



A Self-Guiding Tool to Conduct Research With Embodiment Technologies Responsibly

Laura Aymerich-Franch^{1*} and Eduard Fosch-Villaronga²

¹ Communication Department, Pompeu Fabra University, Barcelona, Spain, ² eLaw - Center for Law and Digital Technologies, Leiden University, Leiden, Netherlands

The extension of the sense of self to the avatar during experiences of avatar embodiment requires thorough ethical and legal consideration, especially in light of potential scenarios involving physical or psychological harm caused to, or by, embodied avatars. We provide researchers and developers working in the field of virtual and robot embodiment technologies with a self-guidance tool based on the principles of Responsible Research and Innovation (RRI). This tool will help them engage in ethical and responsible research and innovation in the area of embodiment technologies in a way that guarantees all the rights of the embodied users and their interactors, including safety, privacy, autonomy, and dignity.

OPEN ACCESS

Edited by:

Beatrice de Gelder,
Maastricht University, Netherlands

Reviewed by:

Michael Madary,
University of the Pacific, United States

Giuseppe Riva,
Catholic University of the Sacred
Heart, Italy

*Correspondence:

Laura Aymerich-Franch
laura.aymerich@gmail.com

Specialty section:

This article was submitted to
Virtual Environments,
a section of the journal
Frontiers in Robotics and AI

Received: 14 August 2019

Accepted: 06 February 2020

Published: 21 February 2020

Citation:

Aymerich-Franch L and
Fosch-Villaronga E (2020) A
Self-Guiding Tool to Conduct
Research With Embodiment
Technologies Responsibly.
Front. Robot. AI 7:22.
doi: 10.3389/frobt.2020.00022

Keywords: embodiment, responsible research & innovation (RRI), body ownership, ethics, virtual reality, social robots, embodiment technologies, avatars

INTRODUCTION

For some time now, there has been an increasing interest in the development of technologies that can couple the human body to a computer interface (Biocca, 1997). Over the years, technology has evolved to the point where it is possible to induce the illusion of embodiment (Madary and Metzinger, 2016) in a virtual (e.g., Slater et al., 2010) or a robotic avatar (e.g., Aymerich-Franch et al., 2017). Specifically, an extensive body of work in the area of virtual reality and social robots has repeatedly demonstrated that, when people embody a virtual or a robotic avatar, they experience body ownership over the body of that avatar (e.g., Slater et al., 2009; Aymerich-Franch et al., 2017) and self-location within its bodily boundaries (e.g., Lenggenhager et al., 2007; Slater et al., 2009). Crucially, during embodiment experiences, users experience the illusion that “what is apparently happening is really happening” (Slater, 2009). Hence, they respond to virtual agents, avatars (Garau et al., 2005), and threats (Slater et al., 2010) as if they were real.

Previous works have drawn attention to the importance of accounting for ethical issues in using immersive virtual reality (Southgate et al., 2017). However, the extension of the sense of self to the avatar during embodiment experiences (Aymerich-Franch, 2018) is a critical aspect that requires special ethical and legal consideration, principally, in light of potential scenarios involving physical or psychological harm caused to, or by, embodied avatars (Aymerich-Franch and Fosch-Villaronga, 2019; Aymerich-Franch et al., 2019).

The likely convergence of social networks and virtual reality represents one of the best examples of these likely scenarios. Potential threats to autonomy (i.e., the capacity to make uncoerced decisions) and privacy arising from this convergence have already been highlighted (O’Brolcháin et al., 2016). These threats, however, do not only apply to scenarios in which the technology reaches

the final user, but also to research contexts. Avatar embodiment experiments are frequently used to recreate dangerous or stressful situations for the user to study human behavior and the fact that users may experience these situations as if they were real entails important ethical challenges (Pan and Hamilton, 2018).

Unfortunately, while the pace of technology development and their applied uses for research dramatically accelerate, the understanding of its implications does not follow in parallel. Thus, the literature falls short in reflecting on the legal and ethical implications of the development and use of embodiment technologies.

The lack of specific regulatory guidelines for embodiment technologies does not help either. Although a vast number of laws and norms might already apply to avatar embodiment, emerging technologies tend to fall into an “institutional void” (Hajer, 2003), challenging the understanding of which and how the regulations apply to a particular technology (Fosch-Villaronga, 2019). The lack of guidance in this respect challenges the legal certainty concerning what boundaries need to be respected, which rights users have, what obligations developers should abide by, and what consequences exist for non-compliance (Stilgoe et al., 2013; Fosch-Villaronga and Heldeweg, 2018; Fosch-Villaronga and Golia, 2019).

Our contribution attempts to dissipate the uncertainty that this scenario raises by creating a self-guiding tool based on the principles of Responsible Research and Innovation (RRI). This tool aims at helping researchers in the field of embodiment technologies engage in ethical and responsible research and innovation processes to develop, test, and implement these technologies guaranteeing the rights of the embodied users and their interactors, including the rights of safety, privacy, autonomy, and dignity.

RRI is an overarching concept that captures crucial aspects concerning what researchers can do to ensure that research and innovation have desirable outcomes (Stahl et al., 2014). It is often the case, however, that such good intentions struggle to translate into specific, practical, and widely adopted actions. The tool that we propose contributes to materialize the principles of RRI in the specific context of research and development of embodiment technologies to overcome this problem.

RESPONSIBLE RESEARCH AND INNOVATION (RRI)

For the European Union, to help innovate responsibly and contribute to ensuring a desirable future for humanity translates into the Responsible Research and Innovation (RRI) framework (European Commission, 2012). The RRI approach provides a suitable framework to guide all the social actors involved in research and innovation (R&I) processes toward this aim. The European Commission (2019) defines RRI as “an approach that anticipates and assesses potential implications and societal expectations concerning research and innovation, intending to foster the design of inclusive and sustainable research and innovation.”

From the lens of RRI, the principles of anticipation, reflection, inclusion, responsiveness, and transparency typically guide R&I processes:

- Anticipation. Anticipation is about encouraging social actors involved in R&I processes to ask “what if” questions so that they envision contingency plans toward potential outcomes, build socially robust, risk-free research, and unveil hidden opportunities (Stilgoe et al., 2013).
- Reflection. Reflexivity encourages researchers to think mindfully about their work. Rethinking prevailing assumptions, values, and purposes in current R&I practices and activities may help raise awareness of the importance of framing issues, problems, and suggested solutions.
- Inclusion. The principle of inclusion is concerned with conducting research not only for society but *with* society and thus involving a wide range of stakeholders from the early stages of the R&I process “both for normative democratic reasons and to broaden and diversify the sources of expertise, disciplines, and perspectives” (Kupper et al., 2015).
- Responsiveness. RRI can reshape R&I processes in response to circumstances that no longer align with the continually evolving needs of society (Stilgoe et al., 2013). Responsiveness alludes to the flexibility and capacity to change R&I processes to ensure the research enforces public values.
- Transparency. Transparency encourages open-access dissemination of the results and conclusions, enabling this way public scrutiny and dialogue.

SELF-GUIDING TOOL TO CONDUCT RESEARCH WITH EMBODIMENT TECHNOLOGIES RESPONSIBLY

RRI promotes reflection upon the consequences of the outcomes of technology and fosters the incorporation of such reflections into the research and design processes. The five principles of inclusion, anticipation, reflection, responsiveness, and transparency that define RRI provide a suitable framework for conducting research and innovating responsibly in any area of R&I, including embodiment technologies. However, one of the most challenging aspects of being able to put into practice these principles is how to implement them in everyday R&I practices, accurately.

Following a basic coaching principle that *finding the right answers is about asking the right questions*, we provide a self-guiding tool with a series of critical questions inspired on Stilgoe et al. (2013), Kupper et al. (2015), and Stahl and Coeckelbergh (2016) concerning each of the five RRI principles (Table 1). Altogether, these questions work as a self-guidance tool that researchers and innovators can use to guide their R&I processes throughout all the stages, from the conception of the project to the final implementation or publication; and throughout all the dimensions, including the process itself, the product, the purposes, and the people involved (Stahl and Coeckelbergh, 2016).

By raising, reflecting, and answering these questions, researchers and developers will equip themselves with the

TABLE 1 | A self-guiding tool to conduct research with embodiment technologies responsibly.**ANTICIPATION**

- What are the psychological, ethical, moral, and legal implications that the embodiment technology I am developing could bring?
- Could someone use the embodiment technology that I am developing for unintended practices in my or other fields? Who and how? If so, what measures have I developed to mitigate this?
- What are the risks and benefits of the embodiment technology that I am developing? How will they be distributed in society if the technology gets implemented?
- If I am designing avatars that will interact with other avatars or real people, do I provide enough protection mechanisms to my users so that they can protect their avatars in case of assault? Which ones?
- If I am working with physical avatars such as robots that could potentially harm someone in case of technical failure, have I considered emergency mechanisms so that the user can completely stop the action of the avatar or alternative control mechanisms to avoid causing harm to others or to the environment? Which ones?
- Who needs to take responsibility if something goes wrong with the embodiment technology that I am developing?
- Have I projected potential scenarios in which different aspects go wrong and have I determined who should take responsibility on each case? Which ones?
- What aspects of the avatars that I am developing could make them not socially desirable? How can I change that?

REFLECTION

- Who could be negatively affected by my research with embodiment technologies? How can I change that?
- How could the embodiment technology I am developing challenge the rights of future users? How can I avoid that?
- Am I equipped with enough knowledge to identify and address the implications of embodiment technologies by myself? If not, what experts could help me and how?
- Do I lead or participate in actions oriented at addressing concerns and fears regarding embodiment technologies for the general public? Which ones? Is there anything else I can do?
- Do my results contribute to providing useful insights for developers regarding how to commercialize embodiment technologies safely for the society from an ethical point of view? How exactly? Is there anything else I can do to ensure that?
- Do my results contribute to providing useful insights for regulators and legislators regarding how to regulate and legislate embodiment technologies? How exactly? Is there anything else I can do to ensure that?
- Do I discuss the ethical implications of my results when I report them?
- Have I conducted an actor analysis to understand which actors play an essential role for the embodiment technology that I am developing and identify people, industries, institutions, or organizations who can affect or be affected by the technology?
- Have I sufficiently reflected on the benefits and risks of the embodiment technology before starting its development: does it honestly and positively contribute to society? How exactly? Is it aimed to resolve a societal challenge? Which one/s?
- Have I organized discussion groups with the different stakeholders involved in the embodiment technology that I am working with to discuss potential ethical implications and create awareness of responsibility and accountability?
- Am I open to receive criticism and consider skepticism about the embodiment technology or experiments using it and integrate it into my research/design process? What could I do differently to encourage more feedback from the relevant stakeholders?
- Does my research with embodiment technologies have long-term consequences that could be potentially negative for society? How can I avoid them?

INCLUSION

- Am I familiar with the embodiment technologies that other researchers, companies, and start-ups are developing and the experimental work that other researchers are conducting with these technologies? What can I do to know them even better?
- Who are the relevant stakeholders in the development of my research in embodiment technologies? Have I included them in the process? How exactly?
- Could the virtual and robot avatars that I create cause gender, race, religious or age discrimination? How can I avoid this?
- Do I respect the principle of diversity when I design avatars so that I have enough avatar choices that can represent well all potential participants or users and do not create conflict in terms of gender, race, ethnicity, religion, and other demographics? (e.g., I have a participant that wears a hijab in real life and does not feel comfortable embodying an avatar that does not, do I have the right type of avatar for her?) How could I increase the range of options to ensure that?
- Have I talked directly to the stakeholders that my embodiment technologies target to enquire about how the technology could really improve their quality of life rather than making assumptions about it? (e.g., I am designing a robot avatar for persons with reduced mobility, but I actually never asked them directly if that is something that could be useful for them and how exactly would they want it) Who else should I talk to?
- Do I engage the target of the embodiment technologies I design in my R&D processes throughout all the stages? How exactly? What else can I do to ensure that?

RESPONSIVENESS

- What do I need to do to ensure social desirability for the research I conduct with embodiment technologies?
- What training am I receiving to conduct research with embodiment technologies responsibly?
- If I encounter ethical conflicts, are there any barriers (e.g., economic interests) that prevent me from changing the course of the development of the embodiment technology? Which ones? What can I do to overcome them?
- Is the technology of embodiment that I am developing prepared to evolve with the constantly evolving technological landscape? How exactly? What else can I do to prepare it better in this regard?
- Is the technology of embodiment that I am developing responding clearly to current societal needs? To which ones specifically?
- If I conduct experiments with embodiment technologies, are they meant to respond to societal needs and challenges? To which ones specifically?

TRANSPARENCY

- Are the motivations of the embodiment technology that I am developing transparent, honest, and geared toward the public interest? (e.g., I write a grant in which I argue that the embodiment technology will be useful for rescue operations just to get the grant, but I already know it will not be useful)
- Am I sharing the results not only with the scientific community but also with the target of the embodiment technology? How am I doing so?
- Am I openly sharing the uncertainties and limitations of the embodiment technology that I have developed with the full range of stakeholders as well as in my publications? All of them? (If this is something difficult for me, ask: What is the worst that could happen if I did so?)
- Do I share the lessons learned of the research with embodiment technologies with my community (including the negative aspects)? How exactly do I do that?

necessary questions to further research with embodiment technologies responsibly. The tool will help researchers and developers integrate the reflections concerning the consequences of their work into their design processes and, hence, foster responsible technology, in line with societal needs and values.

While the tool specifically targets embodiment technologies, it can easily be adapted to other emerging technologies, hence, it is useful to a much wider community, including researchers from cyberpsychology, virtual reality beyond avatar embodiment, human-computer and human-robot interaction, affective computing, and other related fields.

DISCUSSION

Innovating is about creating and transforming the future. However, transforming the future does not necessarily mean for the better. Researchers might not always be able to foresee the potential negative impacts of their research on society. Users might also be more focused on the practical benefits they gain from employing the technology than on reflecting whether that is beneficial for them or not (Carr, 2011). As Parsons notes (Parsons, 2019), an unfortunate limitation of cyberpsychology literature is that “it rarely discusses the ways in which technologies are increasingly part of personhood or the ethical issues that result” (p. XVI-Preface). This work aims at mitigating this lack of reflection, at the same time, offering a tool that allows researchers to take action toward correcting potential bad practices.

The self-guiding tool to conduct research with embodiment technologies responsibly that we provide sets a series of questions taking into account the different principles of RRI to help researchers and developers steer the development of embodiment technologies into a desired and socially accepted direction.

The extension of the sense of self to the avatar implies that if the avatar is harmed, the embodied user experiences psychological harm as a result (Aymerich-Franch and Fosch-Villaronga, 2019). Also, it implies that, if the avatar causes harm due to technical failure, the users can falsely attribute responsibility to themselves, even if it is not their fault (Aymerich-Franch et al., 2019). A mindful reflection of the potential implications of embodiment technologies may prevent the occurrence of undesired outcomes such as the exacerbation of existing behaviors, sexual harassment of the avatar, or the wrong self-attribution of responsibility.

As McBride and Stahl (2014) highlight, RRI has developed at the governance level, but this does not ensure that practitioners will follow it, as governance procedures usually end up being considered as “hurdles to be jumped over and administrative rituals to be fulfilled.” In this respect, it is essential to promote, in parallel to actions at the governance level, institutional change

REFERENCES

- Aymerich-Franch, L. (2018). Is mediated embodiment the response to embodied cognition? *New Ideas Psychol.* 50, 1–5. doi: 10.1016/j.newideapsych.2018.02.003
- Aymerich-Franch, L., and Fosch-Villaronga, E. (2019). What we learned from mediated embodiment experiments and why it should matter to policymakers. *Presence Teleoperat. Virtual Environ.* 27, 63–67. doi: 10.1162/pres_a_00312

to foster reflection on the consequences of the technology among researchers. To this end, the effective integration of all these reflections into the R&I process implies, as a first step, inviting researchers to ask themselves a series of questions focused on understanding whether the technological development that they carry out aligns with the RRI principles. The answer to these questions may trigger the researcher to take action (e.g., by establishing dialogues with different stakeholders or by seeking further guidance and additional training to equip themselves with the adequate tools to better understand how to frame possible concerns and how to mitigate them).

In this paper, we have taken one step forward toward bridging the gap between conceptual and applied RRI by translating the general principles of the RRI framework into a practical tool for conducting research with embodiment technologies. This self-guiding tool is intended to help researchers give careful thought to the consequences of the technology they develop in an anticipatory, inclusive, reflective, responsive, and transparent manner. That said, the tool we present is by no means to replace regulatory and ethical compliance processes, but an invitation to reflect deeply and consciously about the societal implications of work with embodiment technologies. It aims to the realization of the RRI goals by creating a practical, integrative, reflective mechanism geared toward addressing embodiment technologies’ societal implications.

To conclude, innovating in a responsible manner contributes to ensuring a desirable future for humanity. By carrying out the self-assessment, we hope researchers in this field will become more sensitive to how essential it is to steer their research in a responsible direction. On most of the occasions, this process will push the boundaries toward a more interdisciplinary, integrative, and thoughtful model of conducting research that may also be more beneficial for the society in the long run.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

LA-F was supported by Programa de Ayudas Ramón y Cajal (Ref. RYC2016-19770), Agencia Estatal de Investigación, Ministerio de Ciencia, Innovación y Universidades, y Fondo Social Europeo. EF-V was supported by the LEaDing Fellows Marie Skłodowska Curie COFUND fellowship, a project that has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 707404.

- Aymerich-Franch, L., Kishore, S., and Slater, M. (2019). When your robot avatar misbehaves you are likely to apologize: an exploration of guilt during robot embodiment. *Int. J. Soc. Robot.* doi: 10.1007/s12369-019-00556-5. [Epub ahead of print].
- Aymerich-Franch, L., Petit, D., Ganesh, G., and Kheddar, A. (2017). Object touch by a humanoid robot avatar induces haptic sensation in the real hand. *J. Comput. Mediat. Commun.* 22, 215–230. doi: 10.1111/jcc.4.12188

- Biocca, F. (1997). The cyborg's dilemma: progressive embodiment in virtual environments. *J. Comput. Mediat. Commun.* 3:JCMC324. doi: 10.1111/j.1083-6101.1997.tb00070.x
- Carr, N. (2011). *The Shallows: What the Internet Is Doing to Our Brains*. New York, NY: WW Norton & Company.
- European Commission (2012). *Options for Strengthening Responsible Research & Innovation*. Available online at: https://ec.europa.eu/research/science-society/document_library/pdf_06/options-for-strengthening_en.pdf (accessed December 20, 2019).
- European Commission (2019). *Responsible Research & Innovation*. Available online at: <https://ec.europa.eu/programmes/horizon2020/en/h2020-section/responsible-research-innovation> (accessed December 20, 2019).
- Fosch-Villaronga, E. (2019). *Robots, Healthcare, and the Law. Regulating Automation in Personal Care*. New York, NY: Routledge.
- Fosch-Villaronga, E., and Golia, A. Jr. (2019). Robots, standards and the law: rivalries between private standards and public policymaking for robot governance. *Comput. Law Security Rev.* 35, 129–144. doi: 10.1016/j.clsr.2018.12.009
- Fosch-Villaronga, E., and Heldeweg, M. (2018). "Regulation, I presume?" said the robot—towards an iterative regulatory process for robot governance. *Comput. Law security Rev.* 34, 1258–1277. doi: 10.1016/j.clsr.2018.09.001
- Garau, M., Slater, M., Pertaub, D.-P., and Razaque, S. (2005). The responses of people to virtual humans in an immersive virtual environment. *Presence*. 14, 104–116. doi: 10.1162/1054746053890242
- Hajer, M. (2003). Policy without polity? Policy analysis and the institutional void. *Policy Sci.* 36, 175–195. doi: 10.1023/A:1024834510939
- Kupper, F., Klaassen, P., Rijnen, M., Vermeulen, S., and Broerse, J. E. W. (2015). *Report on the Quality Criteria of Good Practice Standards in RRI*. Amsterdam: RRI Tools; Athena Institute, VU University Amsterdam.
- Lenggenhager, B., Tadi, T., Metzinger, T., and Blanke, O. (2007). Video ergo sum: manipulating bodily self-consciousness. *Science* 317, 1096–1099. doi: 10.1126/science.1143439
- Madary, M., and Metzinger, T. K. (2016). Real virtuality: a code of ethical conduct. Recommendations for good scientific practice and the consumers of VR-technology. *Front. Robot. AI* 3:3. doi: 10.3389/frobt.2016.00003
- McBride, N., and Stahl, B. (2014). "Developing responsible research and innovation for robotics," in *Proceedings of the IEEE 2014 International Symposium on Ethics in Engineering, Science, and Technology* (Chicago, IL), 27.
- O'Brolcháin, F., Jacquemard, T., Monaghan, D., O'Connor, N., Novitzky, P., and Gordijn, B. (2016). The convergence of virtual reality and social networks: threats to privacy and autonomy. *Sci. Eng. Ethics* 22, 1–29. doi: 10.1007/s11948-014-9621-1
- Pan, X., and Hamilton, A. F. D. C. (2018). Why and how to use virtual reality to study human social interaction: the challenges of exploring a new research landscape. *Br. J. Psychol.* 109, 395–417. doi: 10.1111/bjop.12290
- Parsons, T. D. (2019). *Ethical Challenges in Digital Psychology and Cyberpsychology*. Cambridge; New York, NY: Cambridge University Press.
- Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 3549–3557. doi: 10.1098/rstb.2009.0138
- Slater, M., Perez-Marcos, D., Ehrsson, H. H., and Sanchez-Vives, M. V. (2009). Inducing illusory ownership of a virtual body. *Front. Neurosci.* 3, 214–220. doi: 10.3389/neuro.01.029.2009
- Slater, M., Spanlang, B., Sanchez-Vives, M. V., and Blanke, O. (2010). First person experience of body transfer in virtual reality. *PLoS ONE* 5:e10564. doi: 10.1371/journal.pone.0010564
- Southgate, E., Smith, S. P., and Scevak, J. (2017). "Asking ethical questions in research using immersive virtual and augmented reality technologies with children and youth," in *2017 IEEE Virtual Reality (VR)*. (IEEE), 12–18.
- Stahl, B. C., and Coeckelbergh, M. (2016). Ethics of healthcare robotics: Towards responsible research and innovation. *Robot. Autonomous Syst.* 86, 152–161. doi: 10.1016/j.robot.2016.08.018
- Stahl, B. C., McBride, N., Wakunuma, K., and Flick, C. (2014). The empathic care robot: a prototype of responsible research and innovation. *Technol. Forecast. Soc. Change* 84, 74–85. doi: 10.1016/j.techfore.2013.08.001
- Stilgoe, J., Owen, R., and Macnaghten, P. (2013). Developing a framework for responsible innovation. *Res. Policy* 42, 1568–1580. doi: 10.1016/j.respol.2013.05.008

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Aymerich-Franch and Fosch-Villaronga. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.