Philosophy of Science (2022), **89**, 1075–1084
doi:10.1017/psa.2022.44

PSA PHILOSOPHY OF SCIENCE ASSOCIATION
PHILOSOPHY OF SCIENCE

SYMPOSIA PAPER

Responsible Dissemination in Sexual Orientation Research: The Case of the AI "Gaydar"

Andreas De Block* and Stijn Conix

Centre for Logic and Philosophy of Science, Institute of Philosophy, KU Leuven, Leuven, Belgium
*Corresponding author. Email: andreas.deblock@kuleuven.be

(Received 13 October 2021; revised 16 February 2022; accepted 21 April 2022; first published online 30 May 2022)

Abstract

A recent controversy about neural networks allegedly capable of detecting a person's sexual orientation raises the question of whether all research on homosexuality should be permitted. This paper considers two arguments for limits to such research, and concludes that there are good reasons to limit at least the dissemination of applied research on the etiology of homosexuality. The paper then briefly sketches how this could work, and looks at three objections against these limitations.

1. Introduction

Prewar genetic research has been used in various countries to justify the sterilization and even killing of thousands of homosexuals (Plant 2011). And until quite recently, even in the West, different theories on homosexuality's etiology were instrumental in setting up programs for "curing" homosexuals (De Block and Adriaens 2013).

Partially inspired by this horrific history of the gay sciences, many have also criticized current research about the causes of homosexuality. Some critics denounced particular studies, for instance, studies focused on so-called conversion therapies, while others argued for a skepticism toward specific research domains, such as heritability studies, or approaches, such as animal models (Stein 1999; Fausto-Sterling 1995). These criticisms sometimes include allusions to the ulterior motives of the scientists involved. For example, David Hull (1998, 390) claims that homonegativity and heteronormativity are among the mainsprings of contemporary evolutionary theories about homosexuality.

Of course, these concerns need not be limited to the intentions of individual researchers. Rather, critics seem especially concerned "that the very motivation for seeking the 'origin' of homosexuality has its source within social frameworks that are pervasively homophobic" (Schüklenk et al. 1997, 9).

The aim of this paper is to investigate the claim that research into the etiology of homosexuality is morally objectionable and should be forbidden. To do this, section 2 sketches the argument that such research is problematic because it strengthens or generates harmful ideas and attitudes about homosexuality. We argue that this argument does not hold, and in section 3 consider the argument that research into the etiology of homosexuality is problematic because of its dangerous applications. This argument does not entail that we should forbid research on the origins of sexual orientation, yet it provides good reasons to put certain restrictions on at least applied research. Section 4 then argues that these restrictions should take the shape of procedures for responsible disclosure of research results. Section 5 gives a rough sketch of what such procedures should look like. In the last section, we discuss three objections to our proposal.

2. Kitcher's argument and homosexuality

A first argument for banning research on the causes of homosexuality is analogous to Kitcher's (2001) argument against research on race differences in cognitive ability. Kitcher's argument goes that one should not pursue a line of research if that line of research is likely to give rise to beliefs that strengthen or stabilize convictions with morally undesirable social consequences. According to Kitcher, this is likely to be the case for research on race differences in cognitive ability regardless of the findings it would yield. Evidence that there are no differences would not harm the underprivileged, but due to existing biases is unlikely to be taken seriously or cause substantial changes. Evidence that the underprivileged have lower cognitive abilities, on the other hand, would strengthen the attitude that the underprivileged group is naturally unsuited for cognitively demanding social roles. Finally, indecisive evidence is likely to be harmful too, because human cognitive biases are such that we tend to interpret indecisive evidence as evidence that supports our prejudices. If society is already biased to believe that there are race differences in cognitive ability, indecisive evidence is likely to be taken as evidence for this hypothesis.

Kitcher's argument is controversial (Talisse and Aikin 2007). However, even if we assume that it holds, it is unclear whether it is applicable to research on the etiology of sexual orientation. Surely there are interesting similarities between such research and what Kitcher considers to be impermissible research. First, homosexuals are still an underprivileged group, especially in countries that continue to criminalize or stigmatize their sexual orientation. Second, harmful beliefs about the inferiority of homosexuals are still legion, for instance that they make bad parents, or that they are attention-seeking narcissists (Margolin 2021). Third, etiological research on homosexuality is such that it will often yield indecisive and hard-to-interpret evidence. For instance, heritability estimates vary dramatically between studies (Långström et al. 2010), and it is notoriously difficult to causally interpret genome wide association studies (Tam et al. 2019).

However, it is not clear whether another crucial condition is met; namely, the condition that directly links the research findings (or their interpretation) with undesirable attitudes and beliefs. With regard to cognitive differences research, it is obvious that if a person thinks that cognitive differences between ethnic groups exist, and that such differences are hard to alter through policies, this will often lead to the

https://doi.org/10.1017/psa.2022.44 Published online by Cambridge University Press

strengthening of that person's harmful beliefs. However, the link between homonegativity and beliefs about the origin and nature of homosexuality is far less clear. Findings about the genetic or endocrine roots of homosexuality are regularly interpreted as indicating that homosexuals should not be held responsible for their sexual orientation. And indeed, Oldham and Kasser (1999) found that reading a vignette that focuses on the genetic and neuroendocrine causes of homosexuality, led to less homonegativity in students. Other studies, however, did not find this influence of etiological beliefs on homonegativity, and suggest that causality might go in the opposite direction (Bailey and Knobe 2021): People often seem to rationalize their pre-existing attitudes by endorsing certain statements about the nature and origin of homosexuality. Moreover, even if it would turn out that findings about the genetic or endocrine roots of homosexuality would be instrumental in combatting homonegativity, this reasoning could itself be branded as a subtle form of homonegativity. After all, it does seem to suggest that we would blame homosexuals if they had chosen their sexual orientation, and that homosexuals would certainly have preferred heterosexuality if they had been given a choice.

Therefore, it is far from obvious whether, from a social justice perspective, it would be better or worse for homosexuals if people believed that their sexual orientation is under strict biological control. In short, Kitcher's imperative not to pursue certain lines of research may work for cognitive differences research, but not for studies into the etiology of homosexuality.

3. Dangerous applications

Even if research into the etiology of homosexuality is not likely to produce beliefs that harm the underprivileged, it may still harm them in other ways. In particular, such research could form the basis of technology that could be used against the underprivileged. Take, for instance, research on how deep neural networks can be used to detect a person's sexual orientation. This technology is at the heart of a recent paper by psychologists Wang and Kosinski (2018), published in *Journal of Personality and Social Psychology*. Wang and Kosinski's research was inspired by the prenatal hormone theory of sexual orientation. Part of the evidence for this theory is that there is some correlation between sexual orientation and traits that are thought to be determined by intra-uterine hormone levels. For example, the exposure of androgens in utero is believed to have an influence on the so-called digit ratio, i.e., the length of the index finger divided by the length of the ring finger. This would then explain why, on average, male homosexuals tend to have a slightly higher digit ratio than male heterosexuals.

The prenatal hormone theory has also produced hypotheses that correlate sexual orientation with visible facial features. After all, if prenatal exposure to hormones influences both sexual orientation and facial morphology, it seems reasonable to expect some kind of correlation between the two. Building on this idea, Wang and Kosinski used deep neural networks, i.e., artificial neural networks able to learn to detect patterns often missed by human observers, to predict one's sexual orientation solely on the basis of one's facial features. Their database consisted of tens of thousands of pictures found on dating websites and on Facebook, and they established the sexual orientation of the individuals in these pictures by means of their profile or

https://doi.org/10.1017/psa.2022.44 Published online by Cambridge University Press

their having liked pages such as "I Love Being Gay" and "Gay and Fabulous." According to Wang and Kosinski, the neural network they developed was able to accurately predict one's sexual orientation based on facial features alone. Moreover, the accuracy of their "artificial gaydar" was much higher than the accuracy of human judges, who did barely better than chance.

Wang and Kosinski's research ran into fierce opposition. First, there were concerns about methodology and data interpretation. Critics objected to the authors' reliance on pictures from dating sites and from openly gay Facebook users because these pictures probably contain hints about the person's sexual orientation (Miller 2018). A replication study confirmed that hunch: the success of Wang and Kosinski's gaydar may have been not so much based on its interpretation of facial features as on its interpretation of other aspects of the pictures, such as colors and lighting (Leuner 2019). Other critics pointed out that the authors overstated both the robustness of the evidence for the prenatal hormone theory and the accuracy of their neural network (Gelman, Marrson, and Simpson 2018).

Secondly, and more importantly for our purposes, there were moral concerns. Some of these concerns related to privacy and consent. Are Facebook pictures public data, for example, to be used at will by scientists (Miller 2018)? The main moral objections, however, concern the potential abuse of Wang and Kosinski's neural network. Critics feared that homonegative individuals, organizations, and governments could use it to identify and trace homosexuals, with the express purpose of harming them. This criticism echoes earlier criticisms of biological research on homosexuality because it could lead to screening procedures that would allow prospective parents to select for heterosexual children (Schüklenk et al. 1997).

There is an asymmetry at work in these research projects similar to the one Kitcher points to in his argument: Depending on the outcome of the research it may be neutral or harmful to homosexuals, but it is unlikely to ever benefit them. Unlike Kitcher's argument discussed in the previous section, this seems to provide a good reason not to allow such research. However, this argument only targets applied research on the etiology of homosexuality, and leaves more basic research on the etiology of homosexuality untouched. There are three reasons for this.

First, it may be that imposed restrictions on basic research are counterproductive. When Kitcher argued for moral constraints on the freedom of scientific inquiry, he explicitly warned against censoring scientific research by governments, funding agencies, or academic journals. Such censorship, he says, will often backfire: "In a world where (for example) research into race differences in IQ is banned, the residues of the belief in the inferiority of the members of certain races are reinforced by the idea that official ideology has stepped in to conceal an uncomfortable truth" (Kitcher 2001, 105).

If research on the etiology of homosexuality is forbidden, chances are that anti-gay voices will spin this as an attempt to hide "inconvenient truths" about homosexuality.

Second, the ship sailed on this a long time ago. There is already such an enormous amount of published research on the etiology of homosexuality that a ban would be pointless. And if we were to restrict all new research on this topic, we would make it much more difficult to criticize the extant body of pretty-robust-but-still-fallible knowledge on the etiology of sexual orientation. The current state of this field would then be the final state. Of course, this may make it more difficult to design well-

https://doi.org/10.1017/psa.2022.44 Published online by Cambridge University Press

functioning applications than would be the case if we had a better theoretical grip on the origins of homosexuality. But applications that are functionally suboptimal need not be less harmful than applications that function better because they are based on better basic research.

Third, the link between basic research and the misuse of technology resulting from such research is obviously more indirect than between the development of this technology and its misuse. This distance is relevant because it entails that the danger of misuse is not so imminent for basic research. Furthermore, the greater the distance is between research and its potential misuse, the more potentially useful research lines we'd also have to limit. There are usually many more research lines that have an indirect link with a misused technology than there are research lines directly linked to such misuse. For instance, if it is decided to forbid research on the etiology of homosexuality because the results can be used for developing "gaydar" technology, one could argue that any research on the intra-uterine effects of androgens (or of all hormones) should be banned as well.

Thus, the applications-argument for limiting research into the etiology of homosexuality only works for applied research. Still, it establishes a substantial and particularly high profile-group of research projects that should be subject to limitations. Three years after its publication, Wang and Kosinksi's paper has already been cited more than five hundred times. Vox, Quartz, the *New York Times*, and the *Financial Times* quickly mediatized the findings, and the *Guardian* published a long interview with Kosinski on the study. Because the topic is salient and controversial—and thus likely to be highly cited and published in high-profile journals—scientists are incentivized to engage in such research. Therefore it is important to know which research should be limited, and what these limitations should be. We turn to these questions in the next sections.

4. Limitations to research or limitations to dissemination?

Nonconsensual medical experimentation on a (vulnerable) population, as happened in the infamous Tuskegee syphilis study, is widely agreed to be unethical, regardless of its actual or expected results. Moreover, the scientific community is well aware that applied research with potentially harmful uses should sometimes be subject to limitations. Indeed, most funding applications require that researchers flag the dual-use potential of their research. However, because of the asymmetry discussed above, applied research on the etiology of homosexuality is not a typical dual-use case: While the research may be harmful, it is unlikely to have any direct benefits. The main reason to permit such research, then, is to enable us to mitigate the potential harms that could ensue if someone with bad intentions does the same research. Indeed, Wang and Kosinski claim that this was the main motivation for their study: "[s]ome people may wonder if such findings should be made public lest they inspire the very application that we are warning against. We share this concern. However, as the governments and companies seem to be already deploying face-based classifiers aimed at detecting intimate traits [. . .], there is an urgent need for making policymakers, the general public, and gay communities aware of the risks that they might be facing already. Delaying or abandoning the publication of these findings could deprive

https://doi.org/10.1017/psa.2022.44 Published online by Cambridge University Press

individuals of the chance to take preventive measures and policymakers the ability to introduce legislation to protect people" (Wang and Kosinski 2018, 255).

Research like Wang and Kosinski's should be permitted if the potential benefits of taking preventive measures and developing legislations outweigh the potential harms of providing ill-intentioned people and governments with these technologies. However, this harm/benefits calculus is particularly hard to make in these asymmetrical dual-use cases, as it is often impossible to predict how soon ill-intentioned people will come up with a technology that doesn't yet exist, how much we speed up that process by doing the research for them, and how much we gain from taking preventive measures. Unlike in traditional dual-use cases, where there are often direct benefits to the research that are reasonably easy to estimate, the benefits of asymmetrical cases like Wang and Kosinski's are nearly impossible to judge.

Fortunately, such a harms/benefits calculus is not necessary here. This is because the potential harm (ill-intentioned people using the technology) can be minimized while still enjoying the potential benefits (taking preventive measures). This is because instead of limiting *research* into the etiology of homosexuality, we can limit the *dissemination of this research*: if the research-outcomes are only communicated to the groups of people that are in a position to take preventive measures, then no harm ensues and future harm can be mitigated. We therefore propose that limitations to applied research on homosexuality should take the shape of *responsible disclosure procedures*.

Limitations on the dissemination of research findings are already common in various fields. In cyber security, vulnerabilities in security systems are studied and documented, but the researchers initially only share the results with the developer, the vendor, or the consumer, so that these stakeholders are given the opportunity to find countermeasures first. Once the safety issues have been sufficiently resolved or mitigated, vulnerabilities can be shared with a broader community of stakeholders to avoid their future reoccurrence (Householder et al. 2017). Similarly, restrictions on the dissemination of research are already broadly accepted in biotechnology and the biomedical sciences, and one of the main recommendations of the influential Fink report was to review research at the publication stage (NRC 2004). For example, there is a lot of research on the genomics of deadly viruses, but the viruses' genomes are not made publicly available because there is a real danger that terrorists would use this information to develop a very effective bioweapon (Kourany 2016).

Of course, individual social scientists that are aware of the risks of their research may already be careful about how they disseminate their work. For example, Wang and Kosinski did not publish the full source code and the learned data structures, and collaborated with stakeholders such as the American Civil Liberties Union in an early stage of their research. Still, a formal and institutionalized framework that oversees such responsible dissemination is lacking in the social sciences (and much of academia). Spelling out such a framework requires answering a range of difficult practical questions and is beyond the scope of this paper. Instead, the remainder of this paper considers two crucial questions that have to be answered for responsible dissemination: Who should regulate dissemination, and what limits to dissemination should we consider? We offer tentative answers to these questions, based on best practices in fields that already have mature procedures to deal with the responsible disclosure of research that poses potential risks (see also Brundage et al. 2018).

https://doi.org/10.1017/psa.2022.44 Published online by Cambridge University Press

4.1. Who should regulate dissemination of potentially harmful research?

Once it is agreed that some research should not be openly accessible upon completion, the question of who should be responsible for making decisions about limiting its dissemination arises. Options here include the individual researcher, the institution that hosted the research, journals, a dedicated independent authority, or a governmental organization. There is likely to be a tradeoff here between safety on the one hand and openness and freedom of research on the other hand: Researchers, journals, and institutions are likely to favor wide dissemination of research, while governmental organizations or a dedicated independent authority are more likely to value safety (Selgelid 2009).

In biotechnology, the Fink report recommended self-governance by scientists and scientific journals, and the various screening procedures that journals in this field have implemented suggest that this is a viable solution. It limits the practical and bureaucratic costs that would come with a dedicated authority or governmental organization, and protects the freedom of research. However, it is important to keep in mind that, due to the reward structure of academia, individual researchers and journals are strongly incentivized to publish controversial and spectacular research like Wang and Kosinski's. Moreover, individual researchers do not always have the expertise to evaluate whether it is safe to publish their research. For example, various LGBTQ-organizations disagreed with Wang and Kosiniski's judgement that it was safe to publish their study. Thus, self-regulation should probably not be left to individual researchers, but to dedicated commissions such as institutional review boards or editorial committees. To be clear, Stanford University's Institutional Review Board approved Wang and Kosinksi's research. However, such ethics committees currently evaluate research before it takes place, and focus primarily on the permissibility of the methods. Responsible dissemination, on the other hand, requires pre-publication review focused on how and to whom results can and should be communicated, what risks are involved in dissemination, and how they can be mitigated.

In addition to such review boards, instances such as Wang and Kosinski's paper could be avoided by changing the general awareness and safety-culture in the field. In biotechnology and cyber security, researchers are typically well aware of the risks of their research through training, formal safety procedures, review processes, and institutions responsible for responsible disclosure. Similar practices in other fields of academia could make researchers more aware of their responsibility, and more careful when they consider publishing their research.

4.2. How should dissemination be limited?

The publication process of Wang and Kosinski's paper illustrates that responsible disclosure is not simply a matter of deciding whether research should be published or not. There are a whole range of options between open access publication and full censorship, such as limiting the information, limiting the audience, or delaying publication until safety measures are in place. In computer security, there are detailed and clear procedures for deciding which dissemination options are best based on vulnerabilities (Householder et al. 2017). Similar procedures should be developed for the social sciences. We highlight two steps that are crucial in these procedures (for more on this, see Ovadya and Whittlestone 2019).

https://doi.org/10.1017/psa.2022.44 Published online by Cambridge University Press

First, it is necessary to evaluate the risk that the research under consideration entails. For some research, the prime risk might lie in making ill-intentioned parties aware of certain applications or research lines. This was probably the case for Wang and Kosinski's paper, as their methods were easy to use and their data publicly available. For other research, the risk lies in the product itself, or perhaps in the data that was collected for the research. This is particularly the case when the application is easy to use or adapt for undesirable purposes.

Second, the choice of a dissemination option should be attuned to the risk at hand. In particular, there are three variables that can be attuned: What precisely should be disseminated, when should it be disseminated, and who should be the audience? If the safety risk lies only in the product, code, or data, it may be sufficient to refrain from publishing all the materials required for making the product. In that case, immediate and open dissemination of a part of the materials could be the responsible choice. If the application is very unlikely to be on the radar of ill-intentioned parties, then even making them aware of its possibility or existence may be an unnecessary risk. If on top of that the application may be of immediate use and may cause substantial harm, like in the Wang and Kosinski case, then publication should be limited to certain audiences only until mitigating measures are in place. Once these are in place, publication of a part of the materials could be warranted.

5. Objections

These brief remarks obviously fall short of fully developed procedures for responsible disclosure that can deal with complex cases like Wang and Kosinski's research. However, we hope they are sufficient to illustrate the complexity of the decisions at hand, and the need to have a formal framework in place to deal with them. While Wang and Kosinski may not have made the best decision, at least they were aware of the problem and took some mitigating measures. It would be unwise to assume that all researchers in the future will be equally (but probably still insufficiently) careful.

We are aware that the limits to free publication of research that we propose here are not uncontroversial, and subject to various problems and objections. We finish this paper by briefly considering some of these.

First, one may argue that our proposal is at odds with the rapidly growing open science movement. And just like openly publishing research like an "AI gaydar" may entail risks, there are risks to limiting open access to research findings. For one, it could decrease inclusivity and lead to a concentration of power in certain research groups (Whittlestone and Ovadya 2020). This objection shows that the decisions at stake here are value-laden decisions, trading off openness, scientific progress, and potential harms to minorities. Rather than showing that there should be no limits to free publication, we think this objection strengthens our proposal to take this decision out of individual researchers' hands, and develop procedures that are supported and evaluated by a broader community.

Second, one may worry that review procedures are slow and, if some research gets censored or communicated to a limited audience, may sometimes be at odds with the professional ambition of researchers (Ovadya and Whittlestone 2019). Thus, there would be strong incentives for researchers to defect, and use different channels of

https://doi.org/10.1017/psa.2022.44 Published online by Cambridge University Press

publication to benefit from the attention that spectacular and controversial research is likely to get. While the risk of defectors is real, it may be possible to minimize it if procedures of responsible disclosure are implemented and supported across the relevant field of research. Even if researchers get their research out in ways that avoid the proposed pre-publication review, they could be punished by not citing such research or not accepting such research at conferences.

Finally, one could object that for almost all research in the social sciences, the risks are not immediately visible. Instead, research sometimes only entails risks in combination with other, potentially future research that in itself also does not entail risks either. Alternatively, it may be that research that is harmless now could have safety risks at some point in the future when it becomes more refined. This objection ties back into our remarks about fundamental research, which we think should not be subject to these measures of responsible disclosure. However, even if we assume that there should be no restrictions to fundamental research, difficult questions remain. For example, there is not always a clear difference between fundamental and applied research, and it remains important to be mindful of research with future potential of misuse: it may be much harder to mitigate risks if one waits too long. For example, Kourany (2016) argues that research on cognitive differences between genders or ethnic groups has done much harm in the past, and that the harms of the continuation of this research tradition are foreseeable and should trump concerns about the freedom of research. Others are less sure about how foreseeable scientific results and the associated harms are (Carrier 2021). However, the difficulty of these decisions is no reason not to try and tackle them. Instead, it highlights the need for strong and clear procedures to deal with them.

Acknowledgments. The authors would like to thank Jacob Stegenga for organizing the symposium on "the sciences of sexual desire," and the attendees of that symposium for their interesting comments and suggestions on this paper. Stijn Conix gratefully acknowledges funding from the Research Council – Flanders (grant 3H200026).

References

Bailey, April, and Joshua Knobe. 2021. "Biological Essentialism Correlates with (But Doesn't Cause?) Intergroup Bias." Preprint, submitted September 17, 2021. https://psyarxiv.com/rx8jc/.

Brundage, Miles, Shahar Avin, Jack Clark, Helen Toner, Peter Eckersley, Ben Garfinkel, Allan Dafoe, et al. 2018. "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation." ArXiv:1802.07228 [Cs]. Preprint, submitted February 20, 2018. http://arxiv.org/abs/1802.07228.

Carrier, M. 2021. "How to conceive of science for the benefit of society: prospects of responsible research and innovation." *Synthese* 198 (19):4749–68.

De Block, Andreas, and Pieter R. Adriaens. 2013. "Pathologizing sexual deviance: A history." *Journal of Sex Research* 50 (3-4):276–98.

Fausto-Sterling, Anne. 1995. "Animal Models for the Development of Human Sexuality: A Critical Evaluation." *Journal of Homosexuality* 28 (3-4):217–35.

Gelman, Andrew, Greggor Marrson, and Daniel Simpson. 2018. "Gaydar and the fallacy of objective measurement." Preprint, retrieved from http://www.stat.columbia.edu/~gelman/research/unpublished/gaydar2.pdf.

Householder, Allen D., Garrett Wassermann, Art Manion, and Chris King. 2017. *The cert guide to coordinated vulnerability disclosure*. Government Technical Report. Carnegie Mellon University.

Hull, David. 1998. "On Human Nature." In *The Philosophy of Biology*, edited by David Hull and Michael Ruse, 383–98. Oxford: Oxford University Press.

Kitcher, Philip. 2001. *Science, Truth, and Democracy*. Oxford: Oxford University Press.

https://doi.org/10.1017/psa.2022.44 Published online by Cambridge University Press

Kourany, J. A. 2016. "Should some knowledge be forbidden? The case of cognitive differences research." *Philosophy of Science* 83 (5):779–90.
Långström, Niklas, Qazi Rahman, Eva Carlström, and Paul Lichtenstein. 2010. "Genetic and environmental effects on same-sex sexual behavior: A population study of twins in Sweden." *Archives of Sexual Behavior* 39 (1):75–80.
Leuner, John. 2019. "A Replication Study: Machine Learning Models Are Capable of Predicting Sexual Orientation From Facial Images." Preprint, submitted February 27, 2019. arXiv:1902.10739.
Margolin, Leslie. 2021. "The Third Backdoor: How the DSM Casebooks Pathologized Homosexuality." *Journal of Homosexuality*. https://doi.org/10.1080/00918369.2021.1945340.
Miller, Arianne. 2018. "Searching for gaydar: Blind spots in the study of sexual orientation perception." *Psychology & Sexuality* 9 (3):188–203.
National Research Council (NRC). 2004. *Biotechnology research in an age of terrorism*. Washington, DC: National Academies.
Oldham, Jeff, and Tim Kasser. 1999. "Attitude change in response to information that male homosexuality has a biological basis." *Journal of Sex & Marital Therapy* 25 (2):121–24.
Ovadya, Aviv, and Jess Whittlestone. 2019. "Reducing Malicious Use of Synthetic Media Research: Considerations and Potential Release Practices for Machine Learning." ArXiv:1907.11274 [Cs], http://arxiv.org/abs/1907.11274.
Plant, Richard. 2011. *The Pink Triangle: The Nazi War Against Homosexuals*. New York: Holt paperbacks.
Selgelid, Michael. 2009. "Dual-Use Research Codes of Conduct: Lessons from the Life Sciences." *NanoEthics* 3 (3):175. https://doi.org/10.1007/s11569-009-0074-y.
Schüklenk, Udo, Edward Stein, Jacinta Kerin, and William Byne. 1997. "The ethics of genetic research on sexual orientation." *Hastings Center Report* 27 (4): 6–13.
Stein, Edward. 1999. *The Mismeasure of Desire*. Oxford: Oxford University Press.
Talisse, Robert, and Scott Aikin. 2007. "Kitcher on the ethics of inquiry." *Journal of Social Philosophy* 38 (4):654–65.
Tam, Vivian, Patel, Nikunj, Michelle Turcotte, Yohan Bossé, Guillaume Paré, and David Meyre. 2019. "Benefits and limitations of genome-wide association studies." *Nature Reviews Genetics* 20 (8):467–84.
Wang, Yilun, and Michal Kosinski. 2018. "Deep neural networks are more accurate than humans at detecting sexual orientation from facial images." *Journal of Personality and Social Psychology* 114 (2):246–57.
Whittlestone, Jess, and Aviv Ovadya. 2020. "The Tension between Openness and Prudence in AI Research." ArXiv:1910.01170 [Cs], http://arxiv.org/abs/1910.01170.

Cite this article: De Block, Andreas and Stijn Conix. 2022. "Responsible Dissemination in Sexual Orientation Research: The Case of the AI "Gaydar"." *Philosophy of Science* 89 (5):1075–1084. https://doi.org/10.1017/psa.2022.44