



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
**PROGRAMA DE POSGRADO EN**  
**FILOSOFÍA DE LA CIENCIA**



<b>Actividad Académica: Inteligencia Artificial, Subjetividad y Consciencia</b>				
<b>Clave:</b>	<b>Semestre:</b>	<b>Campo de conocimiento: FCC</b>		
<b>Carácter: Obligatoria ( ) Optativa ( X ) de Elección ( )</b>		<b>Horas por semana</b>		<b>Horas al semestre</b>
<b>Tipo:</b>		<b>Teóricas:</b>	<b>Prácticas:</b>	<b>No. Créditos:</b>
		3	1	64
<b>Modalidad: Presencial</b>		<b>Duración del programa: 1 semestre</b>		

**Seriación:** Si ( ) No ( x )      **Obligatoria** ( )      **Indicativa** ( )

**Inteligencia Artificial, Subjetividad y Consciencia**

**Objetivo general:**

El curso de Inteligencia Artificial, Subjetividad y Consciencia está diseñado para brindar a los estudiantes de posgrado una comprensión profunda de los conceptos y desafíos relacionados con la intersección entre la inteligencia artificial (IA), la subjetividad humana y la consciencia.

El objetivo principal ofrecer al alumno una aproximación rigurosa al debate actual sobre la naturaleza de la consciencia y el papel que esta juega en la cognición, en la agencia y en general en nuestras intuiciones morales. Para aproximarnos a estas cuestiones se analizará en detalle la relación entre la perspectiva de primera persona, información y representación mental. Se discutirá el papel que juega la consciencia en la atribución de capacidades cognitivas y el rol de la subjetividad y la perspectiva de primera persona a la hora de evaluar si los sistemas artificiales pueden ser sujeto y objeto de consideraciones morales.

**Objetivos específicos:**

Conocer el debate entorno al problema de la consciencia y las representaciones mentales, así como el papel que estas juegan en nuestras intuiciones morales y su extensión a sistemas artificiales.

Desarrollar la capacidad argumentativa en filosofía

Desarrollar la capacidad de entender y analizar los textos contemporáneos en filosofía analítica y más particularmente en el área de filosofía de la mente.

Desarrollar la capacidad de revisar la literatura empírica relevante.

Contenido Temático			
Unidad	Temas	Horas	
		Teóricas	Prácticas
1	El problema duro de la consciencia. Desafíos metafísicos en el estudio de la consciencia	8	4
2	Consciencia y Ciencia Cognitiva. Desafíos empíricos en el estudio de la consciencia.	6	3
3	Consciencia e Información: Teorías representacionales de la consciencia.	10	6
4	Consciencia, subjetividad y perspectiva de primera persona	10	6
5	El punto de vista de las máquinas: Inteligencia Artificial y moralidad	8	3
<b>Total de horas:</b>		<b>42</b>	<b>22</b>
<b>Suma total de horas:</b>		<b>64</b>	

#### Bibliografía y actividades:

Obligatoria:

- Block, N (2002): “Concepts of Consciousness”, en Chalmers 2002
- Block, "Perceptual consciousness overflows cognitive access", *Trends in Cognitive Sciences* 12 (2011), pp. 567-57
- Chalmers, D (2010): *The Character of Consciousness*. Oxford University Press. (Capítulo 5 y 6)
- Egan, Andy, "Appearance properties?", *Noûs* 40, 3 (2006), pp. 495-521.
- Jackson, F (1982): “Epiphenomenal Qualia”, *Philosophical Quarterly* 32 (reimpreso en Chalmers 2002)
- Kriegel, Uriah, *Subjective Consciousness: A Self-Representational Theory*. Oxford University Press, USA, 2009. (capítulos 1 y 4)
- Levy, N. (2014). *Consciousness and moral responsibility*. Oxford University Press.
- Lewis, D. (1979). Attitudes de dicto and de se. *Philosophical Review*, 88(4), 513–543.
- Millikan, R. G. (1989). Biosemantics. *Journal of Philosophy*, 86, 281–297.
- Neander, K. (1991). Functions as selected effects: The conceptual analyst’s defense. *Philosophy of*

Science, 58(2), 168–184.

- Nagel, T (1974): “What is it like to be a Bat?”, *Philosophical Review* 83, (reimpreso en Chalmers ed (2002): *Philosophy of Mind: Classical and Contemporary Readings*, OUP)
- Rosenthal, David (2002) "Explaining Consciousness", in *Philosophy of Mind: Classical and Contemporary Readings*, ed. David J. Chalmers, New York: Oxford University Press, 2002, pp. 406-421
- Rudy-Hiller, Fernando, "The Epistemic Condition for Moral Responsibility", *The Stanford Encyclopedia of Philosophy* (Winter 2022 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/win2022/entries/moral-responsibility-epistemic>
- Searle, John (1980), "Minds, Brains and Programs", *Behavioral and Brain Sciences* 3 (3): 417–457
- Sebastián, MA (2022) *El problema de la Consciencia: una Introducción Crítica a la Discusión Filosófica Actual*. Madrid, Ediciones Cátedra.
- Sebastián, MA (2022b) *First-Person Perspective in Experience: Perspectival De Se Representation as an Explanation of the Delimitation Problem*. *Erkenntnis*.
- Sebastián, MA (2021) *First-person representations and responsible agency in AI*. *Synthese*.
- Shoemaker, Sydney, "Phenomenal character", *Noûs* 28, 1 (1994), pp. 21-38.
- Siegel, Susanna (2010) *Do Visual Experiences Have Contents?* En ed. Bence Nanay *Perceiving the World*. Oxford University Press.
- Michel Tye (2002) *Consciousness, Color and Content*. MIT Press. (Capítulo 3. Representationalism: The Theory and its Motivations)

#### Complementaria:

- Baars, Bernard, "Global Workspace theory", in Tim Bayne, Axel Cleeremans and Patrick Wilken, ed., *The Oxford Companion to Consciousness* (Oxford University Press, 2009).
- Block, Ned, "Wittgenstein and Qualia", *Philosophical Perspectives* 21, 1 (2007), pp. 73-115.
- Block, Ned, "Mental paint", in Martin Hahn and B. Ramberg, ed., *Reflections and Replies: Essays on the Philosophy of Tyler Burge* (Mit Press, 2003), pp. 165-200.
- Carruthers, Peter, "Higher-Order Theories of Consciousness", *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2011/entries/consciousness-higher/>.
- Fischer, J., & Ravizza, M. (1998). *Responsibility and control: A theory of moral responsibility*. Cambridge.

bridge University Press.

- Fodor, J. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. MIT Press.
- Krishnan, A. (2009). *Killer robots: Legality and ethicality of autonomous weapons*. Routledge.
- García-Carpintero, M. (2017). The philosophical significance of the de se. *Inquiry*, 60(3), 253–276.
- Kripke, S (1980): *Naming and Necessity*, tercera conferencia, Harvard University Press.
- Kriegel, Uriah and Kenneth Williford, ed., *Self-Representational Approaches to Consciousness* (The MIT Press, 2006).
- Macpherson, Fiona, "Ambiguous figures and the content of experience", *Noûs* 40, 1 (2006), pp. 82-117.
- Müller, V. (2020). Ethics of artificial intelligence and robotics. In: Zalta, E. N. (Ed.) *The stanford encyclopedia of philosophy* (Fall 2020 Edition). <https://plato.stanford.edu/archives/fall2020/entries/ethics-ai/>.
- Neander, K. (2017). *A mark of the mental. In defense of informational teleosemantics*. MIT Press.
- Sebastián (2014a) "Dreams an Empirical way to settle the discussion between cognitive and non-cognitive theories of consciousness". *Synthese*, Vol. 191, Issue 2: 263-285
- Shoemaker, Sydney, "Introspection and Phenomenal Character", *Philosophical Topics* 28, 2 (2001), pp. 247-273.
- Shoemaker, "Phenomenal character", *Noûs* 28, 1 (1994), pp. 21-38.
- Weisberg, "Abusing the notion of what-it's-like-ness: A response to Block", *Analysis* 71 (2011), pp. 438-443.

**Nota:** (en caso que exista alguna)

<b>Medios didácticas:</b>	<b>Métodos de evaluación:</b>
Exposición profesor(a) ( X)	Exámenes o trabajos parciales ( X)
Exposición alumnos ( )	Examen o trabajo final escrito ( X )
Ejercicios dentro de clase ( )	Trabajos y tareas fuera del aula ( )
Ejercicios fuera del aula ( )	Exposición de alumnos ( )
Lecturas obligatorias (X)	Participación en clase(X)
Trabajo de investigación (X)	Asistencia (X)
Prácticas de campo ( )	Prácticas ( X)
Otros: ( )	Otros: ( )

### **Evaluación y forma de trabajo**

**Participación (10%):** Se espera la participación activa de los estudiantes en los diálogos que constituirán el curso. Los estudiantes serán evaluados por su participación.

**Examen (20%):** Habrá dos exámenes tipo test a lo largo del curso. El objetivo de estos exámenes es simplemente asegurar que las nociones clave quedan fijadas.

**Comentarios (25%):** Los estudiantes deberán enviar antes de cada sesión un brevísimo comentario de unas 300-600 palabras sobre las lecturas. Los comentarios consistirán exclusivamente en objeciones y dudas concretas para su debate en clase.

**Ensayo Final (60%):** Al final del semestre los alumnos deberán presentar un ensayo de unas 4.000 palabras discutiendo alguno de los temas tratados en clase. Antes del 20 de noviembre (7 noviembre recomendado) los alumnos deberán enviar una propuesta (15%) de máximo dos páginas resumiendo el ensayo final. Esta propuesta incluirá (1) una exposición clara del problema que se desea aclarar y posiblemente resolver, o la tesis que se va a defender; (2) una breve explicación de por qué el problema/tesis es interesante; (3) un esquema de cómo pretender llevar a cabo la tarea (resumiendo de forma anticipada el proceso argumentativos que se seguirá, etc.) (4) biografía que se pretende emplear (que no deberá ser muy extensa, uno o dos textos primarios y otros dos auxiliares o secundarios es suficiente).

Los criterios de evaluación del ensayo (45%) serán:

- a) Originalidad
- b) Rigor filosófico (fuerza argumentativa, consistencia lógica, precisión conceptual, rigor descriptivo, etc.)
- c) Organización, claridad estructural y estilística.

d) Evidencia de haber entendido el material discutido en clase.

**Imparte:** Dr. Miguel Ángel Sebastián González

**Mail:** msebastian@gmail.com

**Día y hora del curso o seminario (dos propuestas):** Lunes 16-20