SCIO. Revista de Filosofía, n.º 21, Noviembre de 2021, 113-148, ISSN: 1887-9853

FROM NEUROETHICS TO NEO-ROMANTICISM
ALDOUS HUXLEY IN RESPONSE TO CURRENT PROPOSALS FOR ETHICAL AND LEGAL REGULATION OF NEUROSCIENCE

DE LA NEUROÉTICA AL NEO-ROMANTICISMO
LA RESPUESTA DE ALDOUS HUXLEY A LAS PROPUESTAS ACTUALES PARA LA REGULACIÓN ÉTICA Y LEGAL DE LA NEUROCIENCIA

Luis Enrique Echarte Alonso[a*]

Fechas de recepción y aceptación: 19 de enero de 2021 y 20 de octubre de 2021

DOI: https://doi.org/10.46583/scio_2021.21.783

Abstract: The neuroethics field emerged in the early 2000s in an effort to face important philosophical dilemmas and anticipate disruptive social changes linked to the use of neurotechnology (Safire, 2002). From very early on, this field grew out of two core issues, namely inquiries into the ethics of neuroscience –concerning the moral use of knowledge and technology– and inquiries into the neuroscience of ethics –on how new brain function evidence can change human self-understanding (Roskies 2002). Similarly, neurolaw is now on a parallel path with two main pillars as Chandler (2018) suggested, (1) "self-reflexive inquiry" (the neuroscience of law) and (2) "inquiry into the development and use of brain science and technologies" (the law of neuroscience). In this paper, I suggest that these two lines of research are still excessively disconnected from one another and, to support this claim, I analyze the three potential point-of-no-return risks that Aldous Huxley associated with technological challenges, namely centralization of power, bureaucratic alienation, and scientific idealism. In addition, I show how Huxley shifted analysis of

[a] Unidad de Humanidades y Ética Médica. Facultad de Medicina. Universidad de Navarra.

[*] Correspondencia: Universidad de Navarra. Facultad de Medicina. Calle Irunlarrea s/n. 31008 Pamplona (Navarra). España.

E-mail: lecharte@unav.es

technological problems from a focus on the rights of potential victims to the duties of potential aggressors. Finally, I argue that Aldous Huxley's view on how to build a bridge that brings pillars 1) and 2) closer together also helps prevent the technological point-of-no-return. According to Huxley, the key is found in paying particular attention to understanding contemplative activity, reinforcing its role in the study of reality, and, eventually, returning the romantic gaze updated to academia.

Keywords: Neurolaw, Aldous Huxley, Neurotechnology, Centralization of power, Bureaucracy, Scientific idealism, Contemplation, Romanticism

Resumen: El campo de la neuroética surge a principios de la década de 2000 en un esfuerzo por enfrentar dilemas filosóficos relevantes y anticipar cambios sociales disruptivos vinculados al uso de la neurotecnología (Safire, 2002). Desde muy pronto, este campo fue constituido en torno a dos cuestiones centrales, a saber, las investigaciones sobre la ética de la neurociencia –en relación con el uso moral del conocimiento y la tecnología– y las investigaciones sobre la neurociencia de la ética –sobre cómo la nueva evidencia de la función cerebral puede cambiar la autocomprensión humana (Roskies 2002)–. De manera similar, el campo del neuroderecho se encuentra ahora en un camino paralelo y, como señala Chandler (2018), se asienta en dos pilares principales: (1) "la indagación autorreflexiva" (la neurociencia del derecho) y (2) "la indagación sobre el desarrollo y uso del cerebro ciencia y tecnologías" (la ley de la neurociencia). En este artículo, sugiero que estas dos líneas de investigación todavía están excesivamente desconectadas entre sí y, para respaldar esta afirmación, analizo los tres riesgos potenciales de *punto de no retorno* que Aldous Huxley asocia a los desafíos tecnológicos, a saber, la centralización de poder, la alienación burocrática y el idealismo científico. Además, muestro cómo Huxley cambia el enfoque habitual: traslada el análisis de los derechos de las potenciales víctimas al análisis de los deberes de los potenciales agresores. Finalmente, sostengo que la visión de Aldous Huxley sobre cómo construir un puente que acerque los pilares 1) y 2) también ayudaría a prevenir el *punto de no retorno* tecnológico. Según Huxley, la clave está en comprender mejor la actividad contemplativa, en reforzar su papel en el estudio de la realidad y, finalmente, en devolver una renovada visión romántica a la academia.

Palabras clave: neuroderecho, Aldous Huxley, neurotecnología, centralización del poder, burocracia, idealismo científico, contemplación, romanticismo.

The two most powerful warriors are patience and time.

Leo Tolstoy

§1. Introduction: Is a more preventive approach necessary?

At present, neuroscientific development is challenging our understanding of human behavior. New findings and technological advances will soon play an important role in the promotion of lifestyles and social policies, and may already be taking place, especially in advertising (i.e., neuromarketing) and education (for example, the new high-impact field of educational neuroscience). One sign of the transformative power of neuroscience is increased use of neuroimaging evidence in court. As seen in Chandler's paper (2015), Canadian judges seem increasingly persuaded by the idea that, "Brain damage mitigates moral blameworthiness" and, as a result, calls for criminal justice system reform are on the rise there. This context helps explain why neurolaw, as an interdisciplinary research area, is now attracting investment, significant media attention and on its way to institutionalization through monographs dedicated to the topic in academic journals and training courses at universities. Although it started as a subspecialty of neuroethics, neurolaw has transitioned to a relatively independent sister field.

In this context, many claims have emerged about the necessity for legal protection against misuse and abuse of neurotechnology, for instance, from the Morningside Group, which is made up of neuroscientists, clinicians, ethicists, artificial intelligence (AI) engineers, and others. Because of the importance of the issues addressed, they aim to add neurorights to the World Medical Association's Declaration of Helsinki in five specific areas as follows: privacy, consent, agency and identity, medical enhancement, and cognitive bias (Yuste et al. 2017). The Morningside Group seeks to protect citizens against the misuse of neuroscience from a mostly reactive stance after detecting the

presence of threats. In this paper, I argue that it is better to privilege a preventive approach over a corrective one and, with this in mind, I will use Aldous Huxley's approach to neurolaw since it is distinctly integrative and challenges current political and academic views.

§2. The point of no return

Beyond the recent boom of neurolaw, it is fruitful to look back at its forerunners. Well before neologism was created in 1991 (in a text from J. Sherrod Taylor and collaborators about traumatic brain injuries in criminal litigation), British philosopher and writer Aldous Huxley first offered profound insight related to it. The most explicit reference appears in his last novel, *Island*, published in 1962, in which Dr Robert, one of the main characters, reveals the central idea that runs through Huxley's utopia: "I'd like to write a little book on human physiology in relation to ethics, religion, politics and law" (Huxley 1962, 155). However, it is in his most famous text, *Brave New World*, where Huxley sketches his main ideas on this interdisciplinary topic. There, he imagines a dystopian society that hardly even needs laws given the control that biotechnological weapons exert over the population. Citizens live happily in obedience to pure power because, as Aroso (2018) recalls, no alternative possibility for confronting wishes and expectations exists, at least on topics that matter because, even if there were numerous laws designed to protect individual rights, citizens would not feel the need to appeal to them.

Fourteen years later, in *Science, Liberty and Peace*, Huxley related the political reality at the time: "In the past, personal and political liberty depended to a considerable extent upon governmental inefficiency. The spirit of tyranny was always more than willing; but its organization and material equipment were generally weak. Progressive science and technology have changed all this completely. Today, if the central executive wishes to act oppressively, it finds an almost miraculously efficient machine of coercion standing ready to be set in motion" (Huxley 1947a, 6). More specifically, another twelve years later, he consolidated his ideas about the two main political dangers contained in expansive scientific research. The first relates to propaganda in a mass communications industry. "A society, most of whose members spend a great part

of their time, not on the spot, not here and now and in their calculable future, but somewhere else, in the irrelevant other worlds of sport and soap opera, of mythology and metaphysical fantasy, will find it hard to resist the encroachments of those would manipulate and control it" (Huxley 1958a, 56). Indeed, many of the current worst-case scenarios around neuroscience have to do with its use to influence people. As Taber et al. (2013) note, progress is being made in this direction. "The explosion of findings in automaticity and brain-imaging research relevant to opinion formation represents a strong step in the direction of explaining what we believe with how we think" (Taber et al. 2013). In line with this, many of the neurorights proposed today aim to protect vulnerable groups, like consumers, electorates, employees, scholars, prisoners, etc. As a counterpoint, however, Huxley thinks that it is more important to develop laws to prevent neuro-propaganda from turning politicians, businessmen and teachers into the worst type of despots.

The second danger relates to government control of emotions through psychopharmacology. "The dictatorships of tomorrow will deprive men of their freedom, but will give them in exchange a happiness nonetheless real, as a subjective experience, for being chemically induced. The pursuit of happiness is one of the traditional rights of man; unfortunately, the achievement of happiness may turn out to be incompatible with another of man's rights –namely, liberty" (Huxley 1958b, 152). Huxley argues that scientific-technological development related to propaganda and emotional control should be subject to strict legal control, similar to monitoring of nuclear energy internationally. Nonetheless, Huxley doubts that any measure will be effective enough without paying attention, first of all, to three factors related to the aforementioned point of no return, namely the centralization of political power, bureaucratization of society, and scientific idealism. More importantly, he attributes, as we will see below, an important role to science in the prevention of such evils.

§3. Centralization and bureaucracy

If the mandate of democratic institutions is to mitigate the accumulation (centralization) of power, they would do well to monitor and regulate the development of neurotechnology. Scientists can help in this endeavor, for

example, by studying the level of power at which temptation towards corruption becomes irresistible and, therefore, help avoid developing a breeding ground for tyrants in all hierarchically organized groups. Empirical research on tolerance for increased power would be, without a doubt, one of the main topics of such research, and far less hypothetical than previously believed. Suffice it to explore, for example, Guinote's 2017 article, or Maasen and Sutter's larger 2007 monograph.

Secondly, Huxley warns against the bureaucratization of society– the greatest cradle of tyrants, "at the head of a highly trained and perfectly obedient bureaucracy" (Huxley 1938, 64)– but also, and more specifically, against the bureaucratization of science. Because scientific and technological development demands increasingly complex and hierarchical organizations, the pursuit of knowledge and technology may pose a danger in terms of disconnection between fields, as well as of alienation from their final practical applications. In turn, scientific work becomes meaningless "and entirely divorced from the world of nature" (Huxley 1947a, 23). This scenario encourages politicians and businessmen to establish scientific monopolies that reinforce their power and, what is worse, to mask the use of science in the pursuit of partisan interests.

Today, *Research Ethics Committees* and, more broadly, *Science Quality Systems* represent the strongest legal response to the divorce between science, industry, and government. In many countries, their existence is mandatory. However, in Huxley's view, they are only useful to the extent that they train researchers on ends and means and on good and bad uses of potential findings. Otherwise, these sorts of institutions end up making matters worse. Committees become corrupted and corrupting when they singularly focus on monitoring abuses and resolving ethical and legal conflicts because they increasingly tempt researchers both to surrender their moral judgment to the committee's assessment and/or to view that assessment as a merely bothersome formality to overcome through whatever means necessary. Not surprisingly, for some years now, critical voices have emerged against the bureaucratization of committees and the paralysis and centralization of science (Wald 2004, Charlton 2010, van den Hoonaard & Hamilton 2016).

Huxley repeatedly insists that scientists cannot afford to delegate the responsibility they have when it comes to the use of scientific knowledge. On

the contrary, they should use science to, for example, research the impact of bureaucracy on the human mind—i.e., how it affects decision-making, the levels needed to avoid passive or uncommitted attitudes, individualistic habits, antisocial behavior, etc. This is no longer just science fiction. Preliminary studies on this topic include, for instance, Bolfikova et al. 2010 and Nørgaard 2018. Yet, there are easier and faster ways to help governments promote safe levels of bureaucracy. Huxley encourages scientists to devote time and effort to making knowledge more intelligible to the overall population, an initiative that should be a professional duty regulated by law (Huxley 1963, 61-65). Bringing science closer to society in all possible ways avoids the unpleasant surprises associated with the development of technology. "This has happened again and again in history. Technology has advanced and this changes social conditions, and suddenly people have found themselves in a situation which they didn't foresee and doing all sorts of things they didn't really want to do" (Huxley 1958c). But as noted, things are different with respect to the past. We are now nearer to the point of no return and measures to help prevent a new *Tower of Babel* (among scientists, but also among media and the general public) seem more urgent than ever. At the end of this article, I will go into more detail about Huxley's idea on the role of popular science in scientific and human progress.

§4. Idealism, dogma and violence

In *Brave New World*, Huxley aptly captured the image of a human-machine society—individuals immersed in an ocean of bureaucracy with repetitive and rootless jobs– that lacks freedom and, curiously, lacks tyrants. This point leads us to the third technological danger I want to point out here, namely scientific idealism. The argument goes as follows: scientific rationality is essential for improving and managing human life. However, this tool can turn against society if, as Huxley warns, the empirical approach becomes the only way to understand reality. Excessive compartmentalization of science, which, as we have seen, is caused, among other factors, by bureaucracy, is a defining feature of hyper-specialists, namely scientists exceptionally learned in a particular

scientific field, but quite ignorant otherwise. This new type of scientist, characterized by neglecting anything beyond method, eventually moves science from methodological reductionism to ontological reductionism. For Huxley, it seems inevitable that excessively partial views of human reality generate increasingly fictitious coherences that lead to self-sufficient, satisfying truths. These coherences are, he jests, better than *soma*, "a holiday from reality whenever you like" (Huxley 1932, 54). All other forms of practical reasoning that fall outside the scope of the scientific paradigm are maligned as meaningless babble that must be silenced. In turn, the temptation to use propaganda, mental conditioning, and other methods of social engineering to homogenize critical thinking inevitably grows. Here, in the technologized society of Huxley's tomorrow, the presence of social dogma, rather than the will of a tyrant, sustains the prison walls.

Many prominent intellectuals have warned against the advent of scientific idealism. For example, two years before the publication of *Brave New World*, the Spanish philosopher José Ortega y Gasset wrote, in *The Revolt of the Masses,* about a similar relationship between hyper-specialization and violence: "It is not a question of the mass-man being a fool. On the contrary, today he is more clever, has more capacity of understanding than his fellow of any previous period... This is what in my first chapter I laid down as the characteristic of our time; not that the vulgar believes itself super-excellent and not vulgar, but that the vulgar proclaims and imposes the rights of vulgarity, or vulgarity as a right… European history reveals itself, for the first time, as handed over to the decisions of the ordinary man as such. Or to turn it into the active voice: the ordinary man, hitherto guided by others, has resolved to govern the world himself. This decision to advance to the social foreground has been brought about in him automatically when the new type of man he represents had barely arrived at maturity. If the psychological structure of this new type of mass-man is studied from the viewpoint of what concerns public life, what we find is as follows: (1) An inborn, root-impression that life is easy, plentiful, without any grave limitations; consequently, each average man finds within himself a sensation of power and triumph which, (2) invites him to stand up for himself as he is, to look upon his moral and intellectual endowment as excellent, complete. This contentment with himself leads him to shut himself off from any

external court of appeal; not to listen, not to submit his opinions to judgment, not to consider others′ existence. His intimate feeling of power urges him always to exercise predominance. He will act then as if he and his like were the only beings existing in the world and, consequently, (3) will intervene in all matters, imposing his own vulgar views without respect or regard for others, without limit or reserve, that is to say, in accordance with a system of 'direct action'" (Ortega y Gasset 1930, 77 and 107).

While long, this quote is worth including in its entirety because, in addition to being in line with Huxley's ideas, it is among the most lucid texts on the consequences of hyper-specialization. Both authors fear that perverted specialists will become incapable of understanding discourse outside of their field, especially if the discourses are interdisciplinary, like the many that include ethics and law. Worse still, these failed researchers experience difficulties in following wider thought as clear evidence of wisdom. To them, non-analytical ethicists and lawyers are nothing but incompetent scientists who jump too much between different ideas without sufficiently connecting them. Therefore, introducing ethical and political discourse in their forums is seen as damaging to science's reputation– a superficial and ideology-free area. Ortega y Gasset makes it clear that hyper-specialist attitudes– by definition, violent and gregarious–increase the gap between the interests of society and those of narcissistic academia. To avoid this dark, but not so distant, horizon, he invites scientists to defend true science, individual freedom, and social progress with weapons like the promotion of history, culture and open dialogue.

Huxley, for his part, points to three signs of transformation toward a scientific mass society: (1) the replacement of pure science with applied science in academia, (2) the hegemony of utilitarianism in the management of social regulation, and (3) the reduction of human progress to social stability, a political objective that begins to subordinate all other political aims. "The people who govern the *Brave New World* may not be sane (in what may be called the absolute sense of the word); but they are not madmen, and their aim is not anarchy but social stability. It is in order to achieve stability that they carry out, by scientific means, the ultimate, personal, really revolutionary revolution" (Huxley [1932] 1946, xii). By contrast, and as a remedy, Huxley makes an appeal to scientists' duty to acquire interdisciplinary training, especially

in the humanities (Echarte 2020a), which would help them avoid ontological reductionism and, in particular, hasty or naïve neuroessentialism.

In short, international and national policies would do well to help reinforce professional commitment to interdisciplinary training. Complementary to that, scientists could advise lawmakers by studying specialization tolerance levels, in a way similar to that mentioned regarding bureaucracy. Last but not least, specific regulatory policies could reinforce scientific research on how to avoid the point of no return. This support is particularly important in research where consistent results are, in most cases, only obtained in the long term.

§5. Putting machines in charge

Despite the above, Huxley likely had some qualms about the introduction of these preventive laws. He was largely influenced by Henry Thoreau's thought, which conceived of legal systems as a matter of education and, therefore and eventually, largely expendable, at least in a correctly educated, utopian society. For both authors, too many legal initiatives lead to overly restricted and inefficient societies, even when done in the name of happiness and freedom. Thoreau and Huxley share the same suspicions of governments that seek to replace people's understanding with laws. They are also suspicious of the citizen wishing to "resign his conscience on the legislator… The mass of men serve the state thus, not as men mainly, but as machines, with their bodies" (Thoreau 1849, 5-6). Of note, Thoreau also uses, like Huxley and Ortega y Gasset, the terms mass and machine to criticize the new man born of runaway technological progress.

Huxley is far less radical than Thoreau when it comes to the desirability of laws. This is seen in the fact that, unlike in *Brave New World*, *Island*'s utopia contains a legal code. "–Law... I was just going to ask you about law. Are you absolutely swordless and punishmentless? Or do you still need judges and policemen? –We still need them... But we don't need nearly so many of them as you do" (Huxley 1962, 155). He refers here to laws that make it possible for citizens to freely live in harmony with one another, with nature, and with themselves. Besides, for Huxley, these laws should enable them to understand

the entire context in which rules are framed, truly capturing the beauty of binding acts and, thus, to observe the law in spirit and truth. The opposite scenario lacks both understanding and beauty, and instead touts a society governed by purely processual laws with citizens who obey them with machine-like detachment. In this sort of state, the government can be made up of men and robots indistinctly and, for that very reason, the latter would probably end up prevailing because of their reliability and ever-increasing computation capacity. However, as Huxley cautions, "every gain has to be paid for. The automatic machine is fool-proof; but just because it is fool-proof it is also grace-proof" (Huxley 1947b, 196). The worst consequence of this loss would, therefore, be the enactment of laws and policies that encourage a kind of computational view in all spheres of social life. "The man who tends such a machine is impervious to every form of aesthetic inspiration, whether of human or of genuinely spiritual origin" (Idem). The transformation from human being to machine would reach its final stage without ever having to introduce a single screw or microchip into the organic body.

Controversies over the limits and differences between man and machine have grown over the last decade, and not just at the theoretical level. Autonomous weapons, carebots (robots specifically designed for healthcare purposes), or driverless cars are current examples of machines that, thanks to AI, can operate without full control on the part of humans. These robots' presence is increasing in everyday life and, with it, creating new problems to be addressed. The field of *robot ethics* aims to protect humans from the potential associated risks. For example, one potential and crucial threat involves unexpected circumstances and tough decisions: injuring a passenger, exposing a patient's medical records, lying to an Alzheimer's patient, etc. Because it is impossible to foresee all the decisions that a robot will have to make, various authors have proposed that the best and easiest way to design a "good" robot is by creating "a machine that itself follows an ideal ethical principle or set of principles" (Anderson and Anderson 2007). Indeed, designing robots capable of applying ethical principles in new situations is a challenge addressed in the field of *machine ethics*. However, its horizons are not limited to human safety. Other fascinating inquiries arise like whether these machines acquire full moral status by the fact of having such capacities. In other words, can they count

as persons? If the answer is positive, then the satisfaction of human desires should not be their only priority and, for the same reason, acts of rebellion against machine slavery would be morally justified, going from pure means to ends in themselves.

We thus arrive at a paradox: by creating safer and more efficient tools, we may divest ourselves of our master status. There are even bigger issues. Would robots have more or less concern for progress and peace than humans? As Nick Bostrom and Eliezer Yudkowsky warn, they may be very different "persons" from those who exist now and "perhaps governed by different rules" (Bostrom and Yudkowsky 2014). Many supporters of the *Technological Singularity Hypothesis* (the belief that accelerating progress in disruptive technologies will soon lead to radical change in human civilization) are optimistic in asserting that dialogue will be possible and fruitful. Furthermore, some claim that the day may come when robots will teach ethics to humans. If Anderson and Anderson are right when claiming that "ethics must be made computable in order to make it clear exactly how agents ought to behave in ethical dilemmas" (Idem), then robots embody the purest and best way of thinking since they are free of biological limitations and bothersome emotions. This might make them the best teachers of ethics or even the best rulers, although, as we will see next, Huxley makes arguments to the contrary.

§6. Men losing their bodies

In the computational theory of mind (CTM), one of the most accepted neuroscientific theoretical frameworks, human consciousness is a form of computation in which subjective experiences (mental phenomena) are folk psychology categories (that is, destined to be eliminated with the progressive advance of science). They are also sometimes called epiphenomena, namely, sterile phenomena that are caused by physical phenomena, but lack significant physical power. Neurophilosopher Paul M. Churchland claims that AI delimits what is most specific about the human being, what he calls "*the seat of the soul and also the engine of reason"* (Churchland 1995, 227-244). He also argues that we should not fear the fact that machines are devoid of grace, aesthetic

inspiration or even phenomenal consciousness because, in his opinion, so are human beings, or at least we do not possess them in the way we usually think of these supposed advantages. Similarly, Guy Kahane and Julian Savulescu attribute to AI a particular type of non-phenomenal consciousness, namely *access consciousness*, which they define as the ability to manage "global access to information" (Kahane and Savulescu 2009). For both authors, that consciousness, rather than the phenomenal one, places human beings at the apex of the evolutionary pyramid.

Savulescu uses this notion of consciousness to narrow the gap between man and machine, and does so from beyond the theoretical point of view. He has no major objections to incorporating technological gadgets to replace or even improve upon human functions. The *Mind Uploading Project* (the study of the hypothetical process of moving individual consciousness to a computer), so popularized in contemporary media, takes this idea of replacement to its logical extreme. Their followers, like Savulescu and Bostrom, argue for not only a robotic model of ethical behavior, but also the final transformation of man into machine. Worryingly, this current of thought is being strengthened in the social imaginary through fully extended medical interventions, such as surgical transplants or, more importantly, prosthetic implants. Human beings can now somehow manage their bodies similarly to how they manage machines, making it easier for us to believe that we too are machines. The problem here is not the use of such always welcome therapeutic technology itself, but what such use induces in our way of understanding ourselves.

Suffice all of this to say that Huxley's warning against the *mechanizing drift* of modern society should be taken seriously. However, on this issue, he advocates for solutions that go to the heart of the problem; rather than enacting laws against dehumanizing theories, he promotes those that encourage habits linked to the development of aesthetic sensitivity. Huxley's proposal also involves the promotion of neuroscientific research related to perception, attention and contemplation– including its obstacles and potential enhancers. Not coincidentally, in the model of society presented in Huxley's *Island*, more importance is given to the education of the senses than to affectivity, memory and reasoning. For example, the effect of *moksha-medicine*, which Huxley described in the novel as a tool for human enlightenment and moral enhancement, is located

at the gates of perception (Huxley 1962, 159, and 184). *Moksha-medicine favors* enhanced contemplative meditation practices, which are essential for dissolving fictitious truths, pretend freedom, and superficial happiness, as well as for distinguishing what is beautiful, real and worthy of love and pursuit.

Experimental sciences study phenomena in their temporal line, while spiritual sciences (Huxley also calls them sciences of consciousness) focus on their present instantiations, where contemplation reveals the immanent values of existence and the root of all creativity. "To be shaken out of the ruts of ordinary perception, to be shown for a few timeless hours the outer and inner world, not as they appear to an animal obsessed with survival or to a human being obsessed with words and notions, but as they are apprehended, directly and unconditionally, by Mind at Large –this is an experience of inestimable value to everyone and especially to the intellectual" (Huxley 1954, 73). Stability and persistence should be, he concludes, secondary objectives. In this regard, *Island* is also a plea against the idolatry of sustainability, conveying to readers that a real utopian society chooses meaningful progress even though it leads to its own destruction. It is striking that, consequently, the author claims that thanatology is the ultimate science, which is relevant for people obsessed with material immortality, whether organic or silica, for themselves or their communities (Huxley 1962, 141-142).

§7. Disembodiment, denaturalization and disenchantment

The Huxlerian *machine-man* and the *mass-man* share the worst kind of blindness, that which prevents them from finding meaning in life and motivation to continue living. *Machine-men*, like the most sophisticated AI robots imaginable, can acquire and process knowledge in the most effective way. However, for Huxley, this knowledge fails to be significant to them; it does not impel them from a subjective, embodied state towards horizons of astonishment and wonder, or horror and dread. Logical conclusions and physical and chemical causes are mainly taken into consideration to explain their decision-making. In contrast, it is the perception of beauty or ugliness that justifies, for Huxley, *organic* people's behavior. These experiences allow people

to engage in reality, to affirm it, to care for it, to pursue ideals, to fight, etc. In short, he distinguishes between (objective-subjective) *understanding* and mere (objective) *knowledge* (Huxley 2003, p. 193) to emphasize the importance of complementing scientific methodologies with other types of approaches to reality, especially those that involve first-person points of view.

Technological dreams, such as mind uploading, involve one of Huxley's worst fears, namely perfect disembodiment that leads to the disenchantment of the natural world. However, he realizes that there is a similar, but more familiar, way of undermining a meaningful, significant life. It is found in modern ways of working that lead to increased loss of contact with reality. "[I]n modern industrial societies vast numbers of men and women pass their whole lives in hideous cities… have to perform manual or clerical work that is repetitive, mechanical and intrinsically meaningless, are rootless, propertyless and entirely divorced from the world of nature… The reason for this dismal state of things is the progressive application of the results of pure science for the benefit of mass-producing and mass-distributing industry, and with the unconscious or conscious purpose of furthering centralization of power in finance, manufacture and government" (1947a, p. 23). In this quote, Huxley denounces work that ignores the end or meaning of action, fixating instead on the means, and the quantity and quality of production. "Man as a moral, social and political being is sacrificed to *homo faber*, or man the smith, the inventor and forger of new gadgets" (Idem). The overly industrialized production model takes human beings away from direct experience with reality, which induces, for Huxley, false images of a denaturalized (objectified) world and then leads to disenchantment itself.

Huxley warns us against a future in which people treat not only their body, but reality as a whole as if it were a mere machine, namely, a device that is the sum of its parts. Alasdair MacIntyre has used similar terms to describe the risk of modernity, although he goes deeper into the idea of the analytical approach, which hyper-specialists so commonly employ, and its hypertrophy. In particular, he expands it from alienated work to alienated lifestyles: "Modernity partitions each human life into a variety of segments, each with its own norms and modes of behavior. So work is divided from leisure, private life from public, the corporate from the personal. So both childhood and old

age have been wrenched away from the rest of human life and made over into distinct realms. And all these separations have been achieved so that it is the distinctiveness of each and not the unity of the life of the individual who passes through these parts that we are taught to think and feel" (MacIntyre 2007, 204). Mass production and the happiness industry are not possible in old-fashioned, small-scale markets, but, in exchange, they can and do experience internal, close and personal commercial relationships, which amount to, from Huxley's point of view, making things easier for human intuitions and higher goals. It involves returning to the source, nature, recovering human experience, and bringing back authentic life and beauty to the world.

Despite the foregoing, Huxley does not defend primitivistic technophobic positions, that is to say, he does not herald a return to the Middle Ages. Reality has a dual dimension that is both objective and subjective and, therefore, he understands the fruitfulness of using mechanistic approaches to solve many human issues: from agricultural production systems to hip prosthetics. However, the related methods become an obstacle when their triumph prevents us from contemplating reality in its irreducible duality. Danger is not found in technological development itself, but rather in the myopia that it produces if it is not accompanied by an adequate education in sensitivity. On this particular matter, Huxley shares Husserl's idea on the paradox of human subjectivity: "being a subject for the world and at the same time an object in the world" (Husserl 1936, 53). That is why Huxley also agrees with phenomenology's *motto* against scientism, that is, philosophers must go *back to the things themselves*. Every kind of knowledge begins with an experience and, hence, with the first-person point of view. The great challenge of science is, according to both Huxley and Husserl, found in making the leap from subjective to objective knowledge, which cannot be done without adequately thinking of the initial phenomenon. However, as explained above, Huxley goes one step further and demands that, before thinking about the phenomenon, we have to learn to adequately perceive it.

Lastly, Huxley identifies one last type of emergent malaise in a disenchanted, blind society. Since personal progress does not depend on closeness to neighbors (whose thoughts and actions one knows), but rather to distant leaders (or even worse, to anonymous shareholders), a pathological feeling

of dependency and vulnerability arises locally. In addition, an excessively changing world does not foster feelings of empowerment, belonging, mutual trust and continuity. Experiencing the future as in the hands of uncontrollable factors induces, Huxley concludes, high levels of uncertainty and insecurity. "At the present time the horrors of insecurity, as exemplified above all in mass unemployment, have impressed themselves so deeply upon the popular mind that, if offered the choice between liberty and security, most people would almost unhesitatingly vote for security" (1947a, p. 20).

Choosing security implies seeking help from the state; namely, it gives new impetus to the centralization of power, which is the main root of insecurity in the first place. The vicious circle is complete: It starts with uncontrolled development of technology and then excessive centralization of power, bureaucratization of society, human desensitization (*robotization*), denaturalization of reality, the horrors of insecurity, and then back again to attempts to overcome these evils with more technological developments and laws that authorize their use, including public video surveillance, collection of digital metadata, social credit systems, etc. At the same time, these measures are accompanied by requests for parallel laws to protect against abuse and limit privacy invasions, for example, the use of neuroimaging to read minds while interrogating a detained person (Greely 2015). However, as already mentioned, Huxley argues that these requests are only palliative measures to delay the inevitable. Laws would do better to focus on avoiding the root causes that generate widespread feelings of insecurity, that is to say, a perception of reality that reduces everything to antecedent causes and resultant effects and thus prevents contemplation of what is in between and its relation to the whole picture.

§8. One may smile, and smile, and be a villain

Huxley's worst-case scenario does not lead to a blue society with authoritarian leaders, at least as far as its external manifestations are concerned. In *Brave New World Revisited*, he writes, "All the traditional names, all the hallowed slogans will remain exactly what they were in the good old days. Democracy and freedom will be the theme of every broadcast and editorial

– but democracy and freedom in a strictly Pickwickian sense" (Huxley 1958a, 155-156). There, all social values are reduced to their objective dimension, although, of course, feelings of freedom, equality, fraternity, etc. persist. Yet, these feelings are triggered by genetic engineering, propaganda and drugs, and not by the actual experiences that people have when they think of themselves as agents of their own lives and committed actors in the world. Concepts lack meaning just as individuals lack inner life. *One may smile, and smile, and be a villain*. Shakespeare's famous words, on Hamlet's lips, do not adequately define the psychological life Huxley depicts. It is emptiness rather than hypocrisy that dwells in the heart of these brave new men, where human acts are reduced to behavior. As a counterpoint, this is essential for understanding Huxley's idea of contemplation. It also brings up the necessity of clarifying some general notions about the difference between action and behavior.

Elizabeth Anscombe defines a) intentional actions as those in which an agent acts for reasons. In particular, reason is conceptually linked to other reasons. In contrast, b) non-intentional behavior is that in which an agent acts for b1) external (physical) or b2) internal (mental) causes (Anscombe 1963, 9-13). Causes are related to each other by way of empirical relationships, for example, falling to the ground due to b1) slipping on ice or b2) excessive alcohol intake or a panic attack. Anscombe also distinguishes between c) motives and d) desires: the former are related to the reasons for action, while the latter are related to causes of behavior. Intentional actions have to do with goals (better future situations), and motives are feelings of attraction or repulsion towards the future that an agent contemplates. By contrast, a mere desire is an experience of attraction or repulsion that is unconnected from an act of contemplation and instead involves a trigger. For example, a murder can be due to c) a decision made after contemplation of the victim's future state (the agent sees the world as better off without the victim) or d) to a fit of rage in which the agent does not think (or feel anything) about the beauty or ugliness of his behavior's ultimate consequences. Here, the agent is not attracted by any future, but rather pushed by a force from behind. It only disappears by satisfying it since desire is at the center of the agent's decision and not the goal towards which it leads.

Anscombe's conceptual framework of human action is useful for understanding the main difference between two objectively similar but subjectively

different behaviors, which is, in turn, helpful for appreciating the risks associated with the *unmanifest* emptiness that Huxley warns about. The democratic emotions that prevail in *Brave New World* are not c) motives, but rather d) desires. Its citizens are blind to the future and, therefore, democracy cannot be meaningfully understood or felt. They might express strong democratic wishes, and even initiate activism in defense of democratic values, but said behavior cannot be described as intentional or rational because it is caused by a physical or mental trigger that maintains an accidental relationship with the democratic ideal. That corresponds to an ethically relevant difference because such a relationship is fragile, and the greatest enthusiasm for civil rights can easily be redirected toward other political regimes without notable resistance from the actors. Quite to the contrary, contemplation of the democratic ideal is less susceptible to interference and manipulation, both in terms of content and motivation.

The latter consideration is also of particular importance for comprehending why Huxley attaches so much value to the matter that makes up reality. *Matter,* and not just form, *matters*. Abstraction of forms provides objective knowledge, but it is only when forms are accompanied by matter, that is when forms are materialized, that agents are able to apprehend the ideal of which every natural reality seems to be impregnated. "Words are not the same as things and that a knowledge of words about facts is in no sense equivalent to a direct and immediate apprehension of the facts themselves" (1947b, 145). According to the author, it is in this formal-material apprehension of reality where agents are able to perceive the dialectic of reality's temporal dimension, and where present and future are laced, showing each being how close it has come to its perfection, that is, to its rightful place in the universe.

This formal-material apprehension, which defines the aesthetic experience, is, for Huxley, the heart and root of all ethical behavior. "Beauty is an immediate experience and this immediate experience is identical with Beauty-as-Principle, Beauty-as-Primordial-Fact. The first of these statements is fully in accord with the doctrines of the Perennial Philosophy… We perceive beauty in the harmonious intervals between the parts of a whole" (1947b, 158). With this approach, he articulates two ideas that have been echoed, with their defenders and detractors, throughout the history of Western philosophy: 1) Beauty is not

a construct of the mind, but rather a sign of the natural world– its realities are subject to a luminous order or cosmos (κόσμος) and 2) aesthetic experience is not just the main driving force behind ethical behavior (Anscombe's motives), but is also the first beacon towards which the agent decides to move. As we will see next, these are the two pillars of Aldous Huxley's theory of contemplation.

§9. Moral perception

Aristotle was one of the first philosophers to sustain that the good does not precede its practice, but that it is learned by doing (Aristotle, *Nicomachean Ethics*, 1105b 10 and 1103b 20-35). Like Huxley, he claims that reflection has to be accompanied by first-person experiences, namely, agent-environment interaction, in order to do the right– to know what is best and to choose it. Therefore, the great Greek philosopher based his ethical realism on this two-dimensional intuition about the apprehension of beauty or on what today might be called objective/subjective intuition.

Huxley, however, is closer to romantic intuitionism than to the classic brand. First-generation romantics, like William Wordsworth and Samuel T. Coleridge, also believed in the *perennial* idea that passions have importance in the moral life; feelings facilitate our first apprehension of the good and evil that is inherent to reality. Passions reveal, therefore, the level of constitutive beauty in the natural world. However, these romantics reject the idea that the human world is simply waiting to be perceived in all its goodness and beauty. On the contrary, beforehand, reality is inchoate rather than fully formulated and complete. It is in the human heart where romantics think reality can take a specific shape and take on a *form* or *eidos*. In *Ode: On the Intimations of Immortality,* Wordsworth writes, "I was often unable to think of external things as having external existence, and I communed with all that I saw as something not apart from, but inherent in, my own immaterial nature" (Wordsworth 1859, 413). However, in these verses, he does not propose a kind of horizontal dynamic alterity, but rather a hierarchical one, in which human actions not only make manifest and take care of reality, but also express it, namely, by completing the work of creation.

Contemplation is, in this context, not just an active reaction to external inputs, but also a transformative look inward. Nature is expressed (completed) in human feelings and individuals find in them the beginning of the road to their rightful place in the universe. Thus, this romantic expressivist approach implies, first, a call to keep closer (ecological) ties to reality and, second, more importance given to feelings– the eyes of the heart– in the search for knowledge. This is also the reason why romantics place individuals at the center of ethical decision-making— feelings have to do both with the material world (that fuels them) and the material body (that channels them). In other words, the first-person perspective is irreplaceable. Only the individual as such, as materialized form, can truly understand reality. Without taking into account each person's difference and originality, the voice of nature, the voice of all its internal and external values, is silent. Matter is inseparably linked to Anscombe's a) actions and c) motives.

About such romantic understanding, Wordsworth writes, "Its object is truth, not individual and local, but general and operative; not standing upon external testimony, but carried alive into the heart by passion; truth which is its own testimony" (Wordsworth 1802, xxxii). Like Huxley, he rejects purely objective knowledge, not only because it is unable to motivate individuals, but also because it fails to move them to the right place. Indeed, romantic movements first arose as a reaction against two Enlightenment dogmas, namely, a) the belief that feelings and valuations are merely capricious or vestigial projections of the human mind toward a neutral world and b) human flourishing and scientific progress are closely related and they have to do with reason's despotic dominance over desire, matter and nature. Wordsworth's words are antithetical to the Enlightenment ideal of the *disengaged, dispassionate, self-controlled self*. His poems represent an attempt to tune human feelings into constitutive beauty, allowing nature to resonate within the heart. There is no better way to unravel the mysteries of reality than through the contemplation of nature, which implies, according to Wordsworth, summoning stronger and nobler feelings, thoughts and actions in order to have a meaningful, good life.

Natural values are not exoterically available, namely, they cannot be grasped by disengaged reason but, as Wordsworth intones, they demand collaboration, understanding and a certain commitment to good actions– the equivalent of

Aristotelian *learning by doing* or Huxlerian *love in action*. In contrast, the romantics see the Enlightenment ideal as promoting an objective and neutral way of looking at reality that leads to the inevitable instrumentalization of nature. "The world is too much with us; late and soon, Getting and spending, we lay waste our powers; Little we see in nature that is ours" (Wordsworth 1807, 122). The Enlightenment's extreme materialism offers, according to romantic opinion, not only a partial view of the world, but also prevents one from tuning into nature. In this context, it makes sense that the romantics did not intend to complement discourses from new science, but, rather, to move away from them. To their taste, loneliness is not the worst of fates. Paradoxically, the romantic idea of escaping from materialism and saving nature implies returning to matter, which means, among other things, discovering its depth through feelings. Only with this discovery is it possible to use reason properly, as well as to forge a new, true science characterized by the integration of feelings and reason. In this framework, feelings and reasons do not maintain a symmetrical relationship; feelings provide moral intuitions to reason in order to confirm, correct, or even defeat them with reflection.

Consistent with the above, the possibility of attentively listening to nature's inner voice, following the romantic spirit, opens up a new form of government where, as Huxley proposes, laws mainly promote the education of sensitivity rather than guaranteeing the fulfillment of citizens' duty. "My youth here witnessed, in a prouder time; the senselessness of joy was then sublime! Happy is he, who caring not for Pope, Consul, or King, can sound himself to know the destiny of man, and live in Hope" (Wordsworth 1845, 237). The influence of Rousseau's philosophy on romantic ideas regarding the State is undeniable, although the author of *The Prelude* seems to apply this political view indistinctly to democracy, soft feudalism, and monarchy. Wordsworth identifies this indistinction in the fact that loving subjects are better than loyal ones.

The first instantiation of romanticism very soon branched out into various currents–some of them radically anti-realistic, anarchist and, ultimately, anti-romantic. However, many of its original ideas about nature, feelings and subjects managed to take root in Western Modern thought. *Transcendentalism*, a movement to which Thoreau belongs, is a clear example of this. Even today, not a few authors are aligned with the élan of nature's expressivist view that

nature is a source of ideas and existential drives. Iris Murdoch, Charles Taylor, Bennet Helm, Michael Stocker, and Lawrence Blum are among the most widely recognized thinkers that support theses similar to Huxley's notion of contemplation in scientific and moral life. For example, Taylor writes about his main keys of thought, saying, "My claim is that the idea of nature as an intrinsic source goes along with an expressive view of human life. Fulfilling my nature means espousing the inner élan, the voice of impulse. And this makes what was hidden manifest for both myself and others. But this manifestation also helps to define what is to be realized. The direction of this élan wasn't and couldn't be clear prior to this manifestation" (Taylor 1989, 374-375). Taylor describes this *espousing* as a moral intuition in which perception and cognition collaborate in its genesis. In other words, moral intuition involves a) the external senses (what Huxley would call *sensorial contemplation*), as well as b) memory, imagination, reflection, etc. (Huxley's *spiritual contemplation*), such that the experience of the good always appears interwoven with the "understanding of my life as an unfolding story" (Ibídem, 47). In other words, while Huxley presents two different intuitional acts (the higher one requires the lower one, but not the other way around), Taylor sees them as one and the same.

Taylor also warns that the current scientific-technological approach puts at risk the sensorial dimension of contemplation. "[It] has gone along with an irretrievable loss in our attunement to our natural surroundings and our sense of community" (Ibidem, 61). However, this issue also reveals a significant difference with Huxley. Taylor argues that an ethical decision does not always need to be accompanied by the perception (the experience) of the good. "We may accept something as a good although we are relatively unmoved by it, because at the lowest, we think very little about it and glide along in conformity with our milieu; or because we revere and look up to established authority; or perhaps best, because we as authoritative for us, sensing in them that they are authentic and great, even though we don't fully understand it or feel it ourselves. But through all these complex chains of intermediation, the connection between seeing the good and being moved by it cannot be broken. Our authorities, or the founders of our traditions, those who give these goods their energy and place in our lives, they felt them deeply" (Ibidem, 73-74). Thus, Taylor attaches less importance to the perennial presence of *bliss* than Huxley

seems to claim, although, for both, the loss or forgetting of moral experiences sooner or later brings about similar consequences, i.e., inert action that leaves behind moral criteria, and finally, total passivity.

§10. A serene zombie apocalypse

Huxley's idea of moral perception leads us to the main obstacle found in designing a machine for learning ethics from a robotic ethics teacher. AI robots could come to imitate humans' calculating reason– the objective dimension of any ethical judgment. It makes sense here to claim that we can learn from machines to make better ethical decisions, just as a computer can help biologists solve the most complex biological calculations. *AlphaFold* is a good example of how an artificial intelligence program is accelerating new scientific discoveries related to protein structure predictions (Senior, Evans, Jumper, et al. 2020). Similarly, it does not seem too difficult to design autonomous robots capable of dealing with new problems, at least to the extent that the solution can be deduced from the program's fundamental framework (basic laws or rules). However, for Huxley, a machine is not capable of taking the leap toward designing basic programming in a way that is not completely arbitrary, i.e., of founding ethical principles. To put it another way, we cannot create an intuitive robot, namely, an AI machine, capable of apprehending the ideals inherent in reality and, therefore, of feeling attracted to them.

Robots cannot, by definition, contemplate reality and, consequently, make the information they handle meaningful. Being blind to beauty (lacking sensitivity and an emotional life) makes them indifferent to their own and others' existence. Machines do not have the interiority in which nature can express itself– there are no feelings that allow machines to go beyond themselves. Hence, using Huxley's terminology, the most sophisticated robot cannot learn or teach ethics (or any other discipline) because they know, but do not understand, the information they process and, therefore, do not use it to their advantage. Robots are tools that expand and extend their users' capabilities, but lack any of the references that are basic to human users, including goals,

autonomy, and existence itself. Anthropomorphic projections as such are also a form of pareidolia.

From a practical point of view, the absence of intuitive capacity is relevant not only in the resolution of new moral problems, but also in the evaluation of new solutions to classic problems. As Huxley claims, the solution to a problem frequently lies not in thinking more, but in trying to hear better– looking deeper into– what nature is calling for. Furthermore, a chain of thoughts, even one starting from correct intuition (from an optimal act of sensitive contemplation), can deviate and produce immoral judgments. Intuition (and especially, for Huxley, contemplative intuition) helps humans monitor abstract thought and anchor it to reality. Otherwise, like Goya's famous painting, reason is likely to fall asleep and produce monsters. The evil of idealization, into which objective knowledge can fall, arises anew with a silica appearance.

Even those who deny human interiority and, therefore, the substantial difference between human and machine rationality, warn of future dangers related to autonomous machines. This is the case of Stephen Hawking, who went so far as to say that, "The development of full artificial intelligence could spell the end of the human race" (Hawking 2014). However, it is difficult to argue why this end is worth preventing if super-intelligent machines are truly superior and their decisions are able to surpass that of humans in terms of objectivity– no matter how cruel (heartless) they may seem to us. Perhaps, at this point, our duty is only to accept our role as Nietzschean *rope-dancers* and prepare the way for the supermen. Nietzsche is an anti-romantic, although it should be noted that he also criticizes the idea of disengaged reason. Enlightenment philosophers are, for him, the fathers of mediocre, short-sighted and selfish *last men*. Nietzsche would agree with Huxley that, rather than protectionist laws, those that foment education in aesthetic intuitions counteract the spread of robotic lifestyles, which, if Huxley is right, will be characterized by an unprecedented degree of idealism, dogma and violence.

Other authors defend more optimistic positions about the end of our species. For example, in *Novacene: The Coming Age of Hyperintelligence*, James Lovelock imagines an idyllic future where silica bodies, and eventually virtual bodies, are the final solution against hunger, war, and death (Lovelock 2019, 117). Human moral views– solidarity, justice, freedom, science, beauty– do

not disappear, but, on the contrary, grow freely once rid of the bonds of the flesh. Lovelock argues that, if we trust in the goodness of reason, we should not fear that smarter machines will be ethically inferior to us.

Even if Lovelock is right, an objection related to the inner world remains. He assigns to these self-sufficient and self-aware intelligent machines a mere *access consciousness*– a non-phenomenal consciousness as defined above. Such definitive replacement of human beings by androids, beyond the practical problems that may arise, would be the worst of dystopias imaginable and, although Huxley does not address that possibility, is in line with the catalog of horrors found in *Brave New World*. A peaceful machine world is the ultimate example of superficiality and false appearances. It would be akin to the most sophisticated amusement park, a universe of smiles, music and candy that is always in operation, but lacks customers. We face the ventriloquist's dream, i.e., perfectly hiding the doll's voice and movements; so too will human programmers aim when creating silica puppets capable of surviving them. If our inner life is real, the monstrosity here is found in using technology to eliminate man and then handing that technology over to nothing. In short, in a universe emptied of interiority, two machines dancing merrily with each other make as much sense as two books supported by white porcelain elephants at the bottom of the sea.

David Chalmers is one of the best-known thinkers to use the term *philosophical zombie* to criticize behavioristic approaches to the mind. With this epithet, he proposes the hypothetical existence of a zombie, namely, an individual who lacks first-person experience, but who is capable of acting like a human being (Chalmers 1996, 94). Beyond the controversy, it is a well-chosen term from an ethics-aesthetic point of view because it expresses the ugliness of a serene technological apocalypse so well.

§11. Emotional self-deception

Artificial intelligence is all about objective intelligence, third-person point of view language, and behavioral imitation. The latter applies to a wide spectrum of intelligent behaviors, including emotional ones. Machine learning, closely

related to computational statistics, provides computers with the ability to autonomously pattern abstraction in order to classify and predict consequence by means of algorithms. With sufficient processing capacity and data storage, a machine may be able to predict and imitate with great accuracy a person's rational and emotional response to a particular situation, without having to give an account of her intentions. In Anscombe's terms, this implies replicating a) intentional acts by b1) external causes. Such a predictive ability may even surprise the agent on whom it is applied, seeing it as an oracle or prophet or, at least, as a machine able to read the mind and reach the inner world.

In other words, there are two ways to respond to the question of how a college student will act when asked how much two plus two is: a) by recognizing the question, solving it, and showing herself to be an adult, rational agent or b) by observing the behavior of individuals in similar circumstances. Learning machines can do both types of tasks, although the second mode is key to the simulation of human feelings, at least if we accept Huxley's non-cognitivist theory of emotions, where, as noted, this particular type of mental event uniquely expresses the experience of constitutive values and ideals– the most intimate human way of engaging with reality. If machines can process the cognitive dimension of emotions, but cannot apprehend experiential values, then there is more truth in the claim that a machine understands students' mathematical knowledge (in the a) and b) sense) than in the claim that a machine understands his or her feelings (only in the b) sense). At most, machines can understand, following Anscombe's distinction, desires, and not motives.

This is a very limiting factor in AI because it reduces machines' predictive capacity. In Taylor's words, "with terms like 'courage' or 'brutality' or 'gratitude', we cannot grasp what would hold all their instances together as a class if we prescind from their evaluative point. Someone who had no sense of this point wouldn't know how to 'go on' from a range of sample cases to new ones. This means… that the 'descriptive' meaning cannot be separated from the 'evaluative'" (Taylor 1989, 54). This reasoning brings us back to the argument of how the emotional life can stave off the risk of idealism associated with *calculating reason* and its worst consequences found in the not-at-all serene apocalypse.

I now turn to a final risk related to emotion-simulating machines in a disenchanted, blind society. For almost two decades, research in AI technology has led to advancements in entertainment and social robots. Today, more and more laboratories are interested in designing machines with *emotional intelligence*, namely, with the ability to handle and respond to human emotions (Zheng et al. 2019, Rincon et al. 2019 and Piccard 2003). Devices with a more human-like interface have multiple uses, most of which have to do with sales strategies and attractive robotic faces. In a way, these machines are much like a commercial agent where customers do not have to take their smiles very seriously, no matter how well they can fake it. Of course, this claim is compatible with saying, as argued above, that, at certain levels of refinement, propaganda techniques can be extremely coercive, and that emotional robots represent a further step in tyrannical influence over minds. However, a completely different matter is the use of such technology for intimate attachment, from elderly care givers to servile friends and romantic partners (Kolling et al. 2013, Cheod et al. 2016). Japanese society is a pioneer in the demand for this type of service, and, if Huxley is right, it is likely to spread globally very soon due to modern societies' decreasing capacity to sustain real engagement.

The use of emotional robots requires a tacit agreement on the need for the fictitious nature of the communication and relationship to disappear (be forgotten) in order to be effective. These machines no longer represent the friendly face of a corporation or state service, but rather simulate individual agents with intuitive capacity and subjectivity to share– behavior with real motivations, in Anscombe's terms. In this kind of relationship, consumers believe or want to believe that machines not only *know,* but also *understand,* them, namely, that they are able to notice when a person's ideal is fulfilled, to value his or her constitutive beauty, and to feel love in the act of accompanying, listening, hugging, etc.

These emotional services presuppose an inter-subjective relationship, that is to say, reciprocal first-person approaches– which is what intimacy means here. Nobody feels loved by a book or by its author; it does not matter how many times the author writes nice statements about his readers, even using the most reasonable arguments and intuitive poetic expressions, like that from Josef Pieper, "It is good that you exist, that you are!" (Pieper 1989, 42-43).

Similarly, it is very rare to feel loved by a public or private institution, or by its members. That is why these robots work best if they use personal pronouns to refer to themselves. *I am with you, I listen to you I love you, let me hug you.* Emotional machines have to be sophisticated enough to make their users believe (or self-deceive) that this type of personal relationship truly exists. Currently, they are not that sophisticated; the ones that do exist depend on their users' capacity for fantasy– loneliness leads to desperate attitudes (Sharkey and Sharkey 2020). The other possibility is that these users truly believe that intimacy can be reduced to its external expression, to mere objectivity. In this case, users are acting as *practical behaviorists*. However, I think this second possibility is, to date, less likely. Indeed, we find here the last risk I want to mention in this paper. There is a slippery slope between the use of emotional machines for intimate attachment within a self-deceptive imaginary and their use within a behaviorist imaginary. If people get used to this kind of substitute for intimacy, they are more likely to think that intimacy does not go beyond what robots can offer.

The last step in this emotional transformation toward superficiality involves despising one's own inner life and becoming one more robot among happy robots… a smiling, unassailable bright wall for others and, more importantly, for one's self. It is more than plausible that these types of users will welcome, on the one hand, a world where robots are in charge (even when it comes to teaching ethics), and, on the other hand, projects to achieve immortality like *Mind Uploading*. Emotional robots belong to the same kinds of manipulative weapons that readers find in *Brave New World*, such as soma, compulsive consumption habits, TV Shows, orgy-porgy and violent passion surrogate treatments. They consolidate, like no other technology, Huxley's dehumanized technological dystopia, where apparent happiness prevails over explicit violence and suffering. Emotions lose their transcendent meaning, ceasing to be motives in Anscombe's sense, because they no longer express intuitive acts that launch humans toward internal and external reality. They are only *desires* and, as such, are not satisfied by what triggers them, but rather by pleasant, self-referential stimuli. In fact, questions and answers about the origin and nature of these types of chemical or silica emotions make it difficult for consumers to fully enjoy the experience. Blindness is a necessary condition for happiness

in *Brave New World* societies, and recovering the *art of seeing* is Huxley's main remedy against instruments of tyranny that end in social imprisonment.

Huxley does not reject technology or pleasure, just as he does not spurn drugs for symptomatic or palliative relief. He rejects technology as an end in itself, and symptomatic treatment as the main way to deal with suffering. Perhaps today some people rely on reality television, happy pills and emotional machines to alleviate their loneliness, but that should not lead us to think that better, more human and real alternatives do not exist. It would be unreasonable, therefore, to legislate against drugs and machines, but ushering in laws that promote their proper understanding and use is critically important.

§12. Final remarks: Turning neurorights into neuroduties

In this paper, I suggest that Huxley's approach is valuable in two ways. First, he shows us that, in the hopes of helping people both defend themselves from tyrants and, most importantly, avoid becoming one, it is better to assume a few duties (some of which are mentioned here) than to develop a set of bans and restrictions to protect citizens against the misuse of science and technology. From a practical point of view, of course, both strategies seem necessary and complementary. Yet, Huxley's warning about the point of no return helps us to understand that it is essential to prevent short-term corrective measures (rights-focused measures) that overshadow duties-focused measures. Second, because Huxley effectively narrows the gap between the neuroscience of law and the law of neuroscience, he addresses the policy problem of controlling neurotechnology in the most comprehensive way possible. This implies, among other things, recognizing that the main ethical issue swirling around technology is not the absence of law, nor is his solution detailed in terms of a constitution of rights.

I have also explained here what Huxley's solution, prioritizing sensitivity education, means. More and more scientists, he writes, "accept the world picture implicit in the theories of science as a complete and exhaustive account of reality; they tend to regard those aspects of experience which scientists leave out of the account, because they are incompetent to deal with them, as

being somehow less real than the aspects which science has arbitrarily chosen to abstract from out of the infinitely rich totality of given facts" (1947a, 29). If scientists had better training in the development of intuitive ability, if they experienced wonder more often, they could more easily resist reductionist temptations. Besides, after being enchanted by their own field, it would be easier for them to appreciate the beauty and ideals that other methodologies grasp. Therefore, beauty is, in Huxley's cosmology, the sacred and ultimate bridge among fields and researchers. Only in this context does Huxley's invitation to reintroduce early romantic ideas into academia make sense.

My last comments refer to two clues that Huxley offers us for making good on this invitation. The first step is found in taking advantage of universal artistic works and learning to contemplate reality with an artistic eye. Of course, the arts are not part of the sciences in the sense that objective knowledge is not their target. Artists work with subjective experiences, which does not mean that artists are not interested in reality. At least for artists like Wordsworth, they only differ in method– one person's perspective. Scientists can learn from artists by discovering and going into the part of reality that, as inherent to matter, cannot be objectified. Artists may teach scientists to use methods and language differently— not to represent nature and objective causes, but to be transported to its wonder-filled depths. "The purified language of science is instrumental, a device for making public experiences understandable by fitting them into an existing frame of reference, or into a new frame of reference that can take its place among the old. The purified language of literary arts is not the means to something else; it is an end itself, a thing of intrinsic significance and beauty, a magical object endowed… with mysterious power" (Huxley 1963, 34-35). With the language of the arts, scientists can look at nature as an artist looks at artwork, namely, seeing matter and form as completely interpenetrated, such that the meaning of art cannot be separated from the unity in which it is constituted. For artists like Wordsworth or Coleridge, art demands to be seen, heard and felt to be understood. This is also why, for them, one artistic genre is not completely reducible to another –perfect *ekphrasis* is not possible. On this matter, Taylor also adds that, "[t]he Romantic order… was not organizer on principles which could be grasped by disengaged reason. Its principle of order was not exoterically available. Rather it was itself an enigma, and one could

only understand it fully by participating in it. The love is such that one has to be initiated into it to see it" (Taylor 1989, 380). Sought-after and acclaimed scientific neutrality becomes a stumbling block in the search for a safe science. Depressing and violent conformism, cowardice over losing professional status, as well as economic interests, selfish ambition, arrogance and excessive desire for power, which are often disguised as ardent zeal for the truth, human flourishing and social progress, are, as Huxley demonstrates, smaller obstacles.

In the second part of *Literature and Science*, Huxley offers excellent advice for putting his suggestions into practice— through the delightful duty of writing popular science. "Popular science is a new art form, partaking simultaneously of the text book and the reportage, the philosophical essay and the sociological forecast" (Huxley 1963, 34-35). The scientist-artist has an excellent opportunity to express for laymen, as well as other scientists and for themselves, what is beautiful about their field and, with that, all of its ideals and bridges. This is the easiest way to transform academia. The hard way is to choose a more daring path and, *by letting the dead bury the dead*, leave academia, with all its interdisciplinary research groups, behind. Aldous Huxley took the latter path by becoming an almost full-time artist and art critic. He knew how to surround himself with influential Hollywood screenwriters and stars and, not by chance, he published one of the most influential and quoted books on bioethics in our age– and it was not an essay or a treatise, but a novel.

Some forms of blindness are extremely contagious but, if Huxley is right, light is always within reach. Maybe then the future would seem less inexorable, and his dystopic prophecies would turn out to be false. Only time will tell whether neuroscience is able to offer a better understanding of the nature of the romantic gaze and, therefore, to better advise educators and legislators about the best ways to favor, among other things, popular science. Perhaps neuroscience will be the first field to encourage a small number of academic defections and foment an objective search for knowledge from outside of science, that is, from art. But to do so, it seems inevitable to bury the dead child that neuroscience has spawned, neuroethics, and to do it inside and with the white coffin that is the current neurolaw –all under a tombstone with no name. We need new ethical and legal terms in this Huxley's romantic way of understanding moral intuitions, growing in sensitivity and, finally, interdisciplinary

dialogue (Echarte 2020b). As Huxley would say, it is not a matter of forgetting the old paths, but of leaving behind the paths that human beings are not capable of loving, or the paths that prevent them from loving. Only time will tell but I am afraid that we will have to delegate the choice of such neologisms to the young and brave researchers of the next century.

References

Anscombe, G. E. M. (1963) *Intention*. New York: Cornell University Press.

Aroso, J.M. (2018). From Brave New World to Island: Huxley's Tales of Alternatives To Law? In Kabashima, H. et al (eds). *The idea of justice in literature*. Wiesbaden: Springer, 79-100.

Bolfikova, E. Hrehova, D., Frenova J. (2010). Manager's Decision-Making in Organizations. Empirical Analysis of Bureaucratic vs. Learning Approach. *Journal of Economics and Business*, 28, 135-163.

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.

Charlton, B.G. (2010). The cancer of bureaucracy: how it will destroy science, medicine, education; and eventually everything else. *Med Hypotheses*, 74, 961-5.

Chandler, J.A. (2018). Neurolaw and Neuroethics. *Camb Q Healthc Ethics*. 27, 590-598.

Cheok, A.D., Levy, D., Karunanayaka, K. (2016) Lovotics: Love and Sex with Robots. In: Karpouzis, K., Yannakakis, G. (eds) Emotion in Games. Socio-Affective Computing, vol 4. New York: Springer.

Churchland P. (1985). *The engine of reasons, the Seat of the Soul*. Cambridge, Massachusetts: MIT Press.

Echarte LE (2020a). Rehabilitar la subjetividad en el diálogo interdisciplinar. La solución de Aldous Huxley al problema de las dos culturas. *Philosophia* 80 (1): 107-160.

Echarte LE (2020b). Two habits of the heart: a bridge-building proposal for professionalism, medical ethics and bioethics. *MEDIC 2020*; 28(1): 83-99.

Greely H.T. (2015). Neuroscience, Mindreading, and the Courts: The Example of Pain. *Journal of Health Care Law & Policy*, 18(2): 171-206.

Guinote, A. (2017). How Power Affects People: Activating, Wanting, and Goal Seeking. *Annu. Rev. Psychol*., 68, 353–81.

Hawking. S. (2014). Will AI kill or save humankind? BBC News. 20 October. https://www.bbc.com/news/technology-30290540

Husserl E. [1936] (1970). *The Crisis of European Sciences and Transcendental Phenomenology*. Evanston, IL: Northwestern University Press.

Huxley, A. (1938). *Ends and Means. An inquiry into the nature of ideals*. London: Chatto and Windus.

Huxley, A. (1947a). *Science, liberty and science*. New York: Harper & Row.

Huxley, A. [1958a] (1966). *Brave New World Revisited*. London: Chatto and Windus.

Huxley, A. (1958c). *The Mike Wallace Interview. Digital Collections. Harry Ramson Center*. This quoted from following link: https://hrc.contentdm.oclc.org/digital/collection/p15878coll90/id/61/rec/4

Huxley, A. (1962). *Island*. London: Chatto & Windus.

Huxley, A. (1963). *Literature and Science*. New York: Harper & Row.

Huxley, A. [1932] (1946). *Brave New World*. New York: Harper Perennial.

Huxley, A. [1954] (1970). *The doors of perception*. New York: Harper & Raw.

Huxley, A. [1958b] (1977). *Moksha: writings on psychedelics and the visionary experience (1931-1963)*. New York: Stonehill.

Huxley A. (2003). *Huxley and God. Essays on Religious Experience*. New York: Crossroad Book.

Huxley, A. (1947b). *The perennial philosophy*. London: Chatto & Windus.

Kahane, G., Savulescu, J. (2009). Brain damage and the moral significance of consciousness. *Journal of Medicine and Philosophy* 34 (1): 6-26.

Kolling, T., Haberstroh J., Kaspar R., et al. (2013). Evidence and Deployment-Based Research into Care for the Elderly Using Emotional Robots. *GeroPsych*, 26: 83-88.

Lovelock, J. (2019). *Novacene: The Coming Age of Hyperintelligence.* London: Penguin Books.

Maasen, S., Sutter B., eds. (2007). *On Willing Selves: Neoliberal Politics and the Challenge of Neuroscience.* New York: Palgrave MacMillan.

MacIntyre, A. (2007). *After Virtue: a study in moral theory*. Notre Dame: University of Notre Dame Press.

Nørgaard, A.S. (2018). Human behavior inside and outside bureaucracy: Lessons from psychology. *Journal of Behavioral Public Administration* 1, 1-16.

Ortega y Gasset, J. [1930] (1932). *The revolt of the masses*. London: George Allen & Unwin LTD.

Piccard, R.W. (2003). Affective Computing Challenges. *International Journal of Human-Computer Studies*, 59 (1), 55-64.

Pieper, J. (1989). *An Anthology*. San Francisco: Ignatius.

Rincon, J.A., Angelo Costa, A., Novais, P., et al. (2019). A new emotional robot assistant that facilitates human interaction and persuasion. *Knowledge and Information Systems*, 60 (1): 363–383.

Roskies, A.L. (2002). Neuroethics for the New Millenium. *Neuron*, 35, 21-23.

Safire, S. (2002). Visions for a new field of neuroethics. In Marcus S.J. editor. *Neuroethics: mapping the field.* San Francisco: Dana Press, 3-9.

Senior, A.W., Evans, R., Jumper, J., et al. (2020). Improved protein structure prediction using potentials from deep learning. *Nature* 577, 706-710.

Sharkey, A., Sharkey, N. (2020). We need to talk about deception in social robotics! Ethics Inf Technol.

Taber, C. S., Young, E. (2013). Political information processing. In Huddy, L., Sears, D. O., Levy, J. (eds.), *The Oxford handbook of political psychology.* New York, NY: Oxford University Press, 525–558.

Taylor C. (1989). *Sources of Self. The Making of the Modern Identity*. Melbourne: Cambridge University Press.

Taylor J. Sherrod; J. Anderson Harp; Tyron Elliott (1991). Neuropsychologists and Neurolawyers. *Neuropsychology* 5(4): 293-305. doi:10.1037/0894-4105.5.4.293

Thoreau, H.D. [1849] (1946). *Civil Disobedience*. Boonton, N.J: Liberty Library.

Yuste R, Goering S, Arcas BAY, Bi G, Carmena JM, Carter A, Fins JJ, et al. (2017). Four ethical priorities for neurotechnologies and AI. *Nature*. 2017 Nov 8; 551,159-163.

van den Hoonaard, W.C., Hamilton, A., eds. (2016). *The Ethics Rupture: Exploring Alternatives to Formal Research-Ethics Review*. University of Toronto Press.

Wald, D.S. (2004). Bureaucracy of ethics applications. *BMJ*, 329:282-284.

Wordsworth, W. (1802). Preface to the "Lyrical Ballads, with Pastoral and other poems". Vol. I. London: T.N. Longman and O. Rees.

Wordsworth, W. (1807). *Poems in two volumes*. Vol. I. London: T.N. Longman and O. Rees.

Wordsworth, W. (1845). *The poems of William Wordsworth*. D.C.L. London: Edward Moxon.

Wordsworth, W. (1859). *Poetic words. Vol VII*. Boston: Little Brown and Company.

Zheng, X., Shiomi, M., Minato, T., Ishiguro, H. (2019). What Kinds of Robot's Touch Will Match Expressed Emotions? *IEEE Robotics and Automation Letters*, 5 (1), 127-134.