

Consciousness and Free Will*

Fall 2016 — Mondays, 2:00–4:25pm

prerequisite: PSYCH 3140 or permission of instructor

Shimon Edelman[†]
Department of Psychology
Cornell University
Ithaca, NY, USA

<http://kybele.psych.cornell.edu/~edelman>

Version 2.1 — August 29, 2016

Abstract

This advanced seminar focuses on the most fundamental type of consciousness: phenomenal awareness, commonly referred to as feeling or sentience. We shall begin with some basic psychological findings and a key philosophical synthesis (Theme I below) and proceed to discuss the most plausible functional-neuroscience theory (Theme II); two computational theories (Themes III and IV); the main unresolved issues (Theme V); and free will (Theme VI).

Much of the material is complex and technical, requiring prior exposure to concepts from computational cognitive science and neuroscience. It is aimed at advanced undergraduate students, as well as graduate students from psychology, neurobiology, computer science, and other cognitive sciences.

The readings consist of chapters 9 and 10 from *Computing the Mind: How the Mind Really Works* (Oxford University Press, 2008), as well as 50 or so papers from the primary literature (listed below and available on the course Blackboard site as a zipped collection of PDFs).

Additional papers and changes to this syllabus may be occasionally posted on Blackboard as needed.

*PSYCH 4320; COGST 4310; LING 4310, BIONB 4330

[†]Email: edelman@cornell.edu.

Administrative details

Learning objectives

The students will be acquainted with the latest thinking and research concerning certain aspects of consciousness, with a special focus on neurobiological and cognitive-computational theories of phenomenal awareness and the freedom of will.

To get credit

- **Read** (20% of the course grade). For each week, there's a list of references (all the papers are also listed at the end of the syllabus, alphabetically by first author). These include primary readings (required) and other readings (optional, but highly recommended).

To be able to follow and participate in the discussion in class, read each week's materials *ahead of the meeting*. To get credit for reading, before each class submit (via Blackboard) a question about the weekly assignment (1 or 2 points each; at least 10 questions must be turned in).

- **Attend**. Attendance will be monitored; missing a class for whatever reason, including travel, will result in a 4-point penalty on the final grade (except when approved ahead of time by the instructor).
- **Participate** (20%). Contribute to the discussion, even if you are not the presenter. Following each class, submit a question or comment regarding some issue(s) that came up in the discussion (1 or 2 points each; at least 10 questions must be turned in).
- **Present** (40%). Prepare a presentation and lead the discussion at one of the meetings. Arrange to meet with the instructor ahead of the class to discuss your plans and preparation.
- **Write** (20%). Compose a final essay and submit it (in pdf, via email to the instructor) by noon on Monday, December 5. Detailed instructions will follow.

Code of integrity

The Cornell University code of academic integrity applies.

Students with disabilities

If you have a disability-related need for reasonable academic adjustments in this course, provide the instructor with an accommodation letter from Student Disability Services. Meeting with the instructor during office hours will help ensure confidentiality. Students are expected to give two weeks notice of the need for accommodations. If you need immediate accommodations, please arrange to meet with the instructor within the first two weeks of classes.

Preliminaries and overview — 8/29

- 8/29**
- computing the mind [14, ch.1-8]¹
 - key aspects of [vertebrate] phenomenality: selfhood, ownership, agency, perspectivalness, free will [24]; [14, ch.9]
 - an overview of the themes and the readings

Primary readings

- [14] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008 ch. 9
- [24] S. Gallagher. Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences*, 4:14–21, 2000

Other readings

- [14] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008 ch. 1-8

¹If you have taken the prerequisite, Psych 3140 Computational Psychology, you are all set with regard to this item. If you haven't, please contact the instructor before enrolling.

Theme I: psychological findings, philosophical insights — 9/12, 9/19

9/12 some findings regarding phenomenality

- selfhood dissolved [8, 40]
- ownership subverted [5]
- agency hijacked [9, 47]
- perspectivalness displaced [3, 4, 19]
- free will debunked [28, 29]

9/19 the minimal self and bare awareness

- a phenomenological blueprint for [human-type vertebrate-type] consciousness minimal self [14, ch.9]
- a précis of *Being No One* [37]
- subjective experience [36]

Primary readings

- [40] L. Sass, E. Pienkos, and B. Nelson. Introspection and schizophrenia: A comparative investigation of anomalous self experiences. *Consciousness and Cognition*, 22:853–867, 2013
- [8] R. L. Carhart, R. Leech, P. J. Hellyer, M. Shanahan, A. Feilding, E. Tagliazucchi, D. R. Chialvo, and D. Nutt. The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. *Frontiers in Human Neuroscience*, 8:20, 2014
- [5] M. Botvinick and J. Cohen. Rubber hands ‘feel’ touch that eyes see. *Nature*, 391:756, 1998
- [47] A. van der Weiden, H. Aarts, and K. I. Ruys. Prime and probability: Causal knowledge affects inferential and predictive effects on self-agency experiences. *Consciousness and Cognition*, 20:1865–1871, 2011
- [9] E. A. Caspar, A. Cleeremans, and P. Haggard. The relationship between human agency and embodiment. *Consciousness and Cognition*, 33:226–236, 2015
- [19] H. H. Ehrsson. The experimental induction of out-of-body experiences. *Science*, 317:1048, 2007
- [4] O. Blanke and T. Metzinger. Full-body illusions and minimal phenomenal selfhood. *Trends in Cognitive Sciences*, 13:7–13, 2009
- [3] O. Blanke. Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience*, 13:556–571, 2012
- [29] B. Libet. Do we have free will? *Journal of consciousness studies*, 6:47–57, 1999
- [14] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008 ch. 9
- [37] T. Metzinger. Précis: Being No One. *Psyche*, 11(5), 2005. URL <http://psyche.cs.monash.edu.au/symposia/metzinger/precis.pdf>

- [36] T. Metzinger. The subjectivity of subjective experience: A representationalist analysis of the first-person perspective. *Networks*, 3-4:33–64, 2004

Other readings

- [38] T. Metzinger. Self models. *Scholarpedia*, 2(10):4174, 2007. URL http://www.scholarpedia.org/article/Self_models
- [11] M. Chadha. No-Self and the phenomenology of agency. *Phenomenology and Cognitive Science*, 2016. doi: 10.1007/s11097-016-9455-1
- [30] J. Limanowski and F. Blankenburg. Minimal self-models and the free energy principle. *Frontiers in Human Neuroscience*, 7:547, 2013. doi: 10.3389/fnhum.2013.00547
- [25] S. Gallagher. Multiple aspects in the sense of agency. *New Ideas in Psychology*, 30:15–31, 2012
- [28] B. Libet. Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8:529–566, 1985
- [1] V. Ainley, L. Maister, J. Brokfeld, H. Farmer, and M. Tsakiris. More of myself: Manipulating interoceptive awareness by heightened attention to bodily and narrative aspects of the self. *Consciousness and Cognition*, 22:1231–1238, 2013

Theme II: what is it like to be a bat vertebrate? — 9/26, 10/3

9/26 You're a vertebrate. Now what? [31, 32, 33]

10/3 midbrain, thalamus, and the global best estimate [34, 35]; what is still missing [18]

Primary readings

- [31] M. A. MacIver. Neuroethology: From morphological computation to planning. In P. Robbins and M. Aydede, editors, *The Cambridge Handbook of Situated Cognition*, pages 480–504. Cambridge University Press, New York, NY, 2009
- [32] B. Merker. The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution. *Consciousness and Cognition*, 14:89–114, 2005
- [33] B. Merker. Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. *Behavioral and Brain Sciences*, 30:63–81, 2007
- [34] B. Merker. From probabilities to percepts: A subcortical ‘global best estimate buffer’ as locus of phenomenal experience. In S. Edelman, T. Fekete, and N. Zach, editors, *Being in Time: Dynamical Models of Phenomenal Experience*, pages 37–80. John Benjamins, 2012
- [35] B. Merker. The efference cascade, consciousness, and its self: naturalizing the first-person pivot of action control. *Frontiers in Psychology*, 4(501):1–20, 2013
- [18] S. Edelman, R. Moyal, and T. Fekete. To bee or not to bee? *Animal Sentience*, 2016. In press. A commentary on *Insects have the capacity for subjective experience*, C. Klein & A. B. Barron, *Animal Sentience* 2016:100

Other readings

- [15] S. Edelman. The minority report: some common assumptions to reconsider in the modeling of the brain and behavior. *Journal of Experimental and Theoretical Artificial Intelligence*, 27:1–26, 2015. doi: 10.1080/0952813X.2015.1042534
- [43] G. F. Striedter. *Principles of Brain Evolution*. Sinauer, Sunderland, MA, 2005, ch. 3

Theme III: the Information Integration Theory — 10/17, 10/24

10/17 an information integration manifesto [44]

10/24 the IIT, v.3 [39]

Primary readings

[44] G. Tononi. Consciousness as integrated information: a provisional manifesto. *Biol. Bull.*, 215:216–242, 2008

[39] M. Oizumi, L. Albantakis, and G. Tononi. From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0. *PLoS Computational Biology*, 10(5):e1003588, 2014. doi: 10.1371/journal.pcbi.1003588

Other readings

[2] D. Balduzzi and G. Tononi. Qualia: the geometry of integrated information. *PLoS Computational Biology*, 5(8):1–24, 2009

[27] C. Koch, M. Massimini, M. Boly, and G. Tononi. Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience*, 17:307–321, 2016

[45] G. Tononi, M. Boly, M. Massimini, and C. Koch. Integrated information theory: from consciousness to its physical substrate. *Nature Reviews Neuroscience*, 17:450–461, 2016

Theme IV: the Geometric Theory — 10/31, 11/7

10/31 consciousness explained? [21]

11/7 immortality hopes shattered [22]; being in time [16]

Primary readings

[21] T. Fekete and S. Edelman. Towards a computational theory of experience. *Consciousness and Cognition*, 20:807–827, 2011

[22] T. Fekete and S. Edelman. The (lack of) mental life of some machines. In S. Edelman, T. Fekete, and N. Zach, editors, *Being in Time: Dynamical Models of Phenomenal Experience*, pages 95–120. John Benjamins, 2012

[16] S. Edelman and T. Fekete. Being in time. In S. Edelman, T. Fekete, and N. Zach, editors, *Being in Time: Dynamical Models of Phenomenal Experience*, pages 81–94. John Benjamins, 2012

Other readings

[20] T. Fekete. Representational systems. *Minds and Machines*, 20:69–101, 2010. doi: 10.1007/s11023-009-9166-2

[10] M. Chadha. Time-series of ephemeral impressions: the Abhidharma-Buddhist view of conscious experience. *Phenomenology and Cognitive Science*, 2014. doi: 10.1007/s11097-014-9354-2

[13] D. C. Dennett. *Consciousness explained*. Little, Brown & Company, Boston, MA, 1991

Theme V: levels and boundary problems — 11/14, 11/21

11/14 “the” system? question, and a cop-out [17]; causal emergence [26]

11/21 system, subsystem, hive [23]

Primary readings

- [17] S. Edelman and T. Fekete. I am what I am. In *Proc. Annual Conference of the Association for Scientific Study of Consciousness (ASSC-17)*, San Diego, CA, July 2013. URL <http://kybele.psych.cornell.edu/~edelman/Archive/Edelman-Fekete-ASSC17-extended-abstract.pdf>. Extended abstract
- [26] E. P. Hoel, L. Albantakis, and G. Tononi. Quantifying causal emergence shows that macro can beat micro. *Proceedings of the National Academy of Science*, 110:19790–19795, 2013
- [23] T. Fekete, C. van Leeuwen, and S. Edelman. System, subsystem, hive: boundary problems in computational theories of consciousness. *Frontiers in Psychology*, 7(1041), 2016. doi: 10.3389/fpsyg.2016.01041. URL http://www.frontiersin.org/consciousness_research/10.3389/fpsyg.2016.01041/abstract

Other readings

- [41] E. Schwitzgebel. If materialism is true, the United States is probably conscious, 2012. URL <http://www.faculty.ucr.edu/~eschwitz/SchwitzPapers/USAconscious-140130a.htm>. Unpublished ms

Theme VI: free will — 11/28

11/28 basics [14, ch.10]; a two-stage biological model [6]; attributable agency [7]

Primary readings

- [14] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008, ch. 10
- [6] B. Brembs. Towards a scientific concept of free will as a biological trait: Spontaneous actions and decision-making in invertebrates. *Proceedings of the Royal Society London (B)*, 278:930–939, 2011. doi: 10.1098/rspb.2010.2325
- [7] H. J. Briegel and T. Müller. A chance for attributable agency. *Minds & Machines*, 25:261–279, 2015

Other readings

- [48] D. M. Wegner. The mind's best trick: how we experience conscious will. *Trends in Cognitive Sciences*, 7:65–69, 2003
- [46] M. Usher. Control, choice, and the convergence/divergence dynamics: a compatibilistic probabilistic theory of free will. *The Journal of Philosophy*, 103:188–213, 2006. doi: 10.2307/20619931
- [42] C. Sripada. Free will and the construction of options. *Philosophical Studies*, 2016. doi: 10.1007/s11098-016-0643-1
- [12] J. F. Cryan and T. G. Dinan. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nature Reviews Neuroscience*, 13:701–712, 2012

References

- [1] V. Ainley, L. Maister, J. Brokfeld, H. Farmer, and M. Tsakiris. More of myself: Manipulating interoceptive awareness by heightened attention to bodily and narrative aspects of the self. *Consciousness and Cognition*, 22:1231–1238, 2013.
- [2] D. Balduzzi and G. Tononi. Qualia: the geometry of integrated information. *PLoS Computational Biology*, 5(8):1–24, 2009.
- [3] O. Blanke. Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience*, 13:556–571, 2012.
- [4] O. Blanke and T. Metzinger. Full-body illusions and minimal phenomenal selfhood. *Trends in Cognitive Sciences*, 13:7–13, 2009.
- [5] M. Botvinick and J. Cohen. Rubber hands ‘feel’ touch that eyes see. *Nature*, 391:756, 1998.
- [6] B. Brembs. Towards a scientific concept of free will as a biological trait: Spontaneous actions and decision-making in invertebrates. *Proceedings of the Royal Society London (B)*, 278:930–939, 2011. doi: 10.1098/rspb.2010.2325.
- [7] H. J. Briegel and T. Müller. A chance for attributable agency. *Minds & Machines*, 25:261–279, 2015.
- [8] R. L. Carhart, R. Leech, P. J. Hellyer, M. Shanahan, A. Feilding, E. Tagliazucchi, D. R. Chialvo, and D. Nutt. The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. *Frontiers in Human Neuroscience*, 8:20, 2014.
- [9] E. A. Caspar, A. Cleeremans, and P. Haggard. The relationship between human agency and embodiment. *Consciousness and Cognition*, 33:226–236, 2015.
- [10] M. Chadha. Time-series of ephemeral impressions: the Abhidharma-Buddhist view of conscious experience. *Phenomenology and Cognitive Science*, 2014. doi: 10.1007/s11097-014-9354-2.
- [11] M. Chadha. No-Self and the phenomenology of agency. *Phenomenology and Cognitive Science*, 2016. doi: 10.1007/s11097-016-9455-1.
- [12] J. F. Cryan and T. G. Dinan. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nature Reviews Neuroscience*, 13:701–712, 2012.
- [13] D. C. Dennett. *Consciousness explained*. Little, Brown & Company, Boston, MA, 1991.
- [14] S. Edelman. *Computing the mind: how the mind really works*. Oxford University Press, New York, NY, 2008.
- [15] S. Edelman. The minority report: some common assumptions to reconsider in the modeling of the brain and behavior. *Journal of Experimental and Theoretical Artificial Intelligence*, 27:1–26, 2015. doi: 10.1080/0952813X.2015.1042534.
- [16] S. Edelman and T. Fekete. Being in time. In S. Edelman, T. Fekete, and N. Zach, editors, *Being in Time: Dynamical Models of Phenomenal Experience*, pages 81–94. John Benjamins, 2012.

- [17] S. Edelman and T. Fekete. I am what I am. In *Proc. Annual Conference of the Association for Scientific Study of Consciousness (ASSC-17)*, San Diego, CA, July 2013. URL <http://kybele.psych.cornell.edu/~edelman/Archive/Edelman-Fekete-ASSC17-extended-abstract.pdf>. Extended abstract.
- [18] S. Edelman, R. Moyal, and T. Fekete. To bee or not to bee? *Animal Sentience*, 2016. In press. A commentary on *Insects have the capacity for subjective experience*, C. Klein & A. B. Barron, *Animal Sentience* 2016:100.
- [19] H. H. Ehrsson. The experimental induction of out-of-body experiences. *Science*, 317:1048, 2007.
- [20] T. Fekete. Representational systems. *Minds and Machines*, 20:69–101, 2010. doi: 10.1007/s11023-009-9166-2.
- [21] T. Fekete and S. Edelman. Towards a computational theory of experience. *Consciousness and Cognition*, 20:807–827, 2011.
- [22] T. Fekete and S. Edelman. The (lack of) mental life of some machines. In S. Edelman, T. Fekete, and N. Zach, editors, *Being in Time: Dynamical Models of Phenomenal Experience*, pages 95–120. John Benjamins, 2012.
- [23] T. Fekete, C. van Leeuwen, and S. Edelman. System, subsystem, hive: boundary problems in computational theories of consciousness. *Frontiers in Psychology*, 7(1041), 2016. doi: 10.3389/fpsyg.2016.01041. URL http://www.frontiersin.org/consciousness_research/10.3389/fpsyg.2016.01041/abstract.
- [24] S. Gallagher. Philosophical conceptions of the self: implications for cognitive science. *Trends in Cognitive Sciences*, 4:14–21, 2000.
- [25] S. Gallagher. Multiple aspects in the sense of agency. *New Ideas in Psychology*, 30:15–31, 2012.
- [26] E. P. Hoel, L. Albantakis, and G. Tononi. Quantifying causal emergence shows that macro can beat micro. *Proceedings of the National Academy of Science*, 110:19790–19795, 2013.
- [27] C. Koch, M. Massimini, M. Boly, and G. Tononi. Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience*, 17:307–321, 2016.
- [28] B. Libet. Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8:529–566, 1985.
- [29] B. Libet. Do we have free will? *Journal of consciousness studies*, 6:47–57, 1999.
- [30] J. Limanowski and F. Blankenburg. Minimal self-models and the free energy principle. *Frontiers in Human Neuroscience*, 7:547, 2013. doi: 10.3389/fnhum.2013.00547.
- [31] M. A. MacIver. Neuroethology: From morphological computation to planning. In P. Robbins and M. Aydede, editors, *The Cambridge Handbook of Situated Cognition*, pages 480–504. Cambridge University Press, New York, NY, 2009.
- [32] B. Merker. The liabilities of mobility: A selection pressure for the transition to consciousness in animal evolution. *Consciousness and Cognition*, 14:89–114, 2005.

- [33] B. Merker. Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. *Behavioral and Brain Sciences*, 30:63–81, 2007.
- [34] B. Merker. From probabilities to percepts: A subcortical ‘global best estimate buffer’ as locus of phenomenal experience. In S. Edelman, T. Fekete, and N. Zach, editors, *Being in Time: Dynamical Models of Phenomenal Experience*, pages 37–80. John Benjamins, 2012.
- [35] B. Merker. The efference cascade, consciousness, and its self: naturalizing the first-person pivot of action control. *Frontiers in Psychology*, 4(501):1–20, 2013.
- [36] T. Metzinger. The subjectivity of subjective experience: A representationalist analysis of the first-person perspective. *Networks*, 3-4:33–64, 2004.
- [37] T. Metzinger. Précis: Being No One. *Psyche*, 11(5), 2005. URL <http://psyche.cs.monash.edu.au/symposia/metzinger/precis.pdf>.
- [38] T. Metzinger. Self models. *Scholarpedia*, 2(10):4174, 2007. URL http://www.scholarpedia.org/article/Self_models.
- [39] M. Oizumi, L. Albantakis, and G. Tononi. From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0. *PLoS Computational Biology*, 10(5):e1003588, 2014. doi: 10.1371/journal.pcbi.1003588.
- [40] L. Sass, E. Pienkos, and B. Nelson. Introspection and schizophrenia: A comparative investigation of anomalous self experiences. *Consciousness and Cognition*, 22:853–867, 2013.
- [41] E. Schwitzgebel. If materialism is true, the United States is probably conscious, 2012. URL <http://www.faculty.ucr.edu/~eschwitz/SchwitzPapers/USAconscious-140130a.htm>. Unpublished ms.
- [42] C. Sripada. Free will and the construction of options. *Philosophical Studies*, 2016. doi: 10.1007/s11098-016-0643-1.
- [43] G. F. Striedter. *Principles of Brain Evolution*. Sinauer, Sunderland, MA, 2005.
- [44] G. Tononi. Consciousness as integrated information: a provisional manifesto. *Biol. Bull.*, 215:216–242, 2008.
- [45] G. Tononi, M. Boly, M. Massimini, and C. Koch. Integrated information theory: from consciousness to its physical substrate. *Nature Reviews Neuroscience*, 17:450–461, 2016.
- [46] M. Usher. Control, choice, and the convergence/divergence dynamics: a compatibilistic probabilistic theory of free will. *The Journal of Philosophy*, 103:188–213, 2006. doi: 10.2307/20619931.
- [47] A. van der Weiden, H. Aarts, and K. I. Ruys. Prime and probability: Causal knowledge affects inferential and predictive effects on self-agency experiences. *Consciousness and Cognition*, 20:1865–1871, 2011.
- [48] D. M. Wegner. The mind’s best trick: how we experience conscious will. *Trends in Cognitive Sciences*, 7:65–69, 2003.