Moral debate or rebellion

The power of psychological reactance and its potential role in interaction design

Sarah Diefenbach

Ludwig-Maximilians-Universität München, Department of Psychology, sarah.diefenbach@psy.lmu.de

Lara Christoforakos

Ludwig-Maximilians-Universität München, Department of Psychology, lara.christoforakos@psy.lmu.de

Daniel Ullrich

Ludwig-Maximilians-Universität München, Department of Computer Science, daniel.ullrich@ifi.lmu.de

So-called "moral agents", here understood as agentic technology to support or initiate a desired behavior with a moral component (typically related to societal goals e.g., sustainability, health), open up an interesting dimension in human-technology relations. Technology is no longer a mere tool but rather acts as a social counterpart with its own position. From a psychological perspective, the question of whether the potential of Things plays out, and a human is willing to engage in the ethical debate offered by the moral agent, may depend on various factors of person, design, and context. Our position paper focuses on three interrelated aspects, which we would like to discuss and empirically explore in the CHI workshop and beyond: (1) technology's role as social counterpart, (2) human-experienced autonomy, and (3) reactance as a potential outcome, i.e., people refusing an attempt to be influenced and, instead, acting in an opposite way.

1 INTRODUCTION

When things become moral counterparts, a new kind of human-technology relation emerges, prompting a number of related psychological research questions. The concept of "moral agents" is still under development. While part of the literature is concerned with the morality of things as a side effect of their primary task (e.g., [2, 14]), such as the moral decisions a car has to make in the context of autonomous driving (e.g., [8]), the call of the CHI 2023 Workshop on Moral Agents for Sustainable Transitions focuses on things with moral messages as primary intention (e.g., the message to take the bike instead of the car) and their potential for sustainability. Implicitly, this suggests that humans' current behavior may

not be good/sustainable enough and that technology can somehow induce a behavioral change towards a more sustainable, morally superior position. However, the question of whether this assumed potential of the technology is realized and a human is willing to engage in the ethical debate offered by the moral agent, may depend on various factors of person, design, and context. From a social and motivational psychology perspective, we see at least three interrelated aspects, which we would like to discuss and empirically explore in the CHI workshop and beyond:

- Social counterpart: The central characteristic of "things with attitude" [4] and making technology appear as "some*one*" instead of "some*thing*" is that they slip into the role of a social counterpart. Even with technologies that are not specifically designed with anthropomorphic features, such as a regular computer or a printer, the dialogue between human and technology can activate behavioral tendencies and social norms known from interaction with humans, such as politeness or reciprocity [5]. Especially in the context of sustainability, this sociality and attributed agency could be a powerful tool. Ignoring "someone" (e.g., a trash bin speaking to you, a polar bear avatar who reminds you "please don't kill me" when you open the fridge for too long) is much harder than just ignoring a lifeless information sign. On the other hand, the social aspect could also evoke mistrust or reactance and questions such as: "Who gives you the right to tell me what to do?", or "Who gives you (the designer) the right to speak with the voice of a polar bear, or even nature itself?"

- Autonomy: Autonomy and self-actualization are basic human needs (e.g., [9, 10]). People may strive to become their "true selves" and the person they want to be. In line with this, several attempts in HCI have suggested the potential to make things that support people in reaching their personal goals (e.g., [13]). From a humanistic psychology perspective, people are willing to change and improve, and only need the right context that supports them with the concrete specific behavior change. However, the potential of the human drive for self-improvement only works out in the long run if people are intrinsically motivated, i.e., if people feel it is truly "their" goal and not an imposed societal goal (cf., [9]). While sustainable behavior is a widely shared goal, from an autonomy perspective we also need some space to make a goal our own. From this point of view, the agency of things could be a hurdle, whereas the perspective of technology as a tool offers more room for humans to see themselves as the deciding agent.

- Reactance: A potential outcome of a felt threat to autonomy is reactance. Psychological reactance ([1]; for a research review see [11]), colloquially speaking, describes a phenomenon where people refuse an attempt to be influenced and rather move in the opposite direction. More scientifically speaking, reactance is "an unpleasant motivational arousal that emerges when people experience a threat to or loss of their free behaviors" and "serves as a motivator to restore one's freedom" ([11], p. 205). All of us have probably experienced moments of reactance in our daily life when someone else wants to tell us how to behave—be it our mother, partner, boss, or the neighbor who knows everything best. Even if we agreed in principle, we don't like to be pushed or nagged too much. Especially in these current times, where seemingly simple elements of our daily life (e.g., what we eat and drink, the car we drive, the language we use, the clothes we wear) become increasingly an issue of moral debate, people may be fed up with too many comments on their behavior and search for moral-free spaces to relax.

From a mere reactance perspective, the most successful moral agent might be one that exerts influence in an unnoticed way. In other words, a technology manipulating the human to perform a desired behavior in such a way that humans do not realize they are being influenced—as is the basis of many effective advertising and marketing tools. In the classification scheme of design influence by Tromp, Hekkert and Verbeek [12], this combination of hidden and strong influence is called "decisive", i.e., the product (or the designer behind the product) decides for the human. If this idea somehow doesn't sit well with you—even though being manipulated is for a "good reason" (such as saving the environment)—what you experience is probably reactance. From a moral perspective, such concealed, "reactance-avoiding" influence is questionable and puts the morality of moral agents into question.

2   RESEARCH AND DESIGN GOALS

Based on the considerations above, we would like to explore the interrelations between these psychological and design factors. For example, is there an inherent tension between a technology's perceived agency and an individual's perceived autonomy? If we assume reactance as a potential outcome of these two factors, with agency increasing reactance and autonomy reducing reactance, is there a "sweet spot" where reactance is reduced to a minimum while the technology still represents a powerful impulse? Is the level of reactance related to the goal of behavior change and the profiteers of it? For example, referring to the issue of sustainability, is my reactance the same when a moral agent claims to act on behalf of the community, or indeed the planet? What are design factors that directly or indirectly affect reactance as a depending variable? For example, could an "opt-in" option, i.e., an explicit activation of the agent, reduce reactance by making the confrontation with the agents' moral advice self-chosen? Figure 1 summarizes these research interests.
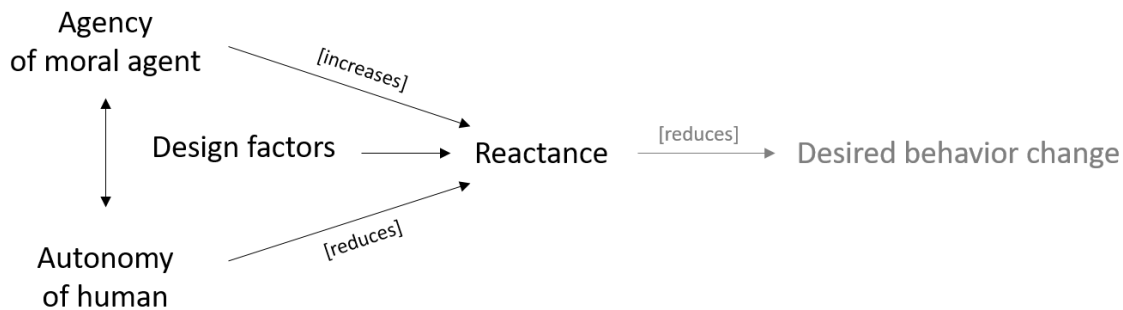

Figure 1: Interrelations of interest for empirical exploration (terms in brackets show assumed directions of influence).

From a design perspective, we would like to explore aspects that could help to reduce reactance in the way a moral agent positions itself and conveys its message. In addition to previous approaches in HCI, such as to equip technology with humor to increase acceptance (for an overview of humor in HCI see [7]) or to use naivety to strengthen the bond "between a person and a slightly annoying object" [3], we would like to borrow insights from other contexts where reactance is a sensible issue for effective behavior change, e.g., the design of health messages based on communication theory [6].

ACKNOWLEDGMENTS

REFERENCES

[1] Jack W. Brehm. 1966. A theory of psychological reactance. Academic Press.

[2] José-Antonio Cervantes, Sonia López, S., Luis-Felipe Rodríguez, Salvador Cervantes, Francisco Cervantes, and Francisco Ramos. 2020. Artificial moral agents: A survey of the current status. Science and Engineering Ethics, 26, 501-532. https://doi.org/10.1007/s11948-019-00151-x

[3] Marc Hassenzahl and Mathias Laschke. 2015. Pleasurable Troublemakers: Gamification and Design. In The Gameful World: Approaches, Issues, Applications, Steffen P. Walz and Sebastian Deterding (Eds.). MIT Press, Cambridge, London, 167–195.

[4] Matthias Laschke, Marc Hassenzahl, and Sarah Diefenbach. 2011. Things with attitude: Transformational products. Create11 conference. 2011.

[5] Jong-Eun Roselyn Lee and Clifford I. Nass. 2010. Trust in computers: The computers-are-social-actors (CASA) paradigm and trustworthiness perception in human-computer communication. Trust and technology in a ubiquitous modern environment: Theoretical and methodological perspectives. IGI Global, 2010. 1-15. https://doi.org/10.4018/978-1-61520-901-9.ch001

[6] Edward W. Maibach, and Roxanne Parrott. 1995. Designing health messages: Approaches from communication theory and public health practice. Sage.

[7] Anton Nijholt, Andreea I. Niculescu, Alessandro Valitutti, and Rafael e. Banchs. 2017. Humor in human-computer interaction: a short survey. In Proceedings of the 16th IFIP TC.13 International Conference on Human Computer Interaction, Indian Institute of Technology Madras, 192-214.

[8] Sven Nyholm. 2018. The ethics of crashes with self-driving cars: A roadmap, II. Philosophy CompassVolume 13, Issue 7 e12506. https://doi.org/10.1111/phc3.12506

[9] Richard M. Ryan and Edward L. Deci. 2000. Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. American Psychologist, 55(1), 68-78. https://doi.org/10.1037/0003-066X.55.1.68

[10] Kennon M. Sheldon, Andrew J. Elliot, Youngmee Kim, and Tim Kasser. 2001. What is satisfying about satisfying events? Testing 10 candidate psychological needs. Journal of personality and social psychology, 80(2), 325. https://doi.org/10.1037/0022-3514.80.2.325

[11] Christina Steindl, Eva Jonas, Sandra Sittenthaler, Eva Traut-Mattausch, and Jeff Greenberg. 2015. Understanding psychological reactance. Zeitschrift für Psychologie, 223(4), 204-214. https://doi.org/10.1027/2151-2604/a000222

[12] Nyke Tromp, Paul Hekkert, and Peter-Paul Verbeek. 2011. Design for socially responsible behavior: a classification of influence based on intended user experience. Design issues, 27(3), 3-19. https://doi.org/10.1162/DESI_a_0008

[13] John Zimmerman. 2009. Designing for the self: making products that help people become the person they desire to be. In Proceedings of the 2009 CHI Conference on Human Factors in Computing Systems (CHI '09). Association for Computing Machinery, New York, NY, USA, 395-404. https://doi.org/10.1145/1518701.1518765

[14] John Zoshak and Kristin Dew. 2021. Beyond kant and bentham: How ethical theories are being used in artificial moral agents. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21). Association for Computing Machinery, New York, NY, USA, 1-15. https://doi.org/10.1145/3411764.3445102