, he noted a certain resistance towards incisiveness. They preferred extensive commentary, avoiding the risk of personal exposure or providing solutions from their perspective. To describe such situations, Kaufmann coined the term "decidophobia" of the philosopher. Unable to use his voice, the philosopher becomes a kind of exegetic device, focusing on minutiae details, but losing sight of the whole: "The exegetic thinker avoids standing alone and saying what he thinks, because he might be wrong and wouldn't know what to say if others followed his example and said what they think".[4] This temptation affects all philosophers, regardless of their disciplinary sector.

These words are not meant as an accusation against contemporary philosophy or its interpreters. The main intention is to outline the field, i.e., to inventory the tools available to the philosopher. From the issues discussed so far, there might be some responsibility on the philosophers' part—especially those in academia. Of course, we cannot exclude that there might be responsibilities in a sort of new version of the *Trahison des Clercs*,[5] or shirking their responsibilities. However, if this is the case, I believe it would involve mainly individual responsibilities or small interest groups.

The real point is to consider the objective conditions in which today's academic philosophers are called to do their work. Therefore, the reluctance towards personal exposure, as mentioned by Kaufmann, indirectly receives reinforcement from the dynamics of evaluating the work of those in universities.

Among others, three issues seem particularly significant to me:

---

[4] Kaufmann, Walter Arnold. 1973. *Without Guilt and Justice: From Decidophobia to Autonomy*. New York: Delta, p. 6.

[5] I am referring to the volume by Benda, Julien. *La Trahison des Clercs*. Paris: Grasset, 1927. Benda's book is a critical examination of the role intellectuals played in shaping the political and social landscape of early 20th-century Europe. Benda argues that intellectuals, whom he refers to as "clerks", have betrayed their duty to remain detached from practical affairs and to uphold universal values and truths. Instead, they have become involved in the nationalist and ideological movements of their time, thereby abandoning their commitment to objectivity and universalism. Benda's work is a passionate call for intellectuals to return to their role as custodians of eternal truths and moral values, rather than becoming servants of political or nationalist causes.

1. How does the contemporary academic environment influence the philosophical approach, particularly in terms of personal exposure and decisiveness in expressing personal viewpoints?
2. Given the influence of evaluation dynamics within universities, how can philosophers avoid the trap of "decidophobia" and regain their unique, personal voices?
3. Is there a potential for philosophers, especially those in academia, to reclaim their roles and responsibilities in the face of societal and institutional pressures, or is this challenge primarily an individual struggle?

5. Unscrewing the Mind: Rethinking the Philosophical Toolkit in the Modern Era

In what precedes, I have tried to highlight how we are currently witnessing a radical transformation of the general context in which philosophical activity takes place, which consequently requires a transformation of the forms of philosophizing and the posture of the philosopher. Choosing not to govern such a change means risking proposing practices no longer in line with the needs of our time. In *Mind the Gap*,[6] I had identified a metaphor that I would like to bring forward here.

Every philosopher, I observed, belongs to a specific scientific-disciplinary field. This is not a mere formal and insignificant membership since the tools of one's trade are provided by the sector in which one is situated. Thus, every philosopher, considering their interdisciplinary specialization, carries with them a toolbox. Until now, when faced with the task of screwing any screw, the philosopher merely had to identify the most suitable screwdriver among their tools. The ongoing scenario changes are so radical that they can be compared to a screw that has a screwing system not present among the screwdrivers available to the philosopher. From this perspective, no technique will suffice. It involves rethinking the entire screwing system, a task that the philosopher can perform if they can entirely rethink their vocation, no longer taking for granted the presumed correspondence between their actions and the reality in which they are or should be carried out.

Today, for philosophy, liberating itself from the risk of technicism does not mean giving up being specialized. On the contrary: the need to achieve genuinely

[6] See Scarafile, Giovanni. 2020. *Mind the Gap*. Pisa: ETS.

specialized outcomes requires a methodical pluralism—the opposite of single-minded thinking—and the philosopher must be free from any pre-existing schema in imagining lines of study, action and interaction. It requires a careful and critically reflective understanding of their toolboxes, a continuous review of their methodological approaches, and a renewed commitment to dialogue and collaboration with other disciplines. As we step into the future, the philosopher's role should not be limited to that of an academic expert or holder of technical knowledge. Instead, they should be a bridge-builder, a critical thinker, an explorer of the unknown, and a seeker of wisdom, harnessing the power of their specialization while retaining a holistic view of knowledge.

In the face of the existential and global challenges of our era, we need philosophers who are ready to contribute to the comprehensive understanding of our reality, engaging with societal and ethical questions with courage and conviction. As Possenti observes, the philosophy scholar's evolution cannot be merely a response to changing contexts or an adaptation to academic environments. It must be an active choice to take up the responsibility to understand, interpret, and guide our collective journey through the complex tapestry of human knowledge and experience. In conclusion, whether the philosophy scholar has become an academic expert or a professional holder of technical knowledge is a question best answered by each philosopher individually, as they carefully assess their relationship to their field, to their society, and to the broader quest for wisdom. The challenge of our era is not just the mastery of technical knowledge, but the reclamation and embodiment of philosophy's original vocation: the love of wisdom.

Tetsugaku
International Journal of
the Philosophical Association of Japan
Volume 8, 2024

II Refereed Articles

Revisiting "*propria uoluntas et liberum arbitrium*" in Augustine's *De Libero Arbitrio*

ISHIKAWA Tomoki
The University of Tokyo

***Abstract**: In the first book of De Libero Arbitrio, Augustine concludes that the origin of evil is "its own will and free decision" (propria uoluntas et liberum arbitrium) (1. 11. 21). Despite its importance for the development of Augustine's theory of will, researchers have largely ignored this statement and failed to provide a detailed analysis of it. This paper will, by contrast, take this claim seriously, aiming to reveal its full philosophical significance by focusing on the function of the concept of will.*

In the first section, I begin by exploring the meaning of propria uoluntas et liberum arbitrium by analyzing it word by word and situating it in its proper context. I show that this expression is composed of words with legal nuances and that, taken as a whole, it expresses the view that nothing other than the mind itself subjugates the mind to desire. In the second section, I focus on Augustine's formulation of philosophical inquiry in his earliest treatises before De Libero Arbitrio, showing that it consists of three steps: the purification of desire, the exercise of reason, and the contemplation of the Truth. In the final section, I demonstrate that propria uoluntas et liberum arbitrium forced Augustine to reorganize the model of the soul and of philosophical inquiry that he had forged since his earliest days. A new program of philosophical inquiry, which was formulated later in De Doctrina Christiana, places the purification at the final stage, suggesting that the concept of will was the driving force in transforming how Augustine engages in philosophy.

Introduction

In the first book of *De Libero Arbitrio* (388, hereafter *lib. arb.*[1]), Augustine reaches the conclusion that the origin of evil is "its own will and free decision" (*propria uoluntas et liberum arbitrium*) (1. 11. 21). This phrase represents the decisive moment when, for the first time, Augustine—one of the most influential figures in the

[1] Quotations from Augustine's works in English are based on the translations listed in the bibliography, with some modifications to reflect the structure and nuance of the original Latin text.

historical development of the concept of will[2]—took up the problem of the will in his writings as being of central importance to his philosophy. I would contend, however, that not enough attention has been paid to this phrase or to the first book of *De Libero Arbitrio*. Peter Brown, in his classic biography, depicted Augustine in 386–391 as a young optimistic philosopher who was not yet mature, or even "more Pelagian than Pelagius" (Brown, 1967/2014, p. 141).[3] In the detailed analyses of the first book of *De Libero Arbitrio* by Paul Séjourné and Robert O'Connell, the significance of *propria uoluntas et liberum arbitrium* was hardly noticed,[4] for their aim was to place this work in a larger context (for Séjourné, the spiritual progress or conversion of Augustine; for O'Connell, the influence of Neoplatonism and Stoicism). Even in Simon Harrison's recent monograph, *propria uoluntas et liberum arbitrium* is passed over and quickly replaced by the shorter term *uoluntas* (Harrison, 2006, p. 68f). For many readers, this phrase has been taken as a slogan to quote rather than a claim to be studied. What, then, did this phrase mean to Augustine when he was writing *lib. arb.*? That is, what is the significance of the concept of will as the origin of evil in the first book of this work?

This paper tackles this question by focusing on the function of the concept of will, in the sense of its influence on Augustine's broader philosophical inquiry. More popular themes, such as the precise definition of the will, the problem of theodicy, and its compatibility with determinism, all fall outside the scope of this paper. Instead, its aim is to examine the notion of the will at a microscopic, practical level. Already before the *lib. arb.*, Augustine had undertaken an extensive investigation of the origin of evil and moral psychology. Thus, the introduction of the concept of will ought to

[2] It is no longer possible to regard Augustine as the privileged "inventor" of the concept of will, given the long debate over Augustine's position in the history of the concept of will that was sparked by Albrecht Dihle's *Theory of Will in Classical Antiquity* (1982). Dihle argues that the concept of will cannot be found in Greek and Latin philosophy before Augustine (Dihle, 1982, pp. 123–144). Michael Frede, on the other hand, criticizes Dihle by highlighting the Stoic essence of Augustine's concept of will (Frede, 2012, p. 173f). In my view, they are both too simplistic to grasp the complex nature and background of Augustine's concept of will. As Richard Sorabji suggested, we should view Augustine as one of many philosophers who contributed to the historical process of "clustering" the diverse mental functions into a single concept of "*uoluntas*" (Sorabji, 2000, p. 337).

[3] For a critique of Brown's influence on the studies of *De Libero Arbitrio*, see Madec, 1994, p. 135, and Harrison, 2006, pp. 28–62.

[4] Séjoruné comments that "il [=Augustin] ne trouve que cette explication [=*propria uoluntas et liberum arbitrium*], ambiguë, s'il en fut, qui était celle de saint Ambroise, il y a trois ou quatre ans, à Milan" (Séjourné, 1951, p. 343, n. 1), as if to say it was a mere borrowing from Ambrose. O'Connell (1970) has no mention of *propria uoluntas et liberum arbitrium*.

have brought about some changes in Augustine's philosophy. Otherwise, the claim that the will is the origin of evil would be almost meaningless, amounting to a mere change of label lacking substantial philosophical novelty. My goal here is to assess the scope of this "change", arguing that the concept of will altered Augustine's approach to philosophy.

This paper proceeds as follows. In the first section, I concentrate on explicating the meaning of *propria uoluntas et liberum arbitrium*. A close analysis of the terms "*proprius*" and "*liber*" will confirm that we should understand *propria uoluntas et liberum arbitrium* as a claim that the submission of the mind to desire results from nothing other than the mind itself. The second section temporarily leaves aside *lib. arb.*, shifting the focus to earlier writings, especially *Soliloquia*. Here, I show that the structure of the philosophical inquiry in the Cassiciacum period consists of three steps: the purification of desire, the exercise of reason, and the contemplation of Truth. The final section will consider the function of the concept of will based on the discussion in the first and second sections. I argue that the introduction of the concept of will brought about a crucial change with regard to the possibility of the purification of the soul, and thus fundamentally transformed the young Augustine's approach to philosophical inquiry.

1. The meaning of *propria uoluntas et liberum arbitrium*

Before attempting to interpret the concept of will, it should be stressed that *propria uoluntas et liberum arbitrium* is a tricky phrase that requires careful attention. It is tempting to think that Augustine is referring here to the notion of "free will", since he himself juxtaposes the words "will" (*uoluntas*) and "free" (*liber*). The term "free will" is so familiar to modern readers that it is easy to assume that we can gain some understanding of *propria uoluntas et liberum arbitrium* by replacing it with the former expression. But this would be a mistake for two reasons: First, there is a danger of anachronism, since the concept of free will as we know it is also a historical product—hence we cannot take it for granted that Augustine had the same concept of free will as we do. Second, and more importantly, the meaning of freedom in the language of young Augustine is far removed from what we understand by the notion of "freedom of the will". If we examine the terms *liber* or *liberare* at the time of *lib. arb.*, we find that, in a philosophical or theological context, their use is limited to the ultimate stage

in philosophy, in which the soul is "liberated" from the condition of this world.[5] There is a clear difference between this "freedom as liberation" and our notion of "freedom of the will", since "freedom of the will" is something we have by default, while "freedom as liberation" is a goal to be achieved in the future (or after death, if it is attainable at all).[6] Therefore, the interpretation of the expression *propria uoluntas et liberum arbitrium* represents a major challenge for us, one that should be resolved by examining Augustine's language and context.

(1) *Voluntas* and *arbitrium*

I propose to read *et* appositionally, such that *propria uoluntas* and *liberum arbitrium* indicate a single state of affairs. Let us start with *uoluntas* and *arbitrium*. Some studies try to make a conceptual distinction between them. For instance, Kirwan and Rist claim that *arbitrium* and *uoluntas* differ in meaning, arguing that *liber* exclusively modifies *arbitrium* to create the technical term *liberum arbitrium* and hence that the expression *libera uoluntas* is merely a "lapse from this correct conclusion" (Kirwan, 1989, p. 86f) or "an alternative for *liberum arbitrium uoluntatis*" (Rist, 1994, p. 186, n. 91). There are, however, 34 appearances of *libera uoluntas* in the second and the third books of *lib. arb.*—a number that is hard to explain on Rist and Kirwan's interpretation. Moreover, as Rist and Kirwan would agree, in all these instances *libera uoluntas* means substantially the same thing as *liberum arbitrium*. Therefore, there is no reason to draw a substantial distinction between *uoluntas* and *arbitrium* in *lib. arb*. While there may still be a difference in nuance, this would have little impact on the argument of *lib. arb.*

Nevertheless, my proposal here should not be confused with that of Johannes Brachtendorf, who reaches a seemingly similar conclusion. Brachtendorf applies the distinction between first-order and second-order will to *uoluntates* and *uoluntas*, concluding that *uoluntas* in the singular is equivalent to *liberum arbitrium*. He argues that "Augustine clearly holds a two plane theory of willing as presupposed by compatibilism" (Brachtendorf, 2007, p. 222), alluding to Harry Frankfurt[7] as a representative of the compatibilists. According to Brachtendorf, Augustine makes explicit use of *uoluntates* (in the plural) and *uoluntas* (in the singular) in writings as *Confessiones*, with the former corresponding to "concrete intentions" and the latter to "the will that confirms these intentions" and the *liberum arbitrium* that "decides over

[5] For example, *sol.* 1. 1. 2; 1. 10. 27; *quant*. 34. 78; *De Musica*. 6. 5. 14; *lib. arb.* 2. 13. 37.
[6] Cf. Holte, 1990, p. 78; den Bok, 1994, p. 243.
[7] Especially Frankfurt, 1971.

uoluntates" (*ibid.*, p. 222f). However, we cannot conclude on this basis that Augustine was a two-plane theorist. In *lib. arb.*, Augustine is undoubtedly aware of the second-order will, saying "Is not taking delight in one's own good will, and valuing it as highly as we described, itself the good will?" (*lib. arb.* 1. 13. 28). Nonetheless, Augustine's intention was not to distinguish between these two orders of will, but rather to *conflate* them, for it is precisely this conflation that enables Augustine to declare: "Anyone who wills to live rightly and honorably, if he wills himself to will this rather than transient goods, acquires so great a thing with such ease (*tantam rem tanta facilitate*) that willing itself is nothing other than having what he willed" (*lib. arb.* 1. 13. 29). This ease (*facilitas*) in attaining virtue is rendered possible by the fact that both first-order and second-order will are referred to using a single term "*uoluntas*". While this reasoning might be philosophically dubious, it reveals Augustine's strategy of unifying all mental phenomena under a single conception of the will, which we will explore in section 3 of this paper. Therefore, at least in *De Libero Arbitrio*,[8] we should not limit the term *uoluntas* (in the singular) and *liberum arbitrium* to the second-order will as conceptualized by modern two-plane theory.

(2) *Proprius* and *liber*

In a similar vein, the two adjectives, *proprius* and *liber,* are meant to have the same philosophical significance. In *lib. arb.*, *proprius* is used consistently to mark a contrast with *communis*.[9] The following passage gives a typical account of *proprius*:

> A proper thing (*proprium*) should be understood, just like a private thing (*priuatum*), as that which each one of us is to ourselves and that which he alone senses in himself because it pertains particularly to his own nature; a common thing (*commune*), on the other hand, also just like a public thing (*publicum*), as what is sensed by all who sense it without destroying or transforming it. (*lib. arb.* 2. 7. 19)

Setting aside the direct context of this passage (namely the *sensus communis* shared by many agents), the characteristics of *proprium* are given by descriptions such as "just like a private thing", "he alone in himself", and "pertains to his own nature".

[8] To be fair, my conclusion does not exclude the possibility that the studies I have mentioned might apply to the later works of Augustine. For a well-ordered account of *uoluntas* and *liberum arbitrium,* particularly in late writings such as *De Gratia et Libero Arbitrio*, see den Bok, 1994, pp. 241–246.

[9] Cf. O'Connell, 1968, p. 53.

This use is consistent in Augustine's early language. In *De Quantitate Animae* (388, hereafter *quant.*), Augustine proposes, "Let us not arrogate (*uindicemus*) anything to ourselves as our own (*nobis proprium*), deceived by a desire of empty glory" (*quant.* 34. 78). Given that the verb *uindicare* signifies a claim to one's own property,[10] what is termed *proprium* here has almost the same legal nuances as *priuatum*. In other words, *proprius* is understood in line with legal vocabulary to signify the possession of what belongs to one.

At the same time, despite my assertion that true freedom was an ultimate goal in the language of the young Augustine, we certainly find some different uses of the adjective *liber* in a non-philosophical context.[11] Let us examine two passages from *lib. arb.*:

> The law does not force them to kill but merely leaves it in their power. Hence, they are free not to kill anyone for things they can lose against their will and thus should not love. (*lib. arb.* 1. 5. 12)

> Freedom […] is genuine only if it belongs to happy people who adhere to the eternal law, but for now I am discussing the "freedom" by which people who have no human masters think of themselves as free and which those who want to be set free from their human masters desire. (*lib. arb.* 1. 15. 32)

In the first example, taken from 1. 5. 12, "being free" is conceived in relation to two operations of law, namely "forcing" (*cogere*) and "leaving in power" (*relinquere in potestate*). Freedom here is contrasted with the former and equated with the latter. The second example, taken from 1. 15. 32, clearly shows that Augustine is not talking about genuine freedom (freedom as liberation), but instead about freedom as a legal

[10] " (leg.) to assert one's title to, claim as one's property (what is in the possession of another)" (OLD, p. 2278).

[11] My discussion of the adjective *liber* is addressed to some researchers who have almost ignored it in Augustine's expression of *liber arbitrium*. Sarah Byers argued that, in Augustine's interchange of *uoluntatis arbitrium* with *liberum uoluntatis arbitrium*, "*liberum* seems to be applied for emphasis, the idea being that since assent is by definition a choice between options (to approve or not approve an impression) " (Byers, 2006, p. 184). Ōnishi Yoshitomo, rightly observing that *liberum arbitrium* does not refer to genuine freedom, suggested that *liberum arbitrium* should just be translated as "will (*ishi* 意志)" without adding the adjective "free (*jiyuu* 自由)" (Ōnishi, 2014, p. 115, n. 3). They both went too far to make the word *liber* substantially meaningless. We should, at the very least, not ignore the fact that Augustine did employ the word *liber*.

status, namely, not being a slave (i.e., "having no human masters"). In Roman law, to be a free person meant to possess a legal personality, which made it possible to exert one's due power and responsibility (*potestas*),[12] a feature that connects 1. 15. 32 with 1. 5. 12. Thus, we can conclude that it is still possible to talk about non-genuine freedom, which is not the ultimate goal of philosophy, in a meaningful way when we talk about freedom in a legal context and register. This concept of freedom in *liberum arbitrium* consists of having one's own legal power as a legal personality, without being subject to compulsion by any external power.

(3) The context and the meaning of *propria uoluntas et liberum arbitrium*

These meanings of *proprius* and *liber* fit well with the immediate context of *propria uoluntas et liberum arbitrium* in *lib. arb.* 1. 11. 21. From 1. 9. 19, Augustine and Evodius discuss how the right order of the soul, in which the mind (*mens*), reason (*ratio*), or spirit (*spiritus*) presides over other "irrational motion of mind" (*irrationales animi motus*) (1. 8. 18), can be disrupted and become occupied by vice. According to Augustine, the mind is naturally stronger than desire (*cupiditas*), and virtue is naturally stronger than vice, as determined by the eternal law. Thus, it should be impossible for the desire to force the mind to fall under its control. Likewise, an equally virtuous mind cannot do so, for this evidently goes against the virtue it possesses. Thus, as a logical consequence, Augustine concludes that what causes the mind to become subjugated to desire is nothing other than the mind itself. The expression *propria uoluntas et liberum arbitrium* was meant to capture this conclusion.

The legal implications of *proprius* and *liber* also reflect this context. It is noteworthy that the whole process of reasoning here presupposes the existence of the eternal law that governs everything. This eternal law provides the legal context for the language of the dialogue that leads to *propria uoluntas et liberum arbitrium*. Before introducing the idea of eternal law, Augustine and Evodius discussed whether there is any particular transgression (such as murder, adultery, or sacrilege) that is permissible under secular (or temporal: in contrast to the eternal) law. Note that what is at stake is not whether the act is good or evil, but whether it is justifiable or not (in the latter case, it is a sin). This naturally requires the notion of law as the criterion of justice. Hence,

[12] Even though *potestas* is often translated as "power" in English, it should be noted that it implies the idea of rightness (*ius*) (Berger, 1953, p. 640). Hence, an excessive use of power beyond one's due limits is not considered to be *in potestate*. Cicero's translation of the Stoic notion of *ἐφ' ἡμῖν* (up to us) by *in nostra potestate* (Kahn, 1988, p. 249) also conveys such implications, for *in nostra potestate* implies an exhortation to recognize and remain within our proper limits, namely, within what is in our *due power*.

the concept of law as a context was already anticipated very early on in the dialogue (at least at 1. 3. 6, where Augustine asks: "Is it because the law forbids it?"). The subsequent discussions on concrete cases might, at first glance, merely look like a succession of "ordinary everyday examples" and "simple concepts" (Harrison, 2006, p. 31). While this is the case, it is not the only issue that the long conversation in 1. 3. 6–1. 6. 14 is about. Instead of making philosophical progress, this part of the dialogue carefully prepares the way for the introduction of a legal *context*, which is suitable for employing *proprius* and *liber* in the legal sense, as seen in (2).

Therefore, *propria uoluntas* refers to a *uoluntas* as the mind's own possession, while *liberum arbitrium* means that it is nothing other than the mind that has the power (*potestas*) to do so.

This analysis helps simplify our interpretation of *propria uoluntas et liberum arbitrium.* By this, Augustine means only that "nothing but the mind itself subjugates the mind to desires". The terms *uoluntas* and *arbitrium* thus do not refer to a particular mental realm in the sense of a substance in the human soul but to an act of mind (*mens*) to will (*uelle*) to be under the control of desires. Nor is Augustine attempting to complicate the meaning of freedom by introducing a brand-new notion of free will; instead, he is merely speaking in legal language.[13] This minimalistic interpretation of *propria uoluntas et liberum arbitrium*, I argue, suffices for our reading of *lib. arb.* 1. 11. 21, providing the basis for a further investigation of the function of the concept of the will in *lib. arb.* Before doing so, however, we need to form an idea of what Augustine's philosophy was like prior to *lib. arb.*, as a necessary step to evaluating its effect.

2. Philosophical Inquiry at Cassiciacum

In his earliest days at Cassiciacum, Augustine was concerned as much with how to conduct his philosophical inquiry as with particular philosophical or theological problems. His struggle offers us ample testimony on this matter and even resulted in the creation of an unprecedented genre. *Soliloquia* (hereafter *sol.*), the title Augustine himself coined to refer to a new form of writing (*sol.* 2. 7. 14), thus provides us with

[13] Brian Stock (2010) is the only study, as far as I have seen, that observed the "legal narrative and literary narrative" intertwined in the text of *lib. arb.* (Stock, 2010, p. 149), but Stock does not develop this idea to analyze the effect that the legal narrative had in *lib. arb.* For the influence of Roman law in the formation of the idea of *uoluntas* in Latin language, see Dihle, 1982, pp. 135–143.

useful material from which to extract information about Augustine's earliest philosophical project. In *Soliloquia,* Augustine breaks the process of philosophical inquiry down into three steps:

> Therefore, the soul needs three things: eyes of which it can now make good use, looking at something (*ut aspiciat*), and seeing (*ut uideat*). (*sol.* 1. 6. 12)

These three things are meant to occur in a certain order, as the following exchange indicates: (Reason) "When, therefore, it has healthy eyes, what is left?" — (Augustine) "That it looks at something" (*sol.* 1. 6. 13). Augustine resorts to visual metaphors, describing the philosophical inquiry as the turning upwards of the eyes of the soul because, for most human souls, the light of Truth[14] is too strong to bear. Let us unpack this metaphor and clarify what each step stands for.

In the first step, "eyes of which it can make good use", refers to healthy eyes (*oculi sani*) that are "removed from and purged of the various desires (*cupiditates*) for mortal things" (*sol.* 1. 6. 12). Prior to this first step, the soul is trapped by desires and in an impure state, which the soul itself sustains by its habits, because it is too deeply immersed into this unhealthy condition to extricate itself. Augustine metaphorically calls this nature of desire "a birdlime" (*uiscum*) or "a cage" (*cavea*) (*sol.* 1. 14. 24). This is why Augustine condemns desires: not because they are morally evil, but because the objects of desires are sought "for their own sake" (*propter seipsa*) (*sol.* 1. 11. 19) and hold us in a state of captivity. Desires are, in essence, binding forces imposed on our souls.

The second step, "looking at something", is tricky, mainly because it sounds almost the same as the "seeing" in the final step. But since the condition of the eye of the soul attained in the first step is described as "[that] which it can make good use of", we can infer that "looking at something" entails using the healthy eyes of the soul. Augustine thinks that the health of the soul's eye is, on its own, an insufficient condition for the contemplation of Truth.

> It's dangerous to want to show them [i.e., healthy eyes] what they don't yet have the strength to see. Therefore, they should be exercised first, and their love should be beneficially delayed and nourished. (*sol.* 1. 13. 23)

[14] "Truth" starting with a capital T is meant to stand for *ueritas* (the transcendent principle that makes something true) as distinguished from *uerum* (specific true things or states of affairs), because the English word "truth" can refer to both.

Even healthy eyes require training to be able to see the dangerously brilliant light of the Truth. *Aspicere* thus refers to this process of training, in order to strengthen the eyes of the soul by passing from things that are illuminated by something else to brighter sources of light. Interestingly, Augustine paraphrases *aspicere*'s nominal form, *aspectus*, by "reason" (*ratio*) (*sol.* 1. 6. 13). This odd juxtaposition suggests that what is trained through the process of *aspicere* is reason. Reason is trained by exercising it in the ascent from material things to intelligible objects. The "order of study" in *De Ordine* (hereafter *ord.*) 2. 7. 24–2. 19. 51, which is described as reason's self-exhibition through the study of liberal arts, corresponds precisely to this second step of *aspicere*. The Latin term *aspicere*, composed of *ad* and *specere*, seems to imply a subtle nuance of this process of exercising reason. It consists of successive attempts to look at particular objects with a range of ontological statuses: It is thus a process of "looking at *x*", while leaving the object *x* undefined.[15]

On the other hand, the final step, "seeing", has a single unique object: Truth. While this object can be referred to in different ways (e.g., as Wisdom or God), "seeing" in all cases denotes the ultimate phase in which the soul gains a "perfect and right look at something (*aspectus rectus atque perfectus*)" (*sol.* 1. 6. 13), namely the light of Truth. Augustine does not dedicate many lines to the description of this final stage, and as long as philosophy is a "*love* of wisdom", this final step should remain outside or at least at the margins of this philosophical inquiry.

In brief, the three steps of philosophical inquiry at Cassiciacum, as Augustine expressed them in *Soliloquia*, are: (1) the purification of desire, (2) the exercise of reason, and (3) the contemplation of Truth.

This scheme of philosophical inquiry also applies to another source for the young Augustine's philosophical program, namely the seven-step ascent in *De Quantitate Animae*. In *quant.* 35. 79, Augustine provides three different formulations of the seven steps. The following table shows the correspondence between the different formulations in the order of appearance:

1	vitalization (*animatio*)	from the body (*de corpore*)	beautifully from another thing (*pulchre de alio*)

[15] This is why I translate *aspicere* by "looking *at something*," adding "at something" to the most recent translation, "looking" (Foley, 2020b, p. 31), following Watson's "looking at something" (Watson, 1990, p. 41).

2	sense (*sensus*)	through the body (*per corpus*)	beautifully through another thing (*pulchre per aliud*)
3	art (*ars*)	about the body (*circa corpus*)	beautifully about another thing (*pulchre circa aliud*)
4	virtue (*uirtus*)	toward the soul itself (*ad seipsam*)	beautifully to the beautiful thing (*pulchre ad pulchrum*)
5	tranquility (*tranquilitas*)	in the soul itself (*in seipsa*)	beautifully in the beautiful thing (*pulchre in pulchro*)
6	initiation (*ingressio*)	toward God (*ad Deum*)	beautifully toward the beauty (*pulchre ad pulchritudinem*)
7	contemplation (*contemplatio*)	with God (*apud Deum*)	beautifully with the beauty (*pulchre apud pulchritudinem*)

While the first formulation might give the impression that the seven steps are independent of each other, the following two formulations indicate that they can be subdivided into three groups, i.e., "*corpus/seipsa (anima)/Deus*" or "*aliud/pulchrum/pulchritudo*". The seven steps should thus be understood in terms of a 3-2-2 formation. I will not go into the details of each step here, for it would be enough to describe them in terms of this 3-2-2 structure. The first three steps concern something that is alien to the soul's nature. Under the influence of Neoplatonism, the early Augustine assumed that the soul is essentially immaterial and beautiful (Harrison, 1992, p. 13f), as its nature should be situated among the intelligible realities (*ea quae uere sunt*) (*quant.* 33. 76). Hence, both the body (*corpus*) and another thing (*aliud*) that is not beautiful are external and inferior to the soul itself. By contrast, the fourth and the fifth steps describe how the return to the soul is deepened, while the sixth and the seventh steps illustrate the final stage in which the soul sees God or beauty itself (*pulchritudo*).

The first three steps correspond to the "purification of desire" in *Soliloquia*, as desire is something alien to the soul's rational nature, just like the *corpus* and *aliud* in *quant.* The fourth and fifth steps also correspond to the "exercise of reason", because the education in the liberal arts, which exemplifies the process of "looking at something" (*aspicere*), is, after all, the soul's self-discovery. As the inquiry

concerning the liberal arts proceeds in *De Ordine*, the soul comes to be convinced "either that the soul itself is reason or that reason belongs to it and either that there is nothing better and more powerful in reason than numbers or that reason is nothing other than number" (*ord.* 2. 18. 48). Since every discipline in the liberal arts is more or less created by reason, reason is the subject and the object of the studies in these disciplines. The soul thus arrives at an awareness of what it is (i.e., reason) through the fourth and fifth steps in *quant.* There is no need to specify that the third step in *sol.* is equivalent to the final two steps in *quant.*, since they both point to the contemplation of God (or Truth) as the final achievement of the philosophical ascent.

The three-step model in *Soliloquia* thus gives the simplest illustration of the essence of philosophical inquiry as Augustine understood it at Cassiciacum. This view remained consistent for two years, from 386 to 388, right before Augustine began working on the first book of *De Libero Arbitrio*.

3. Concept of Will and the Re-organization of Philosophical Inquiry

Let us now turn back to the question posed at the beginning of this article: What is the significance of the concept of will as the origin of evil in the first book of *De Libero Arbitrio*? Does its emergence mean that Augustine discovered a new field of mental phenomena, in place of other potential sources of evil, such as lust (*libido*) and desire (*cupiditas*)? Such a view would be wholly mistaken since it conceives of Augustine's philosophical project as resting on something like multiple-choice questions. *Voluntas* is not one candidate for the origin of evil that stands on a par with the others. Instead, will, for Augustine, differs in one notable way from other movements of the soul (*motus animi*): it pertains to the mind (*mens*). The analysis of *propria uoluntas et liberum arbitrium* (*lib. arb.* 1. 11. 21) in section 1 has already made clear that Augustine means by this that what subjugates the mind to the control of desire is nothing other than the mind itself. Thus, the mind has a will. Desire and lust, by contrast, are something external to the mind. Desire is certainly conceived of as a powerful binding force in the early writings of Augustine, as we saw in section 2, but it is not a movement that belongs to the mind itself. Rather, the very fact of its being a "binding force" demonstrates the externality of desire with regard to the mind. As long as desire is designated as the origin of evil and sin, the mind remains pure and detached from all evil and sin.

Such an understanding of the soul and the mind can be vividly observed in Augustine's passion for demonstrating the immortality of the soul in his earliest writings. A series of demonstrations occupies the whole second book of *Soliloquia* and *De Immortalitate Animae*, but apparently these were unsuccessful, as Augustine himself admits in *Retractationes* (1. 5. 1). The essence of the demonstration of the immortality of the soul is that the human soul is a substratum (*subiectum*) or a seat (*sedes*) of Truth, and this seat "can't be dragged away from it by any bodily death" (*sol.* 2. 19. 33). What is called "seat" here corresponds to the mind (*mens*) in *De Libero Arbitrio*. As long as Truth is imperishable, the mind, as a seat of Truth, cannot perish either and should be immune to desire and lust: the mind is the last hope for us humans that something could lead us up toward divine wisdom. The demonstration of the immortality of the soul was thus a demonstration of the possibility for humans to acquire ultimate wisdom. It was a necessary consequence for Augustine that his interlocutor in *Soliloquia* should be "Reason" (*Ratio*), given the status of the reason (which is often juxtaposed with mind—cf. *lib. arb.* 1. 8. 18) as a guide in philosophical inquiry.[16]

Given this understanding of the human soul, it was impossible for Augustine to imagine that the mind spontaneously submits itself to desire. In *De Immortalitate Animae*, he claims that the mind (*animus*)[17] "never wills" (*imm. an.* 13. 20) itself to become body (*corpus*) and does not suffer such a thing "either by its own will or by another forcing it (*nec propria uoluntate nec alio cogente*)" (*imm. an.* 14. 23). Here, Augustine rejects the possibility that mind could be spontaneously inclined toward evil and sin, which is the same conclusion he reaches in *lib. arb.* 1. 11. 21. A minimalistic reading of *lib. arb.* 1. 11. 21 suggests that Augustine merely admits the possibility that the mind *can* will to be a comrade of desires, but not the essential nature of the mind to will to be so. However, it is the possibility that Augustine refused to accept in *imm. an.* 14. 23, marking a big step forward on Augustine's part. It is now clear that this step was made possible by abandoning the conception of the mind as something free of desire and evil. At the same time, Augustine's adoption of the concept of will explains why he gave up on demonstrating the soul's immortality: the concept of will deprived the mind (or reason) of its privileged position as a pure divine

[16] On the somewhat divine status of Reason as an interlocutor in *Soliloquia*, see O'Connell, 1968, p. 122f; Yamada, 1985, pp. 17–21; Cary, 2000, pp. 82–87.

[17] *Mens* and *animus* often overlap in Augustine's usage as he employs *animus* in a specific way when he is concerned with the soul's rational aspect, while *anima* expresses the vivifying force. Cf. O'Daly, 1987, p. 7f; Cary, 2000, p. 178, n. 2.

realm within the human soul, making it impossible, or even meaningless, to demonstrate the immortality of the soul.[18]

More importantly, this shift in his understanding of mind and soul went beyond changing merely one element in his system of thought. Let us recall the second section of this paper: The second step in the program of the philosophical inquiry in *Soliloquia* was the exercise of reason that has already been purified from desire in the first step. We observe here an implicit assumption that reason is, in its own nature, devoid of desire, and desires and lusts are dirt or stains attached externally to reason. This assumption is, however, invalidated by the concept of will in *De Libero Arbitrio*. If the mind can subjugate itself to the control of desire through its own will, purification of the mind is not an easy task of cleansing the external dirt from reason as in the schema presented in *Soliloquia*. It is not about removing the evil will from the mind. As we saw in section 1, Augustine carefully avoids substantiating the "evil will". What Augustine calls the "evil will" (*mala uoluntas*) is merely the wrong use of the will by the soul. Evil will is not a substantiated realm in the mind but an act of willing of the mind. Thus the evil will cannot be removed from the mind unless the mind itself is completely transformed. In this way, the discovery of the concept of will forced Augustine to reframe or reorganize his earliest program of philosophical inquiry.

The reorganization of philosophical inquiry was not like total destruction and reconstruction from scratch. Augustine's task was to place the remains of his past philosophical efforts and achievements in a new context and elaborate a new philosophical construct. After concluding that *propria uoluntas et liberum arbitrium* is the origin of evil, Augustine undertakes to ground the entire system of moral psychology solely in the concept of will. Augustine starts by asking Evodius whether they have some sort of will (*sitne aliqua nobis uoluntas?*) (*lib. arb.* 1. 12. 25), a point which serves as the starting point or basis for the following discussions (Harrison, 2006, p. 100f). After confirming that they do possess a will, Augustine goes on to define every aspect of his philosophical inquiry in light of this concept of will: virtue, the happy life, eternal law—they all depend on the good will (*bona uoluntas*) "by

[18] Phillip Cary also observes the same "change of mind" that led him to abandon the divinity and immortality of the soul but at the same time claims that this took place in *De Moribus Ecclesiae Catholicae* (388–390), a polemical work written slightly after *De Libero Arbitrio*. Cary's point is that, after leaving Milan, "Augustine realizes that Catholic doctrine does not allow him to conceive of the soul as divine and immutable" (Cary, 2000, p. 112). Even though it is true that the denial of the divinity of the soul fits well with Catholic doctrine, Cary does not seem to explain how it was possible for Augustine to bring himself into conformity with Catholic doctrine. If it lacked an inner motivation, Augustine's "change of mind" would have been mere acquiescence to an externally imposed dogmatic demand.

which we seek to live rightly and honorably and to attain the highest wisdom" (*lib. arb.* 1. 12. 25). Desire and misery, by contrast, can be explained as the lack of a good will. They represent wrong uses of the will, meaning that the will itself is exempt from accusation. Hence, the concept of will, which encompasses every aspect of philosophical inquiry, appears as a promising principle from which to embark on the task of reorganization.

However, my own view is that this task was not fully completed in *De Libero Arbitrio*, because a more difficult problem, that of justifying God's creation of human beings with the will, presented itself to Augustine, a problem that took Augustine seven years to finish *De Libero Arbitrio*. Nor do I mean to argue that Augustine underwent a dramatic change of mind just as Peter Brown depicted in the chapter titled "The Lost Future" of his classic biography (Brown, 1967/2014, pp. 139–150).[19] The introduction of the concept of will was just the beginning or a trigger of the reorganization of his philosophical inquiry. There is, however, evidence that Augustine eventually went in a direction in which this potentiality could be realized. Let us compare the two versions of the "seven steps of inquiry" in *De Quantitate Animae* and *De Doctrina Christiana* (396–426, hereafter *doctr.*), which were published, respectively, before and after *De Libero Arbitrio* (388).

	***quant.* 33. 70–35. 79**	***doctr.* 2. 7. 9–11**
1	*animatio*, *de corpore*, *pulchre de alio*	fear (*timor*)
2	*sensus*, *per corpus*, *pulchre per aliud*	piety (*pietas*)
3	*ars*, *circa corpus*, *pulchre circa aliud*	knowledge (*scientia*)
4	*uirtus*, *ad seipsam*, *pulchre ad pulchrum*	bravery (*fortitudo*)
5	*tranquilitas*, *in seipsa*, *pulchre in pulchro*	counsel of compassion (*consilium misericordiae*)
6	*ingressio*, *ad Deum*, *pulchre ad pulchritudinem*	understanding (*intellectus*)
7	*contemplatio*, *apud Deum*, *pulchre apud pulchritudinem*	wisdom (*sapientia*)

[19] To be fair, Brown later regrets that he overly emphasized Augustine's change of mind at the end of the new edition of this book (Brown, 1967/2014, p. 490).

The respective programs for the ascent of the soul outlined in these two texts have the same number of steps, yet their content differs a great deal. The first difference is that, in *doctr.*, there is no equivalent to the first three steps in *quant.*, namely vitalization, sense, and art, which correspond to the souls of plants, animals, and humans and are described without any moral implications.[20] In *quant.*, it is only after the fourth step (virtue) that the moral or religious aspect comes to the fore, while in *doctr.*, the moral aspect is apparent from the very first step (fear of God). Thus, steps 1–3 in *doctr.* seem to correspond to the fourth step (virtue) in *quant.* (Kato, 1991, p. 242). The second difference is that the purification of desire is achieved in the fourth step in *quant.*,[21] while it is not attained in *doctr.* until the sixth step (understanding). If we combine these two observations, we see that the fourth step in *quant.* overlaps with the first six steps in *doctr.* This odd correspondence can be explained by the emergence of the concept of will in Augustine, which took place between the writing of these two texts. As we have seen, the expression *propria uoluntas et liberum arbitrium* (*lib. arb.* 1. 11. 21) involves a denial of the assumption that the mind can be free of desire. Therefore, the purification of the soul is now the task that must be confronted by the end of the philosophical inquiry. At the same time, the first three steps in *quant.* become meaningless as the object of purification is no longer the desires that are external to the mind and thus that can be removed in only three steps.

Thus, following *lib. arb.*, philosophical inquiry of Augustine went through a significant transformation. The problem of will is not merely one of the questions addressed by philosophical inquiry, but the driving force that changed how Augustine engages in philosophy.

Conclusion

The question I raised at the beginning of this paper was, "What is the function of the concept of will in *De Libero Arbitrio*?" This question was addressed in several steps. I began by analyzing the key phrase from *lib. arb.*, "its own will and free decision" (*propria uoluntas et liberum arbitrium*), concluding that this should be understood in

[20] The third step (art) is a "wealth common to both the goods and the evils" (*quant.* 33. 72), which implies that art is indifferent to morality.

[21] The fourth step is described as "achieving (*efficere*) purity" and the fifth as "keeping (*tenere*) purity" (*quant.* 33. 74).

terms of the spontaneous submission of the mind to the desire. Having elucidated the structure of philosophical inquiry in Augustine's Cassiciacum writings in the second section, I then showed in the third section that this conclusion had two consequences for Augustine's philosophy: The first was a radical change in the theory of the soul in relation to desire and mind. The second was the reorganization of the program of philosophical inquiry, which presupposed the possibility of the purification of the soul at an early stage, followed by a new schema that puts the purification at the final stage of the inquiry. These effects and functions of the concept of the will, I believe, will make it more acceptable to those who do not share the same theological motivation as Augustine and his followers.

The significance of the concept of will can be placed in a broader context. Augustine's trust in the mind or reason prior to *De Libero Arbitrio*, which was bound up with the immortality of the soul, was rooted in an understanding of humans as something mortal—a view that had been dominant since classical antiquity. The concept of will that Augustine presented in *lib. arb.*, on the other hand, opens up another possibility of human finitude. Even the best parts of humans—i.e., mind and reason—are bound to commit sin and evil. We are finite beings, however long we live, however talented we might be, and whatever efforts we make. We are finite beings as long as we remain human—not because we will eventually die, but because the will that pervades our souls is finite. Such a view of humanity cannot be found in the early works of Augustine, who initially thought that wisdom would be attained either in this life or after death (cf. *sol.* 1. 12. 20). Nor can it be found in the Stoic philosophers who supposed that it is impossible for humans to be wise (*sapiens*) in practice, but possible in theory. But if this is the case, how and for what purpose do we live and philosophize, bearing this heavy finitude of our being? Augustine's later philosophy, which was developed in light of the doctrine of original sin and the divine Trinity, can be read as a response to or struggle with this question. The nuanced remark by Evodius by the end of the first book of *De Libero Arbitrio*— "I yield to your will, and quite readily join mine to yours in judgment and in prayer" (*Cedo uoluntati tuae et ei meam iudicio et uoto libentissime adiungo*)" (*lib. arb.* 1. 16. 35)—seems to suggest how central the will (*uoluntas*) has become to their philosophical inquiry.

References

1. Augustine's Works: Critical Editions and Translations

Augustine's works are cited based on the latest editions from either *Corpus Christianorum Series Latina (=CCSL)* or *Corpus Scriptorum Ecclesiasticorum Latinorum* (=CSEL).

Colleran, J. (trans.). (1964). *St. Augustine: The Greatness of the Soul, The Teacher*. Westminster, MD: The Newman Press.

King, P. (ed., trans.). (2010). *Augustine: On the Free Choice of the Will, On Grace and Free Choice, and Other Writings (Cambridge Texts in the History of Philosophy)*. Cambridge: Cambridge University Press.

Foley, M. P. (2020a). *On Order: St. Augustine's Cassiciacum Dialogues, Volume 3*. New Haven: Yale University Press.

Foley, M. P. (2020b). *Soliloquies: St. Augustine's Cassiciacum Dialogues, Volume 4*. New Haven: Yale University Press.

Watson, G. (1990). *Soliloquies and Immortality of the Soul: With an Introduction, Translation and Commentary*. Warminster: Aris & Phillips.

2. Secondary Literature

Berger, A. (1953). Encyclopedic Dictionary of Roman Law. *Transactions of the American Philosophical Society*, 43(2), 333–809.

Brachtendorf, J. (2007). Augustine's Notion of Freedom: Deterministic, Libertarian, or Compatibilistic? *Augustinian Studies*, 38(1), 219–231.

Brown, P. (2013). *Augustine of Hippo: A Biography (A New Edition with an Epilogue)*. Berkley and Los Angeles: University of California Press (First edition 1967).

Byers, S. (2006). The Meaning of *Voluntas* in Augustine. *Augustinian Studies*, 37(2), 171–189.

Cary, P. (2000). *Augustine's Invention of the Inner Self*. Oxford: Oxford University Press.

den Bok, N. W. (1994). Freedom of the Will: A Systematic and Biographical Sounding of Augustine's Thoughts on Human Willing. *Augustiniana*, 44(3/4), 237–270.

Dihle, A. (1982). *The Theory of Will in Classical Antiquity*. Berkeley and Los Angeles: University of California Press.

Frankfurt, H. G. (1971). Freedom of the Will and the Concept of a Person. *The Journal of Philosophy*, 68(1), 5–20.

Frede, M., ed. Long, A. A. (2012). *A Free Will: Origins of the Notion in Ancient Thought*. Berkeley, Los Angeles: University of California Press.

Glare, P. G. W. (ed.). (2012). *Oxford Latin Dictionary. Second Edition.* Oxford: Oxford University Press. (=OLD).

Harrison, C. (1992). *Beauty and Revelation in the Thought of Saint Augustine*. Oxford: Clarendon Press.

Harrison, S. (2006). *Augustine's Way into the Will: The Theological and Philosophical Significance of* De Libero Arbitrio. Oxford: Oxford University Press.

Holte, R. (1990). St. Augustine on Free Will. In Settimana Agostiniana Pavese (ed.). (1990). *Lectio Augustini* «De Libero Arbitrio» *di Agostino D'Ippona* (pp. 67–84). Palermo: Edizioni "Augustinus".

Kahn, C. H. (1988). Discovering the Will: From Aristotle to Augustine. In Dillon, J. M. & Long, A. A. (eds.). *The Question of "Eclecticism": Studies in Later Greek Philosophy* (pp. 234–259). Berkley and Los Angeles: University of California Press.

Kato, T [加藤武]. (1991). *Augusutinusu no gengo ron* [アウグスティヌスの言語論, *Augustine's Theory of Language*]. Tokyo: Sōbunsha [創文社].

Kirwan, C. (1989). *Augustine*. London: Routledge.

Madec, G. (1994). *Petites Études Augustiniennes*. Paris: Institut d'Études Augustiniennes.

O'Connell, R. J. (1968). *St. Augustine's Early Theory of Man, A.D. 386–391*. Cambridge, MA: Harvard University Press.

O'Connell, R. J. (1970). De Libero Arbitrio I: Stoicism Revisited. *Augustinian Studies*, 1, 49–68.

O'Daly, G. (1987). *Augustine's Philosophy of Mind*. London: Duckworth.

Ōnishi, Y. [大西克智]. (2014). *Ishi to jiyuu: hitotsu no keifugaku* [意志と自由―一つの系譜学, *Will and Freedom: a Genealogy*]. Tokyo: Chisen Shokan [知泉書館].

Rist, J. M. (1994). *Augustine: Ancient Thought Baptized*. Cambridge: Cambridge University Press.

Séjourné, P. (1951). Les conversions de saint Augustin d'après le *De libero arbitrio*, livre I (suite). *Revue des Sciences Religieuses*, 25(4), 333–363.

Sorabji, R. (2000). *Emotion and Peace of Mind*. Oxford: Oxford University Press.

Stock, B. (2010). *Augustine's Inner Dialogue: The Philosophical Soliloquy in Late Antiquity*. Cambridge: Cambridge University Press.

Yamada, A. [山田晶]. (1985). *Sorirokuia* ni okeru *Ratio* no imi ni tsuite—*Soliloq.* I, c. I, n. I— [『ソリロクィア』における《ラチオ》の意味について―*Soliloq.*

I, c. I, n. I—, On the meaning of *Ratio* in *Soliloquia—Soliloq.* I, c. I, n. I—], Chūsei Tetsugaku Kai [中世哲学会]. *Chūsei Shisou Kenkyū* [中世思想研究], *27,* 1‐23.

Deux schémas de mouvement chez Henri Bergson: autour de sa théorie de la mémoire

AMANO Emiri
Specially Appointed Assistant Professor, Takasaki City University of Economics

***Résumé**: La présente étude vise à clarifier le rapport entre la théorie bergsonienne de la reconnaissance dans Matière et mémoire (1896) et le « schéma dynamique », terme qu'utilise Bergson dans son article de 1902 intitulé « L'effort intellectuel ». Le « schéma dynamique » désigne une représentation abstraite qui intervient dans l'effort intellectuel, et à partir de laquelle le mouvement de l'esprit se développe. Ce concept joue un rôle majeur dans l'élaboration de la définition bergsonienne de la causalité comme « un passage graduel du moins réalisé au plus réalisé ». L'apparition de cette nouvelle conception de la causalité est d'autant plus importante dans le développement de la philosophie bergsonienne qu'elle est liée à la théorie de la vie de L'évolution créatrice. Il est vrai que le concept de schéma dynamique, avec la notion de la causalité qui en résulte, figure pour la première fois dans « L'effort intellectuel » ; il n'apparaît pas encore dans Matière et mémoire, où Bergson emploie plutôt le terme « schème moteur » pour expliquer la reconnaissance. Néanmoins, les deux ouvrages s'appuient essentiellement sur la même théorie de la mémoire. Il doit donc y avoir une continuité entre Matière et mémoire et « L'effort intellectuel », mais il est difficile d'en déterminer la nature. Certains chercheurs insistent sur le parallélisme entre les deux textes ; d'autres soulignent plutôt leur hétérogénéité. En comparant « L'effort intellectuel » avec Matière et mémoire, nous montrons que la vérité est entre ces deux points de vue : l'idée qui portera le nom de schéma dynamique existe déjà dans la théorie de la reconnaissance de Matière et mémoire, mais l'existence de cette idée naissante n'est appréciable que difficilement à cause de la conception centripète du sentiment de l'effort que Bergson emprunte à W. James.*

Introduction

La présente étude[1] vise à clarifier le rapport entre la théorie bergsonienne de la reconnaissance dans *Matière et mémoire* et le « schéma dynamique », terme qu'utilise Bergson dans son article de 1902 intitulé « L'effort intellectuel ».

Le « schéma dynamique » désigne une représentation abstraite d'idées et d'images, représentation qui intervient dans l'effort intellectuel (par exemple la compréhension d'une phrase), et à partir de laquelle le « mouvement de l'esprit »[2] se développe. Ce concept joue un rôle majeur dans l'élaboration de la définition bergsonienne de la causalité comme « un passage graduel du moins réalisé au plus réalisé ».[3] L'apparition de cette nouvelle conception de la causalité est d'autant plus importante dans le développement de la philosophie bergsonienne qu'elle est liée à la théorie de la vie de *L'évolution créatrice*.

Il est vrai que le concept de schéma dynamique, avec la notion de la causalité qui en résulte, apparaît pour la première fois dans « L'effort intellectuel » et qu'il n'est pas encore présent dans *Matière et mémoire*, où Bergson utilise plutôt le terme « schème moteur » pour expliquer la reconnaissance. Néanmoins, les deux ouvrages s'appuient essentiellement sur la même théorie de la mémoire. Il doit donc y avoir une continuité entre *Matière et mémoire* et « L'effort intellectuel ». La difficulté est cependant d'en déterminer la nature. Certains chercheurs insistent sur le parallélisme entre les deux textes ; d'autres soulignent plutôt leur hétérogénéité. Le commentaire de Camille Riquier dans *L'énergie spirituelle* montre très bien cette difficulté : après avoir nettement séparé les deux textes sur la base de l'hétérogénéité entre le schème moteur (*Matière et mémoire*) et le schéma dynamique (« L'effort intellectuel »), Riquier ne manque pas de suggérer que Bergson pourrait avoir « rencontré la seconde notion par approfondissement de la première ».[4] Nous croyons en effet que la vérité est entre les deux : l'idée qui portera le nom de schéma dynamique existe déjà dans la théorie de la reconnaissance de *Matière et mémoire*, mais Bergson n'était pas en mesure de s'y engager pleinement à cause de sa conception du sentiment de l'effort.

Notre discussion se déroulera selon les étapes suivantes : la section 1 examinera le schéma dynamique dans l'article « L'effort intellectuel » ; la section 2

[1] Je présente ma sincère gratitude à Messieur Takafumi Ishiwatari et Madame Mathilde Tahar, qui ont attentivement lu et corrigé mon manuscrit.

[2] ES, p. 189. Les ouvrages de Bergson seront désignés par les abréviations suivantes :
« MM » : *Matière et mémoire* (1896). Paris, PUF, 2007.
« ES » : *L'énergie spirituelle* (1919), Paris, PUF, 2009.
« M » : *Mélange*, Paris, PUF, 1972.

[3] M, p. 550.

[4] ES, p. 335.

s'occupera ensuite de la théorie de la reconnaissance de *Matière et mémoire* afin d'y repérer, au-delà du schème moteur, le concept de schéma dynamique qui existe pour ainsi dire en germe ; enfin, la section 3 essaiera de montrer pourquoi ce concept naissant a dû rester à l'arrière-plan par rapport au schème moteur.

1. Schéma dynamique et causalité dans « L'effort intellectuel »

Suivant la discussion élaborée dans « L'effort intellectuel », nous expliquerons d'abord brièvement ce qu'est le « schéma dynamique » (1.1), avant d'examiner la distinction que donne Bergson entre l'« intellection [...] automatique »[5] qui n'exige aucune intervention du schéma dynamique (1.2) et « l'intellection vraie »[6] où intervient l'effort, et par conséquent le schéma dynamique (1.3). Finalement, nous constaterons que Bergson définit, à partir de la notion du schéma dynamique, ce qu'est la « causalité réelle ».[7]

1.1 Le principe de « L'effort intellectuel »

La question capitale de l'article de 1902 consiste à savoir quel est le « jeu des représentations » dans l'effort d'attention intellectuelle et volontaire (c'est-à-dire l'effort intellectuel).[8] La psychologie de l'attention de cette époque-là, représentée par Ribot, étudie principalement les mouvements corporels qui accompagnent l'acte d'attention, et non les mouvements des représentations, c'est-à-dire le « contenu » de l'attention.[9] Bergson cherche au contraire à trouver la caractéristique propre à l'effort intellectuel « dans la représentation elle-même ».[10]

Le schéma dynamique est une représentation que Bergson conçoit pour caractériser le jeu des représentations dans l'effort intellectuel : lorsqu'on s'engage dans une activité intellectuelle, il y a un « schéma », c'est-à-dire une représentation abstraite et condensé d'images et d'idées qui sert de source d'inspiration pour la formation ultérieure des images concrètes.

---

[5] ES, p. 168.
[6] ES, p. 169.
[7] M, p. 550.
[8] Cf. ES, p. 153-4.
[9] Cf. *ibid*.
[10] Cf. *ibid.*

Bergson distingue ainsi trois formes de l'effort intellectuel : « l'effort [...] de rappel »,[11] « l'effort de l'intellection »[12] et « l'effort d'invention ».[13] Il conviendra ici d'examiner l'effort de l'intellection puisque celui-ci est directement lié à la théorie de la reconnaissance du 2e chapitre de *Matière et mémoire*.

1.2 L'intellection automatique qui exclut l'effort

En ce qui concerne l'intellection, Bergson oppose l'« intellection [...] automatique » sans effort à l'intellection « qui [...] est généralement observable là où il [=l'effort] se produit», ou « l'intellection vraie ».[14] Il va sans dire que le schéma dynamique ne concerne que le deuxième genre d'intellection.

L'intellection dite automatique consiste à « esquisser machinalement, quand on le [=l'objet] perçoit, l'action que l'habitude a associée à cette perception ».[15] L'intellection de ce genre se retrouve, par exemple, dans la reconnaissance d'un objet usuel. Dès que nous percevons un objet, nous esquissons mécaniquement les actions à travers lesquelles nous pourrions l'utiliser. Il en est de même pour la conversation courante où « la réponse [succède] à la question sans que l'intelligence s'intéresse au sens de l'une ou de l'autre ».[16]

L'intellection de ce genre se fait surtout à travers la coordination sensori-motrice habituelle, sans sortir de ce niveau de sensations et de mouvements.

1. 3 L'intellection vraie qui demande l'effort

Dans intellection vraie, par contraste, nous nous situons d'emblée au niveau plus élevé de la pensée abstraite pour descendre ensuite vers les images perceptives concrètes.

Le premier exemple typique donné par Bergson est l'opération mathématique.[17] Lorsque nous suivons un calcul, nous le refaire pour notre propre compte.[18] Ce processus est, selon Bergson, comme suit :

---

[11] ES, p. 155.
[12] ES, p. 167.
[13] ES, p. 174.
[14] Cf. p. ES, 167-9.
[15] ES, p. 168.
[16] *Ibid.*
[17] Cf. ES, p. 169-170.
[18] Cf. *ibid.*

> « Le long de la démonstration vue ou entendue nous avons cueilli quelques suggestions, choisi des points de repère. De ces images visuelles ou auditives *nous avons sauté à des représentations abstraites de relation*. Partant alors de ces représentations, nous les déroulons en mots imaginés qui viennent rejoindre et recouvrir les mots lus ou entendus».[19]

Regardons de plus près l'action du saut et de la descente.

Tout d'abord, il faut noter qu'il y a un *saut*, c'est-à-dire un passage des sensations et des mouvements « *à des représentations abstraites de relation* » qui sont le schéma dynamique. Ce « saut » ne se fait évidemment pas à l'aveugle, mais d'après les « suggestions » de ce que nous voyons ou entendons. Et cette « relation » est celle que prennent, au niveau abstrait des idées, les suggestions ou les repères cueillis au niveau de sensation. Nous sautons à partir des sensations et des mouvements, qui se déroulent de manière automatique (c'est-à-dire évidente), pour établir « hypothétiquement » [20] des relations entre eux. En d'autres termes, nous reconstruisons activement le « sens »[21] de ce que nous voyons et entendons. Ce saut ou cette construction hypothétique de relation est décisif en tant que point de départ de l'intellection vraie.

Passons ensuite à la descente. Dans l'intellection vraie, la « relation » que nous établissons d'abord est nécessairement vague et inachevée, mais elle se développe au fur et à mesure en « mots imaginés », qui redescendent finalement au niveau des sensations pour se superposer aux « mots lus ou entendus ». Les « mots lus ou entendus », qui nous ont servi de points de repère pour le saut, peuvent, à cet égard, être qualifiés également de « cadre »[22] dans lequel s'insèrent les mots imaginés. En plus, cette superposition réussite, qui s'achève chaque fois, confirme que notre hypothèse de départ était dans la bonne direction.

Bergson complète son argumentation par un autre exemple. Il évoque le cas où « nous conversons dans une langue étrangère que nous connaissons imparfaitement ».[23] Lorsque nous essayons de comprendre un énoncé oral d'une langue étrangère, « *nous nous plaçons d'emblée dans un ordre de représentations plus ou moins abstraites,* suggéré par ce que notre oreille entend » et nous adoptons alors un « ton

[19] *Ibid*., nous soulignons.
[20] ES, p. 172.
[21] ES, p. 171.
[22] ES, p. 170.
[23] Cf. ES, p. 171-2.

intellectuel ».[24] C'est là que se trouve le schéma dynamique. Et « une fois adopté ce ton intellectuel, nous marchons, avec le sens conçu, à la rencontre des sons perçus ».[25] Ainsi, nous descendons, également dans ce cas, du schéma au niveau sensoriel.

Sur la base de ces exemples, Bergson résume l'effort de l'intellection de manière suivante : « L'effort intellectuel pour interpréter, comprendre, faire attention, est [...] un mouvement du « schéma dynamique » dans la direction de l'image qui le développe » et « *le sentiment de l'effort d'intellection se produit sur le trajet du schéma à l'image* ».[26]

1.4 Le schéma dynamique et la causalité

Ce qui est assez remarquable dans « L'effort intellectuel », c'est que Bergson conçoit la causalité comme un mouvement ou une « force » s'opérant ailleurs que dans l'ordre sensoriel auquel appartient, comme nous avons vu ci-dessus, l'intellection automatique. Selon lui, il y a dans l'effort intellectuel une force *sui generis* qui est la « causalité réelle » :

> « Quant à savoir comment elles [=les forces] travaillent, c'est une question qui n'est pas du ressort de la psychologie toute seule : elle se rattache au problème général et métaphysique de la causalité. [...] Cette activité, qui est la *causalité réelle*, consiste dans un passage graduel du moins réalisé au plus réalisé, de l'intensif à l'extensif, d'un état d'implication réciproque des parties à un état de juxtaposition de ces parties les unes aux autres, enfin du schéma à l'image. Or, l'effort intellectuel tel que nous l'avons défini, n'est pas autre chose. En ce sens, il nous présenterait la relation causale à l'état pur».[27]

Ce que Bergson entend ici par « métaphysique », c'est une entreprise dont le but est de ressaisir cette « force » ou « activité » à sa « source » même.[28] La psychologie ne

[24] Cf. ES, p. 171, nous soulignons.
[25] Cf. ES, p. 171-2.
[26] Cf. ES, p. 173-4, soulignage par Bergson.
[27] M, p. 549-550, nous soulignons. L'article « L'effort intellectuel » est paru dans *La Revue philosophique* en 1902. Et en 1919, ce texte est repris mais remanié dans l'*Energie spirituelle*, en « profitant par ricochet des acquis de L'*Évolution créatrice* » (Riquier, Camille, *Archéologie de Bergson. Temps et Métaphysique*, Paris, PUF, 2006, p. 373). La version citée ici est celle de 1902.
[28] Cf. M, p. 698 ; Riquier, *op. cit*., p. 372.

peut pas aller seule jusqu'à la source même.[29] Bref, la « causalité réelle » consiste dans « un passage graduel » du « moins réalisé » en tant que source même au « plus réalisé ». Et nous retrouvons cette causalité par excellence dans l'effort intellectuel, c'est-à-dire dans le passage graduel des « représentations abstraites de relation », à l'image matérielle.

Bergson indique que la descente des « représentations abstraites de relation » vers les sensations se trouve non seulement dans l'effort intellectuel proprement dit, mais aussi dans l'« effort corporel ».[30] « Pour apprendre tout seuls un exercice complexe, tel que la danse », nous commençons par nous servir de « dessin de *relations*, surtout temporelles, entre les parties successives du mouvement à exécuter » et ensuite nous le « converti[ssons] en sensations motrices réelles et par conséquent en mouvements exécutés ».[31] Et Bergson reconnait que cette représentation « ressemble beaucoup à ce que nous appelions un schéma ».[32] Nous pouvons dire de là que la « causalité réelle » se trouve aussi dans l'« effort corporel ».

Nous retrouvons également dans les cours au Collège de France 1900-1 sur l'« Idée de cause » (deux ans avant la première publication de « L'effort intellectuel ») la même conception de la causalité. C'est, dit Bergson, dans « un progrès vers la matérialité, vers la réalisation » que l'on trouve la « quintessence de la causalité ».[33]

Bergson n'utilise cette formule métaphysique de la causalité qu'au début du 20e siècle. En effet, lorsque « L'effort intellectuel » est repris en 1919 dans *L'énergie spirituelle*, il reformule l'activité de la « causalité réelle » que nous avons citée ci-dessus en l'opération « même de la vie ».[34] Bien des chercheurs, en remarquant ce remaniement, sont convaincus que c'est cette conception de la causalité qui a conduit Bergson à la théorie de la vie de *L'évolution créatrice*.[35] Est-il possible, alors, de repérer une idée analogue dans le texte de *Matière et mémoire* ?

Gilles Deleuze reconnaît le même « saut » que nous venons de voir dans l'analyse de la compréhension du 2e chapitre de *Matière et mémoire*. Selon lui, il y a là « le saut dans l'ontologie » pour « sortir de la psychologie ».[36] Nous n'insistons pas

[29] Cf. Riquier, *op.cit*., p. 376.
[30] ES, p. 178.
[31] Cf. ES, 178-9, soulignage par Bergson.
[32] *Ibid.*
[33] Cf. Henri, Bergson, *L'idée de Temps. Cours Au Collège de France 1901-1902*, Paris, PUF, 2017, p. 199.
[34] ES, p. 190.
[35] Cf. par exemple, F. Worms, *Introduction à* Matière et mémoire *de Bergson*, Paris, PUF, 1998, p. 184.
[36] Cf. G. Deleuze, *Le Bergsonisme*, Paris, PUF, 2004, p. 52.

sur l'ontologie de la différence, qui est un concept proprement deleuzien. Ce qui importe dans notre contexte, c'est que Deleuze reconnaît une conception non-psychologique à l'intérieur du 2e chapitre de *Matière et mémoire.*

En revanche, Camille Riquier considère le schéma dynamique ainsi que la « causalité réelle » comme l'élément décisif qui distingue par rapport au texte de 1896 les textes au début du 20e siècle, y compris « L'effort intellectuel ». Il souligne que, dans *Matière et mémoire*, la « vie de l'esprit » ne désigne qu'une vie subsistant grâce à la chaleur que le corps donne à l'esprit,[37] et il en conclut que Bergson n'y reconnaît aucune place pour la « vie de l'esprit » au sens propre.[38] « Bref, *Matière et mémoire* appelle comme son *complément*, en même temps qu'il le laisse en friche, le problème de l'activité causale proprement dite, lequel va devenir de plus en plus aigu à mesure qu'on s'approche de L'*évolution créatrice* ... ».[39] Autrement dit, cette conception métaphysique de la causalité ne se trouve pas dans *Matière et mémoire*. Selon Riquier, « il manque encore à Bergson d'entériner la distance prise avec la psychologie ».[40]

Étant donné que la remarque de Riquier concernant la dépendance de l'esprit au corps s'appuie sur une lecture rigoureuse du texte bergsonien, il nous faut dire que le rapprochement deleuzien de *Matière et mémoire* et de « L'effort intellectuel » apparaît quelque peu naïf en ce qu'il néglige cette dépendance. Néanmoins, à la différence de Riquier, nous croyons qu'il est possible de montrer que l'idée métaphysique de la causalité ou le « schéma dynamique », même s'il n'est pas nommé ainsi, apparaît effectivement, bien qu'en filigrane, au milieu de la théorie de la reconnaissance de *Matière et mémoire*.

2. Explication de la compréhension dans la théorie de la reconnaissance de *Matière et mémoire*

Examinons maintenant la théorie de la reconnaissance qui se trouve dans le 2e chapitre de *Matière et mémoire* afin de constater tout d'abord que le concept de « schème moteur » se situe au niveau sensori-moteur (2.1). Cependant, Bergson dit à propos de la compréhension orale, que le rétablissement des « rapports »[41] s'opère « ailleurs »[42]

[37] Cf. MM, p. 170.
[38] Cf. Riquier, *op. cit*., p. 361.
[39] *Ibid*., nous soulignons.
[40] *Ibid*., p. 367.
[41] Cf. MM, p. 138-9.
[42] MM, p. 108.

que dans le « schéma moteur » (2.2). Nous ajoutons en outre que le même type de rapport abstrait se retrouve également pour l'effort corporel (2.3).

2. 1 Le schème moteur de *Matière et mémoire*, qui consiste dans le niveau sensori-moteur

Le « schème moteur » constitue le fondement théorique à partir duquel Bergson explique le mécanisme de la reconnaissance, c'est-à-dire l'action de mémoire qui nous permet d'identifier l'objet perçu dans le présent comme connu. Bergson commence par critiquer la thèse associationniste. Selon Bergson, en effet, la thèse associationniste, qui postule que la reconnaissance s'explique par la simple juxtaposition de perception et souvenir, est incapable de rendre compte de la « cécité verbale », phénomène pathologique où le patient conserve intacte « la faculté d'écrire sous la dictée ou spontanément » mais reste impuissant « à saisir ce qu'on pourrait appeler le mouvement des lettres », à tel point qu'on ne peut plus dessiner des lettres qu'en suivant le modèle à chaque instant.[43]

Ce qui est perdu dans un tel état, c'est, selon Bergson, l'habitude d'articuler l'objet perçu par l'accompagnement moteur de la perception ou, plus précisément, l'habitude de « démêler tout de suite », en esquissant « les mouvements qui s'y adaptent »,[44] les articulations de l'objet aperçu » pour « en compléter la perception visuelle ».[45] Le « schème moteur » consiste en cette habitude de démêler les « articulations ». Et puisqu'il ne s'agit ici que de ce discernement, il n'est pas besoin d'exécuter effectivement les mouvement ; il suffit d'« y tendre ».[46] Bergson reconnaît, au niveau du mouvement ainsi dessiné, la « base de la reconnaissance ».[47] Les associationnistes qui expliquent la reconnaissance par une juxtaposition pour ainsi dire statique de souvenir et de perception n'ont donc, selon Bergson, « pas démêlé [...] l'élément moteur » du mécanisme de la reconnaissance.[48]

Cet accompagnement moteur de la perception est rendu possible par notre système nerveux. Notre système nerveux est, dit-Bergson, évidemment disposé en vue d'établir des « connexions » automatiques « entre l'impression sensorielle et le mouvement qui l'utilise », de construire des « appareils moteurs, reliés, par

[43] Cf. MM, p. 106.
[44] MM, p. 101.
[45] Cf. MM, p. 106-7.
[46] MM, p. 101.
[47] *Ibid.*
[48] Cf. MM, p. 117.

l'intermédiaire des centres, à des excitations sensibles »,[49] – en un mot, en vue de l'« adaptation motrice ».[50]

De cette analyse de la tendance motrice qui accompagne la perception, Bergson conclut : « Nous jouons [corporellement] [...] notre reconnaissance avant de la penser [intellectuellement] ».[51] Cela revient à dire que le schème moteur fonctionne « tout de suite » au niveau sensori-moteur, exactement de la même manière que l'intellection automatique de l'article de 1902.

2.2 La représentation des rapports qui se trouve ailleurs que dans le « schème moteur »

2.2.1 La nécessité de sauter « ailleurs »

Dans *Matière et mémoire*, Bergson distingue deux sortes de reconnaissance : l'une, où la reconnaissance basique fonctionne seule (reconnaissance automatique) et l'autre, où s'y ajoute l'opération intellectuelle (reconnaissance attentive ou intellectuelle). Il conviendra d'examiner cette dernière en s'appuyant sur l'exemple de « l'audition du langage articulé »[52] que Bergson donne dans la seconde moitié du 2e chapitre.

Le processus d'audition du langage peut, d'après Bergson, être divisé en deux parties. La première étape consiste dans « un processus automatique sensori-moteur ».[53] Il s'agit ici de « reconnaître le son [de la parole] »[54] et dans cette étape, à travers le « schème motrice », c'est-à-dire la tendance motrice corporelle, on discerne les articulations d'un mot dans la « masse sonore » initialement donnée pour que l'oreille entende le mot.[55] A ce niveau, il n'y a aucune intervention de la pensée ou de l'intellect, qui n'est d'ailleurs pas nécessaire pour échanger, par exemple, une salutation toute faite, puisque l'activité linguistique se déroule alors « à la manière d'un réflexe ».[56]

La deuxième étape consiste dans « une projection active et pour ainsi dire excentrique » de « souvenirs-images »[57] emmagasinés dans la « mémoire

[49] Cf. MM, p. 102.
[50] MM, p. 108.
[51] MM, p. 103.
[52] Cf. MM, p. 119.
[53] *Ibid.*
[54] *Ibid.*
[55] Cf. MM, p. 120.
[56] Cf. MM, p. 91-2.
[57] Cf. MM, p. 119.

indépendante [du niveau sensori-moteur] ».[58] Il s'agit cette fois de « reconnaître avec intelligence »[59] ou de « retrouver le sens [du son] » pour « en pousser plus ou moins loin l'interprétation ».[60]

Bergson affirme qu'il faudrait chercher les souvenirs-images « ailleurs » qu'à l'ordre sensori-moteur :

> « [S]i nous établissions [...] que le rôle de l'ébranlement perceptif est simplement d'imprimer au corps une certaine attitude où les souvenirs viennent s'insérer, alors, tout l'effet des ébranlements matériels [cérébraux] étant épuisé dans ce travail d'adaptation motrice, il faudrait chercher le souvenir ailleurs ».[61]

La première moitié de ce passage présente la fonction du schème moteur, que nous avons vue en section 2.1. Dans la reconnaissance attentive, le schème moteur, en restant toujours au niveau sensori-moteur, sert de cadre dans lequel les souvenirs finissent par s'insérer.

Que signifie alors « chercher le souvenir ailleurs » ? À cette question, Bergson donne une réponse plus claire en décrivant ce qui se passe lorsqu'on comprend intellectuellement le discours d'une autre personne :

> « Dans le cas particulier qui nous occupe, l'objet est un interlocuteur dont les idées s'épanouissent dans sa conscience en représentations auditives, pour se matérialiser ensuite en mots prononcés. Il faudra donc [...] *que l'auditeur se place d'emblée parmi des idées correspondantes*, et les développe en représentations auditives qui recouvriront les sons bruts perçus en s'emboîtant elles-mêmes dans le schème moteur ».[62]

L'auditeur doit refaire par lui-même, ou dans sa propre conscience, le mouvement de l'interlocuteur. Il doit, pour ce faire, « se place[r] d'emblée parmi des idées correspondantes ». C'est ce domaine des idées que Bergson entend par l'expression « ailleurs », où il faut, selon lui, chercher le souvenir. Et en plus, puisque les « idées »

[58] Cf. MM, p. 81.
[59] Cf. MM, p. 129.
[60] Cf.MM, p. 119.
[61] MM, p. 108.
[62] MM, p. 129, soulignage par Bergson.

sont, dans ce cas, la cause des « mots prononcés » ou matérialisés, ce déplacement fait d'emblée revient à la reconstruction de la « cause »[63] de ce qui est perçu.

Donc, *Matière et mémoire* ainsi que « L'effort intellectuel » expriment la nécessité de « sauter » du niveau des sensations et des mouvements au niveau des « idées » ou du « sens ». Et le progrès à partir des « idées » ferait écho à ce que Bergson appellera la « causalité réelle » dans « L'effort intellectuel ».

2.2.2 Contraste entre le « schème moteur » et le rétablissement des « rapports » dans le domaine des idées

Il nous faut maintenant montrer que la reconstruction de la « cause » dans le domaine des idées consiste dans le rétablissement fait par l'esprit des « rapports » immatériels, et que ce rétablissement est, en tant qu'action de l'esprit, de nature différente de l'opération du « schème moteur ».

C'est dans la partie consacrée à la « structure d'une phrase »[64] que Bergson évoque éminemment la nécessité pour l ' « esprit » de « rétablir les rapports » dans le domaine des idées, ailleurs que dans le domaine sensori-moteur, où se situe le « schème moteur ». « Certains théoriciens de l'aphasie sensorielle », dit-il, « raisonnent comme si une phrase se composait de noms qui vont évoquer des images de choses », alors que, selon Bergson, « affinée ou grossière, une langue sous-entend beaucoup plus de choses qu'elle n'en peut exprimer » et force « l'esprit » « à rétablir les rapports » inexprimés, c'est-à-dire, à « établir entre les images [exprimées] des rapports et des nuances de tout genre ».[65]

Or, de ce rétablissement, ainsi poursuit Bergson :

> « Essentiellement discontinue, puisqu'elle procède par mots juxtaposés, la parole ne fait que jalonner de loin en loin les principales étapes du mouvement de la pensée. C'est pourquoi je comprendrai votre parole si je pars d'une pensée analogue à la vôtre pour en suivre les sinuosités à l'aide d'images verbales destinées, comme autant d'écriteaux, à me montrer de temps en temps le chemin. [...] Les images ne seront jamais en effet que des choses, et la pensée est un mouvement».[66]

[63] MM, p. 115 ; p.129.
[64] Cf. MM, p. 138-9.
[65] Cf. *ibid.*
[66] MM, p. 139.

Examinons étape par étape cette citation.

D'abord, les « images verbales » qui ne font « que jalonner de loin en loin les principales étapes de mouvement de la pensée » sont équivalentes aux « mots » que le « schème moteur » a distingués (voir 2.2.1).

Ensuite, notons que ce n'est que « de loin en loin » ou « de temps en temps » que les mots, ou le schème moteur derrière eux, servent « à me montrer le chemin ». D'une part, en effet, puisque je ne peux sauter dans le domaine des idées qu'à l'aide de mots entendus, le schème moteur « délimit[e] le champ des images » vers lequel je dois sauter pour chercher ou « cueillir » les images.[67] Mais d'autre part, étant donné que le schème moteur est formé par une coordination entre les impressions sensorielles subies et la tendance motrice corporelle qui les accompagne automatiquement, et que cette coordination a besoin d'être établie chaque fois dans le présent sensori-moteur,[68] l'opération seule du schème moteur ne parvient jamais à reconstruire la continuité du « mouvement de la pensée ». En d'autres termes, les « rapports » inexprimés ou sous-entendus ne peuvent pas être rétablis par le schème moteur.

C'est pourquoi les destinations du schème moteur n'apparaissent que « de loin en loin » ou « de temps en temps ». Bergson décrit une situation similaire de la manière suivante : « Le schème moteur, soulignant ses [=l'interlocuteur] intonations, suivant, *de détour en détour*, la courbe de sa pensée, montre à notre pensée le chemin ».[69] Cette discontinuité revient à témoigner en effet que le schème moteur est attaché à la relation sensori-motrice qui ne peut s'établir qu'à chaque fois, dans le présent.

La continuité propre au mouvement de la pensée devrait donc être reconstruite par l'esprit, au sens où celui-ci embrasse le sous-entendu qui est nécessairement exclu du domaine sensori-moteur.

En outre, puisqu'il s'agit de reconstruire le sous-entendu inexprimé dans la perception, cette reconstruction prend nécessairement le caractère d'une « hypothèse »[70] lancée « au-devant de la perception ».[71] C'est pourquoi cette reconstruction est également qualifiée de « création ».[72]

---

[67] Cf.MM, p. 103.
[68] Cf. MM, p. 153-4.
[69] MM, p. 135, nous soulignons.
[70] Cf. MM, p. 112.
[71] Cf. MM, p. 108 ; p. 142.
[72] Cf. MM, p. 113.

Pour résumer, le schème moteur est pour ainsi dire enfermé dans le présent sensori-moteur et est par conséquent discontinu, tandis que la reconstruction des relations par l'esprit progresse au contraire en avant[73] et qu'elle est un mouvement continu.

On voit donc que Bergson conçoit déjà dans *Matière et mémoire*, bien qu'il n'y utilise pas le terme « schéma dynamique », une activité analogue, c'est-à-dire la reconstruction du mouvement ou de la continuité de la pensée par l'esprit en tant que rétablissement de relations abstraites, fondamentalement distincte du « schème moteur ».

2.3 Des représentations abstraites de relation dans les efforts corporels de *Matière et mémoire*

Nous avons vu en section 1.4 qu'il existe, selon « L'effort intellectuel », dans l'effort corporel aussi, la « représentation » qui « ressemble beaucoup » au schéma dynamique, c'est-à-dire le « dessin de relations, surtout temporelles, entre les parties successives du mouvement à exécuter ». Or, d'après nous, Bergson reconnaît, également dans *Matière et mémoire*, l'existence d'une « représentation » de ce genre à l'égard de l'effort corporel pour contracter l'habitude motrice corporelle : « [D]ans l'opération par laquelle on contracte le souvenir-habitude » tel que l'apprentissage d'une leçon ou la maîtrise d'une dance, se trouve « un certain effort *sui generis* qui nous permet de retenir l'image elle-même », et « nous nous servons de [cette] image fugitive pour construire un mécanisme stable ».[74] La question est de savoir ce qu'est cette « image elle-même ». Bergson dit qu'elle est, dès la première tentative, « dans notre esprit, invisible et présente ».[75] Ce n'est pas quelque chose de clairement saisissable. Si « nous reconnaissons [...] telle erreur que nous venons de commettre » au moment de la première récitation d'une leçon, par exemple, ce n'est qu'« à un vague sentiment de malaise » ou « comme si nous recevions des obscures profondeurs de la conscience une espèce d'avertissement ».[76] Bergson considère cette image invisible également comme « une représentation d'ensemble, une sorte d'idée

[73] Bergson invoque en effet le même argument à propos du mécanisme de la « lecture » : « la lecture courante est un véritable travail de *divination, notre esprit cueillant çà et là* quelques traits caractéristiques et comblant tout l'intervalle par des souvenirs-images. » (MM, p. 113, nous soulignons).
[74] Cf. MM, p. 91.
[75] MM, p. 92.
[76] Cf. *ibid.*

complexe embrassant le tout, et où les parties [ont] une unité inexprimablement sentie ».[77]

Cette « représentation d'ensemble [...] où les parties [ont] une unité inexprimablement sentie » de laquelle nous nous servons pour contracter l'habitude motrice fait écho au « dessin de relations [...] entre les parties successives du mouvement à exécuter » de « L'effort intellectuel ». Cette « représentation » ou « image » nous informe, lors de la première récitation d'une leçon, qu'une erreur a été commise dans la « relations [...] entre les parties successives », et cependant cette représentation est essentiellement une « représentation abstraite de relation ». C'est pourquoi elle ne peut être reconnue que comme « une espèce d'avertissement » depuis les « obscures profondeurs de la conscience ». Et grâce à la « faculté » de cette « représentation », « nous n'avons pas besoin d'attendre du hasard la répétition accidentelle des mêmes situations » pour construire un mécanisme moteur stable.[78] Cela revient à dire, dans les termes de « L'effort intellectuel », que nous pouvons « *apprendre tout seuls* un exercice complexe, tel que la danse » (nous soulignons).

Si notre analyse jusqu'ici est exacte, ce que Bergson appelle « schéma dynamique » dans « L'effort intellectuel » ainsi que l'image qui exprime une idée semblable (à savoir les « représentations abstraites de relations » descendant vers le domaine de sensation), se trouvent déjà dans *Matière et mémoire*, dans la conception de la compréhension intellectuelle, et dans celle de l'acquisition du mouvement habituel. Ces représentations se distinguent du « schème moteur » qui ne relève que de la sensation « brute » dans « L'effort intellectuel ».

Il faut néanmoins admettre que l'existence dans *Matière et mémoire* du « schéma dynamique » demeure à peine visible derrière le schème moteur. Le temps est venu pour nous d'examiner les raisons de cette opacité.

3. Le « schéma dynamique » à peine visible derrière le « schème moteur »

3.1 « Former son oreille aux éléments d'une langue nouvelle »

Analysons d'abord le passage suivant sur la première étape de l'audition d'une nouvelle langue. Dans ce texte, Bergson explique l'acquisition du schème moteur pour « former son oreille aux éléments d'une langue nouvelle » :

[77] Cf. MM, p. 93.
[78] Cf. MM, p. 91.

> « Les impressions auditives organisent des mouvements naissants, capables de scander la phrase écoutée et d'en marquer les principales articulations. Ces mouvements automatiques d'accompagnement intérieur, d'abord confus et mal coordonnés, se dégageraient alors de mieux en mieux en se répétant ; ils finiraient par dessiner une figure simplifiée, où la personne qui écoute retrouverait, dans leurs grandes lignes et leurs directions principales, les *mouvements mêmes* de la personne qui parle. Ainsi se déroulerait *dans notre conscience, sous forme de sensations musculaires naissantes***,** ce que nous appellerons le *schème moteur* de la parole entendue. Former son oreille aux éléments d'une langue nouvelle ne consisterait alors ni à modifier le son brut ni à lui adjoindre un souvenir ; ce serait coordonner les tendances motrices des muscles de la voix aux impressions de l'oreille, ce serait perfectionner l'accompagnement moteur ».[79]

Quels sont les « *mouvements mêmes* de la personne qui parle » que « la personne qui écoute retrouverait » ? S'agit-il du mouvement de sa pensée ou du mouvement au niveau sensori-moteur ? A première vue, il semble que Bergson parle ici du second type de mouvement, c'est-à-dire du mouvement des muscles vocaux chez le locuteur (« Ainsi se déroulerait [...] sous forme de sensations musculaires naissantes »). Cela reviendrait à dire que l'on est dans le domaine du « schème moteur » et non celui du « schéma dynamique ».

Or, Bergson, comme nous le verrons juste après, retourne à nouveau à ce passage lorsqu'il parle du « mouvement de la phrase » dans l'explication de la deuxième étape du processus de l'audition du langage : « Un mot s'anastomose toujours avec ceux qui l'accompagnent, et selon l'allure et le mouvement de la phrase dont il fait partie intégrante, il prend des aspects différents ».[80] A la différence de l'« image [...] isolée », « le mot qui s'organise avec la phrase » participe à « la réalité vivante », « telle, chaque note d'un thème mélodique reflète vaguement le thème tout *entier* ».[81] « L'allure et le mouvement de la phrase » sont donc, selon lui, de nature générale. Et la participation des mots à ce mouvement général est ce qui est sous-entendu par le terme utilisé par Bergson d'« anastomose ».[82]

---

[79] MM, p. 121, nous soulignons.
[80] MM, p. 130-1.
[81] Cf. *ibid*., nous soulignons.
[82] Dix ans plus tard, dans un cours au Collège de France, Bergson oppose de nouveau ce terme à une simple juxtaposition : « En somme, on ne pourrait comprendre le sens d'une phrase, même si chaque mot entendu allait évoquer une idée ou une image correspondante,

Le mouvement général de la phrase est, comme nous l'avons vu en section 2.2.2, reconstruit par l'esprit, ailleurs que dans le « schème moteur ». Toutefois, lorsque Bergson explique plus tard la deuxième étape, il attribue ce mouvement de la phrase au schème moteur en disant que « je comprends fort bien ce commencement de reconnaissance automatique qui consisterait, *comme on l'a vu plus haut*[c'est-à-dire dans le long passage que nous avons cité plus haut], à souligner les principales articulations de cette phrase, à *en adopter ainsi le mouvement* ».[83] Notre analyse en section 2.2.2 nous permet d'indiquer qu'il y a ici une oscillation dans l'explication donnée par Bergson : dans un passage il affirme que le schème moteur « adopte » le mouvement de la phrase, alors qu'il établit dans un autre que la phrase est un mouvement de la pensée qui échappe au schème moteur. En d'autres termes, il n'est pas possible de reconstituer le « mouvement de la phrase », qui est le mouvement général, seulement en « coordonn[ant] les tendances motrices des muscles de la voix aux impressions de l'oreille ».

Nous pouvons donc dire que dans ces descriptions, sous le mot « mouvement », les mouvements sensori-moteurs (ce qui correspond à l'« analyse »[84] par le schème moteur) et le mouvement général de la pensée (ce qui correspond à l'« effort de synthèse »[85] par l'esprit) sont confondus, celui-ci étant à peine visible derrière ceux-là.

Qu'est-ce qui serait alors la source de cette confusion ?

3.2 La raison pour laquelle le « schéma dynamique » est à peine visible derrière le « schème moteur »

Rappelons-nous qu'il faut comprendre l'effort du texte de 1902 comme un « mouvement », ou mieux encore une « force » par laquelle s'opère la « causalité réelle ». Or, c'est le statut donné à ces deux notions dans le 4e chapitre de *Matière et mémoire* qui est à l'origine de la confusion à laquelle nous avons affaire ici.

Dans le 4e chapitre, Bergson discute l'existence du « mouvement réel ».[86] Il soutient que « le mouvement réel se distingue du mouvement relatif en ce qu'il a une

parce qu'une phrase n'est pas une juxtaposition d'éléments, mais une véritable *anastomose*. » (M, p. 701, nous soulignons).

[83] MM, p. 131, nous soulignons.

[84] Cf. MM, p. 111-2.

[85] Cf. *ibid.*

[86] Cf. MM, p. 215.

cause réelle, en ce qu'il émane d'une force ».[87] Mais en même temps, « il faudrait s'entendre sur le sens de ce dernier mot »,[88] car, dans les sciences de la nature, la force « n'est qu'une fonction de la masse et de la vitesse »[89] à laquelle échappe finalement le mouvement réel. « Il faudra donc se rejeter sur le sens métaphysique du mot, et étayer le mouvement aperçu dans l'espace sur des causes profondes, analogues à celles que notre conscience croit saisir dans le sentiment de l'effort ».[90] Le problème de cette méthode, selon Bergson, c'est qu'il est difficile de dire que « le sentiment de l'effort est [...] celui d'une cause profonde » parce que « *les analyses décisives [...] ont [...] montré qu'il n'y a rien autre chose, dans ce sentiment, que la conscience des mouvements déjà effectués ou commencés à la périphérie du corps* ».[91] Ainsi, en fin de compte, « l'analyse nous ramène toujours au mouvement lui-même », qui signifie évidemment le mouvement au niveau sensori-moteur. D'où, conclut Bergson, l'évidence de l'existence du mouvement réel.

En effet, « les analyses décisives » se réfèrent à un article de W. James intitulé « Le sentiment de l'effort »,[92] où le philosophe américain critique la thèse selon laquelle l'acte volontaire s'accompagne toujours du sentiment de l'émission d'une force irréductible au mouvement musculaire périphérique. En ce qui concerne le concept de sentiment de l'effort, on distingue deux interprétations opposées de l'époque : la théorie centrifuge et la théorie centripète. Selon la première position, d'origine essentiellement biranienne, c'est précisément l'effort qui produit le mouvement corporel, et dans le fait de sentir cet effort nous saigissons à sa source « l'émission d'une force » ou « un courant centrifuge d'énergie» (cf. M, 689-690). Selon la deuxième position, au contraire, « le sentiment de l'effort se résout dans un ensemble de sensations périphériques (tension des muscles, leur froissement, frottement de la peau, etc...) » et « un courant centrifuge d'énergie» n'est qu'une illusion (cf. *ibid.*). L'article ci-dessus de James soutient donc nettement une thèse centripète, sur laquelle s'est appuyé Bergson lors de la rédaction de sa *Matière et*

[87] Cf. MM, p. 218.
[88] *Ibid.*
[89] *Ibid.*
[90] *Ibid.*
[91] Cf. *ibid.*, nous soulignons.
[92] Cf. *Matière et mémoire*, trad. en japonais par Naoki Sugiyama, Tokyo, Kodansha, 2019, p. 323-4.

mémoire. Ce qui explique sans doute pourquoi il n'a pas cherché à saisir la « cause » ou la « source » de l'énergie spirituelle associée au sentiment de l'effort.[93]

En revanche, on reconnaît aisément que la position bergsonienne de « L'effort intellectuel » selon laquelle « le sentiment de l'effort d'intellection se produit sur le trajet du schéma à l'image » ne s'accorde plus au point de vue jamesien. Alors qu'il était enclin à s'accorder au point de vue centripète dès l'*Essai sur les données immédiates à la conscience,* Bergson prend résolument ses distances par rapport à celui-ci dans les années 1900.[94] Dans « L'effort intellectuel », Bergson oppose le point de vue centrifuge au point de vue centripète sur « le sens véritable des oscillations de l'attention » et déclare que « même si l'on n'accepte pas la première thèse, il semble bien qu'il faille en retenir quelque chose ».[95] Ce qu'il faut retenir pour Bergson, c'est le progrès « du schéma à l'image perçue ».[96] En outre, un peu plus tard, dans les cours de Collège de France 1906-7 sur « La théorie de la volonté », Bergson critique explicitement le point de vue centripète de James sur le sentiment de l'effort en admettant que la théorie centrifuge « contient un fond de vérité» (M, 695) .[97]

Si notre analyse est exacte, donc, ce qui empêche Bergson de *Matière et mémoire* d'affirmer pleinement l'idée naissante du schéma dynamique ou de la « force » exercée par lui là où il serait légitime pour lui de le faire, c'est son attachement au point de vue centripète fourni par James.[98]

Conclusion

[93] Sans entrer dans le détail, nous nous bornons ici à noter qu'il faudrait un examen approfondi sur James et d'autres théories psychologiques de l'époque pour clarifier de manière satisfaisante la position bergsonienne sur cette question.

[94] Cf. Riquier, *op. cit.*, p. 368-9. Riquier reconnait que Bergson, qui était proche de la perspective centripète dans l'*Essai* et *Matière et mémoire,* se rapproche de la perspective centrifuge dans les années 1900.

[95] Cf. ES, 173.

[96] *Ibid.*

[97] Cf. M, p. 692-4.

[98] Jankélévitch considère la perspective « centrifuge » de « L'effort intellectuel » comme un prolongement de celle de *Matière et mémoire* (cf. Jankélévitch, Vladimir, *Henri Bergson*, Paris, PUF, 2015, p. 112-3.). Il nous faut dire pourtant que, bien que Bergson adopte la thèse centrifuge lorsqu'il explique la deuxième étape de la reconnaissance dans *Matière et mémoire,* elle ne fait pas partie de l'étape du commencement de la reconnaissance. En ce sens, l'évolution de la pensée de *Matière et mémoire* vers « L'effort intellectuel » au sujet de l'action de mémoire ne se réduit pas à un simple prolongement.

En guise de conclusion, il conviendra de récapituler ce que nous avons discuté. En section 1, nous avons examiné « L'effort intellectuel » pour montrer que le « schéma dynamique » consiste dans la reconstruction des relations abstraites et continues qui se forment en dehors du mouvement sensori-moteur. En section 2, nous avons ensuite constaté, en nous appuyons sur *Matière et mémoire*, que le « schème moteur » est au contraire borné à ce qui se passe de façon immédiate au niveau sensori-moteur, alors que Bergson reconnaît en même temps qu'il faut chercher une représentation des relations « ailleurs » que dans ce schème. Nous avons ainsi repéré un contraste entre le « schème moteur » discontinu et le « schéma dynamique » continu. Finalement, en section 3, nous avons essayé de montrer que malgré cet évident contraste, les explications fournies dans *Matière et mémoire* confondent les deux concepts l'un avec l'autre, en conséquence de quoi l'existence d'une idée analogue au schéma dynamique au sein de la théorie de la reconnaissance de *Matière et mémoire* n'est appréciable que difficilement. Nous avons soutenu que cette difficulté est due au fait que Bergson de *Matière et mémoire* s'appuie en grande partie sur la conception centripète du sentiment de l'effort qu'il emprunte à W. James.

Comme le suggère Riquier,[99] Bergson considérera plus tard « une force spécifique, *sui generis* » qui se trouve dans l'« effort », c'est-à-dire dans le développement à partir du « schéma dynamique », comme une « volonté »,[100] et il appellera le lancement hypothétique en avant qui y est à l'œuvre[101] « élan vers le futur ».[102]

Ainsi, dans *L'évolution créatrice* de 1907, Bergson élaborera plus loin l'idée du mouvement de l'esprit que l'on ne pouvait qu'entrevoir dans *Matière et mémoire,* pour concevoir quelques concepts décisifs comme l'« élan vital » ou la « création ». En même temps, cette évolution de la pensée conduira Bergson à utiliser la « vie » au lieu de la « causalité réelle », dans la version reprise de « L'effort intellectuel » lors de la publication de *L'énergie Spirituelle* en 1919.

Bibliographie

[99] Cf. Riquier, *op. cit*., p. 371-2.
[100] M, p. 704.
[101] Cf. M, p. 700.
[102] M, p. 859 ; p. 1083.

Bergson, Henri (1896). *Matière et mémoire*. 9e éd. (Paris: PUF, 2007). « MM » (traduit en japonais par Naoki Sugiyama, Tokyo, Kodansha, 2019.)

Bergson, Henri (1919). *L'énergie spirituelle*. 9e éd. (Paris: PUF, 2009). « ES »

Bergson, Henri (1972). *Mélanges*. 1ère éd. (Paris: PUF). « M »

Henri, Bergson (2017). *L'idée de Temps. Cours Au Collège de France 1901-1902*. 1ère éd. (Paris: PUF).

Deleuze, Gilles (1966). *Le Bergsonisme*. 3e éd. (Paris: PUF, 2004).

Jankélévitch, Vladimir (1930). *Henri Bergson*. 4e éd. (Paris: PUF, 2015).

Riquier, Camille (2006). *Archéologie de Bergson. Temps et Métaphysique*. 1ère éd. (Paris: PUF).

Worms, Frédéric (1998). *Introduction à* Matière et mémoire *de Bergson*. 1ère éd. (Paris: PUF).

Morality as Misconception of Life:
A Defense of Bernard Williams' Critique of the Morality System

WATANABE Kazuki
The University of Tokyo

***Abstract**: Bernard Williams called the core feature of modern morality "the morality system", and proposed a substantive critique of the system that made a strong impact on contemporary ethics. However, any critique of morality, including Williams', is confronted with the pervasive questions, "What morality is being criticized?" and "On what grounds is it being criticized?" So moral critics need to show that their critique is about our morality and that it is justified. The two questions are precisely the objections posed to Williams' critique, and the aim of this paper is to defend it.*

Williams argues that the morality system motivates us to conceptualize obligation and blame in a peculiar way, in which obligation is inescapable and blame only applies to truly voluntary actions, and shows that we can conceptualize them differently and naturalistically. The first objection to this critique, however, concerns the question of "What morality?": the morality system is merely an extreme morality that people with a moral sense do not adopt (The Scope Objection). I will respond to this objection by arguing that it misses the real target of Williams' critique, which is the commitment to voluntariness. Unlike the conception of obligation in the morality system that opponents consider extreme, the commitment is widely presupposed in modern morality. At this point, the second repeated objection comes to the fore, which concerns the question of "On what grounds?": the morality system's commitment to voluntariness is indispensable for a fair morality (The Point of Morality Objection). I will answer this objection by showing the Williamsian consideration against the commitment. The commitment that attempts to realize extreme fairness actually defeats the very workings of blame. I conclude that the problem with the morality system's commitment to voluntariness is a misconception of life: it misconstrues human motivation, the real workings of co-existence, and the required fairness.

Introduction

This paper aims to defend Bernard Williams' critique of the morality system and proposes positive arguments for this critique. Although his critique, which is one of

the severest attacks on modern morality, has had a wide impact on contemporary ethics, it also has some issues. Objections to it have been raised repeatedly, and because Williams' texts are rather elusive, his critique has not been fully explained.[1] It would, therefore, seem incumbent on those sympathetic to Williams to address these objections and clarify Williams' project. This study is one such attempt.

Williams called the core feature of modern morality "the morality system" and proposed a substantive critique of it that surpassed that of any other contemporary philosopher.[2] However, any critique of morality, including Williams', is confronted with pervasive questions. One is "*What morality* is being criticized?" and the other is "*On what grounds* is it being criticized?" This kind of suspicion of moral critics is natural, insofar as morality is at the heart of our values. As Nietzsche said, "One has taken the value of these 'values' for granted, as a fact, as beyond all questioning" (Nietzsche 1998, 5). Thus, moral critics need to show both that their critique is about *our* (common sense) morality and that the critique is *justified*. The above two questions are repeated objections to Williams that I address in this paper.

This study comprises three parts. In the first section, I summarize Williams' critique of the morality system as agreed upon by commentators. In the second section, I address the *scope objection* to Williams that his critique is directed only at extreme morality, arguing that commentators have mistakenly underestimated the target of Williams' critique. In the third section, I address the *point of morality objection* that the morality system is still justified because its nature is essential for any morality; I will argue that this is not the case and conclude that the morality system involves a misconception of life. How it misconceives life will be clarified throughout this paper.

1. Summary of Williams' Critique of the Morality System

[1] The interpretive studies of Williams (Jenkins 2006; Queloz 2022; Chappell and Smith 2023) have elucidated the point of Williams's critique, but they have tended to leave the repeated objections to Williams unanswered, which I think is a problem of the earlier research. By answering these objections, I will make the point and meaning of the critique clearer.

[2] Skorupski takes Williams' critique as representative of the critique of morality (Skorupski 1998).

While Williams' description and critique of the morality system [3] (hereafter abbreviated as MS) were scattered throughout many texts over a period of more than 20 years, commentators on Williams have converged on treating the final chapter of *Ethics and the Limits of Philosophy* (ELP)[4] as the main text in which Williams presented the most comprehensive and extensive discussion of MS. To understand the essence of Williams' critique, I summarize the chapter in this section. I will first extricate the important features of MS, and then show Williams' critique.

Obligation: Practical and inescapable

Williams' discussion of MS is first devoted to "the special notion of obligation it uses" (ELP 193). According to Williams, MS' special notion of *moral obligation* tends to have distinctive features. First, moral obligation is a kind of *practical conclusion*: "an obligation applies to someone with respect to an action" (ELP 194). Thus, a moral obligation is a specific action that an agent must perform. This feature is naturally related to some principles of morality, such as the "*ought* implies *can*" principle, that the obligation that an agent morally ought to do "must be within the agent's power" (ELP 194–5). In addition, the fact that a moral obligation is a practical conclusion relates to the point that moral obligations cannot conflict with each other, since an agent cannot perform several obligatory actions simultaneously; what one must do is only the *actual* obligation.

Second, moral obligation is *inescapable*: it is an action that cannot be dismissed without blame. This is expressed in various moral principles. For example, W. D. Ross tends to adopt the principle that one cannot override one's moral obligation unless one is under a more "general" obligation, which Williams calls the "obligation-out, obligation-in principle" (ELP 201). According to this principle, for instance, if you see an old person in need of help, you can cancel your promise to visit your friend, because the helping is an "actual" obligation. The important point is that the principle requires a general obligation to override a less general obligation, so moral obligation itself is still an action that one must do, regardless of one's desire. Moral obligations are, in this sense, inescapable.

[3] Williams often refers to it simply as "morality," although in everyday terminology we use "morality" to refer to the broad normative understanding (Griffin 2015, 93–4). I will use "the morality system" instead of "morality" to avoid any possible confusion.

[4] Quotations from *Ethics and the Limits of Philosophy* (Williams 1985/2011) are given with the abbreviation (ELP) and page numbers.

Blame: Focused on voluntariness

In addition to the notion of obligation, there is MS' special notion of *blame* that is related to the former. According to Williams, blame, typically a response to a failure to fulfill an obligation, is "the characteristic reaction" of MS (ELP 197). The most striking feature of blame in MS is its *focus on voluntariness*; only truly voluntary actions are blameworthy. This focus on voluntariness is conspicuous in modern moral philosophy's obsessive discussion of *free will*. Modern philosophers, in contrast to the ancients, have tended to claim that moral responsibility and the practice of blame are only tenable if there is an agent's free will—a freely chosen will to act. This idea is motivated by the focus on voluntariness, as free will is conceived as the true voluntariness. Indeed, many philosophers have tried to reconcile the focus on voluntariness and free will with a naturalistic picture of human action, in which one's actions are causally determined.[5]

The important point here is that MS drives philosophers to *purify* voluntariness, as it is a crucial condition in the practice of blame (ELP 216). They are motivated to scrutinize an agent's voluntariness; they have argued that true voluntariness or free will is analyzed as the ability to have acted otherwise. This is the principle of *alternate possibilities* that most moral philosophers share in the discussion of free will (Van Inwagen 1975, 188; Frankfurt 1988, 1).[6] Indeed, it is a moral platitude that, if one could not have done otherwise, one cannot be blameworthy.[7] In summary, MS motivates philosophers to purify voluntariness, and the most typical way of the purification is shown in the principle of alternate possibilities.

Williams' critique: Naturalistic understanding of morality

[5] In this paper, following Williams, the word "naturalistic" is generally used to mean the minimalist understanding of human behavior, in which various human actions are maximally explained by general factors such as desire. In the theory of action, naturalism refers specifically to the scientific and causal view of human action. In ethics, naturalism is related to the understanding of ethical motivation in terms of pre-ethical primitive motivation. For this minimalist interpretation of naturalism, see Williams 1995d and Williams 2000.

[6] Although Williams does not mention this principle directly in ELP, he does mention it in his later paper on the free will debate and MS (Williams 1995a, 5). The relationship between this widely held principle and MS is discussed in more detail in later sections.

[7] Moreover, theorists committed to MS often purify voluntariness in terms of agents' history; they often "raise questions about their freedom to choose a different character" (ELP 216). Alternate possibilities are needed not only at the moment of action, but also at the moment of acquiring the character that causes the action.

The features of MS appear to generate stringent consequences. Most obviously, the inescapability of obligation induces that one cannot escape moral obligations unless one has a general obligation as a pardoner, and this leads to the primacy of obligation in practical moral thinking about what one should do. Accordingly, some moralists would argue that you have a moral obligation to yourself: you have an obligation to put yourself to better use, rather than wasting resources on morally neutral things. They would say that you are guilty of enjoying the opera rather than donating to famine relief. As Williams remarks, if practical and inescapable obligation structures ethical thought, "there are several natural ways in which it can come to dominate life altogether" (ELP 202).

There is a common understanding in the literature that Williams' critique of these features of MS is that we do not need to conceptualize obligation or blame as stringently as MS.[8]

> In order to see around the intimidating structure that morality has made out of the idea of obligation, we need an account of what obligations are when they are rightly seen as merely one kind of ethical consideration among others. This account will help to lead us away from morality's special notion of moral obligation, and eventually out of the morality system altogether. (ELP 202)

Instead of MS' conception of obligation and blame, we can have a more naturalistic understanding of morality—what Williams calls "everyday" notion of obligation and blame (ELP 201, 214)—which does not commit the features of MS.

What is the everyday conception of morality? The basic idea is that obligation and blame are merely institutions "to secure reliability" (ELP 206). To show this, Williams provides a naturalistic explanation of how obligation works as a means of securing mutual reliability. According to Williams, obligation is nothing more than the institution by which we prioritize ethically important actions in our deliberations (ELP 205). Because of the existence of obligations, we can generally rely on each other as agents who prioritize obligatory actions in their deliberations. Obligation typically includes the positive obligation of helping others in an emergency, the negative obligation of not harming others' basic interests, and the general obligation to keep promises (ELP 206–8). The observance of obligations in this sense is insured by certain psychological maneuvers, such as the encouragement of ethical motivations, and the instilling of ethical dispositions, which often enable

[8] See Jenkins 2006, 71; Louden 2007; Chappell and Smith 2023; Queloz 2022.

the feeling of practical necessity that one "must" do the obligatory action (ELP 206, 210).

This naturalistic understanding of obligation helps us escape the stringent moral life motivated by MS. In the naturalistic understanding, obligation is *only one way* to ensure reliability (ELP 208). For example, a person who has been hurt by racism and is therefore committed to the anti-racism movement may, on her way to visit her friend, witness an anti-racist demonstration and consider it necessary to join the demonstration. Suppose that she does not believe that one is obliged to participate. Here, there is no general moral obligation to attend, but we can understand that attending is so ethically important to her that it outweighs the moral obligation to keep her promise, since the explanation for her action is ethically reliable. In this naturalistic understanding of obligation, an agent is not dominated by inescapable obligations.

Not only obligation. Williams also offers a naturalistic account of blame, according to which blame is a means of correcting agents to ensure coexistence. By blaming an agent who has acted unethically, a community instills a disposition to act ethically (ELP 215). To work well as a modification of agents, this practice of blame must be not only intimidating but fair; we need a certain requirement of voluntariness, such that one has really acted, one knows what one was doing, and one intended the action (ELP 215–6). As in legal cases of punishment, we do have certain conditions for attributing responsibility to ensure fairness without committing ourselves to the metaphysical scrutiny of alternate possibilities.

Again, this naturalistic account of blame helps us escape from MS. As we have seen, MS motivates us to purify voluntariness to the extent that moralists seek to establish true voluntariness by securing alternate possibilities of actions. They assume that true voluntariness or free will is in tension with the naturalistic picture of human action; actions are not truly voluntary if they are causally determined by antecedent facts. However, once we have a naturalistic understanding of blame, we do not need to purify voluntariness as MS does. When we blame someone, we do not need to metaphysically examine alternate possibilities for their actions. What we need for the practice of blame to ensure coexistence is a reasonable degree of voluntariness, but never to the degree of voluntariness that MS requires.

In summary, as long as we understand obligation and blame as institutions for securing coexistence, we do not need to assume that obligation is inescapable, nor to purify voluntariness as alternative possibilities. The drive of MS to conceptualize

obligation and blame in a peculiar way is therefore misguided and naturalistically untenable.

2. The Scope Objection to Williams

In this section, I introduce a repeated objection to Williams that the specific MS he criticizes is just an extremely strict form of morality and is different from common sense morality or ordinary moral theories. In short, it is unclear whether Williams' critique covers real morality. I refer to this objection as the *scope objection* and address it accordingly by proposing my interpretation of Williams.

The scope objections

MS seems demanding. This demands that moral obligations are inescapable. Williams also argues that this demanding conception of morality is "not an invention of philosophers" but "the outlook of almost all of us" (ELP 194). So it would seem that, according to Williams, MS demonstrates the essential stringency of modern morality.

However, philosophers have repeatedly objected to Williams because his view of morality seems farfetched. Their objections focus on the MS' notion of obligation. There is no convergence of moral *theories*, they point out, on the inescapability of moral obligations, let alone the fact that pre-theoretical moral *intuitions* always require moral obligations to be inescapable. In short, they argue that the morality Williams criticizes is just a strawman; he simply invents an extremely strict morality and criticizes it for being strict. The complaint is that Williams' critique can, at best, be "a verbal matter—a dispute about how to use the word 'moral" (Clark 2015, 56). Let us call this objection the *scope objection* because it argues that the scope of Williams' critique of MS is too narrow to be a substantive critique of morality.

At the most local level, opponents argue that MS is even more stringent than Kantian morality, although Williams himself claims that "the purest, deepest and most thorough account of morality is Kant" (ELP 194). Kantian philosophers, such as Robert Louden, claim that Kantian morality is not as stringent as MS (Louden 2007, 114–6). Kantian morality can allow for Williams' everyday conceptions of obligations that are overridable because Kant allows for "imperfect obligations" that are overridable.

On a more general level, opponents argue that MS is far more stringent than ordinary moral theory (Darwall 1987, 79–80). Unlike an extreme moral theory,

ordinary moral theories "must be careful not to claim that moral obligations exist beyond what can reasonably be expected of human beings" (Darwall 1987, 80). According to Darwall, Williams' claim that moral obligations are inescapable is too strict to be a statement of an ordinary moral theory. Rather, the claim seems to be based on the confusion between *all things considered* moral obligations, which are inescapable, and *prima facie* obligations, which are overridable. *Prima facie* obligations can be overridden by ethical considerations other than general obligations, such as supererogation (Darwall 1987, 81; Kamm 1985, 119–20). Thus, unambitious moral theories do not require that all moral obligations to be *all things considered* moral obligations, and they do not presuppose the "obligation-out, obligation-in principle". [9] Instead, they limit their scope to avoid being too demanding. Unambitious morality may provide exceptions to moral obligations: if an obligation is trivial, such as a casual promise, it can be overridden by non-obligatory considerations.

Furthermore, there is the objection that even if MS can be a kind of moral theory, it cannot be a moral culture or pre-theoretical moral common sense (Darwall 1987, 81; Leiter 2022, 22–23). Brian Leiter disagrees with Williams, arguing that "it is a pure philosopher's fantasy to think that real people in the moral culture at large find themselves overwhelmed by this burdensome sense of moral obligation" (Leiter 2022, 22). Except for certain moralists, persons with moral common sense do not restrict their lives by accepting the principle that moral obligations are inescapable. People of moral common sense do not feel as obliged by moral obligations as MS requires, but they are well committed to morality and the authority of moral obligations (Clark 2015, 55–6).

Preliminary response: Distinguishing the nature of the morality system from its products

In brief, the scope objection claims that MS is a far-fetched morality, which is different from ordinary morality. Let us uncritically assume that they are correct in that moral theories are not necessarily committed to the inescapability of moral obligation. Nevertheless, I argue that opponents are wrong in assuming that this claim resolves Williams' critique.

[9] Although I do not explore this issue in detail in this paper, MS's conception of obligation seems to differ from that of utilitarianism. Williams argues that utilitarians are usually pre-theoretically committed to MS (ELP 198, 204). According to Williams, utilitarians internalize MS in practice, even though it may conflict with their utilitarianism in theory. Williams concludes that utilitarianism is a "marginal" form of MS (ELP 198).

The basic idea of my response is that Williams' critique is not so much directed at the stringent features that MS suggests, but at the underlying nature of MS. Williams contends: "The important thing about morality is its *spirit*, its *underlying aims*, and *the general picture of ethical life* it implies" (ELP 193; my emphasis). While the "obligation-out, obligation-in principle", or the inescapability of obligations may be the demanding productions of MS, they are not the necessary components of MS. Those features are just not the primary target of Williams' critique.

Ample textual support exists for this interpretation. Williams keeps noting that it is "the *pressure* within the morality system" or "(t)he *pressure* of the demand within the morality system" that requires the inescapability of moral obligations (ELP 194, 201; my emphasis). Also, his careful use of the words in the sentences such as "Morality *encourages* the idea, only an obligation can beat an obligation" (ELP 200; my emphasis) or "It (MS) starkly *emphasizes* a series of contrasts" (ELP 216; my emphasis), supports my reading. The inescapability of obligation is only one of the possible consequences that the MS *encourages* and not the nature of the MS itself. As Williams notes, MS "is not one determinate set of ethical thoughts. It embraces a range of ethical outlooks" (ELP 193).

Response: The nature of the morality system

To properly consider the validity of Williams' discussion, we must identify the underlying *nature* of MS rather than being blinded by its stringent products. What is this? While it is a problem with Williams' presentation of the critique that Williams is not entirely clear about the nature of MS, which is a major cause of the opponents' misunderstanding of the scope of his critique, we can reconstruct the nature by reading Williams' paper on which the discussion of ELP is based.

When we consider what drives us to conceptualize obligation as a practical conclusion, we can see that the MS' notion of *blame* is at work. Consider the feature of MS' obligation that "(a)n obligation applies to someone with respect to an action—it is an obligation to do something—and *the action must be in the agent's power*" (ELP 194; my emphasis). This feature is produced by the thought that one is blameworthy for failing to comply with an obligation and, at the same time, one is blameworthy only for what one is *able* to do. The "obligation-out, obligation-in principle", according to which a more general obligation is required to cancel an obligation, is also motivated by the notion of blame: blame for failing to fulfill an obligation is excusable if one is *unable* to fulfill two obligations, but only the more

general obligation.[10] As these considerations show, the underlying basis of the MS' stringent property of obligation is its notion of blame.[11] In an earlier paper on moral obligation, Williams states the point clearly:

> The important point is that we do not first have a determinate notion of moral obligation (in the wider sense) to which the notions of blame and related reactions are added. The class of moral obligations, in the wider sense, just *is* the class of *oughts* about an agent's actions to which blame and similar reactions are added. (Williams 1981b, 121)

Here the notion of moral obligation "in the wider sense" is precisely the MS' notion of obligation, in which moral obligation is understood as a practical *ought* (Williams 1981b, 118–21). Williams thus clearly claims that the MS' notion of obligation is constituted by the notion of blame.

As seen in the first section, the MS' notion of blame asserts that one is only blamable for what one could have voluntarily done, and that the voluntariness thus given should be philosophically scrutinized as, typically, free will or alternate possibilities. Indeed, this strong focus of voluntariness in blameworthiness motivates features of MS such as the "*ought* implies *can*" principle in which one is not blameworthy for what one cannot do. Hence, the nature of MS is primarily characterized as a *powerful commitment to voluntariness in blameworthiness*.

A paradigm example of MS' commitment to voluntariness is Galen Strawson's argument about the impossibility of moral responsibility. In its simplest form, Strawson's "basic argument" can be summarized as follows (Strawson 1994, 5):

> P1: Nothing can be causa sui *i.e.* nothing can be the cause of itself.
> P2: To be truly morally responsible for one's actions, one must be causa sui, at least in certain crucial mental respects.
> Conclusion: No one is morally responsible.

[10] To fully understand how the principle is based on the MS' notion of blame, we need to consider the connection between blame and reason, as we discuss in detail in the next section. Blame implies that one has a strong reason to have acted otherwise; in the context of obligation, the strong reason to comply with the obligation. This blame is only excusable if one has a stronger reason to have acted in that way, *i.e.* if one had a stronger obligation.

[11] Unlike other commentators who focus on obligation, Jenkins points out this primary role of blame in MS (Jenkins 2006, 70–71).

According to Strawson, causa sui means having brought about one's principles of choice (Strawson 1994, 6–7). In other words, to be truly responsible for our actions, we must freely choose our will to act in this manner. However, this line of thought falls into an infinite regress: to be responsible for the present will, I must have chosen the past will, and to be responsible for the past will, I must have chosen an earlier will, and so on. The deterministic interpretation of the basic argument is as follows: Since we cannot have chosen our first condition of being (e.g., a result of heredity), which conditions our present state, we cannot be held truly responsible for our actions (Strawson 1994, 7). Strawson argues that this view of true moral responsibility has been central to Western moral traditions (Strawson 1994, 8). Although Strawson's "basic argument" may be an exaggerated form of the MS' commitment, it seems true that the commitment to voluntariness underlies most modern moral thought. The obsessive commitment is a characteristic of MS.

Accordingly, Williams and modern philosophers have agreed upon that the *powerful commitment to voluntariness in blameworthiness* is at the heart of the modern morality. Given this point, we can address the scope objections by arguing that opponents missed the essential point of Williams' critique. It is not that MS is problematic *just because* it tends to have certain stringent features regarding obligation, but that MS is problematic because it has a certain commitment to voluntariness that is not tenable in naturalistic terms. What Williams primarily problematizes is the commitment, the nature of MS, which is at the heart of and encourages the stringent features of obligation.

The fact that opponents have missed this point is shown by the way they have failed to examine the MS commitment to voluntariness in blame. While Kantian philosophers have argued that the Kantian notion of obligation is not as stringent as MS, they have ignored the fact that Kant is most powerfully committed to MS' focus on voluntariness and free will. Kant's discussion of obligations (the categorical imperative) is intended to secure the possibility of *free will* outside the realm of causality and inclination.[12] Moral theorists also seem to be pre-theoretically committed to MS focus on voluntariness, although they may have provided a less stringent theory of obligation than the "obligation-out, obligation-in principle". The principle of alternate possibilities has been the shared principle of moral responsibility, and few moral theories have challenged this principle. This also applies

[12] Indeed, Kant argues that will under the categorical imperative and free will are "one and the same" (Kant 2011, 123).

to dominant moral cultures: in pre-theoretical moral cultures, people have a strong intuition that actions out of free will are only blameworthy, whereas they are less often overwhelmed by the inescapability of moral obligation.

In summary, the scope objections missed the real target of Williams' critique, and therefore, Williams' critique is still substantial. We need to critically examine the commitment to voluntariness that underlies modern moral thought, since it naturally leads not only to demanding conceptions of obligation but also, and more importantly, to the naturalistically untenable view of blameworthiness.

3. The point of morality objection to Williams

Although we recognize that MS' commitment to voluntariness is the real target of Williams' critique and can address the scope objections, the next question is whether it is really problematic. Here, Williams' critique faces another repeated objection: opponents argue that it is the *point of morality* to secure fairness that we conceptualize blameworthiness with the commitment. Let us call this the *point of morality objection*, according to which MS' strong commitment to voluntariness is justified as a point of morality to secure fairness. In this section, I address this objection by arguing that the commitment of MS, in reality, defeats the point of blame.

The point of morality objection: The demand of fairness

Robert Louden points out that the core differences between Kant and Williams "come down to Kant's robust commitment to a strong sense of free will and practical reason, and Williams' firm rejection of both" (Louden 2007, 127):

> In my view, Williams' desire to decouple ethics from the presupposition of free will would, if widely adopted, be a disaster, not only for the morality system and Kant's ethics but for any ethical outlook worth taking seriously. (Louden 2007, 127)

Louden points out here that any decent morality must have the MS' commitment to free will; otherwise, it would become exceedingly difficult to convincingly distinguish cases where agents were responsible from those where they were not (Louden 2007,

127–8).[13] Indeed, the demand for fairness in blameworthiness has led many philosophers to purify voluntariness, an important concern of which is that blameworthiness would be a matter of *luck* if not based on strong voluntariness (Sher 2015, 179–180; ELP 217; Queloz 2022, 197–8). To ensure the immunity of luck in blameworthiness, we need a view in which an agent has real choices of action; the necessary condition for blameworthiness must be that an agent has free will, understood as the ability to act otherwise. Thus, the principle of alternate possibilities is justified by the demand for fairness, which ensures luck's immunity to blameworthiness, according to the proponents of MS.

In summary, the *point of morality objection* claims that the commitment of the MS—the purification of voluntariness to the extent that alternate possibilities or free will are conceptually necessary—is justified by the demand for fairness. Without this commitment, morality would never be fair.

Response: Nihilistic ideal of the morality system

The Williamsian response to the *point of morality objection* would be twofold: first, we can make a naturalistic sense of the sufficiently fair practice of blame; and second, MS' commitment to voluntariness is so strong that it defeats the very practice of blame. In other words, the MS' ideal of voluntariness is *nihilistic*; we will deglamorize *adequately-voluntary* actions once we attempt to idealize *ultimately-voluntary* actions as MS demands. This is a problem for defenders of MS.

As for the naturalistic understanding of the practice of blame, we have already seen a summary of Williams' account: the system of blame is a device for securing coexistence which functions as far as we have an adequately fair attribution of blame. This is secured through certain requirements of voluntariness. Indeed, Williams has pointed out that there are requirements of self-knowledge and intention that one knew what one was doing and one intended the action (ELP 215–6), or, in the later writing, Williams asserts that a voluntary action is an intended action "which has no inherent or deliberative defect" (Williams 1995b, 25). In fact, as in political and legal cases, we have a certain sense of freedom, such as a distinction between persuasion and manipulation, without committing ourselves to a metaphysical

[13] Louden also ascribes to Williams a complete skepticism about free will or an agent's ability to act otherwise (Louden 2007, 127–8), based on a reading of Williams' desire-based picture of action. But I will leave this issue aside since Williams never proposes such a view in his text, and certainly attempts to secure *freedom*, though not free will in the strict MS' sense, in various texts (e.g. Williams 2006b).

discussion of free will. What is needed, then, for a Williamsian naturalistic account is an examination of such practical conditions of voluntariness to ensure fairness.[14]

However, the question is, why are these conditions sufficient? Proponents of MS would argue that we need a stronger requirement for voluntariness, such that one must have a real possibility of acting otherwise, if we are to make one blameworthy. Here, Williams needs an *additional explanation* against the MS' requirement. Williams seems to provide this additional explanation hinted at in ELP and argued at length in later writings. This explanation consists of two considerations on the practice of blame and human motivation and amounts to showing that MS' requirement cannot make sense of the proper working of blame by which we *change* through blame.

The first step is to consider the working of *reason-statements* implied in the practice of blame. As Williams points out, blame "seems to have something special to do with the idea that the agent had a reason to act otherwise" (ELP 214). When we blame someone for acting in a certain way, we imply that the person has a *reason* to act otherwise. Indeed, much of the time, this statement is straightforwardly true as they are motivated to act otherwise.[15] I may have been irritated by small things, even though I was motivated not to do so; straightforwardly, I had more reason to act differently than I did. Here, blame acts as a reminder that I had reason to act differently. However, if we carefully examine the workings of reason statements, we find various ways other than this straightforward case. Most importantly, Williams points to the case of "a proleptic invocation of a reason," in which a blamed agent is retrospectively given a reason to have acted otherwise by the very fact of being blamed (Williams 1995c, 41–2). For example, even if I have no direct motivation to hide my irritation, I get a reason for having done so when I am confronted with blame from someone I respect and love.

An important point regarding proleptic blame is that, as a practice of mutual modification of behavior, this kind of *transformation* of agents is essential to blame. Indeed, we *change* ourselves through the experience of being blamed. The most robust case is blaming children, where they had a very strong motivation to behave as they did, but the blame itself provides the reason for behaving otherwise. Not only children but also adults learn and share an ethical understanding through the proleptic work of blame. However, this kind of proleptic working of blame cannot be made sense, given

[14] In fact, Williams commits himself to such a project elsewhere in later writings (Williams 1993; Williams 1995c; Williams 2006a; Williams 2006b).

[15] On the Williams' argument for the necessary connection between the truth of reason-statement and an agent's motivation, see Williams 1981a.

the MS' requirement; we cannot blame someone when, at the very time of action, they have no reason or motivation to act otherwise. One may not have *the capacity to act otherwise* because of a strong motivation to act in a certain way or a lack of cognitive understanding that acting otherwise is good. For example, an ingrained racist may have no reason or ability to stop making racist remarks. Nevertheless, we have a point in blaming the racist because blaming might proleptically provide a reason to stop racist remarks. The MS' commitment to voluntariness overlooks this working of blame.

To generalize the case of the racist, we obtain the second consideration that helps to favor the naturalistic account over MS. This is a consideration of *practical necessity*, where one has such a strong motivation to act in a certain way that no other option is available (ELP 209). We cultivate certain moral characters that often determine our actions, because the actions are essential to the characters (Williams 1995a, 17). For example, a determined liberal's vote against a monstrous anti-liberal president is largely determined by her character and moral identity, and it was not obvious that the liberal had the capacity to act otherwise. Her line of thinking is literally "I cannot do otherwise." However, this moral phenomenon of practical necessity seems compatible with the practice of blame and moral responsibility; the liberal would be morally responsible for her actions. This is also true of blame, as seen in the case of racism. If someone acts unethically out of practical necessity, there is a point of blame. We hope that persons will *change* their character through blame, as seen in cases of proleptic blame. However, the demand of MS is too strong to allow this possibility. On MS' account, if one's action is determined, then that action cannot be considered voluntary and therefore never blameworthy; on the Williams account, however, a proleptic blame for action of practical necessity has the point that one will be motivated to reconsider and *change* one's character.

Blame, like obligation, is essentially an ethical practice of living together in a society in which each member has very different motivations and does not know each other's motivations. That we are different is very important; given one's motivation or character, one's actions can be practically necessary. To make sense of blame as the art of coexistence in such a society, we must accommodate the possibility of ourselves changing through blame. The MS' requirement of voluntariness prevents us from accommodating such a possibility: we are never blameworthy unless we have a

real choice to act differently. The MS' requirement suggests that we never really *change* through blame; blame only reminds us of what we already have.[16]

While making sense of blame as changing each other, as shown in the proleptic blame to practical necessity, already gives us one reason to criticize the MS' requirements, we have another well-known reason to do so. That is, the naturalistic picture of human action suggested by modern science discourages us from presupposing the real choice to act differently all the time.[17] In a naturalist perspective, motivation and character often, if not always, determine one's actions. If we take the naturalist picture seriously and still accept the MS' requirements, the consequence is the impossibility of blame and moral responsibility, as pictured by Galen Strawson. In this context, MS is *nihilistic*; the proponents of MS attempt to secure a fair basis of moral responsibility but destroy moral responsibility altogether. As long as we presuppose MS, this is a real threat. But the good news for those who are concerned about this threat is that there is no need for the practice of blame to be understood with the commitment of MS. As Williams eloquently puts it: "(j)ust as there is a "problem of evil" only for those who expect the world to be good, there is a problem of free will only for those who think that the notion of the voluntary can be metaphysically deepened" (Williams 1993, 68).[18] Instead of such a nihilistic demand on blame, we have a naturalistic account of the practice and a tradition of thinking about the fair condition of blameworthiness.

Conclusion

In this paper, I defend Bernard Williams' critique of modern morality, which Williams calls the morality system, against two major objections. In the first section, I

[16] Williams points out that this line of thinking places the metaphysical demand on Kant and Kantians to show that ethical reasons are necessarily available to every human agent: "if morality is to limit genuinely moral comment to blame, and blame to what is available to the rational deliberator, it is faced with a vast epistemological demand, to show (as Kant appropriately thought) that the correct ethical demands are indeed available to any rational deliberator—as Kant was disposed to put it, 'as such'" (Williams 1995e, 246).

[17] In this sense, Kant has been consistently faithful to MS in attempting the project of conceptualizing an agency outside the natural world, which Williams calls "the more extravagant metaphysical luggage of the noumenal self" (ELP 73).

[18] The problem of free will, which has preoccupied many modern philosophers, is a distinctive product of MS. This Williamsian view of the free will debate is called "free will pessimism" and is defended by Paul Russell. See Russell 2017.

summarize Williams' critique in the way it has been widely understood by commentators. The critique is about the unnaturalistic understanding of obligation and blame. In the second section, I address the *scope objection* to Williams, according to which Williams' critique—especially the critique of the concept of obligation—is only applicable to extreme morality. In response, I argue that the critique is directed at commitment to voluntariness rather than at the conception of obligation, which is only a possible product of commitment. Indeed, the commitment that one can only be blamed for what one is capable of doing otherwise is widespread in moral philosophy, as well as in pre-philosophical modern moral cultures. In the third section, I answer the *point of morality objection* that the commitment of the MS to voluntariness is essential for satisfying the fairness of blameworthiness. I argued that this commitment is so demanding that it defeats one of the important purposes of blame as an ethical practice in agent modification. Although blame should function as a modification of an agent, the MS' requirement for fairness thwarts the functioning of blame. By answering the two objections, I have clarified the nature of the modern morality and why it should be criticized.

Williams concludes his *Ethics and the Limits of Philosophy* by saying that the MS' "philosophical errors are only the most abstract expressions of a deeply rooted and still powerful *misconception of life*" (ELP 218; my emphasis). In this paper, I defended this conclusion by showing specifically how the philosophical errors of MS express the misconception of life; MS contains the misconception of human motivation, the real workings of the arts of coexistence, and the fairness required. The next question is whether we can convincingly reconstruct ethics in a way that is not based on such a misconception but on the naturalistic understanding of the human art of ethics. An example of such a work would be a more detailed theory of blame, in which the conditions of fairness, though not the misdirected fairness of the MS, are sufficiently protected. What we need, I think, is to free ourselves from illusions of the morality system and take this project of naturalistic reconstruction seriously.[19]

References

[19] An earlier version of this paper is a Master's thesis at the University of Edinburgh. I would like to thank David Levy, Anju Nakane, and Yoshiki Yoshimura for their insightful comments on the dissertation. I also thank three anonymous referees for their comments on this paper.

Chappell, Sophie-Grace & Smith, Nicholas (2023). “Bernard Williams.” *The Stanford Encyclopedia of Philosophy* (Summer 2023 Edition), Edward N. Zalta & Uri Nodelman(Eds.),URL=<https://plato.stanford.edu/archives/sum2023/entries/williams-bernard/>. (accessed 2023-08-25).

Clark, Maudemarie (2015). “On the Rejection of Morality Bernard Williams’s Debt to Nietzsche.” In *Nietzsche on Ethics and Politics*. Oxford U. P. USA., 41–61.

Darwall, Stephen (1987). “Abolishing morality.” *Synthese,* 72 (1), 71–89.

Frankfurt, Harry (1988). “Alternate possibilities and responsibility.” In *The Importance of What We Care About*, Cambridge U. P., 1–10.

Griffin, James. (2015). “Rejecting ‘Morality’.” In *What Can Philosophy Contribute To Ethics*? Oxford U. P., 93–104.

Jenkins, Mark (2006). *Bernard Williams,* Routledge.

Kamm, Frances (1985). “Supererogation and Obligation.” *The Journal of Philosophy*, 82(3), 118–138.

Kant, Immanuel (2011). *Groundwork of the metaphysics of morals: a German-English edition*, Mary Gregor & Jens Timmermann (Eds.), New York: Cambridge U. P.

Leiter, Brian (2022). “Bernard Williams’s Debt to Nietzsche: Real or Illusory?” In Andras Szigeti & Matthew Talbert (Eds.), *Morality and Agency: Themes from Bernard Williams*, Oxford U. P., 17–34.

Louden, Robert (2007). “The Critique of the Morality System.” In Alan Thomas (Ed.), *Bernard Williams (Contemporary Philosophy in Focus)*, Cambridge U. P., 104–134.

Nietzsche, Friedrich (1998). *On the Genealogy of Morality*, Maudemarie Clark. & Alan J. Swensen (Trans.), Hackett Publishing.

Queloz, Matthieu (2022). “A Shelter from Luck: The Morality System Reconstructed.” In Andras Szigeti & Matthew Talbert (Eds.), *Morality and Agency: Themes from Bernard Williams*, Oxford U. P., 182–209.

Russell, Paul (2017). “Free Will Pessimism.” In David Shoemaker (Ed.), *Oxford Studies in Agency and Responsibility*, Volume 4. Oxford U. P., 93–120.

Sher, George (2005). “Kantian fairness.” *Philosophical Issues* 15 (1), 179–192.

Skorupski, John (1998). “Morality and ethics.” In *The Routledge Encyclopedia of Philosophy*.
URL=<https://www.rep.routledge.com/articles/thematic/morality-and-ethics/v-1.> (accessed 2023-08-23)

Strawson, Galen (1994). “The impossibility of moral responsibility.” *Philosophical Studies* 75 (1-2), 5–24.

Van Inwagen, Peter (1975). “The Incompatibility of Free Will and Determinism.” *Philosophical Studies*, 27 (3), 185–199.

Williams, Bernard (1981a). “Internal and external reasons.” In *Moral Luck: Philosophical Papers 1973–1980*, Cambridge U. P., 101–113.

Williams, Bernard (1981b). “*Ought* and moral obligation.” In *Moral Luck: Philosophical Papers 1973–1980*, Cambridge U. P., 114–123.

Williams, Bernard (1985/2011). *Ethics and the Limits of Philosophy* (Routledge Classics), Routledge.

Williams, Bernard (1993). *Shame and Necessity*, University of California Press.

Williams, Bernard (1995a). “How free does the will need to be?” In *Making sense of humanity and other philosophical papers 1982-1993*, Cambridge U. P., 3–21.

Williams, Bernard (1995b). “Voluntary acts and responsible agents.” In *Making sense of humanity and other philosophical papers 1982-1993*, Cambridge U. P., 22–34.

Williams, Bernard (1995c). “Internal reasons and the obscurity of blame.” In *Making sense of humanity and other philosophical papers 1982-1993*, Cambridge U. P., 35–45.

Williams, Bernard (1995d). “Nietzsche’s minimalist moral psychology.” In *Making sense of humanity and other philosophical papers 1982-1993*, Cambridge U. P., 65–76.

Williams, Bernard (1995e). “Moral luck: a postscript.” In *Making sense of humanity and other philosophical papers 1982-1993*, Cambridge U. P., 241–247.

Williams, Bernard (2000). “Naturalism and Genealogy.” In Edward Harcourt (Ed.), *Morality, Reflection, and Ideology*, Oxford U. P., 148–159.

Williams, Bernard (2006a). “Values, Reasons, and the Theory of Persuasion.” In *Philosophy as a Humanistic Discipline*, Princeton U. P., 109–118.

Williams, Bernard (2006b). “Moral Responsibility and Political Freedom.” In *Philosophy as a Humanistic Discipline*, Princeton U. P., 119–125

