# Travis LaCroix, PhD | CURRICULUM VITAE

CONTACT  $\bowtie$  Room 3180, Marion McCain Building  $\bowtie$  tlacroix@dal.ca Information Dept. of Philosophy, Dalhousie University 6135 University Avenue, PO Box 15000  $\bowtie$  +1 (902) 494-3628 Halifax Nova Scotia, B3H 4R2  $\bowtie$  +1 (902) 494-3518

CITIZENSHIP Canada

AREAS OF Philosophy and Ethics of AI & Emerging Technologies

Specialisation Social Dynamics, Norms, and Conventions

Autism

AREAS OF Philosophy of Language & Linguistics
COMPETENCE Classical & Evolutionary Game Theory
Traditional, Formal, & Social Epistemology

General Philosophy of Science

APPOINTMENTS Dalhousie University, Halifax, Nova Scotia CAN

Department of Philosophy, Faculty of Arts and Social Science

Faculty of Graduate Studies

Assistant Professor, Ethics and Computer Science, 2021 - 2024

University of Toronto, Toronto, Ontario CAN

Department of Philosophy

Schwartz Reisman Institute for Technology and Society

Postdoctoral Fellow, Social Cognition, 2020 - 2021

Supervisor: Jennifer Nagel

Université de Montréal, Montréal, Québec CAN

Mila - Québec Artificial Intelligence Institute

Department of Computer Science and Operations Research

Postdoctoral Fellow, Ethically Aligned Artificial Intelligence, 2020 - 2021

Supervisor: Yoshua Bengio (Mila / UdeM) Dominic Martin (UQàM)

Mila - Québec Artificial Intelligence Institute, Montréal, Québec CAN

Visiting Researcher / Research Assistant, 2018 – 2020

Supervisor: Yoshua Bengio

Concordia University, Montréal, Québec CAN

Coordinator, Social Justice Centre, 2019 – 2021

### **EDUCATION**

### University of California, Irvine, Irvine, California USA

PhD, Philosophy, March 2020

Department of Logic and Philosophy of Science

Dissertation: "Complex Signals: Modularity, Reflexivity, & Hierarchical Structure"

Committee: Jeffrey A. Barrett (Chair)

Brian Skyrms Simon Huttegger Cailin O'Connor

Master of Arts, Social Science, March 2018 Institute for Mathematical Behavioral Science

### Simon Fraser University, Burnaby, British Columbia CAN

Master of Arts, Philosophy, April 2016

Thesis Topic: "On Signalling Games and Their Models"

Committee: Nicolas Fillion (Chair)

Matt DeVos (External, Mathematics)

### University of British Columbia, Vancouver, British Columbia CAN

Bachelor of Arts (Hons.), Philosophy; English Literature, April 2014

First-class standing

### Camosun College, Victoria, British Columbia CAN

Associate of Arts, English, April 2011

### Canadian Institute For Advanced Research

CIFAR Deep Learning / Reinforcement Learning Summer School, Summer 2020

# EXTERNAL CONSULTING

### Bank of Canada, Ottawa, Ontario CAN

Expert Consultant on AI Ethics and Data Science (2020-2021)

### **PUBLICATIONS**

### Journal Articles (Peer Reviewed)

**Note**: First-listed date refers to online-first publication; Second-listed date refers to assignment of volume.

- 11. Arianna Falbo and Travis LaCroix. 2022. "Est-ce que vous compute? Code-switching, Cultural Identity, and AI." Feminist Philosophical Quarterly (Accepted).
- 9. Travis LaCroix. 2022. "Moral Dilemmas for Moral Machines." AI and Ethics (Forthcoming).

https://doi.org/10.1007/s43681-022-00134-y

- 8. Travis LaCroix and Cailin O'Connor. 2021. "Power by Association." *Ergo* (Forthcoming).
- 7. Travis LaCroix. 2021. "Reflexivity, Functional Reference, and Modularity: Alternative Targets for Language Origins." *Philosophy of Science* (2021) 88(5): 1234–1245. https://doi.org/10.1086/715217
- Travis LaCroix, Anders Geil, and Cailin O'Connor. 2020. "The Dynamics of Retraction in Epistemic Networks." *Philosophy of Science* (2021) 88(3): 415–438. doi.org/10.1086/712817
- Jeffrey A. Barrett and Travis LaCroix. 2020. "Epistemology and the Structure of Language." Erkenntnis (2022) 87(2): 953–967. doi.org/10.1007/s10670-020-00225-4
- Travis LaCroix. 2020. "Communicative Bottlenecks Lead to Maximal Information Transfer." Journal of Experimental and Theoretical Artificial Intelligence (2020) 32(6): 997–1014. doi.org/10.1080/0952813X.2020.1716857
- 3. Travis LaCroix. 2019. "Using Logic to Evolve More Logic: Composing Logical Operators via Self-Assembly." British Journal for the Philosophy of Science (2022) 73(2): 407–437. doi.org/10.1093/bjps/axz049
- 2. Travis LaCroix. 2019. "Evolutionary Explanations of Simple Communication: Signalling Games and Their Models." Journal for General Philosophy of Science / Zeitschrift für allgemeine Wissenschaftstheorie (2020) 51(1): 19–43. doi.org/10.1007/s10838-019-09481-7
- 1. Travis LaCroix. 2018. "On Salience and Signalling in Sender-Receiver Games: Partial Pooling, Learning, and Focal Points." Synthese (2020) 197(4): 1725–1747. doi.org/10.1007/s11229-018-1766-z

### Conference Proceedings (Peer Reviewed)

- 3. Michael Noukhovitch, Travis LaCroix, Angeliki Lazaridou and Aaron Courville. 2021. "Emergent Communication under Competition." In U. Endriss, A. Nowé. F. Dignum, and A. Lomuscio (eds.), *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021)*, London, UK, 3–7 May 2021, International Foundation for Autonomous Agents and Multiagent Systems. 974–982. dl.acm.org/doi/10.5555/3463952.3464066
- 2. Michael Noukhovitch, Travis LaCroix, Angeliki Lazaridou, and Aaron Courville. 2020. "Emergence of Communication with Selfish Agents. (Extended Abstract)" In Andrea Ravignani, Chiara Barbieri, Molly Flaherty, Yannick Jadoul, Ella Lattenkamp, Hannah Little, Mauricio Martins, Katie Mudd, Tessa Verhoef (eds.), The Evolution of Language: Proceedings of the 13th International Conference. 314-318. brussels.evolang.org/proceedings/evolang13 proceedings.pdf
- Travis LaCroix. 2019. "Biology and Compositionality: Empirical Considerations for Emergent-Communication Protocols." NeurIPS 2019 Workshop on Emergent Communication: Towards Natural Language. arxiv.org/abs/1911.11668

### Book Reviews (Peer Reviewed)

1. Travis LaCroix. 2022. 'Review of Ronald J. Planar and Kim Sterelny (2021) From Signal to Symbol: The Evolution of Language, Cambridge, MA: The MIT Press.' Philosophy of Science (Invited).

### Popular Media (Not Peer Reviewed)

- 2. Travis LaCroix. 2022. "A Category Mistake: Benchmarking Ethical Decisions for AI Systems Using Moral Dilemmas" *Blog of the APA*, 16 June 2022. blog.apaonline.org/2022/06/16/a-category-mistake-benchmarking-ethical-decisions-for-ai-systems-using-moral-dilemmas/
- 1. Travis LaCroix. 2022. "Research Summary: 'Moral Dilemmas for Moral Machines'" The AI Ethics Brief #103. Montréal AI Ethics Institute. 14 June 2022. montrealethics.ai/moral-dilemmas-for-moral-machines/

### ARTICLES

# Papers Under Review

(Drafts available upon request)

Revise & Resubmit Decision Received

Under Review

"Information and meaning in the evolution of compositional signals"

"The linguistic blind spot of value-aligned agency, natural and artificial"

"Metaethical perspectives on 'benchmarking' AI ethics" (w/ A. S. Luccioni)

### Conference Presentations

### Refereed Talks (Peer Reviewed)

- 19. "Information and meaning in the evolution of compositional signals" **Philosophy of Science Association**Pittsburgh, USA, 10–13 November 2022.
- 18. "Moral dilemmas for moral machines"

  American Philosophical Association, Pacific Division
  Vancouver, Canada, 13–16 April 2022.
- 17. "The linguistic blind spot of value-aligned agency, natural and artificial" Workshop on Normative Theory and Artificial Intelligence Online, 10 March 2022.
- 16. "Reflexivity, functional reference, & modularity: Targets for language origins" **Philosophy of Science Association**Baltimore, USA, 11–14 November 2021.
- "Est-ce que vous compute? Code-switching, cultural identity, & AI"
   (A. Falbo and T. LaCroix)
   Workshop on Feminism, Social Justice, and AI
   Online, 26-28 July 2021.

14. "If gradualism is the correct approach to language origins, then compositionality is not a plausible explanatory target"

### Canadian Philosophical Association

London, Canada, 1–4 June 2020. [Cancelled due to COVID-19]

13. "Learning from learning machines: Optimisation, rules, & social norms" (T. LaCroix and Y. Bengio)

### Canadian Society for the History & Philosophy of Science

London, Canada, 30 May - 1 June 2020. [Cancelled due to COVID-19]

12. "Emergent communication under conflict of interest" (M. Noukhovitch, T. LaCroix, and A. Courville)

EvoLang XIII: Evolution of Language International Conferences

Brussels, Belgium, 14–17 April 2020. [Cancelled due to COVID-19]

11. "Learning from learning machines: Optimisation, rules, & social norms" (T. LaCroix and Y. Bengio)

Artificial Intelligence and Moral Learning, Symposium at AISB-2020:

Society for the Study of AI and Simulation of Behaviour

London, UK, 6–9 April 2020. [Cancelled due to COVID-19]

10. "Learning from learning machines: Optimisation, rules, & social norms" (T. LaCroix and Y. Bengio)

### Canadian Society for Epistemology

Montréal, Québec, 14–16 November 2019.

9. "Accounting For role-asymmetries in the evolution of compositional signals"

### Canadian Philosophical Association

Vancouver, Canada, 1–4 June 2019.

8. "Using logic to evolve more logic: Composing logical operators via self-assembly" Society for Exact Philosophy

Toronto, Canada, 17–19 May 2019.

7. "Using logic to evolve more logic: Composing logical operators via self-assembly" American Philosophical Association, Pacific Division Vancouver, Canada, 17–20 April 2019.

6. "Less is more: Degrees of compositionality for complex signals"

### Philosophy of Science Association

Seattle, USA, 1–4 November 2018. Symposium Contribution.

5. "Reference by proxy and Truth-in-a-model"

### Western Canadian Philosophical Association

Calgary, Canada, 26–28 October 2018.

4. "On the role of power in the evolution of inequitable norms"

(T. LaCroix and C. O'Connor)

### L'Association Canadienne de Philosophie

Montréal, Québec, 4–7 June 2018.

3. "The role of power in the evolution of inequitable norms" (T. LaCroix and C. O'Connor)

# Latin American Association for Analytic Philosophy, and Colombian Conference on Logic, Epistemology, and Philosophy of Science Villa de Leyva, Colombia, 16–18 May 2018.

2. "On salience and signalling in sender-receiver games"

### Western Canadian Philosophical Association

Regina, Canada, 13-15 October 2017.

1. "Signalling games & their models"

Colombian Conference on Logic, Epistemology, & Philosophy of Science Bogotá, Colombia, 17–19 February 2016.

### Invited Talks (Not Peer Reviewed)

8. "On being in (autistic) community"

### Philosophy Colloquium, Dalhousie University

Halifax, Canada, 26 November 2022

7. "Moral dilemmas for moral machines: Perspectives on 'benchmarking' ethics for AI"

Philosophy Colloquium, Dalhousie University

Halifax, Canada, 07 October 2022 [Cancelled due to power outage]

6. "Workshop Panel on Embedded EthiCS in FCS"

### Philosophy Colloquium, Dalhousie University

Halifax, Canada, 24 June 2022

5. "The linguistic blind spot of value-aligned agency, natural and artificial"

# Philosophy Colloquium, Dalhousie University

Halifax, Canada, 26 November 2021

4. "Power by association"

(T. LaCroix and C. O'Connor)

# Durham University, Department of Philosophy

Online, 12 May 2021

3. "AI human compatibility and Stuart Russell's approach: What is the right proxy for artificial moral agency?"

(D. Martin and T. LaCroix)

### Journée d'étude sur l'éthique des algorithmes (CRÉ-OBVIA)

Montréal, Québec, 12 February 2021

2. "The tragedy of the AI commons"

(T. LaCroix and A. Mohseni)

# Weekly Seminar, Schwartz Reisman Institute for Technology & Society

Toronto, Canada, 9 December 2020

1. "Epistemology and the structure of language"

(J. A. Barrett and T. LaCroix)

Concordia Research Symposium in Philosophy (CRSP) Winter 2020 Montréal, Québec, 7 February 2020.

### Poster Presentations (Peer Reviewed)

 "Cheap-talk declarations and the tragedy of the AI commons" (T. LaCroix and A. Mohseni)

### Philosophy of Science Association

Virtual Poster Forum: January, February 2021

In-Person Poster Forum: Baltimore, USA, 11–14 November 2021.

 $2.\ {\rm ``Emergent\ communication\ under\ competition''}$ 

(M. Noukhovitch, T. LaCroix, A. Lazaridou, and A. Courville)

Autonomous Agents and Multiagent Systems (AAMAS)

London, UK, 3–7 May 2021.

 "Biology & compositionality: Empirical considerations for 'Eme-Comm' protocols" Neural Information Processing Systems (NeurIPS) 2019 Workshop on Emergent Communication

Vancouver, Canada, 8-14 December 2019.

### Invited Commentaries (Not Peer Reviewed)

- 3. "A pragmatic-semiotic defence of bivalent logic" (Marc Champagne) CPA/ACP 2019 Annual Congress, University of British Columbia, Vancouver, Canada, 1–4 June 2019.
- 2. "Responsibility for saying and asserting" (Henry Schiller) CPA/ACP 2018 Annual Congress, Université du Québec à Montréal, Montréal, Québec, 4–7 June 2018.
- 1. "Simplicity and a priori probability principles" (Noa Latham) CPA/ACP 2018 Annual Congress, Université du Québec à Montréal, Montréal, Québec, 4–7 June 2018.

### TEACHING EXPERIENCE

### Instructor of Record

 $Dalhousie\ University$ 

Social, Ethical and Professional Issues in Computing Science Winter 2024

Department of Philosophy (2490) (scheduled)

Faculty of Computer Science (3101)

Case Studies in Computing and Society Winter 2024
Faculty of Computer Science (1801) (scheduled)

Faculty of Arts & Social Science (1801)

Topic - Artificial Intelligence and the Value-Alignment Problem

Philosophy of Social Science Fall 2023
Department of Philosophy (3520 / 5520) (scheduled)

Department of Political Science (3496/5496)

Logic: Understanding of Scientific Reasoning Fall 2023
Department of Philosophy (2660) (scheduled)

Social, Ethical and Professional Issues in Computing Science

Winter 2023

Department of Philosophy (2490)

Faculty of Computer Science (3101)

Case Studies in Computing and Society

Winter 2023

Faculty of Computer Science (1801) Faculty of Arts & Social Science (1801)

Topic - Artificial Intelligence and the Value-Alignment Problem

Topics in Philosophy of Psychology

Fall 2022

Department of Philosophy (4070 / 5070)

Topic - The Philosophy of Autism and Autistic Philosophy

Logic: Understanding of Scientific Reasoning

Fall 2022

Department of Philosophy (2660)

Social, Ethical and Professional Issues in Computing Science

Department of Philosophy (2490) Faculty of Computer Science (3101)

Case Studies in Computing and Society

Winter 2022

Winter 2022

Faculty of Computer Science (1801)

Faculty of Arts & Social Science (1801)

 $\boldsymbol{Topic}$  - Artificial Intelligence and the Value-Alignment Problem

Topics in Ethics I

Fall 2021

Department of Philosophy (4114 / 5115)

Topic - Evolutionary Ethics

University of Toronto

Seminar in Philosophy of Science

Summer 2021

Department of Philosophy (455H)

*Topic* - Artificial Intelligence and the Value-Alignment Problem

### TA Supervision

Dalhousie University

Case Studies in Computing and Society

Winter 2022

Faculty of Arts and Social Science (1801)

7 teaching assistants

Social, Ethical, and Professional Issues in Computing Science

Winter 2022

Department of Philosophy (2490)

Faculty of Computer Science (3101)

2 teaching assistants

### Teaching Assistant

University of California, Irvine

Inductive Logic, Simon Huttegger  Department of Logic and Philosophy of Science (31)  Department of Philosophy (31)	Spring 2018
Introduction to Linguistics, Arunima Choudhury Department of Linguistics (3)	Winter 2018
Acquisition of Language, Lisa Pearl Department of Linguistics (51) Department of Cognitive Sciences (56)	Fall 2017
Inductive Logic, Simon Huttegger  Department of Logic and Philosophy of Science (31)  Department of Philosophy (31)	Spring 2017
Probability and Statistics for Economics I, Kent Johnson Department of Economics (15A)	Winter 2017
Simon Fraser University	
Critical Thinking, Jillian McIntosh Department of Philosophy (xx1)	Spring 2016
Critical Thinking, Jillian McIntosh Department of Philosophy (xx1)	Fall 2015
Critical Thinking, Jillian McIntosh Department of Philosophy (xx1)	Spring 2015
Introduction to Ethics, Evan Tiffany Department of Philosophy (120W)	Fall 2014
uest Lectures	

### **Guest Lectures**

University of California, Irvine

"Language and Cognition"  $Acquisition\ of\ Language\ (Linguistics/Psychology),\ 1$  Dec. 2017.

### RESEARCH ACTIVITIES

# Research Assistant

University of California, Irvine

Social Dynamics and Diversity in Epistemic Communities

Cailin O'Connor (NSF Grant 1535139) 2017 – 2019

Simon Fraser University

Constructing Questions for Critical Thinking

Jillian McIntosh Spring 2015

University of British Columbia

Aristotle's Earlier Logic

John Woods Summer 2014

GRANTS AND FELLOWSHIPS

### Research Awards

Fonds de Recherche du Québec – Société et Culture (FRQSC) **Bourse postdoctorale (B3Z)** [Postdoctoral fellowship] University of Calgary (CAD 110,000), DECLINED, 2020-2022

Social Sciences and Humanities Research Council (SSHRC) **Joseph-Armand Bombardier Canada Graduate Scholarships** Simon Fraser University (CAD 17,500), 2015-2016

Social Sciences and Humanities Research Council (SSHRC)

Joseph-Armand Bombardier Canada Graduate Scholarships
University of British Columbia (CAD 17,500), DECLINED, 2014-2015

### Academic Awards

University of California, Irvine

Associate Dean's Graduate Fellowship (USD 6,900), 2020, DECLINED.

Chancellor's Club for Excellence Fellowship (USD 12,000), 2019, NOMINATED.

Justine Lambert Graduate Prize in the Foundations of Science, for "On Salience and Signaling: Partial Pooling, Learning, and Focal Points" (USD 1000), 2018. Social Science Merit Fellowship (USD 257,818), 2016–2022.

Simon Fraser University

Graduate Fellowship (CAD 6250), 2015

University of British Columbia

John Alexander Scholarship in Humanities (CAD 1500), 2014 Robert and Kazuko Barker Award (CAD 225), 2014

Professional Service Journal Referee. British Journal for the Philosophy of Science (2022, 2021, 2020, 2019); Episteme (2022); Dialogue: Canadian Philosophical Review / Revue canadienne de philosophie (2021); Synthese (2021, 2020); American Philosophical Quarterly (2021); Proceedings of the Royal Society B: Biological Sciences (2021); Erkenntnis (2020); Journal of Experimental and Theoretical Artificial Intelligence (2020); Games (2020); International Studies in the Philosophy of Science (2020); Inquiry: An Interdisciplinary Journal of Philosophy (2019).

Book Referee. The MIT Press (2021).

Conference Referee. CoopAI2020: Cooperative AI - NeurIPS 2020 Workshop (2020); EvoLang XIII: Evolution of Language International Conferences (2020); The Vancouver Summer Philosophy Conference (2019).

Ethics Reviewer. Thirty-Sixth Conference on Neural Information Processing Systems (NeurIPS, 2022).

Programme Committee Member. Workshop on Gamification and Multiagent Solutions International Conference on Learning Representations (ICLR) (2022); AAAI / ACM Conference on Artificial Intelligence, Ethics, and Society - AIES (2022, 2021); ACM Conference on Fairness, Accountability, and Transparency - ACM FAccT (2022); CoopAI2021: NeurIPS 2021 Cooperative AI Workshop (2021).

Workshop Organisation. Scientific Advisor and Senior Programme Committee Member, Emergent Communication Workshop (EmeComm 2023); Scientific Advisor and Senior Programme Committee Member, Emergent Communication Workshop (EmeComm 2022), International Conference on Learning Representations (ICLR, 2022).

Symposium Organisation. "Evolutionary Explanations of Compositional Communication" Philosophy of Science Association 2018 Biennial Meeting, November, 2018.

Graduate Supervision. Graduate Fellows Cohort Lead - Schwartz Reisman Institute for Technology and Society (15 Graduate Students, 2020-2021); Peer advisor - Mila Québec AI Institute Peer Advisor Program (2021).

Undergraduate Supervision. Delia Topitzer, Dalhousie University (2022), honour's thesis advisor; Amy Tenenbaum, Dalhousie University (2022); Ariel Rozinov, University of Toronto (2021).

Colloquium Chair. Department of Philosophy, Dalhousie University (Fall 2022).

Conference Chair. Philosophy of Science Association, (10–13 November 2022, Pittsburgh, USA); The Vancouver Summer Philosophy Conference (19–23 August 2018, Vancouver, CAN); F-SEW: Formal Social Epistemology Workshop (25–26 May 2018, Irvine, USA); The Western Canadian Philosophical Association (13–15 October 2017, Regina, CAN); American Philosophical Association, Pacific Division (30 March – 3 April 2016, San Francisco, USA).

**Public Philosophy.** United Nations Educational, Scientific and Cultural Organisation (UNESCO) global consultation on the ethics of artificial intelligence (2020).

IN THE MEDIA

### Interviews

Interview with Ryan Bellefontaine. 2021. 'Should you delete Facebook? It's complicated' *The Signal*.

Engagement with Research

Article discussing LaCroix (2022): Jonny Thomson, 'Whose ethics should be programmed into the robots of tomorrow?' Big Think. 11 Apr 2022.

https://bigthink.com/thinking/robots-ethical-decisions/

White paper on Falbo and LaCroix (2022): 'Keeping with the status quo: Epistemic calcification and testimonial smothering in AI' Feminism, Social Justice, and AI Workshop.

Write up on LaCroix, Geil, O'Connor (2021): Rebecca Lea Morris, 'Why retracted papers continue to be cited and what we can do about it' *Medium* 22 Sep 2021. https://medium.com/@rlmorris/why-retracted-papers-continue-to-be-cited

SPECIALISED TRAINING	University of Toronto		
	Course Instructor Training	Summer 2021	
	Basic Occupational Health & Safety Awareness Training Accessibility for Ontarians with Disabilities Act Training	Summer 2021 Summer 2021	
	·	Summer 2021	
	Mila - Québec AI Institute	III: 1 0001	
	Equity, Diversity and Inclusion in Academic Research	Winter 2021	
	Concordia University		
	Sexual Violence Awareness & Prevention Training	Fall 2019, 2020	
	University of California, Irvine		
	Sex Offence Prevention Training	Fall 2017	
	Teaching Assistant Professional Development Program	Fall 2016	
	Simon Fraser University		
	Teaching Orientation Program	2014,2015,2016	
	TA/TM Day: Teaching Orientation Program	Fall 2015	
	TA/TM Day: Teaching Orientation Program TA/TM Day: Teaching Orientation Program	Spring 2015 Fall 2014	
		1811 2014	
	University of British Columbia		
	Student Directed Seminar, Training Workshop	Fall 2012	
Languages	English (native), French (intermediate)		
TECHNICAL	Mathematics: ODEs / PDEs, Linear Algebra, Multivariate / Vector Calculus		
Capacities	Logic: Set Theory, Modal Logic, Predicate Logic		
	Programming Languages: Python, Java, Javascript		
	Markup: LaTeX, HTML, CSS, Markdown		
	Computation and Graphing: R, MatLab, Excel		
Affiliations	Canadian Philosophical Association		
	Philosophy of Science Association Society for Exact Philosophy		
	Canadian Society for the History and Philosophy of Science		
	American Philosophical Association		

 $Available\ upon\ request.$ 

References