LINUS T. HUANG

Curriculum Vitae

Division of Humanities & Centre for AI Research
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong

LinusHuang@ust.hk
Philpeople.org/profiles/Linus-Huang
+1 (646) 685 4687

EDUCATION

Ph.D.   Unit for the History and Philosophy of Science, University of Sydney, 2017

M.A.   Department of Philosophy, University of California, Riverside, 2008

M.S.   Cognitive Neuroscience Program, National Yang Ming University, Taiwan, 2006

B.S.   School of Medicine, National Yang Ming University, Taiwan, 2004

PROFESSIONAL APPOINTMENT

2023-present   Research Assistant Professor, Division of Humanities & Centre for AI Research, Hong Kong University of Science and Technology

2024-2024   StarTrack Scholar, Social Computing Lab, Microsoft Research Asia

2020-2023   Postdoctoral Fellow, Society of Fellows in the Humanities, University of Hong Kong

2018-2020   Postdoctoral Fellow, Department of Philosophy, University of California, San Diego

2017-2018   Postdoctoral Scholar, Institute of European and American Studies, Academia Sinica, Taiwan

2015-2016   Australian Endeavour Scholar, School of Philosophy, Psychology and Language Sciences, University of Edinburgh & Department of Philosophy and Centre for Theoretical Neuroscience, University of Waterloo

RESEARCH SPECIALTIES

AOS: Philosophy of Artificial Intelligence, Philosophy of Cognitive Science, Philosophy of Mind

AOC: Philosophy of Science, Chinese Philosophy, Moral Psychology, Philosophy of Biology

GRANTS, AWARDS, AND FELLOWSHIPS

2024   StarTrack Scholars Fellowship, Microsoft Research Asia

2023   General Research Fund, Hong Kong SAR Research Grants Council

2021   Short-Term Research Fellowship, Virtual Clinic on Increasing Fairness in Targeted Advertising, Alexander von Humboldt Institute for Internet and Society

2018   UCSD-Taiwan Postdoctoral Fellowship, Academia Sinica

2018   Short-Term Research Collaboration Fellowships, Academia Sinica

2015   Australian Endeavour Award, Department of Education, Australia (69,000 AUD)

2015   Finalist, Postgraduate Presentation Prize, Australasian Association of Philosophy Annual Conference

2014   Royston George Booker and Herbert Johnson Travel Grant, University of Sydney

2014   University of Sydney Postgraduate Award, University of Sydney

2013   Postgraduate Research Support Scheme, University of Sydney

2012   Postgraduate Research Prize for Outstanding Academic Achievement, University of Sydney

2012   Summer Seminar Fellowship, Big Questions in Free Will Project, Florida State University

2012   Summer School Fellowship, Tubingen International Summer School: How do we make decisions? Perspectives from philosophy and science, University of Tubingen.

2011   University of Sydney International Scholarship, 2011-2014, University of Sydney

2007   Taiwan Merit Scholarship, 2007-2009, National Science Council, Taiwan (70,000 USD)

2006   Dean's Distinguished Fellowship Award, University of California, Riverside

PUBLICATIONS

Books

2022   Bechtel, W., & Huang, L. T. *Philosophy of Neuroscience*. Cambridge: Cambridge University Press.

Refereed Journal Articles

accepted   Sechman, M., & Huang, L. T.* Towards a Mechanistic Account of Embodied Implicit Bias: Heterarchical Control Network and Its Implications for Intervention. *Topoi: An International Review of Philosophy.* (* corresponding author, accepted with minor revisions).

accepted   Huang, L. T., Hung, T.-W., & Lin, Y.-T. The Dynamics of Personal Autonomy: A Comprehensive Framework for Evaluating Human-Carebot Interactions. *Asian Journal of Philosophy.* (Accepted with minor revisions).

2024   Huang, L. T., Papyshev, G, & Wong, J. Democratizing Value Alignment: From Authoritarian to Democratic AI Ethics. *AI and Ethics.*

2022   Huang, L. T., Chen, H.-Y., Lin, Y.-T., Huang, T.-R., Hung, T.-W. Ameliorating Algorithmic Bias, or Why Explainable AI Needs Feminist Philosophy. *Feminist Philosophy Quarterly.*

2021   Chen, H.-Y., Yu, L.-A., & Huang, L. T.* To Mask or Not to Mask: Epistemic Injustice in the COVID-19 Pandemics. *Techné: Research in Philosophy and Technology.* (* corresponding author)

2021   Huang, L. T. More Dynamical and More Symbiotic: Cortico-Striatal Models of Resolve,

Suppression, and Routine Habit. *Behavioral and Brain Sciences.*

2021   Huang, L. T., Bich, L, & Bechtel, W. Model Organisms for Studying Decision-Making: A Phylogenetically Expanded Perspective. *Philosophy of Science.*

2021   Huang, L. T. Can massive modularity explain human intelligence? Information control problem and implications for cognitive architecture. *Synthese.*

2020   Lin, Y.-T., Hung, T.-W., & Huang, L. T.* Engineering Equity: How AI Can Help Reduce the Harm of Implicit Bias. *Philosophy & Technology.* (* corresponding author)

2020   Carruthers, G., Carls-Diamante, S., Huang L. T., Rosen, M., & Schier, E. How to operationalise consciousness. *Australian Journal of Psychology*, 71(4), 390-410.

2018   Schwitzgebel, E., Huang, L. T., Higgins, A., & Gonzalez-Cabrera, I. The insularity of Anglophone philosophy. *Philosophical Papers,* 47(1), 21-48.

2012   Schwitzgebel, E., Rust, J., Huang, L. T., Moore, A. T., & Coates, J. Ethicists' courtesy at philosophy conferences. *Philosophical Psychology*, 25(3), 331–340.

Refereed Conference Proceedings

2020   Huang, L. T., & Bechtel, W. A Phylogenetic Perspective on Distributed Decision-Making Mechanisms. In *Proceedings of the 42th Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

2020   Bechtel, W., & Huang, L. T. Decentering Cognition. In *Proceedings of the 42th Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

2015   Huang, L. T. Reconceptualizing the computational problem of response selection. In *Proceedings of the 10th Conference of the Australasian Society of Cognitive Science.* Melbourne: Australasian Society of Cognitive Science.

2014   Huang, L. T. The nativist input problem. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Web-Based Writing

2021   Eren, E., Sekwenz, M-T., Huang, L. T., & Rzepka, S. Improving Fairness in Ad Display to Users [Blog Post]. In Birner, N., Hod, S., Kettemann, M. C., Pirang, A., & Stock, F. (Eds.), *Increasing Fairness in Targeted Advertising* [The Ethics of Digitalisation Clinics Report, 1]. Alexander von Humboldt Institute for Internet and Society. https://graphite.page/fair-targeted-ads/

Work In Progress

No Bias, No GAN: Generative Models Necessarily Produce Widespread, Unexpected, and Uninterpretable Biases (under review).

Hegemonic Machines (under preparation).

Neurodemocracy (under preparation).

PRESENTATIONS

Refereed Conference Oral Presentations

2024   Towards a Mechanistic Account of Embodied Implicit Bias, European Society for Philosophy and Psychology Annual Meeting, Grenoble, France, July 2-5.

2023   The Dynamics of Personal Autonomy: A Comprehensive Framework for Evaluating Human-Carebot Interactions, 10th Biennial Meeting of the Asia-Pacific Philosophy of Science Association, Hanoi, July 21-22

2021   Model Organisms for Studying Decision-Making: A Phylogenetically Expanded Perspective, Biennial Meeting of the Philosophy of Science Association, Baltimore, MD, November 11-14.

2021   The Promise and Perils of Explainable AI: Lessons from Feminism, Workshop on Feminism, Social Justice and AI, Waterloo, ON, July 26-28.

2019   Against the "Pyramid Scheme" of Executive Function: Empirical and Conceptual Challenges, Society for Philosophy and Psychology Annual Meeting, La Jolla, CA, July 10-13.

2018   Neurodemocracy: Self-Organization of the Embodied Mind, Australasian Society for Philosophy and Psychology 2018 meeting, Sydney, NSW, December 5-7

2018   Neurodemocracy: Self-Organization of the Embodied Mind, the 26th Biennial Meeting of the Philosophy of Science Association, Seattle, WA, November 1-4.

2014   Reconceptualizing the Computational Problem of Response Selection. Conference of the Australasian Society of Cognitive Science, Melbourne, VIC, December 8-10.

2014   The Nativist Input Problem. Annual Conference of the Cognitive Science Society, Québec, QC, July 23-26.

2013   The Nativist Input Problem. Society for Philosophy and Psychology Annual Meeting, Providence, RI, June 12-15.

2008   Moral Psychology and Moral Development in Xunzi. American Philosophical Association Central Division Meeting, Chicago, IL, April 16-19.

Selected Invited Talks

2025   Hegemonic Machines: How Generative AI Normalizes Oppression, Departmental Colloquium, Department of Cultural and Religious Studies, Chinese University of Hong Kong, January 9.

2024   No Bias, No GAN: Generative Models Produce Widespread, Unexpected, and Uninterpretable Biases, 2024 International Symposium on AI and Humanities Research, Taiwan, November 29.

2024   Dynamic Value Alignment: Enhancing User Autonomy Through Multi-agent, Moral Foundations Theoretical Framework, Microsoft Research Asia TAB Workshop: Shaping the Future with Societal AI, November 14.

2024   Hegemonic Machines: How Generative AI Normalizes Oppression, Departmental Colloquium, Department of Philosophy and Religion, Peking University, China, November 5.

2024   No Bias, No GAN: Generative Models Produce Widespread, Unexpected, and Uninterpretable Biases, International Workshop on The Philosophy & Ethics of AI, Taiwan, September 7.

2024   No Bias, No GAN: Generative Models Produce Widespread, Unexpected, and Uninterpretable Biases, International Workshop on The Philosophy & Ethics of AI, Taiwan, September 7.

2023   Why Explainable AI needs Feminist Philosophy, Department of Philosophy, University of Adelaid, August 29.

2024   No Bias, No GAN: Generative Models Produce Widespread, Unexpected, and Uninterpretable Biases, Philosophy of AI in Asia Conference, Hong Kong, March 21.

2023   Why Explainable AI needs Feminist Philosophy, Department of Philosophy, Chulalongkorn University, December 18.

2023   Why Explainable AI needs Feminist Philosophy, Department of Psychology, National Taiwan University, April 11.

2021   Why Explainable AI needs Feminist Philosophy, Hong Kong Institute for the Humanities and Social Sciences, University of Hong Kong, September 7.

TEACHING EXPERIENCE

Courses Taught as a Sole Instructor

Introduction to Ethics of AI and Data Science (Hong Kong University of Science and Technology, Fall 2023)

Philosophy and Cognitive Science (University of Hong Kong, Fall 2022)

Philosophy of Artificial Intelligence (University of Hong Kong, Fall 2021)

Courses Taught as a Teaching Assistant

Introduction to Critical Thinking (University of California, Riverside, Fall 2009, Winter 2010)

Reason, Belief, and Truth (University of California, Riverside, Spring 2008)

Contemporary Moral Issues (University of California, Riverside, Winter 2008)

Evil (University of California, Riverside, Fall 2007)

Teaching Certificate

Building AI Knowledge and Skills for Curriculum Development, City University of Hong Kong, 2023

Teaching Development Program, Faculty of Arts and Social Sciences, University of Sydney, 2014

COURSES PREPARED TO TEACH

Lower Level Courses

Logic

Critical Thinking

Introduction to Ethics of AI and Data Science

Philosophical Introduction to AI

Philosophy and Human Nature

Introduction to Chinese Philosophy

Upper Level Courses

Philosophy of Mind

Philosophy of Science

Philosophy of Cognitive Science

Philosophy of Neuroscience

Philosophy of Artificial Intelligence

Philosophy of Biology

Moral Psychology

Special Topics Courses

Implicit Bias

SERVICE

Service to Department

Co-Chair, Meaning and Reality in Social Context Conference, Academia Sinica, January 2019

Equity Officer and HPS Liaison, Philosophy Student Committee, University of Sydney 2014-5

Organizer, Minorities and Philosophy Initiative University of Sydney Chapter, 2014-5

Service to Profession

Steering Committee Member, ACM Fairness, Accountability, and Transparency Conference, 2023-Present

Advisor, Academic Placement Data and Analysis, 2020-Present

Peer Reviewer, MIT Press, 2024

Peer Reviewer, Routledge, 2024

Peer Reviewer, ACM Fairness, Accountability, and Transparency Conference, 2022

Peer Reviewer, Society for Philosophy and Psychology Annual Meeting, 2020

Symposium Organizer, Biennial Meeting of the Philosophy of Science Association, 2018

Graduate Mentor, Minorities and Philosophy Initiative, 2014-2015

Peer Reviewer, *Philosophical Psychology,* 2014

Peer Reviewer, *Journal of Chinese Philosophy,* 2010

Peer Reviewer, *Philosophy of Science,* 2009

LANGUAGES

Mandarin Chinese, native fluency in reading, writing, and speaking

Classical Chinese, intermediate fluency in reading

PROFESSIONAL MEMBERSHIPS

Association for Computing Machinery, 2023-Present

Society of Philosophy and Psychology, 2008-Present

Cognitive Science Society, 2013-Present

American Association of Philosophy Teachers, 2015-Present

Philosophy of Science Association, 2018-Present

REFERENCES

William Bechtel
UC Distinguished Professor
Department of Philosophy
University of California
San Diego, CA 92093, USA
Phone: +1 (858) 822-4461
Email: wbechtel@ucsd.edu

Eric Schwitzgebel
Professor, Philosophy
Department of Philosophy
University of California
Riverside, CA 92521, USA
Phone: +1 (951) 827-4288
Email: eric.schwitzgebel@ucr.edu

Xing Xie
Partner Research Manager
Social Computing Lab

Microsoft Research Asia
Beijing 100080. China
Phone: +86 010-59178888
Email: xing.xie@microsoft.com

Dominic Murphy
Professor, History and Philosophy of Science
Unit for the History and Philosophy of Science
University of Sydney
Sydney, NSW 2006, Australia
Phone: +61 2 9351 3762
Email: dominic.murphy@sydney.edu.au

Robert Rupert
Professor, Philosophy
Department of Philosophy
University of Colorado
Boulder, CO 80309, USA
Phone: +1 (303) 735-0988
Email: robert.rupert@colorado.edu

Bryce Huebner
Associate Professor, Philosophy
Department of Philosophy
Georgetown University
Washington DC, 20057, USA
Phone: +1 (202) 687-7487
Email: bryce.huebner@georgetown.edu

Peter J. Graham (Teaching Reference)
Associate Dean for Student Academic Affairs
College of Humanities, Arts, and Social Sciences
Professor, Philosophy
Department of Philosophy
University of California
Riverside, CA 92521, USA
Phone: +1 (323) 252-7063
Email: peter.graham@ucr.edu

Graduate Courses Taken

Philosophy	
Metaphysics and Epistemology	
Philosophy of Science	Philosophy of Psychology
Philosophy of Cognitive Science	Philosophy of Mind
Philosophy of Neuroscience	Consciousness
Theory of Mind	Perception
Philosophy of Language	Epistemology
Naming and Necessity	Logic
Philosophy of Biology*	Philosophy of Psychiatry*
Value Theory	
Free Will and Responsibility	Habermas
Caring and Practical Reason	Velleman
Multiculturalism	Moral Psychology*
Empirical Approach to Moral Psychology	
History of Philosophy	
Moral Philosophy of Hume and Kant	Aristotle's Ethics
20th-Century Analytic Philosophy	Asian Philosophy* (Offered by Comp. Lit. Dept.)
Asian Philosophy (Offered by Philosophy Dept.)	Classical Chinese* (Language Course)
Freud*	Existentialism*
Cognitive and Computational Neuroscience	
Cognitive Neuroscience	Theory of Computability
Intentionality and Perception	Consciousness and Perception
Computational Neuroscience*	Social Psychology*
* Audit	