

## **Projeto de Pesquisa de Pós-doutorado:**

### **Melhoramento moral, status moral, complexidade e teoria das virtudes**

#### **1. Introdução**

A questão central que desejo investigar com minha pesquisa é a tensão entre a desejabilidade e os riscos do uso da tecnologia para melhorar traços humanos fundamentais – isto é, o melhoramento humano. Na minha dissertação de mestrado, discuti o uso da tecnologia para alterar traços humanos tais como inteligência, memória de longo prazo e preferências sociais. Em 2013, fui estudante visitante no Uehiro Centre for Practical Ethics, em Oxford, no início do meu mestrado, e minha pesquisa permaneceu intimamente conectada com essa instituição. Em 2014, iniciei meu doutorado na Universidade de Oxford sob a supervisão do Professor Julian Savulescu. Durante minha tese de doutorado, concentrei-me nos riscos de longo prazo do melhoramento das disposições morais – isto é, do melhoramento moral – e desenvolvi diretrizes para evitar esses riscos. Pretendo detalhar ainda mais essas diretrizes práticas e expandir minha pesquisa sobre riscos a longo prazo para outras formas de melhoramento humano.

#### **2. Pesquisas anteriores**

A minha tese de doutorado argumenta que tentar melhorar as disposições morais com a tecnologia<sup>1</sup> está sujeito a consequências inesperadas. Essas disposições são complexas e frágeis quando sujeitas a intervenções tecnológicas. Por exemplo, uma droga que aumenta a disposição de cooperar entre indivíduos pode levar inesperadamente a um conflito entre grupos porque pode aumentar o paroquialismo<sup>2</sup>. Além disso, o melhoramento moral pode alterar as relações psicológicas nas quais se baseiam nossa identidade pessoal. Portanto, pode minar

---

<sup>1</sup> Tal como inicialmente proposto em Persson & Savulescu (2012).

<sup>2</sup> Como sugerido em De Dreu (2012).

interesses individuais cuja satisfação requer a permanência de nossa identidade pessoal no tempo (e.g., o interesse que eu coma sorvete amanhã será satisfeito somente se minha identidade pessoal permanecer amanhã, pois este interesse não pode ser satisfeito por outra pessoa)<sup>3</sup>. De forma análoga, melhoramentos morais podem ser prejudiciais aos interesses que requerem a continuidade de gerações humanas<sup>4</sup>. Não obstante esses riscos, existem fortes razões em favor de melhorar as disposições morais. O melhoramento moral pode diminuir significativamente uma ampla gama de riscos extremos que surgem de nossas falhas morais (por exemplo, guerra nuclear por falta de cooperação internacional)<sup>5</sup>. Usar uma teoria das virtudes como guia para o melhoramento moral evita esses riscos e outros perigos expressos na literatura<sup>6</sup>. Se aplicarmos tal guia, as intervenções tecnológicas destinadas a melhorar as disposições morais são extremamente desejáveis. A fim de desenvolver ainda mais essa pesquisa, pretendo escrever publicações para periódicos de revisão por pares inspiradas pelo trabalho que fiz em minha tese de doutorado. Planejo ao menos três publicações advindas de minha tese, além de uma quarta que irá expandir uma pesquisa já publicada.

### **3. Status moral e interesses individuais**

Minha investigação acerca do impacto do melhoramento moral em interesses individuais será desenvolvida em um artigo de filosofia moral. Argumentarei que quando o melhoramento moral leva à elevação do status moral, um exame dos interesses individuais mostra que negligenciar efeitos sobre nossa identidade pessoal pode levar a consequências catastróficas. O status moral é a propriedade de um ente que faz seus interesses moralmente

---

<sup>3</sup> Tal visão de identidade pessoal baseada em relações psicológicas foi classicamente defendida por Lewis (1983) e também por Parfit (1986).

<sup>4</sup> A continuidade de gerações humanas como moralmente relevante em si mesma é uma posição defendida por Scheffler (2018), dentre outros.

<sup>5</sup> Um levantamento desses riscos é feito em Bostrom & Ćirković (2008).

<sup>6</sup> Tais como em Hughes (2015) e Snow (2010).

relevantes pelo bem do próprio ente<sup>7</sup>. Os interesses de entes de maior status moral importam mais que os interesses de entes com menor status moral. Pessoas possuem maior status moral do que animais não-humanos. O que confere status moral a um ente é possuir certas capacidades psicológicas tais como razão prática, empatia e cognição complexa. Hipoteticamente, podem existir entes com status moral maior que seres humanos, *supra-pessoas*. Irei argumentar que o simples aumento de nosso status moral por meio do melhoramento moral não implica que nossos interesses serão satisfeitos. Na realidade, esse aumento provavelmente prejudicará a realização de certos tipos de interesses individuais. Irei me concentrar em cenários nos quais o aumento do status moral produz mudanças significativas na identidade pessoal. Considerarei três cenários: a criação de supra-pessoas *de novo* dentro de uma geração de forma a coexistirem com as pessoas atuais, o melhoramento moral de pessoas para se tornarem supra-pessoas e a substituição geracional de pessoas por supra-pessoas. Mencionarei brevemente as preocupações que esses cenários geram na literatura e suas refutações por outros autores<sup>8</sup>. Argumentarei que uma consideração detalhada dos interesses individuais em jogo irá enfraquecer essas refutações. Portanto, o aumento do status moral requer uma atenção especial à preservação da identidade pessoal. Muitas formas de melhoramento humano radical provavelmente envolvem uma mudança no status moral e são frequentemente evocadas para lidar com os riscos extremos enfrentados pela humanidade neste século. No entanto, essa mudança também pode implicar riscos extremos. Qualquer forma de melhoramento humano radical precisa levar em conta seu impacto sobre a identidade pessoal para ser considerada segura. Recebi amplos comentários sobre este manuscrito de muitos especialistas na área, como os professores Roger Crisp, Jeff McMahan e Allen Buchanan, portanto acredito que se tornará uma publicação importante.

---

<sup>7</sup> Definição corrente usada em, dentre outros, DeGrazia (2008), Korsgaard (2018) e McMahan (2002).

<sup>8</sup> Tais como Buchanan (2009), Douglas (2013), McMahan (2010) e Powell (2013).

#### **4. Complexidade e fragilidade dos traços morais**

A investigação sobre a complexidade e a fragilidade dos traços morais será transformada em um artigo de pesquisa nos campos da psicologia moral e da ética aplicada. Argumentarei que (1) qualquer teoria adequada dos traços morais humanos, em um sentido descritivo, enfrenta grandes dificuldades epistêmicas, e (2) a modificação substancial dos traços morais é relativamente propensa a consequências inesperadas. Para suportar a primeira afirmação irei lançar mão da etiologia complexa dos traços morais, sua alta suscetibilidade a contingências e dificuldades epistêmicas decorrentes da visão em primeira pessoa da moralidade humana. Para apoiar a segunda afirmação, primeiro argumentarei que mesmo um exemplo aparentemente incontestado de melhoria moral, o aumento da cooperação humana, pode levar a efeitos prejudiciais inesperados. Em segundo lugar, vou generalizar o exemplo e argumentar que a intervenção tecnológica em traços morais individuais levará muitas vezes a efeitos paradoxais no nível do grupo. Em terceiro lugar, sustentarei que, na medida em que os traços morais estão ligados à motivação, tentar modificá-los leva à retroalimentação e à irreversibilidade.

#### **5. Teoria das Virtudes e Melhoramento Moral**

A terceira publicação derivada diretamente da minha pesquisa de doutorado será no campo da ética aplicada. Vou delinear algumas prescrições positivas para evitar os riscos que investiguei acima. Estes riscos requerem o uso de um guia de segurança para o projeto de melhoria das disposições morais. Tal guia deve garantir o seguinte: robustez à incerteza moral, adequação empírica, equilíbrio correto entre disposições, preservação da identidade pessoal e sensibilidade a considerações práticas (tais como efeitos paradoxais). Argumentarei que alguma forma de teoria da virtude cumpre essas prescrições e resolve outras preocupações presentes na literatura. Compararei tentativas anteriores de desenvolver uma teoria da virtude

como guia para o melhoramento moral e outras sugestões possíveis. Elaborarei uma nova teoria guia provisória e oferecerei algumas diretrizes práticas originais que resultam do uso de tal teoria. Por exemplo, o melhoramento da individualidade pode, na verdade, proporcionar melhores resultados para promover a cooperação em larga escala do que o próprio melhoramento da cooperação.

## **6. Simulando a difusão do melhoramento moral**

Trabalharei também escrevendo uma publicação que expande uma publicação anterior em coautoria com Anders Sandberg<sup>9</sup>. Desenvolvi simulações sociais que tentam modelar a disseminação do melhoramento moral na sociedade. A próxima publicação examinará os efeitos em nível de grupo dos melhoramentos morais, aprimorando nosso modelo computacional para incluir também interações entre grupos e reputação.

## **7. Direções futuras**

Após essa primeira etapa da pesquisa, que deve durar um ano, planejo expandir minha pesquisa em duas direções. Em primeiro lugar, estenderei minha teoria da virtude de modo a ser aplicável ao melhoramento humano em geral. Também irei desenvolver essa teoria mais detalhadamente a fim de oferecer uma agenda de pesquisa prática para os cientistas investigando métodos de melhoramento moral. Espero que tais resultados possam ser aplicados a uma ampla gama de casos, inclusive talvez oferecer uma solução alternativa para o problema do alinhamento de valores no campo da ética da inteligência artificial. Em segundo lugar, desenvolverei mais as considerações que fiz sobre o conflito entre interesses pessoais e interesses impessoais no caso do melhoramento do status moral. Trabalharei em uma análise filosófica mais profunda de como lidar com tais casos com o objetivo de fornecer conclusões

---

<sup>9</sup> A saber: Sandberg & Fabiano (2017).

gerais sobre como lidar com quaisquer casos conflitantes para além do melhoramento do status moral.

### 8. Cronograma de Pesquisa e Atividades Institucionais

Meses	Pesquisa	Atividade Institucional
1-3	Escrever e submeter publicação “Moral Status and Individual Interests”.	Evento acadêmico nacional e seminário em Ética Aplicada
4-6	Escrever publicação “Complexity and Fragility of Moral Traits” e “Virtue Ethics and Moral Enhancement”	Seminário em Ética Aplicada
6-9	Submeter “Complexity and Fragility of Moral Traits” e “Virtue Ethics and Moral Enhancement”. Escrever “Simulating the Diffusion of Moral Enhancement”.	Evento acadêmico internacional e seminário em Ética Aplicada
9-12	Submeter “Simulating the Diffusion of Moral Enhancement”. Planejar direções futuras.	Seminário em Ética Aplicada
12-24	Extensão da teoria das virtudes e aprofundamento da investigação sobre interesses pessoais e impessoais.	Seminário em Ética Aplicada

### 9. Referencias e Bibliografia Fundamental Planejada

- Agar, N. (2013). Why is it possible to enhance moral status and why doing so is wrong? *Journal of Medical Ethics*, 39(2), 67–74. <https://doi.org/10.1136/medethics-2012-100597>
- Baertschi, B. (2014). Neuromodulation in the service of moral enhancement. *Brain Topography*, 27(1), 63–71. <https://doi.org/10.1007/s10548-012-0273-7>
- Bostrom, N., & Ćirković, M. M. (2008). *Global Catastrophic Risks*. Oxford University Press.
- Buchanan, A. (2009). Moral Status and Human Enhancement. *Philosophy & Public Affairs*, 37(4), 3–35.
- Casebeer, W. D., & Churchland, P. S. (2003). The Neural Mechanisms of Moral Cognition: A Multiple-Aspect Approach to Moral Judgment and Decision-Making. *Biology & Philosophy*, 18(1), 169–194. <https://doi.org/10.1023/A:1023380907603>
- Dahlsgaard, K., Peterson, C., & Seligman, M. E. P. (2005). Shared virtue: The convergence of valued human strengths across culture and history. *Review of General Psychology*, 9(3), 203–213. <https://doi.org/10.1037/1089-2680.9.3.203>
- DeGrazia, D. (2008). Moral status as a matter of degree? *Southern Journal of Philosophy*, 46(2), 181–198. <https://doi.org/10.1111/j.2041-6962.2008.tb00075.x>
- DeGrazia, D. (2014). On the Moral Status of Infants and the Cognitively Disabled: A Reply to Jaworska and Tannenbaum. *Ethics*, 124(3), 543–556. <https://doi.org/10.1086/675077>

- Douglas, T. (2013). Human enhancement and supra-personal moral status. *Philosophical Studies*, 162(3), 473–497. <https://doi.org/10.1007/s11098-011-9778-2>
- Driver, J. (1998). The Virtues and Human Nature. In R. Crisp (Ed.), *How Should One Live?* <https://doi.org/10.1093/0198752342.003.0007>
- Goette, L., Huffman, D., & Meier, S. (2006). The impact of group membership on cooperation and norm enforcement. *American Economic Review*, 96(2), 212–216. <https://doi.org/10.1257/000282806777211658>
- Hughes, J. (2015). Moral Enhancement Requires Multiple Virtues. *Cambridge Quarterly of Healthcare Ethics*, 24(01), 86–95. <https://doi.org/10.1017/S0963180114000334>
- Hursthouse, R. (2001). *On Virtue Ethics*. <https://doi.org/10.1093/0199247994.001.0001>
- Korsgaard, C. M. (2018). Fellow Creatures: Our Obligations to the Other Animals. In *Uehiro Series in Practical Ethics*. <https://doi.org/10.1093/oso/9780198753858.001.0001>
- Lewis, D. (1983). Survival and Identity. In *Philosophical Papers Volume I*. <https://doi.org/10.1093/0195032047.003.0005>
- McMahan, J. (2010). Cognitive Disability and Cognitive Enhancement. *Cognitive Disability and Its Challenge to Moral Philosophy*, 345–367. <https://doi.org/10.1002/9781444322781.ch20>
- McMahan, J. (2002). *The Ethics of Killing*. <https://doi.org/10.1093/acprof:oso/9780199396078.001.0001>
- Parfit, D. (1984). *Reasons and Persons*. <https://doi.org/10.1093/019824908X.001.0001>
- Persson, I., & Savulescu, J. (2012). *Unfit for the Future*. <https://doi.org/10.1093/acprof:oso/9780199653645.001.0001>
- Persson, I., & Savulescu, J. (2013). Getting moral enhancement right: the desirability of moral bioenhancement. *Bioethics*, 27(3), 124–131. <https://doi.org/10.1111/j.1467-8519.2011.01907.x>
- Peterson, C., & Park, N. (2012). Classifying and Measuring Strengths of Character. In S. Lopez & C. R. Snyder (Eds.), *The Oxford Handbook of Positive Psychology*, (2 Ed.). <https://doi.org/10.1093/oxfordhb/9780195187243.013.0004>
- Powell, R. (2013). The biomedical enhancement of moral status. *Journal of Medical Ethics*, 39(2), 65–66. <https://doi.org/10.1136/medethics-2012-101312>
- Sandberg, A., & Fabiano, J. (2017). Modeling the Social Dynamics of Moral Enhancement: Social Strategies Sold Over the Counter and the Stability of Society. *Cambridge Quarterly of Healthcare Ethics*, 26(3), 431–445. <https://doi.org/10.1017/S0963180116001109>
- Savulescu, J. (2009). Moral Status of Enhanced Beings: What Do We Owe the Gods? In J. Savulescu & N. Bostrom (Eds.), *Human Enhancement*. Oxford University Press.
- Sherman, N. (2013). *Moral Psychology and Virtue*. <https://doi.org/10.1093/oxfordhb/9780199545971.013.0035>
- Shook, J. R. (2012). Neuroethics and the Possible Types of Moral Enhancement. *AJOB Neuroscience*, 3(4), 3–14. <https://doi.org/10.1080/21507740.2012.712602>
- Van Hoof, S. (2014). Virtue and identity. In S. Van Hoof (Ed.), *The Handbook of Virtue Ethics* (pp. 153–162). Retrieved from <https://www.cambridge.org/core/books/handbook-of-virtue-ethics/virtue-and-identity/A48F35337D52FD3A29919D50F5E388A6>
- Walker, L. J. (1982). The Sequentiality of Kohlberg's Stages of Moral Development. *Child Development*, 53(5), 1330. <https://doi.org/10.2307/1129023>
- Warren, M. A. (2000). The Concept of Moral Status. In *Moral Status*. <https://doi.org/10.1093/acprof:oso/9780198250401.003.0001>