



Delft University of Technology

Emotions and Digital Well-Being

The Rationalistic Bias of Social Media Design in Online Deliberations

Marin, L.; Roeser, S.

DOI

[10.1007/978-3-030-50585-1_7](https://doi.org/10.1007/978-3-030-50585-1_7)

Publication date

2020

Document Version

Final published version

Published in

Ethics of Digital Well-Being

Citation (APA)

Marin, L., & Roeser, S. (2020). Emotions and Digital Well-Being: The Rationalistic Bias of Social Media Design in Online Deliberations. In B. Christopher, & F. Luciano (Eds.), *Ethics of Digital Well-Being: A Multidisciplinary Approach* (Vol. 140, pp. 139-150). (Philosophical Studies Series; Vol. 140). Springer. https://doi.org/10.1007/978-3-030-50585-1_7

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

<https://www.openaccess.nl/en/you-share-we-take-care>

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

Chapter 7

Emotions and Digital Well-Being: The Rationalistic Bias of Social Media Design in Online Deliberations



Lavinia Marin and Sabine Roeser

Abstract In this chapter we argue that emotions are mediated in an incomplete way in online social media because of the heavy reliance on textual messages which fosters a rationalistic bias and an inclination towards less nuanced emotional expressions. This incompleteness can happen either by obscuring emotions, showing less than the original intensity, misinterpreting emotions, or eliciting emotions without feedback and context. Online interactions and deliberations tend to contribute rather than overcome stalemates and informational bubbles, partially due to prevalence of anti-social emotions. It is tempting to see emotions as being the cause of the problem of online verbal aggression and bullying. However, we argue that social media are actually designed in a predominantly rationalistic way, because of the reliance on text-based communication, thereby filtering out social emotions and leaving space for easily expressed antisocial emotions. Based on research on emotions that sees these as key ingredients to moral interaction and deliberation, as well as on research on text-based versus non-verbal communication, we propose a richer understanding of emotions, requiring different designs of online deliberation platforms. We propose that such designs should move from text-centred designs and should find ways to incorporate the complete expression of the full range of human emotions so that these can play a constructive role in online deliberations.

Keywords Online emotions · Online deliberation · Text communication · Social media · Well-being · Deliberation platforms

L. Marin (✉) · S. Roeser

Ethics and Philosophy of Technology Section, Department of VTI, Faculty of TPM, TU Delft, Delft, The Netherlands

e-mail: l.marin@tudelft.nl

7.1 Introduction

In the emerging debate concerning the multiple facets of digital well-being (Burr et al. 2020, p. 1), one important aspect still in need of research is the inter-personal dimension of well-being. For example, the interpersonal relations established between friends, family or work colleagues contribute to the individual sense of well-being, suggesting that, when it comes to well-being, we are not the masters of our individual happiness, but that we rely on others and on the quality of our relations with others. For a long time, these inter-personal relations have taken place offline, however, with the advent of the digital society, people also interact with each other online increasingly. Concerning the personal aspect of inter-relational well-being, some researchers have argued that there can be genuine online friendships (Kaliarnta 2016) hence opening up the possibility for other personal relations to be also mediated online. There is however another aspect of inter-personal well-being that has been less studied when it comes to its online dimension: the public aspect of relating to others as a member of a community with particular interests and values. The membership to an online community can happen in formal ways such as being part of a closed group, or by following of a public figure or topic on social media, but also informally – someone finds oneself situated on the same side with other strangers when engaging in a debate around a matter of concern.

This public aspect of online well-being has been related to governance and social development by Burr et al. (2020) and it refers to the fact that most people aim to have a good life not just in the private sphere of their homes, but also in the public realm. This public dimension of inter-personal well-being deserves further exploration given the recent turn toward digital citizenship (Mossberger et al. 2008; Isin and Ruppert 2015) and online platforms for civic participation. While the discussion concerning e-democracy has been going on for a while, we want to engage it from a different angle: can we genuinely pursue public well-being to the same extent online as offline? In this paper, we are not asking whether e-democracy and online civic participation are possible or even effective, but whether participants can achieve some level of well-being as a result of their public engagement online as it is presumably the case with offline democratic participation. We have focused the scope of our question to the possibility of creating meaningful deliberations on social media.

Online social media is chosen here as an object of inquiry because digital online platforms dedicated specifically to deliberation are still in their infancy (Verdiesen et al. 2016) and because, as we will claim, the design principles that make social media inefficient in channelling online debates are the same design principles also used on digital deliberation platforms.¹ Our main claim in this paper is that online debates on social media do not mediate the full range of human emotions and thus are an impediment for successful deliberations online. We argue that we need to

¹ Examples of such platforms are LiquidFeedback <https://liquidfeedback.org/>, Debate Hub <https://debatehub.net/>, DemocracyOS <https://democracyos.org/>

rethink how we design for online deliberation, by keeping emotions in mind. We base this claim on four observations:

1. We need to take into account emotions in deliberation because they point out what matters to people.
2. Online platforms tend to mediate user's emotions in an impoverishing way, making visible only a narrow range of emotions, mostly expressed in radical terms.
3. Online social media platforms facilitate debates by emphasising the mediation of text-centred messages; however, textual communication has a hidden rationalistic bias, downplaying emotions and embodied interactions by design.
4. Paradoxically, this rationalistic bias gives room for specific kinds of emotions, namely often hostile responses that are not corrected due to lack of non-verbal communication.

These observations will be illustrated and developed in this chapter.

7.2 The Contribution of Emotions to Deliberative Processes

Public debates about controversial topics, taking place online or offline, are frequently heated and end up in stalemates, for example debates about potentially risky technological and scientific developments such as climate change, vaccination or genetic modification. This is due to the scientific and moral complexities of these risks, which lead to strong emotional responses by people (Slovic 2010). This effect is exacerbated by social media: the way emotions are typically treated in online debates increase estrangement and polarization. People from different informational 'bubbles' blame each other for seeing the world in an irrational and lop-sided way. Such hostile online interactions have the potential to affect people's wellbeing severely. In other words, online environments can be a platform for deliberation on technological risks, but they can themselves also give rise to negative impacts or risks.

However, we would like to point out that the role of emotions is usually misunderstood in public deliberations. Rather than seeing emotions as irrational states, we will argue in what follows that emotions can contribute to emotional-moral reflection and public deliberation on controversial topics such as technological risks.

Emotions are often seen as a threat to rationality, in public exchanges but also in academic research, for example in empirical decision theory (Dual Process Theory, e.g. Kahneman 2011) as well as in moral philosophy, where the opposition between 'rationalism' and 'sentimentalism' has dominated the metaethics debate. However, emotion researchers in psychology and philosophy have argued over the last decades that emotions are intertwined with or part of rationality and cognition. For example, the neuropsychologist Antonio Damasio (1994) has shown that people who lack emotions due to a brain defect (in their amygdala) lose their capacity to be practically rational and to make concrete moral judgments. Psychologists and philosophers have developed so-called cognitive theories of emotions (Lazarus 1994;

Scherer 1984; Solomon 1993). Emotions play an important role in moral wisdom and in forming moral judgments (Little 1995; Nussbaum 2001; Zagzebski 2003; Roberts 2003; Roeser 2011; Roeser and Todd 2014). These insights can be expanded to discussions about risky and controversial technologies. Emotions are crucial to debates about technological risks, because emotions can point out what morally matters. Conventional, quantitative approaches leave out important ethical considerations such as justice, fairness, autonomy and legitimacy (Roeser 2006, 2018). In this chapter we will argue that addressing emotions in a different way can help to overcome stalemates in deliberations about controversial topics: emotions can contribute to sympathy and understanding of shared values, which can in turn contribute to finding commonly shared solutions, thereby also contributing to people's wellbeing.

In what follows, we will argue that, when emotions are not properly included in ways online deliberative platforms are designed, this leads to impoverished and lopsided interactions in which nuanced emotions get lost while harmful emotions tend to prevail. These impoverished communications online can be harmful to people's wellbeing. We will argue that paradoxically, this is due to the rationalistic, text-based bias of such platforms, as they leave out emotions and embodied, non-verbal communications. Our main claim is that online social media platforms rely heavily on text-based communication and miss important nonverbal aspects of communication. We will argue that this also has an effect on how emotions are perceived and expressed.

7.3 Online Emotions and the Tendency for Extreme Emotions to Prevail

Is there a specific mode in which emotions appear when debating with other users on social media? In order to tackle this issue, we start from the observation that the ways in which emotions are mediated on social media are already leading to the expression of a narrow range of emotions in specific ways. In other words, the landscape of online emotions is rather barren, dominated by a few main emotions to the detriment of emotional diversity and complexity. The case of moral outrage will illustrate this claim as outrage seems to be emphasised to the detriment of other emotions in online deliberations.

It seems that the already polarised emotional responses in public debates are exacerbated when taking place in online media. On social media, one commonly encounters extreme negative reactions such as venting of anger, blaming or shaming which lead to the dominance of extreme viewpoints. One of the most visible emotions online is outrage, and this is also one of the most studied emotions in online contexts. The Internet has been deemed the medium of outrage (Han 2017, p. 8), as it is conducive to the expression of waves of outrage in a visible manner such as group bullying, harassment, and online mobbing. Outrage, although it has a

negative valence, can still have constructive civic uses as for example leading to mobilisations for action (Spring et al. 2018, 1068). Spring et al. (2018) have argued that outrage can be used more effectively for mobilising groups of people into political action, as compared to empathy or reappraisal (Spring et al. 2018, S. 1067). However, Brady and Crockett (2019) have pointed out that, at least in online environments, the expression of outrage does not lead to social mobilisation, but rather has mostly negative effects (Brady and Crockett 2019, S. 79). Brady and Crockett identify at least two problems with online outrage: first, it reduces the effectiveness of collective action - since there are so many themes to be outraged about, the moral anger tends to dissipate instead of coalescing among online users (Brady and Crockett 2019, S. 79); secondly, the participation of certain marginalised groups is discouraged - especially of minorities - via “coordinated harassment” (Brady and Crockett 2019, S. 79), hence outrage becomes effectively an anti-democratic tool. Spring et al. (2018) have noticed that outrage is seen as morally permissible only for majority groups who will tend to deem the expression of outrage as inappropriate when it comes from marginalised groups: ‘only certain groups are “allowed” to express outrage. For example, stigmatized group members are often held to higher moral standards (e.g., accused of expressing inappropriate emotions, especially anger, at greater rates than majority group members’ (Spring et al. 2018, p. 1069).

Both problems are related to the architecture of the online environment which, according to Brady and Crockett, makes it too easy to express outrage as a reaction to anything as it does not incur any costs (Brady and Crockett 2019, S. 79). This means that it can be the case that people manifest outrage in their messages without even feeling it because they may have other benefits, for example as a way of virtue-signalling to the group (Spring et al. 2018, p. 1067). This could also be the case in offline situations, but the offline outrage is easier to detect by the audience because emotions are harder to fake in real life.

The previously cited work on outrage showcases a general problem with online emotions: while a particular emotion can have a social role and be useful in certain contexts, it may become toxic when mediated via online social platforms. As Brady and Crockett (2019) rightly point out, the costs of expressing any emotion online – hence also outrage – are quite small and thus the sheer quantity of online outrage seems to be overwhelming. While we agree with Brady and Crockett in general, we think that we need to revisit the link between the design of a platform and the emotions it allows to be expressed. If expressing outrage is cost-free, why is it not the case that all emotions are equally expressed? What makes outrage so special in its flourishing in online media?

Empirical studies have shown that both negative and positive emotions flourish in online debates just as is the case in offline debates (Wojcieszak et al. 2009, p. 1082). Several researchers have concluded that the online medium as such is not an impersonal medium, no less devoid of emotions, and that, even if text-based communication makes it harder to convey emotional cues, users will compensate for this feature “by the use of emoticons, or by verbalizing emotions in a more explicit way” (Derks et al. 2008, p. 780). We do not contest that emotions can be expressed effectively in online communications, but rather we want to question the

quality of the expression of such emotions and the effect this has on online users. When online representations of emotions are based either on self-reports or on contextual information such as usage of expressions, emoticons, and typography, the effect is not quite the same. We think that emotions which are not fully and accurately expressed online do not achieve the same effect as the emotions in offline, real-life scenarios, as we will argue in what follows.

7.4 The Rationalistic Bias of Text-Based Online Communications

The misrepresentation of emotions online can be traced to the design choices made by social media platforms. One salient design feature is the heavy reliance on texts to convey messages among users. We only need to look back at the history of the text as medium to understand its rationalistic bias. The invention of the optical texts² in the twelfth century is tied by media historians to the need to design texts as a tool for the purpose of quick intellectual appraisal of complex arguments, enabling to see and comprehend arguments at a glance (Illich 1993). Thus, text was not initially used to express emotions, but to convey complex ideas to a wide audience in the form of books, journals, novels, etc. An exception is of course written literature (novels, poetry etc.). But until the advent of the Internet, most people (except for professional literary, academic or journalistic writers) did not use text as a regular medium of communication, except for the occasional letter writing. The epistolary novel shows how the genre of letter writing could be used to express rich emotions. However, the success of this genre was due to the talent and training of the authors. The fact remains that most people do not have the writing skills needed to express a full range of emotions in writing. Furthermore, even letters were long texts with delayed delivery, forcing the correspondents to write in a different way than they would have spoken, conveying thoughts and ideas with a timelier effect. By contrast, online instant messages were designed specifically with the purpose of replacing fully the need for face-to-face communication. Social media posts function similarly to instant messaging: they allow for the immediate publishing of updates, and of comments of quick responses. The assumption that we can communicate just as effectively via text messages as we can do in speech is buried deeply in the design of social media platforms. This poses several problems for those wishing to engage in online deliberation

A first problem is that online text-messages lack certain meta-communicational features which are essential for a successful act of communication. Offline real-life interactions involve nonverbal communication conveyed via tone of voice, gestures,

²Optical texts are texts written in such a way as to be readable at a glance, in silence. Before the twelfth century, most manuscripts were written in *scripta continua*, demanding users to read them out loud so to understand the content (see Marin et al. 2018 for a more comprehensive discussion).

and facial expressions. Such non-verbal communication can provide for essential clues as how to understand someone's words: the same words, for example a simple expression such as 'poor you!', can have completely different meaning depending on whether they are expressed in a caring, ironic, mocking or even aggressive way. How could one compensate for the lack of the meta-communicational features? Emoticons could in principle accommodate that to some degree, but even they can be ambiguous. Another solution would be more text devoted specifically to explaining the emotions of the users. This could work in a lengthy text such as a novel, where a detailed description of the characters might also work to render explicit the emotions of the characters. However, because instant messaging and social media favour short messages, there is not enough time and space to write lengthy descriptions of feelings. An additional impediment is that most online users are not professional writers, and may not be able to use words to convey something usually left to gestures, tone of voice and facial expressions. This makes it that much of the emotions and subtext are often lost in the process of online communication.

Furthermore, a second problem is that in such heavily text-based forms of communication, we need to rely on the honesty of the users when they report their emotions, and we need to take their emotional reports at face value. But online, behind the veil of anonymity, deception may be more likely to happen than sincerity. Already pointed out by Brady and Crockett, people may express outrage online even when they are not feeling it – because it signals their virtue. This could happen just as well with other emotions besides outrage.

A third problem can be that users do not report all their emotional states. For users to report an emotion in a text medium, it should be powerful enough to disturb them. Thus, mild emotions such as boredom, curiosity, amusement, annoyance, etc. may pass by unnoticed unless their subjects will take the time and effort to state clearly what they are feeling. But online users have no incentives to continuously report their feelings via text updates, while their emotions would often get noticed in offline interactions via non-verbal communication such as bodily or facial expressions, and these could be relevant for interactions between people. Meanwhile, more extreme emotions such as anger, excitement or hatred will get the spotlight much easier. For extreme emotions, one does not even need to self-report, sometimes the way a message is written is enough to tell something about the emotional state of the user: exclamation marks and using Caps Lock are indicative of more extreme emotions.

Hence, when compared to actual face-to-face communication, text-based messages are typically poor indicators of the full range of emotions felt by users. Certain emotions manage to prevail in online interactions while others do not get expressed and shared properly. To get back to our previous question as to why outrage seems to be a more prevalent emotion online than in offline interactions: this could be due to the poor expression of the full range of emotions on social media, thus leading to a distorted view of the other users and of the online environment. The visibility of online outrage may be due to its multiple modes of expression, which are not shared by other, more subtle emotions. To use a fortunate expression of Brady and Crockett, online we are confronted with "emotional noise" (p. 79) meaning that the heavy

expression of certain (extreme) emotions tends to drown other emotions, making their signalling invisible. Furthermore, this may also be due to the urgency of outrage and the fact that it cannot be controlled easily, and that outrage can be used for ‘virtue signaling’ more than other emotions. And since the Internet is a medium of self-presentation facilitating spontaneous interactions (Nadkarni and Hofmann 2012), people might choose easy ways to present a virtuous self. These are hypotheses which would require further empirical and conceptual research.

The text-centred way in which online social media platforms are designed is focused on the informational content of what is said, thereby often obfuscating the emotional and expressive context. Such design ignores the possibility that not all the messages we communicate online are meant to inform, and that many times we say things just to vent out, to express emotions, or to signal to others our allegiance to a community. This design underlies the rationalistic assumption previously exposed when we discussed the pitfalls of offline deliberation. Online deliberation seems to assume that users are rational epistemic agents whose emotions do not matter. The online platforms thus repeat the same mistakes as in-real-life deliberation: focusing on what is said and ignoring the context and personal meaning of a message. Thus, context and subtext are regularly downplayed as insignificant add-ons to the message’s meaning. The paradox of online communication is that its rationalistic bias promoted by the text-based communication creates room for populist emotions, but not for more nuanced, sympathetic and reflective moral emotions which are essential in a deliberation. This poses a problem for online deliberation because people will not be able to express their full range of emotions.

7.5 Designing Online Environments for more Emotionally Fine-Grained Expression

In this section we propose several design features meant to foster well-being through emotionally rich deliberation environments online. We think that current designs of social media platforms do not take the mediation of emotions explicitly into account, with possible detrimental effects for the user’s well-being. The design features proposed below are intended as a starting point for a wider debate concerning the emotional environments of social media. Should the emotional environments be designed with a particular emotional state in mind (e.g. related to well-being)? To what extent are the emotional reactions of users on social media the effect of the misuse of these media for other purposes than they were intended? These questions cannot be expanded on here, but deserve further elaboration. Furthermore, this debate on the limits of designing social media affordances for emotional well-being cannot be answered by philosophers alone, but rather deserve an interdisciplinary approach.

Concerning online deliberation – be it on dedicated platforms or on social media groups – we think of several design features that could foster a more emotionally

rich environment. We propose that the first stage of an online debate should be dedicated to choosing what to debate: for example, concerning a policy proposal, users could vote what they find the most interesting policies. The choice option is already implemented in the existing deliberation platforms, but it is done only via text bits which get up-voted. However, we propose to complement the text-based explanations of policies with video clips or audio clips in which proponents explain their policy proposals. This would give a more human touch to the debate, by making known the faces of the proponents and their own emotions attached to these proposals. A danger – as always in video based content – is that some proposals might get voted because their speakers are charismatic, and not for the content itself. This could be partially by-passed by having the users first read the policy proposal in a text snippet, and then asking them to click on a video recording of the expanded proposal.

In a next stage, users could have the opportunity to comment on these proposals and explain why they support a certain policy. In this phase, users could be encouraged to also post video or audio clips with their comments of the policy. The audio clips would probably work better since these would foster anonymity, while also allowing for a personal touch – the voice of the user, their emotions being discernible from the voice recording.

In the final stage, when a policy has been proposed for a general vote, the final debate could take place in a video conference format with each side designating representatives to speak for it. Thus, users could watch the debate – either live or recorded – and then cast their vote. Again, the danger of charisma-based votes needs to be averted. One possibility would be to combine video with text transcripts. Thus, before casting their votes, the users would be asked to read the transcript of the debate. The reading would ask of the users to focus on the content of the debate, while the emotions expressed in the debate would still be in their minds.

For regular discussions among social media users, no formal constraints can be imposed on the users. After all, nobody can be hindered from starting a deliberation in a group or in a private chat. However, we propose that the text comments and messages to be screened for sensitive words. Once users type in a negative message containing certain trigger-words, an AI algorithm could detect it and it would pop up the question “Are you sure you want to send this?”. This feature is already being experimented with by Instagram³ but we think that it would also help to make users understand the emotional consequences of their actions, not just by asking them if they are sure, but for example, also showing them the image of a suffering face, or an emoticon.

Overall, we suggest moving away from the asynchronous text message conversation to a format more expressive of the interlocutor’s bodily presence: their tone of voice, face and movements could be recorded – either via video or at least audio. This can help interlocutors to remember that the other online user is a human being who may feel affected by their messages, and that everyone matters equally. If

³ See <https://www.bbc.com/news/technology-48916828>

people cannot hide behind the anonymity of a nickname but see each other as real human beings with feelings to be acknowledged, then we hope that such online debates could become more meaningful for all participants, possibly leading to a form of digital civic well-being.

On a final note, we are aware that video alone cannot solve all the problems of text-based communication without adding its own problems. The raising popularity of vloggers on YouTube has not always led to more emotionally aware users, rather new phenomena such as self-radicalization after watching videos on YouTube (Alfano et al. 2018) became possible. In this paper we wanted to draw attention to the hypothesis that text-based communication undercuts emotional expression in a very particular way which imposes its own media-logic on existing emotions. The solution proposed here would not be to entirely give up text-based messages, nor to replace text with another medium,⁴ but to look at the possible convergence of multiple media: text and video, text and sound clips, text and emoticons, text and images, etc.

7.6 Conclusion

In this paper we have suggested that emotions are mediated in an incomplete way in online social media by the heavy reliance on textual messages which fosters a rationalistic bias and a bias towards less nuanced emotional expressions. This incompleteness can happen either by obscuring emotions, showing less than the original intensity, misinterpreting the emotion, or eliciting emotions without feedback. Online interactions and deliberations tend to contribute rather than overcome stalemates and informational bubbles, partially due to prevalence of anti-social emotions. It is tempting to see emotions as being the cause of the problem of online verbal aggression and bullying. However, we argue that social media are actually designed in a too rationalistic way, because of the reliance on text-based communication, thereby filtering out social emotions and leaving space for easily expressed antisocial emotions. Based on research on emotions that sees these as key ingredients to moral interaction and deliberation as well as on research on text-based versus non-verbal communication, we propose a richer understanding of emotions, requiring different designs of online deliberation platforms. We propose that such designs should move from text-centred designs and should find ways to incorporate the complete expression of the full range of human emotions so that these can play a constructive role in online deliberations.

⁴There are currently other solutions being investigated by tech companies such as Apple – for example dynamic avatars – but we do not have the space to go into these here.

References

- Alfano, Mark, J. Adam Carter, and Marc Cheong. 2018. Technological Seduction and Self-Radicalization. *Journal of the American Philosophical Association* 4 (3): 298–322.
- Brady, William J., and Molly J. Crockett. 2019. How Effective Is Online Outrage? *Trends in cognitive sciences* 23 (2): 79–80. <https://doi.org/10.1016/j.tics.2018.11.004>.
- Burr, C., M. Taddeo, and L. Floridi. 2020. The Ethics of Digital Well-Being: A Thematic Review. *Science and Engineering Ethics*. <https://doi.org/10.1007/s11948-020-00175-8>.
- Damasio, Antonio R. 1994. *Descartes' Error: Emotion, Reason and the Human Brain*. New York: G.P. Putnam.
- Derks, Daantje, Agneta H. Fischer, and Arjan E.R. Bos. 2008. The Role of Emotion in Computer-Mediated Communication: A Review. *Computers in Human Behavior* 24 (3): 766–785. <https://doi.org/10.1016/j.chb.2007.04.004>.
- Han, Byung-Chul. 2017. *In the Swarm: Digital Prospects*; Translated by Erik Butler. Cambridge, MA: MIT Press.
- Illich, Ivan. 1993. *In the Vineyard of the Text: A Commentary to Hugh's Didascalicon*. Chicago: University of Chicago Press.
- Isin, Engin F., and Evelyn Sharon Ruppert. 2015. *Being Digital Citizens*. London/ Lanham: Rowman & Littlefield International.
- Kahneman, Daniel. 2011. *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kaliarnta, Sofia. 2016. Using Aristotle's Theory of Friendship to Classify Online Friendships: A Critical Counterview. *Ethics and Information Technology* 18: 65–79.
- Lazarus, Richard S. 1994. *Emotion and Adaptation*. New York/Oxford: Oxford University Press.
- Little, Margaret Olivia. 1995. Seeing and Caring: The Role of Affect in Feminist Moral Epistemology. *Hypatia* 10 (3): 117–137. <https://doi.org/10.1111/j.1527-2001.1995.tb00740.x>.
- Marin, Lavinia, Jan Masschelein, and Maarten Simons. 2018. Page, Text and Screen in the University: Revisiting the Illich Hypothesis. *Educational Philosophy and Theory* 50 (1): 49–60. <https://doi.org/10.1080/00131857.2017.1323624>.
- Mossberger, Karen, Caroline J. Tolbert, and Ramona S. McNeal. 2008. *Digital Citizenship: The Internet, Society, and Participation*. Cambridge, MA/London: MIT.
- Nadkarni, Ashwini, and Stefan G. Hofmann. 2012. Why Do People Use Facebook? *Personality and Individual Differences* 52 (3): 243–249. <https://doi.org/10.1016/j.paid.2011.11.007>.
- Nussbaum, Martha C. 2001. *Upheavals of Thought*. Cambridge: Cambridge University Press.
- Roberts, Robert Campbell. 2003. *Emotions: An Essay in Aid of Moral Psychology*. Cambridge/ New York: Cambridge University Press.
- Roeser, Sabine. 2006. The Role of Emotions in Judging the Moral Acceptability of Risks. *Safety Science* 44 (8): 689–700.
- . 2011. *Moral Emotions and Intuitions*. Basingstoke: Palgrave Macmillan.
- . 2018. *Risk, Technology, and Moral Emotions*. London: Routledge.
- Roeser, Sabine, and Cain Samuel Todd. 2014. *Emotion and Value*. 1st ed. Oxford: Oxford University Press.
- Scherer, Klaus R. 1984. On the Nature and Function of Emotion: A Component Process Approach. In *Approaches to Emotion*, ed. Klaus R. Scherer and Paul Ekman, 293–317. Hillsdale/London: Lawrence Erlbaum Associates.
- Slovic, Paul. 2010. *The Feeling of Risk: New Perspectives on Risk Perception*. London: Earthscan.
- Solomon, Robert C. 1993. *The Passions: Emotions and the Meaning of Life*. Indianapolis/ Cambridge: Hackett Publishing.
- Spring, Victoria L., C. Daryl Cameron, and Mina Cikara. 2018. The Upside of Outrage. *Trends in Cognitive Sciences* 22 (12): 1067–1069. <https://doi.org/10.1016/j.tics.2018.09.006>.
- Verdiesen, E. P., M. V. Dignum, M. J. van den Hoven, Martijn Cligge, Jan Timmermans, and Lennard Segers. 2016. *MOOD: Massive Open Online Deliberation Platform-A Practical Application*. In ECAI 2016: 22nd European Conference on Artificial Intelligence, 29 August-2

- September 2016, The Hague, The Netherlands, Including Prestigious Applications of Artificial Intelligence (PAIS 2016): Proceedings, ed. Gal A. Kaminka et al. Amsterdam: IOS Press.
- Wojcieszak, Magdalena E., Young Min Baek, Michael X. Delli, and Carpini. 2009. WHAT IS REALLY GOING on? Structure Underlying Face-to-Face and Online Deliberation. *Information, Communication & Society* 12 (7): 1080–1102. <https://doi.org/10.1080/13691180902725768>.
- Zagzebski, Linda. 2003. Emotion and Moral Judgment. *Philosophy and Phenomenological Research* 66 (1): 104–124. <https://doi.org/10.1111/j.1933-1592.2003.tb00245.x>.

Lavinia Marin is a Postdoctoral Researcher in the Ethics and Philosophy of Technology Section at TU Delft, the Netherlands.

Her current research investigates conditions of possibility for online critical thinking by looking into how thinking is mediated by technologies and shaped by embodied actions and emotions. In addition, she is involved in a research project at TU Delft focused on identifying best practices for ethics education in engineering.

l.marin@tudelft.nl

Sabine Roeser is Professor of Ethics at TU Delft (distinguished Antoni van Leeuwenhoek Professor) and Head of the Ethics and Philosophy of Technology Section. She is Integrity Officer of TU Delft. Roeser has been working at Delft University of Technology since September 2001.

Roeser's research covers theoretical, foundational topics concerning the nature of moral knowledge, intuitions, emotions and evaluative aspects of risk, but also urgent and hotly debated public issues on which her theoretical research can shed new light, such as nuclear energy, climate change, biotechnology and public health issues. Roeser frequently is member of governmental advisory boards. She has given numerous interviews and public lectures. Her most recent book is *Risk, Technology, and Moral Emotions* (2018, Routledge).

Roeser has led various research projects with highly competitive funding from the EU and the Dutch Research Council (NWO). She is Co-chair of the Ethics of Socially Disruptive Technologies project, a 27 million Euro, 10 year multi-university research programme.

Also see: <http://www.tbm.tudelft.nl/sroeser>

Research Interests: Moral Emotions and Intuitions, Metaethics, Risk Ethics, Energy Ethics, Climate Ethics, Digital Ethics, Public Health Ethics and Ethics of Medical Technologies.

S.Roeser@tudelft.nl