

Responsibility and Human Enhancement

Simone Arnaldi

The idea of a special section on ‘human enhancement’ (HE) and responsibility originated from observing how the public debate on HE, i.e. the intentional effort to improve individuals’ performance with the help of technical or biomedical interventions [1], has developed over the past decades. More specifically, we noted two trends in the media and academic debate on this subject.

Firstly, there has been an increasing public attention to and interest in the subject of HE which is no longer confined to the realms of science fiction and fringe science. Enhancement has become instead a ‘mainstream’ subject, with highly visible applications such as performance enhancing drugs in professional and amateur sports, the workplace, and education, exoskeletons in logistics and the military, or neural implants for boosting memory and communicating with external digital devices. Secondly, this rapid growth of mundane enhancement applications has been met by a remarkable stability of the arguments and positions represented in the public debate. Indeed, the discussion on HE has been, and still is, primarily organized around the two questions of HE legitimacy and desirability, on the one hand, and of its technical feasibility, on the other hand. The advocates of HE answer positively to the question of legitimacy and desirability, insisting on the benefits for individuals [2] and, ultimately, for society as a whole [3]. The critics of HE draw opposite conclusions,

maintaining that HE will lead to greater inequalities [4], sow divisions [5], and exacerbate existing risks or create new ones [6]. In a similar fashion, futuristic visions of enhancements have been criticized as deterministic and misleading [7], and it has been maintained that options to enhance human performance on an unprecedented scale is all but assured by current and future technological trends [8]. Proponents of HE often fail to recognize that technologies do not develop in isolation from society, but are socially embedded and influenced by social, economic, and political conditions [9]. Critics frequently neglect the fact that enhancement technologies are already widespread in use and, depending on how enhancement is defined, have been commonplace for centuries [10].

Dissatisfied with this state of affairs, the contributors to this special section started looking for a way out of the impasse by drawing upon another increasingly important concept in technology and innovation studies and policy making: the notion of responsibility. While responsibility is not the only possible answer to the discontent with the debate on HE (see, e.g., the attempt to assess whether enhancement is permissible on the basis of the ‘nature’ of human activities which would be altered by HE interventions [11]), this notion has drawn our attention for its remarkable absence in the HE debate (see Shelley-Egan et al. [12] for an exception). At least in Europe, this is particularly surprising given the recent diffusion of the Responsible Research and Innovation (RRI) approach. Centred on the idea of aligning technology and scientific knowledge with societal goals by way of the mutual responsabilisation of social actors

S. Arnaldi (✉)
Department of Political and Social Sciences, University of Trieste,
Piazzale Europa 1, I-34127 Trieste, Italy
e-mail: sarnaldi@units.it

([13–15]), RRI requires researchers and practitioners alike to reflect and deliberate on the purposes of enhancements and their possible contributions concerning societal challenges. Inspired by the more pragmatic orientation of RRI, we decided to shift our focus of attention away from the prior assessment of the ethical admissibility and the technical possibility of enhancement interventions. Our efforts were instead directed to explore what the place of ‘responsibility’ is in the HE debate, and what conditions qualify HE as ‘responsible’. The contributions to this Special Section, which are the outcome of the research project “Responsibility and human enhancement. Concepts, implications and assessments” funded by the Independent Social Research Foundation (ISRF), address these aspects on different levels.

The article by Darian Meacham and Miguel Prado Casanova investigates the implications of the ‘extension’ of the human mind into and through hybrid, human-artifacts cognitive systems. Meacham and Prado explore a particular form of relation between humans and cognitive artifacts: interaction-dominance. In interaction-dominant systems, components cannot be isolated to determine exactly what their contribution to the system’s behaviour is. The measured behaviour of an interaction-dominant system is an emergent property of the system itself, as it “reflects the *coordination* of many componential processes” [16]. In interaction-dominant systems, it is therefore “difficult, and often impossible, to assign precise causal roles to particular components. It is also difficult, and often impossible, to predict the behaviour of components within interaction-dominant systems from their behavior in isolation” ([17], 41). Meacham and Prado argue that this condition negatively affects our possibility to identify agency and assign responsibility for the consequences of the functioning of interaction-dominant systems, including hybrids of humans and cognitive artifacts. Detecting whether hybrid cognitive systems behave according to an interaction-dominant logic can have therefore important ethical and political implications. The authors introduce ‘pink noise’, a neither random nor predictable, correlated fluctuation in the system’s behaviour over time, as a ‘signature’ of interaction-dominant systems and, therefore, as a “potentially useful heuristic, [and] a possible canary in the proverbial mine” to detect interaction-dominant dynamics (on ‘pink noise’ as a universal feature of the emergent coordination among system components see, e.g., van Orden, Kloos, and

Wallot [18]). Meacham and Prado remind us that the “seamless integration” of humans and cognitive artifacts in hybrid, interaction-dominant systems, which pink noise signals, may hamper our capacity to morally, politically, or legally evaluate an action or behaviour, as the delineation of causal roles and functional competence, which is central to this assessment, is utterly difficult or ultimately impossible in these systems. As a conclusion, the authors suggest that noise can do more than signaling the characteristics of the system’s behaviour, and that, as a disturbance and temporary disruption of tightly coupled elements of the system, it may turn to be virtue in the design of increasingly pervasive hybrid cognitive systems, creating the distance and demarcation between the elements of the system which is necessary to adjudicate and assign responsibility.

The article by Guido Gorgoni focuses on the notion of responsibility and on the implications which different definitions of responsibility have for the governance of HE. Gorgoni notes that the concept of responsibility is multifaceted and that this notion has been differently interpreted in theory and dissimilarly applied in practice. The author summarizes this variety by distinguishing four responsibility paradigms: Fault, Risk, Precaution, and Responsible Research and Innovation (RRI). Each of them has specific features in terms of temporal orientation (prospective/retrospective), agency (active/passive), guiding principles and means of responsible action. Comparing these paradigms, Gorgoni argues that RRI, “if taken seriously”, has specific characteristics which makes it uniquely positioned to tackle the challenges of HE governance. Taking RRI seriously, however, means to recognize that this governance approach is much more than engaging social actors in science, technology and innovation. In Gorgoni’s opinion, questions about ethical acceptability and social desirability cannot be separated from human rights which are essential for what he calls “the constitutional identity” of RRI. This perspective makes explicit the inherently normative and political nature of this approach which it shares with HE, an equally ideologically and normatively committed concept. According to Gorgoni, it is this similarity that gives RRI its potential, as an explicitly normative governance mechanism, to steer the debate on the fundamental, normative assumptions of HE. Reorganizing the debate on HE around the question of how it contributes to the fulfillment and implementation of human rights, this ‘constitutional approach’ to RRI may also help overcome the dilemmatic choice between

discussing “enhancements” with a “small e” (those already existing) or addressing Enhancement with a “capital E” (speculation about the future, mainly those envisioned by transhumanist discourse) [12]. From Gorgoni’s point of view, these dimensions are necessarily complementary, as they share the same fundamental assumptions about individuals and society.

The article by Toni Pustovrh examines the social embeddedness of HE (which Gorgoni highlights) by observing the technological and socio-cultural trends in the contemporary workplace and their effects on the motivations for using cognitive enhancement. In Pustovrh’s view, HE technologies provide new possibilities for adaptation, as individuals can directly restrict the human body and mind, adjusting them to specific environments, niches and demands. In this context, pharmacological cognitive enhancement (PCE) has emerged as a tool available for improving our chances of adaptation to the ongoing changes affecting the workplace. The wicked question is what human traits could or should be modified, as different environments promote distinct types of capabilities and adaptations, while they simultaneously diminish others. Efficiency and performance-related values, attitudes, expectations, and norms are surely predominant in the workplace niche, but, as Pustovrh argues, they could result as ineffective or even detrimental to individual adaptation in other social contexts, such as the family, friendship, and community where cooperation and solidarity, empathy, and emotional connectedness are socially valuable and desirable. While helping us work under more stressful and fatiguing demands, PCE can paradoxically create a positive loop intensifying performance-related expectations, thus worsening the already fragile balance between the different spheres of social life. From this point of view, Pustovrh concludes, the debate on what responsibility means for HE should not primarily look at optimizing individual adaptation to single socio-cultural niches, considered in isolation. Precisely the opposite should be the case, and the logic of *balanced integration* between the different societal domains, as well as the differentiated positions and roles individuals have within and across such domains, should guide decisions and policies on cognitive enhancement in the workplace and beyond.

The final article (by the author) presents a public engagement technique which can be used to anticipate and explore the interdependencies, alignments, and conflicts concerning HE technologies, moral principles and

practices, and responsibility paradigms and arrangements. This technique draws upon the techno-moral scenario (TMS) approach, which is a scenario method aimed at stimulating public reflection on the moral consequences of technological change [19]. This version of techno-moral scenarios, called ‘rTMS’ in the article, examines the outcome of the original technique (an account of the hypothetical evolution of ethical controversies following the introduction of new technologies) through the lenses of Gorgoni’s classification of responsibility paradigms. In this way, rTMS generate four, alternative “responsibility scenarios”, each of them based on a specific paradigm. The intent in creating alternative responsibility scenarios is not to decide beforehand what responsibility approaches are relevant or legitimate to manage the ethical controversies stirred by HE technologies, but to examine how each responsibility paradigm differently shapes the institutional arrangements that preside over the assumptions and assignment of responsibility concerning moral consequences of technological transformations. Finally, the (in-)compatibilities of the elements in these ‘ideal-typical’ responsibility scenarios are assessed, with the purpose of designing more realistic ‘meta-scenarios’ which can have a better adherence to socio-technical trajectories that are more likely to develop.

The articles all propose alternative ways to reflect on HE, in order to escape the impasse that currently characterizes the public debate on this subject. To do so, all the articles took the vantage point of responsibility. Our understanding of this notion is certainly influenced by the academic literature and policy discussion about RRI, but it cannot be reduced to it alone. Retrospectively, responsibility emerges from these article as the individual and collective disposition to anticipate and, reflect upon, and deliberate about the consequences of HE on society, and the embedding of such a disposition into organizational configurations and policy alternatives, institutions and norms, regulations and procedures, technical solutions, and cultural adaptations, so that the moral and social conflicts related to HE discourse and interventions can be addressed and the consistency between technological options, normative orientations, and social formations can be established or maintained.¹

¹ This formulation is indebted to the definition of Responsible Research and Innovation (RRI) suggested by Jeroen van den Hoven ([20], 82).

Crucially, the articles refrain from setting substantive and procedural standards to establish this consistency. Instead, they focus on delineating the conditions of allowing anticipation, reflection, and deliberation, proposing: (1) the design of deliberately “noisy” extended cognitive systems including humans and artifacts, so that agency and responsibility can still be assigned in the system (Meacham and Prado Casanova), (2) the use of human rights as a yardstick to orient science, technology, and innovation (Gorgoni), (3) the preference for integration over optimization as a point of reference for evaluating cognitive enhancement in the workplace and beyond (Pustovrh), and (4) the development of public engagement techniques and tools to explore the enmeshed relations between HE technologies, morality and responsibility (Arnaldi).

Overall, this collection does not provide a comprehensive appraisal of the links between HE and responsibility. Yet, the articles suggest three areas for further work which can advance a more inclusive understanding of this connection. Conceptually, a more complex interpretation of responsibility can improve our capacity to examine, assess, and design governance frameworks for enhancement technologies. Theoretically, acknowledging that human identity is constructed across different social domains can broaden our characterization of the relevant HE consequences for which we have to assume responsibility, while recognizing that agency is transformed by the emergent behaviour of hybrid, human-artifact systems emphasizes the importance of design requirements to enable the exercise of responsibility. Methodologically, the involvement of citizens scrutinizing HE technologies, their moral consequences, and their social implications, and subsequently deliberating about the forms of their responsible governance, require the development of public engagement techniques which can explore the co-evolution of these tightly knotted dimensions.

We dedicate the last words of this introduction to our colleague Toni Pustovrh who suddenly passed away during this project. Publishing his last work in this collection is a way to remember a young and gifted scholar, and a dear friend.

Acknowledgements The author gratefully acknowledges the funding from the Independent Social Research Foundation (ISRF), Flexible Grants for Small Research Groups program, which made it possible to conduct the project “Responsibility and Human Enhancement. Concepts, implications and assessments”, of which

the contributions to this Special Section represent one of the outcomes. My gratitude is due to the Jacques Maritain Institute (Trieste, Italy) for hosting the project and to the University of Padova (Italy) which has offered its premises and facilities for the project events and meetings. Finally, I would like to thank Arianna Ferrari for participating in the research and Franc Mali who kindly agreed to co-edit, with me, the draft of Toni Pustovrh’s paper.

References

1. Sauter A, Gerlinger K (2014) The pharmacologically improved human: performance-enhancing substances as a social challenge. BoD - Books on Demand, Norderstedt
2. Harris J (2011) *Enhancing evolution: the ethical case for making better people*. Princeton University Press, Princeton
3. Bostrom N, Roache R (2011) Smart policy: cognitive enhancement in the public interest. In: Savulescu J, ter Meulen RHJ, Kahane G (eds) *Enhancing human capacities*. Wiley, Chichester, pp 138–149
4. Garcia T, Sandler R (2008) Enhancing justice? *NanoEthics* 2:277–287. <https://doi.org/10.1007/s11569-008-0048-5>
5. Fukuyama F (2003) *Our posthuman future: consequences of the biotechnology revolution*. Profile Books, London
6. McVeigh J, Evans-Brown M, Bellis MA (2012) Human enhancement drugs and the pursuit of perfection. *Adicciones* 24:185–190
7. Nordmann A (2007) If and then: a critique of speculative nanoethics. *NanoEthics* 1:31–46. <https://doi.org/10.1007/s11569-007-0007-6>
8. Canton J (2004) Designing the future: NBIC technologies and human performance enhancement. *Ann N Y Acad Sci* 1013:186–198. <https://doi.org/10.1196/annals.1305.010>
9. Swierstra T, Stemerding D, Boenink M (2009) Exploring techno-moral change: the case of the ObesityPill. In: Sollie P, Düwell M (eds) *Evaluating new technologies*. Springer, Dordrecht, pp 119–138
10. Meacham D (2015) The subject of enhancement: augmented capacities, extended cognition, and delicate ecologies of the mind. *The New Bioethics* 21:5–19. <https://doi.org/10.1179/2050287715Z.00000000063>
11. Santoni de Sio F, van Wynsberghe A (2016) When should we use care robots? The nature-of-activities approach. *Sci Eng Ethics* 22:1745–1760. <https://doi.org/10.1007/s11948-015-9715-4>
12. Shelley-Egan C, Hanssen AB, Landeweerd L, Hofmann B (2018) Responsible research and innovation in the context of human cognitive enhancement: some essential features. *Journal of Responsible Innovation* 5:65–85. <https://doi.org/10.1080/23299460.2017.1319034>
13. Stilgoe J, Owen R, Macnaghten P (2013) Developing a framework for responsible innovation. *Res Policy* 42: 1568–1580. <https://doi.org/10.1016/j.respol.2013.05.008>
14. von Schomberg R (2013) A vision of responsible research and innovation. In: Owen R, Bessant J, Heintz M (eds) *Responsible innovation*. Wiley, Chichester, pp 51–74

15. Forsberg E-M, Quaglio G, O’Kane H et al (2015) Assessment of science and technologies: advising for and with responsibility. *Technol Soc* 42:21–27. <https://doi.org/10.1016/j.techsoc.2014.12.004>
16. Washburn A, Coey CA, Romero V et al (2015) Interaction between intention and environmental constraints on the fractal dynamics of human performance. *Cogn Process* 16:343–350. <https://doi.org/10.1007/s10339-015-0652-6>
17. Richardson MJ, Chemero A (2014) Complex dynamical systems and embodiment. In: Shapiro L (ed) *The Routledge Handbook of Embodied Cognition*. Routledge, Abingdon, pp 39–50
18. Van Orden GC, Kloos H, Wallot S (2011) Living in the pink. In: Hooker C (ed) *Philosophy of complex systems*. Elsevier, Amsterdam, pp 629–672
19. Boenink M, Swierstra T, Stemerding D (2010) Anticipating the interaction between technology and morality: a scenario study of experimenting with humans in bionanotechnology. *Studies in Ethics, Law, and Technology* 4(2). <https://doi.org/10.2202/1941-6008.1098>
20. van den Hoven J (2013) Value sensitive design and responsible innovation. In: Owen R, Bessant J, Heintz M (eds) *Responsible Innovation*. Wiley, Chichester, pp 75–83