

## EDITORIAL

**JIMENA YISEL CABALLERO CONTRERAS** | Virtualidad y ciencia

## ARTÍCULOS

**ALDANA MASTRANDREA** | Aportes teórico-conceptuales para el abordaje de la percepción y representaciones sociales del riesgo hídrico

**KARLA CASTILLO VILLAPUDUA** | De la posible renovación del pensamiento crítico ante la filosofía emancipadora de J. Rancière

**FREDERICO AGUILAR TAMAYO et al.** | Sacralidad y leyenda popular. Imagen, mito y relato en el personaje de la llorona

**ALBA MASSOLO; MARIEL TRAVERSI** | ¿Es posible mitigar sesgos cognitivos? Un análisis crítico de diversas propuestas para reducir el sesgo a mi favor

**MARIANO PETTARIN** | La Historia como ciencia Autónoma de los Asuntos Humanos: Collingwood y la experiencia como experiencia histórica

**CARLOS MAXIMILIANO SENCI** | Revisitando críticamente nociones de similitud en la teoría de juegos comportamental

**ENRIC TRILLAS; ALEJANDRO SOBRINO CERDEIRIÑA** | Alegato por el control del razonamiento

**LEOPOLDO TILLERÍA** | Una aproximación al problema del mal en Schelling

**MANAS SAHU** | The Kantian notion of freedom and Autonomy of Artificial Agency

## ENTREVISTAS

**CRISTINA PONTES BONFIGLIOLI** | Entrevista a Nicola Liberati: desafíos metodológicos para una nueva filosofía de la tecnología

## RESEÑAS

**BRENO BENEDYKT** | Um apartamento em Urano: crônicas da travessia

**JUAN CRUZ APCARIÁN** | El asedio a la imaginación

**Editor en jefe**

Emiliano Aldegani (Universidad Nacional de Mar del Plata, Argentina)

**Editores Adjuntos**

Cristina Pontes Bonfiglioli (Universidade de São Paulo, Brasil)

Flaminio de Oliveira Rangel (Universidade Federal de São Paulo, Brasil)

Ivy Judensnaider (Universidade Paulista, Universidade Estadual de Campinas, Brasil)

Jimena Yisel Caballero Contreras (Universidad Nacional Autónoma de México, México)

**Secretaria de redacción**

Jimena Yisel Caballero Contreras (Universidad Nacional Autónoma de México, México)

**Comité editorial**

Agustin Adúriz Bravo (Universidad de Buenos Aires, Argentina), Alberto Clemente De La Torre (Universidad Nacional de Mar del Plata, Argentina), Ana Paula Bispo (Universidade Estadual da Paraíba, Brasil), Arindam Bose (Tata Institute of Social Sciences (TISS), India), Charbel El-Hani (Universidade Federal da Bahia, Brasil), Fernando Santiago dos Santos (Instituto Federal de São Paulo, Brasil), Xavier Ruiz Collantes (Universitat Pompeu Fabra de Barcelona, España), Jimena Yisel Caballero Contreras (Universidad Nacional Autónoma de México, México), Lucas Emmanuel Misseri (Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina), Maria Elice Brzezinski Prestes (Universidade de São Paulo, Brasil), Mariano Nicolás Hochman (Universidad de Buenos Aires, Argentina), Renato Marcone José de Souza (Universidade Federal de São Paulo, Brasil), Silvia Dotta (Universidade Federal do ABC, Brasil), Thais Cyrino de Mello Forato (Universidade Federal de São Paulo, Brasil), Vasil Gluchman (University of Prešov, Eslováquia), Waldmir Nascimento de Araujo Neto (Universidade Federal do Rio de Janeiro, Brasil).

**Asesores académicos externos**

Abigail Vital – Centro Federal de Educação Tecnológica – RJ, Brasil, Alexandre Bagdonas – Universidade Federal de Lavras, Brasil, André Noronha – Instituto Federal de São Paulo, Brasil, Boniek Venceslau da Cruz Silva – Universidade Federal do Piauí, Brasil, Breno Arsioli Moura – Universidade Federal do ABC, Brasil, Carlos Eduardo Ribeiro – Universidade Federal do ABC, Brasil, Carlos Roberto Senise Júnior – Universidade Federal de São Paulo, Brasil, Daniel Quaresma Figueira Soares – Universidade de São Paulo, Brasil, Danilo Cardoso – Universidade de São Paulo, Brasil, Denilson Cordeiro – Universidade Federal de São Paulo, Brasil, Esdras Viggiano – Universidade Federal do Triângulo Mineiro, Brasil, Evaldo Araujo de Oliveira Filho – Universidade Federal de São Paulo, Francisco Ângelo Coutinho – Universidade Federal de Minas Gerais, Brasil, Guilherme Brockington – Universidade Federal do ABC, Brasil, Helio Elael Bonini Viana – Universidade Federal de São Paulo, Brasil, Isabelle Priscila Carneiro de Lima – Instituto Federal da Bahia, Brasil, Ivã Gurgel – Universidade de São Paulo, Brasil, Jose Antonio Ferreira Pinto – Universidade Estadual da Paraíba, Brasil, Leonardo André Testoni – Universidade Federal de São Paulo, Brasil, Luciana Caixeta Barboza – Universidade Federal do Triângulo Mineiro, Brasil, Luciana Monteiro de Moura – Universidade Federal de São Paulo, Brasil, Luciana Zaterka – Universidade Federal do ABC, Brasil, Lúcio Costa – Universidade Federal do ABC, Brasil, Marco Braga – Centro Federal de Educação Tecnológica – Rio de Janeiro, Brasil, Maria Inês Ribas Rodrigues – Universidade Federal do ABC, Brasil, Maria Luiza Ledesma Rodrigues – Universidade Estadual Paulista, Brasil, Marlon Cesar de Alcântara – Instituto Federal Sudeste de Minas Gerais, Brasil, Nadja Magalhães – Universidade Federal de São Paulo, Brasil, Nei de Freitas Nunes Neto – Universidade Federal da Grande Dourados, Brasil, Renato Kinouchi – Universidade Federal do ABC,

Brasil, Renato Marcone José de Souza – Universidade Federal de São Paulo, Brasil, Simone Alves de Assis Martorano – Universidade Federal de São Paulo , Simone Nakaguma – Universidade Federal de São Paulo, Winston Schmiedecke – Instituto Federal de São Paulo, Brasil.

### **Formato de la publicación**

Digital: Portable Document Format (PDF), Hyper Text Markup Language (HTML), Extensible Markup Language (XML), Eletronic Publication (ePub).

### **Idiomas aceptados**

Castellano, portugués e inglés (lenguas de la publicación).

### **Normas de publicación**

<https://periodicos.unifesp.br/index.php/prometeica/about/submissions>

### **Contacto**

[prometeica@unifesp.br](mailto:prometeica@unifesp.br)

### **Responsable**

Emiliano Aldegani (Universidad Nacional de Mar del Plata, Argentina)

### **Diseño de isologo**

Victoria Reyes (<http://www.victoriareyes.com.ar>)

## EDITORIAL

**5-6 | JIMENA YISEL CABALLERO CONTRERAS** | Virtualidad y ciencia

## ARTÍCULOS

**7-24 | ALDANA MASTRANDREA** | Aportes teórico-conceptuales para el abordaje de la percepción y representaciones sociales del riesgo hídrico

**25-35 | KARLA CASTILLO VILLAPUDUA** | De la posible renovación del pensamiento crítico ante la filosofía emancipadora de J. Rancière

**36-59 | FREDERICO AGUILAR TAMAYO et al.** | Sacralidad y leyenda popular. Imagen, mito y relato en el personaje de la llorona

**60-76 | ALBA MASSOLO; MARIEL TRAVERSI** | ¿Es posible mitigar sesgos cognitivos? Un análisis crítico de diversas propuestas para reducir el sesgo a mi favor

**77-91 | MARIANO PETTARIN** | La Historia como ciencia Autónoma de los Asuntos Humanos: Collingwood y la experiencia como experiencia histórica

**92-107 | CARLOS MAXIMILIANO SENCI** | Revisitando críticamente nociones de similitud en la teoría de juegos comportamental

**108-122 | ENRIC TRILLAS; ALEJANDRO SOBRINO CERDEIRIÑA** | Alegato por el control del razonamiento

**123-135 | LEOPOLDO TILLERÍA** | Una aproximación al problema del mal en Schelling

**136-149 | MANAS SAHU** | The Kantian notion of freedom and Autonomy of Artificial Agency

## ENTREVISTAS

**150-154 | CRISTINA PONTES BONFIGLIOLI** | Entrevista a Nicola Liberati: desafíos metodológicos para una nueva filosofía de la tecnología

## RESEÑAS

**155-157 | BRENO BENEDYKT** | Um apartamento em Urano: crônicas da travessia

**158-161 | JUAN CRUZ APCARIÁN** | El asedio a la imaginación



## VIRTUALIDAD Y CIENCIA

Dentro del estudio de las comunidades virtuales, se reconoce a la comunidad científica como la primera. Y no se refiere a la concepción basada en la comunicación digital; la virtualidad del trabajo del científico radica en la forma moderna de realizar investigación y la importancia de la divulgación de la ciencia.

El científico investiga, analiza y publica sus resultados con la finalidad de que otros científicos los lean y puedan continuar con el proceso de generación de conocimiento; aceptarlos, refutarlos, complementarlos; todo ello sin siquiera conocerse, sólo al leerse mutuamente. Así viaja el conocimiento, sin fronteras, ideologías o dogmas.

Proceso nacido y ubicado en la Escuela Invisible, (Palo Alto, Frankfurt), donde el conocimiento viajaba de aquí para allá sin la necesidad de tener una localización física. El desarrollo del conocimiento en la actualidad, gracias a las nuevas plataformas digitales, es más visible, más abarcador y más trascendental. Aquí está la importancia de la labor de los journals científicos y académicos de todo el mundo, que naturalmente evolucionaron a un formato digital.

De eso se trata la virtualidad, retomando su apreciación ontológica de potencia de ser. En el ámbito científico sabemos que el conocimiento existe hasta que es reconocido por la otredad, un texto académico es virtualidad al publicarse, citarlo lo hace potencialmente real y los siguientes pasos (refutarlo, aceptarlo, etc.) lo hace trascendental.

Ese movimiento natural de la difusión de la ciencia sentó las bases para comprender posteriormente lo que las comunidades virtuales generan: comunicación, inteligencia colectiva, cultura participativa, etcétera. La apropiación veloz de las Tecnologías de la Comunicación y la Información, principalmente el inicio de la web, aceleró el proceso comunicativo a escala mundial, logrando que el proceso de difusión de conocimiento incrementara.

La información existe, de forma desmesurada y caótica, como se ha estudiado los últimos años, lo cual genera inclusive desinformación (titulado “infodemia” actualmente). La regulación de dicho conocimiento se logró gracias a la institucionalización de entidades académicas. Al menos en el área científica. Es ahí donde los journals académicos tienen el trabajo de conciliar el proceso de la comunidad virtual académica.

El proceso de conciliación del conocimiento se da en un alcance nunca antes visto. Un tema presentado es dictaminado por científicos expertos en el área, no importa de qué nacionalidad sean, o en qué espacio físico se encuentren, el alcance de la comunicación digital rompe las fronteras permitiendo acelerar el proceso de difusión de conocimiento y de rigurosidad metodológica.

A su vez, la publicación en journals digitales de acceso libre pone en común los contenidos, donde gracias a los buscadores web y bases de datos, los interesados en el tema pueden encontrar los textos de forma rápida, tanto expertos en esas áreas como al público en general, desde un estudiante universitario, hasta un adolescente curioso. La información está ahí, abalada, cuidada y demostrada.

Una de las premisas del conocimiento científico es que nunca se detiene, la potencialidad (virtualidad) de seguir transformándose puede llegar a ser infinita. Texto publicado, texto citado, texto analizado, texto transformado.

Cuántas críticas trajo consigo la concepción de posmodernidad atravesada por el fenómeno de la globalización. Fueron pocos los que reconocieron la potencialidad que traería aquello al desarrollo de la ciencia.

Las comunicaciones digitales expandieron el mundo de la vida cotidiana, y no sólo eso, abrieron un mar de posibilidad que permite la interconexión de las personas, las comunidades, y por supuesto, los conocimientos.

Actualmente vivimos en carne propia el ejemplo más claro de todo lo anterior; pese a la llegada de una pandemia mundial, gracias a la virtualidad pudimos adaptarnos y lograr continuar nuestras vidas cotidianas y no sólo eso, la comunidad científica no detuvo su labor logrando conseguir y difundir conocimiento sobre el nuevo virus haciéndole frente.

El mundo virtual hoy en día es parte del mundo vital.

***Jimena Yisel Caballero Contreras***

*Editora Adjunta*

*(Universidad Nacional Autónoma de México)*

*jimenayisel@politicas.unam.mx*

## APORTES TEÓRICO-CONCEPTUALES PARA EL ABORDAJE DE LA PERCEPCIÓN Y REPRESENTACIONES SOCIALES DEL RIESGO HÍDRICO

*THEORETICAL-CONCEPTUAL CONTRIBUTIONS TO ADDRESSING THE PERCEPTION AND SOCIAL REPRESENTATIONS OF WATER RISK* TÍTULO DEL

*CONTRIBUIÇÕES TEÓRICO-CONCEITUAIS PARA ABORDAR O PERCEPÇÃO E REPRESENTAÇÕES SOCIAIS DO RISCO DA ÁGUA*

**Aldana Mastrandrea**

(Universidad Nacional del Sur, Argentina)

[aldana.mastrandrea@uns.edu.ar](mailto:aldana.mastrandrea@uns.edu.ar)

Recibido: 09/03/2021

Aprobado: 04/04/2021

### RESUMEN

La sociedad actual experimenta un conjunto de transformaciones que se manifiestan en la compleja relación que establece con la naturaleza. Dicha relación presenta diferentes grados de incertidumbre según la probabilidad de ocurrencia de fenómenos que signifiquen una amenaza o peligro. Por ello, es imprescindible conocer las características de este vínculo para generar adecuados mecanismos de actuación ante la ocurrencia de eventos potencialmente peligrosos para la población, como inundaciones, anegamientos y desbordes de cursos de agua. El artículo, de carácter teórico, adopta una estrategia de investigación cualitativa y se centra en el abordaje de la Teoría Social del Riesgo desde la perspectiva de la Geografía de la Percepción y la Teoría de las Representaciones Sociales. El mismo, a través de la revisión de la literatura, se propone realizar un recorrido teórico-conceptual mediante la selección de conceptos, categorías y dimensiones de análisis que ofrezcan un marco contextual para conocer y analizar la configuración de escenarios de criticidad ambiental en relación con los recursos hídricos. Dicho abordaje, constituye un enfoque enriquecedor para incrementar los niveles de conocimiento sobre el riesgo y orientar las acciones que, desde diferentes ámbitos, debieran implementarse en la gestión integral y participativa del riesgo hídrico.

Palabras clave: percepción de riesgos. criticidad ambiental. gestión y ordenamiento territorial.

### ABSTRACT

Today's society experiences transformations that are manifested in the complex relationship it establishes with nature. This relationship presents different degrees of uncertainty according to the probability of phenomena that represent a threat or danger. Therefore, it is essential to know the characteristics of this link to generate adequate mechanisms of action in the event of potentially dangerous events for the population, such as floods, flooding, and overflows of watercourses. The article, of a theoretical nature, adopts a qualitative research strategy and focuses on the approach to the Social Theory of Risk from the perspective of

the Geography of Perception and the Theory of Social Representations. Through the literature review, it proposes to carry out a theoretical-conceptual journey through the selection of concepts, categories, and dimensions of analysis that offer a contextual framework to know and analyze the configuration of environmental criticality scenarios concerning water resources. This approach constitutes an enriching approach to increase the levels of knowledge about risk and guide the actions that, from different areas, should be implemented in the comprehensive and participatory management of water risk.

Keywords: risk perception. environmental criticality. management and land use planning.

## RESUMO

A sociedade atual vive um conjunto de transformações que se manifesta na complexa relação que estabelece com a natureza. Esta relação apresenta diferentes graus de incerteza dependendo da probabilidade de ocorrência de fenômenos que representam uma ameaça ou perigo. Por isso, é fundamental conhecer as características desse elo para gerar mecanismos de ação adequados em caso de eventos potencialmente perigosos para a população, como enchentes, inundações e transbordamentos de cursos d'água. O artigo, de natureza teórica, adota uma estratégia de pesquisa qualitativa e enfoca a abordagem da Teoria Social do Risco na perspectiva da Geografia da Percepção e da Teoria das Representações Sociais. Por meio da revisão da literatura, propõe-se realizar um percurso teórico-conceitual utilizando da seleção de conceitos, categorias e dimensões de análise que oferecem um arcabouço contextual para conhecer e analisar a configuração de cenários de criticidade ambiental em relação aos recursos hídricos. Esta abordagem é enriquecedora para aumentar os níveis de conhecimento sobre o risco e orientar as ações que, a partir de diferentes áreas, devem ser implementadas na gestão integral e participativa do risco hídrico.

Palavras-chave: percepção de risco. criticidade ambiental. gestão e ordenamento do território.

## Introducción

En la actualidad uno de los mayores desafíos para los actores sociales provenientes del ámbito científico-técnico y gubernamental consiste en poder incrementar los niveles de conocimiento para implementar adecuados mecanismos de actuación y reducir así, los potenciales daños ambientales derivados de la ocurrencia de diferentes riesgos. La gestión de riesgos se trata de una cuestión de desarrollo que debiera centrarse en la disminución de la incertidumbre y en el control de la vulnerabilidad con el fin de reducir los problemas futuros asociados con eventos potencialmente peligrosos, por ejemplo, los derivados de la alternancia de eventos de exceso y déficit hídrico. Según Lavell (1996):

El enfoque de la Gestión de Riesgo se refiere a un proceso social complejo a través del cual se pretende lograr una reducción de los niveles de riesgo existentes en la sociedad y fomentar procesos de construcción de nuevas oportunidades de producción y asentamiento en el territorio de condiciones de seguridad y sostenibilidad aceptables... se refiere al proceso por medio del cual un grupo humano o individuo toman conciencia del riesgo que enfrenta, lo analiza y lo entiende, considera las opciones y prioridades en términos de su reducción... (Lavell, 1996: 9).

Los procesos, como expresión directa del dinamismo que caracteriza el comportamiento de los sistemas ambientales, incluyen las acciones de diversos actores, cuyas prácticas y representaciones sociales contribuyen a construir y organizar socialmente el lugar que habitan. Santos (1996) expresa que el espacio geográfico constituye un conjunto de relaciones realizadas a través de las funciones y de las formas que se presentan como testimonio de una historia escrita por procesos de lógicas temporales distintas. Es decir, el espacio se define como un conjunto de formas que expresan motivaciones y acciones de las sociedades y por las relaciones producidas a través de dichas formas y funciones sociales.

Las circunstancias a partir de las cuales se producen las condiciones de riesgo hídrico responden a las características de la relación sociedad-naturaleza. En dicha relación intervienen factores vinculados con un evento natural desencadenante, como puede ser la ocurrencia de precipitaciones intensas, y con aspectos socioeconómicos, políticos, técnicos, institucionales, ideológicos y culturales (Maskrey, 1993, Natenzon y Ríos, 2015, Lavell *et al.*, 2020). De este modo las condiciones que explican la ocurrencia de riesgos, cuyo detonante puede ser un evento natural, guardan relación directa con los procesos de configuración y organización territorial que definen la magnitud del impacto, en tal sentido, condiciones multiescalares y multidimensionales (Arteaga, 2020). Puede afirmarse que el riesgo deriva de una construcción social por parte de los diferentes actores que, en un escenario espaciotemporal concreto y de acuerdo con diferentes mecanismos de percepción, elaboran determinadas representaciones sociales de los acontecimientos que se manifiestan en la vida cotidiana (Mastrandrea y Pérez, 2020).

Para conocer la complejidad que involucran los procesos que se vinculan con los recursos hídricos es necesario considerar la construcción social de escenarios de riesgo (Beck, 1993, Rees, 2006). Esta dimensión de análisis incluye las múltiples perspectivas e interpretaciones del riesgo hídrico de acuerdo con cada persona en función del contexto social e histórico de la comunidad. Así, Herzer (2011) sostiene que existe una cuestión de escalas espaciales y territoriales de análisis que deben tenerse en cuenta para analizar el riesgo:

En los procesos de construcción de riesgo, hay actores sociales involucrados que no concuerdan necesariamente con aquellos que padecen los desastres. Su forma de percibir los procesos, así como la conciencia crítica que desarrollan también varían y se hallan sesgadas por sus diversos posicionamientos sociales (Herzer, 2011: 58).

De allí que, es fundamental tener en cuenta las diferentes percepciones, valoraciones y experiencias de los actores sociales para tener una mirada integral del riesgo hídrico y poder llevar adelante una adecuada gestión del mismo (Cardona, 1993, Carballo *et al.*, 2014). El riesgo, entendido como construcción social, debe ser incluido en la agenda pública a través de un análisis de las categorías de peligrosidad, vulnerabilidad, exposición e incertidumbre (Natenzon *et al.*, 2015). La probabilidad de ocurrencia de un evento peligroso es una condición necesaria para la elaboración de planes de ordenamiento territorial que incluyan programas y proyectos tendientes a conocer la configuración de diferentes escenarios de riesgo (Vega Meléndez, 2017). Para llevar a cabo dicho proceso es necesario promover una cultura preventiva y participativa y articular los intereses de diversos actores sociales entre diferentes sectores gubernamentales y no gubernamentales, académicos, de organizaciones sociales y de la población en general (Medina y Vigo, 2020).

Teniendo en cuenta lo expresado anteriormente, el presente trabajo se propone realizar -desde un enfoque teórico y descriptivo- un recorrido teórico-conceptual sobre el riesgo hídrico desde la Teoría Social del Riesgo, la Geografía de la Percepción y la Teoría de las Representaciones Sociales a fin de ofrecer un marco contextual para el abordaje integral del riesgo hídrico.

## **1. La Teoría Social del Riesgo: hacia una conceptualización**

La sociedad actual experimenta un conjunto de transformaciones que se manifiestan en la compleja relación que establece con la naturaleza. Dicha relación presenta diferentes grados de incertidumbre según la probabilidad de ocurrencia de fenómenos que signifiquen una amenaza o peligro<sup>1</sup>. Por ello, es imprescindible conocer las características de este vínculo para generar adecuados mecanismos de actuación ante la generación de eventos potencialmente peligrosos para la población.

---

<sup>1</sup> Según Natenzon *et al.* (2015) los conceptos de amenaza y peligro refieren al mismo fenómeno, es decir, al proceso o factor disparador de escenarios de riesgo.

Teniendo en cuenta las investigaciones y trabajos que abordan el concepto de riesgo se encuentran las conceptualizaciones propuestas por autores de distintas disciplinas<sup>2</sup>, como por ejemplo: Geografía, Sociología, Filosofía, Ingeniería y Abogacía. En primer lugar, se consideran los aportes de distintos investigadores provenientes del ámbito internacional como: Beck (1993), Maskrey (1993), Wilches Chaux (1993), Giddens *et al.* (1996), Lavell (1996, 1999, 2020) y Rees (2006). Los autores definen el concepto de riesgo y proporcionan los elementos de análisis necesarios para conocer la relación sociedad-naturaleza y los posibles efectos ante la ocurrencia de un desequilibrio que altere dicha relación. En segundo lugar, se tienen en cuenta las conceptualizaciones de riesgo propuestas por investigadores del ámbito nacional, entre quienes se encuentran: Natenzon (1995), Natenzon *et al.* (2003), Cardona (1993), Carballo (2001, 2014), Bertoncetto y Alesandri (2003), García Acosta (2005), Herzer (2011) y Natenzon y Ríos (2015). En sus estudios, sostienen que las primeras líneas de investigación en torno al concepto de riesgo se iniciaron con trabajos de especialistas procedentes del campo de las Ciencias Naturales, los cuales se focalizaron principalmente en la ocurrencia de las amenazas -fenómenos físicos-. Sin embargo, coinciden en que es preciso considerar otras categorías de análisis como, por ejemplo, la vulnerabilidad y la exposición, dimensiones continuas en el tiempo y espacio que se relacionan con aspectos socioculturales y socioeconómicos de las comunidades.

En relación con el primer grupo de autores, se encuentran los aportes de Beck (1993) quien sostiene que la probabilidad de ocurrencia de riesgos constituye situaciones de peligro para la sociedad. Estos se manifiestan tanto en la ocurrencia de complejos problemas ambientales, como en los peligros diarios que afectan la vida cotidiana de las personas, entre los que se encuentran las inundaciones, anegamientos y sequías. Sin embargo, el reconocimiento de los peligros diarios como peligrosos por parte de los distintos actores de la sociedad difiere de acuerdo con los factores que inciden en su percepción del riesgo. De allí que, el análisis de la valoración que las comunidades tienen respecto de los períodos con abundantes precipitaciones y sequías es fundamental para reconocer las concepciones construidas en torno a estos fenómenos y a los problemas que adolecen con el fin de llevar a cabo una adecuada gestión del riesgo hídrico.

Otra visión, es la que brinda Maskrey (1993) quien afirma que la lectura del riesgo que realizan los especialistas provenientes del ámbito científico difiere de la representación social que tienen los habitantes, las comunidades expuestas a un peligro y las autoridades gubernamentales encargadas de la gestión y toma de decisiones. Así, es necesario profundizar en el estudio de la percepción individual y colectiva del riesgo hídrico para conocer las decisiones que se materializan en el espacio, según las concepciones construidas por los habitantes, las cuales inciden en la probabilidad de que se generen escenarios de criticidad ambiental.

Por su parte, Wilches Chaux (1993) sostiene que la generación de escenarios de riesgos y de criticidad ambiental presenta dos orígenes: los procesos intrínsecos de transformación de la naturaleza y la actividad humana. Estos orígenes, a su vez, se encuentran interrelacionados por lo que no es posible establecer una división tajante entre riesgos de origen natural y social, lo que implica tener en cuenta una visión integral de la relación sociedad-naturaleza en la definición del concepto de riesgo. Así, Wilches Chaux (1993) entiende el riesgo como “... cualquier fenómeno de origen natural o humano que signifique un cambio en el medio ambiente que ocupa una comunidad determinada, que sea vulnerable a ese fenómeno” (1993: 17).

Mientras que Giddens *et al.* (1996) expresan que el riesgo es un resultado imprevisto que surge como consecuencia de nuestras propias actividades o decisiones, y no por obra divina, fortuna o fatalidad. Por su parte, Lavell (1999, 2020) plantea que el riesgo se refiere a un contexto caracterizado por la probabilidad de pérdidas y daños en el futuro, que van desde las físicas hasta las psico-sociales y culturales. Al respecto, un aporte significativo con respecto a la concepción del riesgo es el de Rees

---

<sup>2</sup> Natenzon *et al.* (2015) plantean que la Geografía y otras disciplinas han realizado aportes significativos en lo que respecta al conocimiento de los escenarios de construcción social del riesgo, entre las que se encuentran: la Ecología Humana, Economía Política, Sociología, Filosofía, Antropología e Historia. Asimismo, sostienen que no es sencillo concebir un concepto de riesgo aceptado universalmente porque cambia según las disciplinas y el enfoque que adoptan. Sin embargo, consideran que las definiciones, en general, se refieren a probables eventos que pueden producir daños.



(2006) quien afirma que la manera en la cual los diferentes actores sociales perciben los riesgos incide en los modos de actuación frente a un problema y ejerce una notable influencia en el desarrollo de la gestión del riesgo. Según la autora el riesgo no es un fenómeno físico, sino más bien cultural, concebido como aquellos peligros que las propias sociedades definen como preocupantes.

En suma, es posible afirmar que el riesgo es una probabilidad de daños relacionados con la existencia de determinadas condiciones naturales y sociales que se expresa bajo una condición latente y puede medirse en términos cuantitativos y cualitativos. De allí que, es importante considerar de modo integral las variables físico-naturales (clima, relieve, geomorfología, hidrografía, suelos y vegetación) y socioculturales (características sociodemográficas, usos del suelo, prácticas espaciales materiales y percepciones y representaciones sociales) que intervienen en la construcción social de escenarios de riesgo hídrico.

Por su parte, con respecto al segundo grupo de autores -provenientes del ámbito nacional- es preciso mencionar que la conceptualización del riesgo de las últimas dos décadas difiere de la concepción que se tenía de él hace décadas atrás. En consecuencia, los modos en los que se analiza investigan y estudian el riesgo requieren de la consideración de nuevos enfoques de investigación.

Según Natenzon *et al.* (2015), los primeros estudios que trataron las catástrofes y los riesgos<sup>3</sup> surgieron durante la primera mitad del siglo XX a partir de un abordaje fisicalista que, desde las Ciencias Naturales, se focalizaron en la ocurrencia de eventos de naturaleza extrema. Estos acontecimientos, ubicaban a la sociedad en un lugar pasivo ante la magnitud de las fuerzas físicas asociadas con fenómenos meteorológicos, geológicos y biológicos, entre otros. En tal sentido, se otorgaba una especial importancia a la ocurrencia de un evento potencialmente peligroso o a la amenaza, entendida esta como la probabilidad de que ocurra un riesgo frente al cual una comunidad es vulnerable (Natenzon, 1995).

Sin embargo, a partir de la década de 1940, desde la perspectiva de la Ecología Humana, comienzan a cuestionarse las aproximaciones realizadas hasta el momento y se introducen los conceptos de percepción y adaptación humana, de la mano de White<sup>4</sup> (1974, en: Saurí Pujol, 2006) quien le asigna un rol primordial a la sociedad en la participación de los desastres. A partir de la década de 1970, se efectuó una crítica a las premisas de la Ecología Humana, desde la llamada Economía Política de los Desastres que brindaba relevancia a las condiciones estructurales -especialmente económicas y políticas- determinantes en la ocurrencia de dichos eventos. En este contexto, se introduce la noción de construcción social de los escenarios de riesgo que antecede a la ocurrencia de un desastre, expresión que articula la amenaza o peligrosidad con la vulnerabilidad. Así, los estudios de la vulnerabilidad física -relacionada con el grado de exposición de los elementos materiales ante un fenómeno- y las investigaciones relacionadas con la vulnerabilidad global -expresión que considera diversos ángulos, como el político, económico, cultural y ambiental, entre otros- ampliaron el campo de estudio y presentaron una mirada multidisciplinaria para comprender la construcción social del riesgo (Natenzon *et al.*, 2015).

Según Cardona (1993), en los países en desarrollo, la vulnerabilidad social es en la mayoría de los casos, la causa de las condiciones de vulnerabilidad física. A diferencia de la amenaza que actúa como agente detonante, la vulnerabilidad social es una condición que se gesta, acumula y permanece en forma continua en el tiempo y se encuentra íntimamente ligada a los aspectos culturales y al nivel de desarrollo de las comunidades. En relación con el concepto de riesgo, el autor sostiene que es el grado de pérdidas

---

<sup>3</sup> Los conceptos de catástrofe y desastre son tomados como sinónimos por Natenzon *et al.* (2015) y se refieren indistintamente al mismo proceso: ocurrencia de un suceso fatídico que genera gran destrucción y que altera el orden regular de las cosas. De este modo, la diferencia entre estos conceptos es esencialmente etimológica. Los autores sostienen que el concepto de desastre ha alcanzado una gran difusión como consecuencia del accionar del campo académico y de organismos internacionales de cooperación especializados. Uno de los primeros trabajos que abordó el concepto de desastre, impacto y cambio social fue el de Samuel Prince que, desde la Sociología, estudió el impacto causado por una explosión de un buque de municiones en el puerto de Halifax-Canadá sobre familias y comunidades (Lavell, 2004, en: Natenzon *et al.*, 2015).

<sup>4</sup> Gilbert White, en su tesis doctoral, aporta una de las primeras aproximaciones sobre los riesgos naturales desde la perspectiva de la Ecología Humana y desde la Geografía de la Percepción. Sus investigaciones se han focalizado especialmente en las inundaciones que ocurrieron en la cuenca del Mississippi en 1927 acompañando su práctica profesional con la académica (Saurí Pujol, 2006).



esperadas debido a la ocurrencia de un evento particular y lo considera una función de la amenaza y la vulnerabilidad. Así, define el riesgo como la relación entre estas dos dimensiones y, según su visión, puede ser de carácter geológico, hidrológico, atmosférico o tecnológico.

Por su parte, teniendo en cuenta estudios sociológicos y antropológicos sobre los desastres, es importante mencionar que estos últimos se definen como los productos o hechos consumados sobre los cuales es necesario actuar en lo que respecta a la emergencia, rehabilitación y reconstrucción (Lavell, 1999). Es dable destacar que el desastre se diferencia del concepto de riesgo el cual debe concebirse en términos de procesos dinámicos dado que, se construye social, histórica y territorialmente en la articulación de amenazas y vulnerabilidades (Natenzon *et al.*, 2015). Douglas (1996) sostiene que:

... el riesgo es una construcción colectiva y cultural... cada sociedad ordena la forma de percibir, aceptar, tolerar o evitar cierto tipo de riesgo. Son las sociedades, según sus creencias, visiones dominantes e instituciones, las que deciden lo que es riesgoso de lo que no, como también las que aceptan o no determinado tipo de riesgo (Douglas, 1996 en: Natenzon y Ríos, 2015: 5).

De allí que, la construcción y reconstrucción de escenarios de riesgo varía de acuerdo con cada sociedad en un período histórico determinado y según la percepción que las personas elaboran de los eventos que atraviesan en sus vidas diarias. Teniendo en cuenta que la construcción de escenarios de riesgo adquiere diferentes características según: la escala en la que se producen (global, nacional, regional, local y micro-local), la intensidad del fenómeno desencadenante y las particularidades socioeconómicas y socioculturales previas a la ocurrencia del evento es importante considerar su complejidad en las diferentes instancias de la gestión del riesgo. Por ello, es preciso abordar de un modo multidimensional las variables -físico-naturales, socioeconómicas, políticas, culturales, históricas y territoriales- que intervienen en la ocurrencia de riesgos ambientales, mediante la relación entre las vulnerabilidades y amenazas. Considerando dichas aproximaciones se entiende que el riesgo es esencialmente social debido a que responde a una construcción sociohistórica compleja que se expresa en una convivencia vulnerable entre diversos grupos sociales y su medio.

Carballo (2001) sostiene que plantear el riesgo como una complejidad ambiental constituye el primer paso para implementar medidas tendientes a reducirlo. La autora, define el riesgo como las condiciones de las interacciones sociales con el medio y de las relaciones sistémicas caracterizadas por una dinámica abierta e inestable. Por su parte, Bertoncello y Alesandri (2003) consideran el riesgo como la forma de cuantificar la incertidumbre e incorporan el cálculo como medición considerando distintas formas de establecerla. En esta misma línea, Ríos y Natenzon (2015) sostienen que el riesgo se produce cuando se puede establecer una probabilidad cuantificable -es decir, cuando se puede medir- del fenómeno a ocurrir. Los autores definen el concepto como un resultado imprevisto que sucede como consecuencia de las actividades y decisiones antrópicas las cuales implican una reflexión continua frente a los problemas que acontecen.

Al respecto, White (1974, en: Lavell, 1999) y su grupo de investigación fueron los primeros en postular de forma explícita que el riesgo no es sinónimo de las amenazas naturales. El autor expuso que el riesgo a sufrir un desastre dependía no sólo de la magnitud de la amenaza natural como tal, sino también, de la vulnerabilidad de la sociedad expuesta a la amenaza. De allí, se llegó a la fórmula ampliamente difundida para articular las dimensiones consideradas teniendo en cuenta que el riesgo es el resultado de la conjunción entre la amenaza y la vulnerabilidad:

$$RIESGO = AMENAZA * VULNERABILIDAD$$

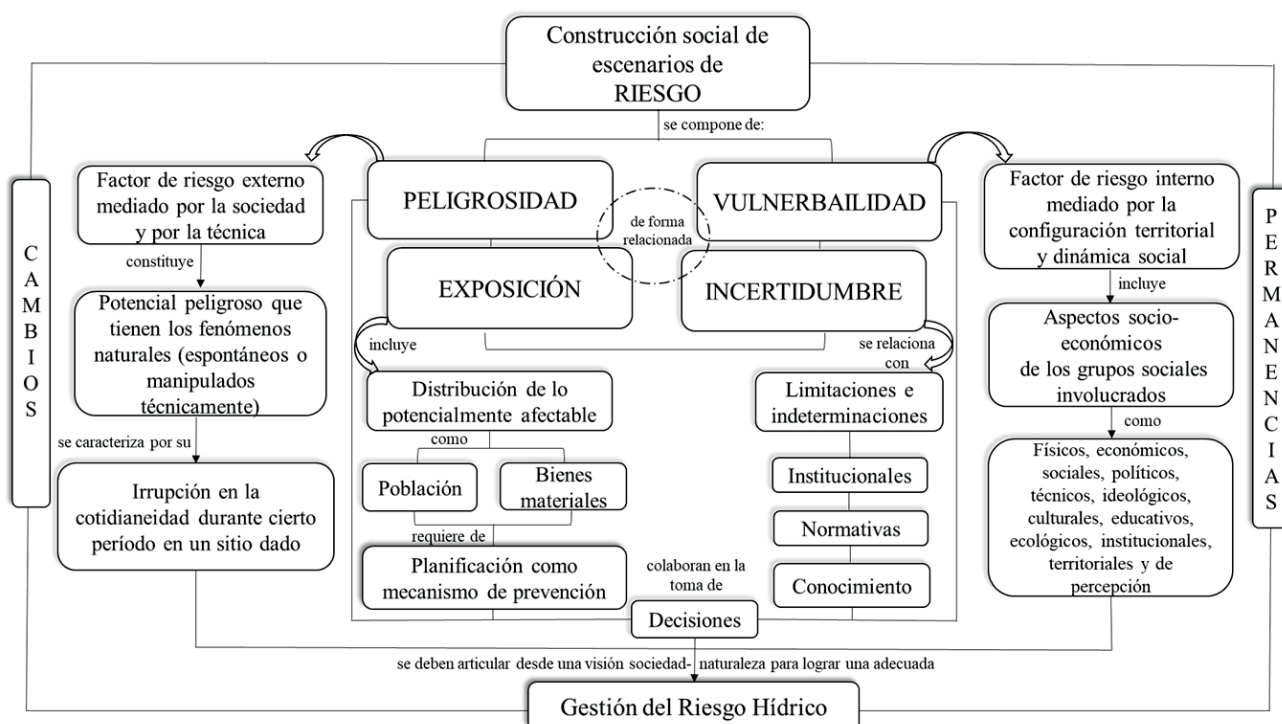
Teniendo en cuenta los posibles eventos que pueden desencadenarse en el futuro y generar condiciones de riesgo Cardona (1993) sostiene que es importante conocer las características de estos fenómenos en un contexto integral para planificar y orientar los procesos de desarrollo. García Acosta (2005) afirma que es necesario considerar los procesos de construcción social de escenarios de riesgo en perspectiva

histórica a través de la percepción que la sociedad construye de ellos. La autora manifiesta que las condiciones de vulnerabilidad definen y determinan la magnitud de los efectos de una amenaza, por ello, es fundamental incluirla en el mismo nivel que las peligrosidades para el abordaje del riesgo.

Por su parte, Herzer (2011) expresa que el espacio urbano ha sido una de los escenarios más frecuentes de generación de escenarios de riesgo. Sin embargo, el análisis del riesgo, ha estado enfocado en los fenómenos físico-naturales dejando de lado el abordaje de los demás componentes de riesgo urbano como la concentración espacial de la población e infraestructura junto con la ausencia de normas y su aplicación.

De este modo, es relevante considerar la existencia de diversas investigaciones que se han llevado a cabo a través de estudios que abordan la Teoría Social del Riesgo, desde diferentes perspectivas de análisis, provenientes tanto de las Ciencias Naturales como de las Ciencias Sociales. Las aproximaciones realizadas manifiestan la complejidad que incluye el concepto de riesgo, generando intercambios y discusiones dentro del ámbito científico. Por lo tanto, su abordaje, requiere de un enfoque sistémico que contemple las categorías de análisis -peligrosidad, vulnerabilidad, exposición e incertidumbre- y las dimensiones que se relacionan con ellas, de modo integral a través de la relación sociedad-naturaleza (figura 1).

**Figura 1.** Características, categorías y dimensiones de análisis de la Teoría Social del Riesgo<sup>5</sup>



Fuente: elaborado por Mastrandrea, sobre la base de Lavell (1999) y Natenzon y Ríos (2015).

## 2. La Geografía de la Percepción. Una perspectiva valiosa para el abordaje de los significados construidos en torno al riesgo hídrico

La Geografía de la Percepción ha realizado aportes significativos acerca de cómo las personas conciben el territorio en términos de vivencias, valoraciones y experiencias construidas individual y colectivamente. Esta perspectiva a partir de la cual se analiza el espacio y las relaciones que en él se producen presenta una significativa potencialidad para conocer las formas en las que los distintos actores

<sup>5</sup> Nota: La figura expresa las relaciones que deben considerarse entre las Ciencias Naturales y Sociales para el conocimiento de la construcción social de escenarios de riesgo. Asimismo, en el esquema se presentan los principales aspectos que han sido seleccionados para el abordaje de las categorías de análisis del riesgo (peligrosidad, vulnerabilidad, exposición e incertidumbre) los cuales debieran orientar la planificación y ordenamiento territorial, desde un enfoque espaciotemporal que incluya tanto los cambios como las permanencias.

sociales construyen -de modo consciente o inconsciente- las percepciones acerca de la realidad en la que desarrollan sus prácticas.

En la presente investigación se abordan, desde la Geografía de la Percepción, las categorías y dimensiones de análisis necesarias para reconocer la percepción social del riesgo hídrico que elaboran diferentes actores sociales (figura 2). En tal sentido, analizar la configuración de escenarios de criticidad ambiental a través de la percepción que las personas tienen y de las representaciones sociales que construyen, especialmente de las inundaciones y sequías, permite reconocer los significados de las prácticas espaciales llevadas a cabo por estos junto con las necesidades y demandas de la población.

En el proceso de conocimiento de la realidad social las personas perciben su entorno próximo, como así también los problemas que enfrentan, de un modo particular que se encuentra íntimamente ligado al contexto social e histórico de la comunidad local. Las experiencias se ordenan, se interpretan y en ciertas ocasiones, se transforman de acuerdo con la incidencia de diversos filtros perceptuales, entre los que se pueden mencionar: la información recibida, las experiencias vividas y las características del contexto -socioeconómico, político, histórico, ambiental y cultural, entre otras.

Con respecto a los primeros estudios relacionados con la percepción social del riesgo hídrico es importante mencionar que, según Wilches Chaux (1993), las investigaciones iniciales tuvieron un gran impulso especialmente en la década de 1960 y 1970 (White, *et al.*, en: Saurí Pujol, 2003). Luego, años más tarde, se incorporaron otros especialistas que coincidían en que las diversas formas en que las poblaciones organizan objetiva y subjetivamente su conocimiento del riesgo influyen en sus decisiones sobre localización, diversificación productiva y formas de autoprotección -estructurales, de comportamiento y de seguridad, entre otras-. En este contexto, Beck (1993) reflexiona acerca de la visibilidad o invisibilidad de los riesgos por parte de la sociedad y sostiene que, en la sociedad del riesgo, lo visible queda a la sombra de las amenazas invisibles siendo estos riesgos negados los que crecen con mayor fuerza.

La perspectiva de Beck (1993) pone de manifiesto que la percepción juega un papel fundamental en la prevención de los riesgos que pueden afectar a una sociedad por tratarse de comunidades que tienen una elevada o baja percepción de él. En tal sentido, la existencia de diversos factores que influyen en los procesos de percepción social del riesgo ejerce una notable influencia en la generación de una conciencia colectiva para conocer y prevenir la probabilidad de ocurrencia de un evento peligroso. Así, para reducir los niveles de incertidumbre es necesario considerar las visiones compartidas de la comunidad ante los fenómenos que se manifiestan y poder implementar una adecuada planificación y gestión del riesgo. Murgida y Gasparotto (en: Ríos y Natenzon, 2015) sostienen que la vulnerabilidad es la dimensión clave para el análisis del riesgo. Esta categoría, implica explorar las concepciones de los distintos actores sociales, sus relaciones, conocimientos, prácticas y tecnologías que determinarán las consecuencias del peligro y los sentidos que le confieren para enfrentarlo.

Es importante destacar que la percepción social del riesgo se construye a partir del conocimiento cotidiano que las personas elaboran en la interacción social y definen la existencia de múltiples realidades en contraposición a una única realidad objetiva. Por ello, la percepción de los problemas que las personas atraviesan en su vida cotidiana difiere de un individuo a otro y, en consecuencia, las respuestas que emiten ante la ocurrencia de un evento potencialmente dañino también son diversas. Siguiendo a Cabré (2008):

Se sabe que los seres humanos perciben la realidad a través de filtros interiorizados. Estos filtros, de carácter psicológico, antropológico y sociológico, actúan de mediadores entre una supuesta realidad por encima de los individuos y la realidad percibida por los seres humanos que pertenecen a comunidades y grupos determinados y se ubican en contextos históricos y sociales (Cabré, 2008: 19).

En este marco, el estudio de la realidad social implica adentrarse en un complejo campo de saberes que aportan elementos significativos y enriquecen los modos de comprender los problemas que enfrentan las personas en su vida cotidiana. Se reconoce entonces, la importancia de considerar los mecanismos a

través de los cuales las personas perciben su espacio más próximo y las relaciones que establecen con el lugar como centro de significado. No obstante, existen dificultades metodológicas vinculadas con la complejidad de estudiar la subjetividad espacial ya que difiere según cada persona en función de la incidencia de filtros perceptuales mencionados anteriormente y del contexto particular de cada comunidad. Teniendo en cuenta dichos aspectos, se reconoce la importancia que reviste el estudio de la percepción en la formación de las imágenes del mundo real que construyen los diferentes actores sociales. Tal perspectiva permite reconocer las decisiones que los mismos toman en función de las concepciones construidas de la realidad y de los modos de interacción social.

## **2.1. El espacio experimentado, percibido e imaginado como categorías de análisis**

El significado que los diferentes actores sociales le otorgan a los acontecimientos de su vida diaria y los modos de actuar en función de ellos se aborda a través de las categorías de análisis de prácticas espaciales propuestas por Lefebvre, (1974, en: Harvey, 1998). En el estudio de la percepción, las prácticas espaciotemporales y las representaciones del espacio construidas por las personas desempeñan un papel fundamental para analizar las transformaciones espaciales que se producen como consecuencia de los sentidos que las personas les asignan a las experiencias personales, situaciones y problemas que atraviesan.

De este modo, Harvey (1998) propuso una grilla de prácticas espaciales considerando tres dimensiones definidas por Lefebvre (1974, en: Harvey, 1998): prácticas espaciales materiales, representaciones del espacio y espacios de representación. El abordaje de lo experimentado, lo percibido y lo imaginado, según Harvey, permite interpretar las prácticas espaciales desde una dimensión integral que se encuentran relacionadas dialécticamente.

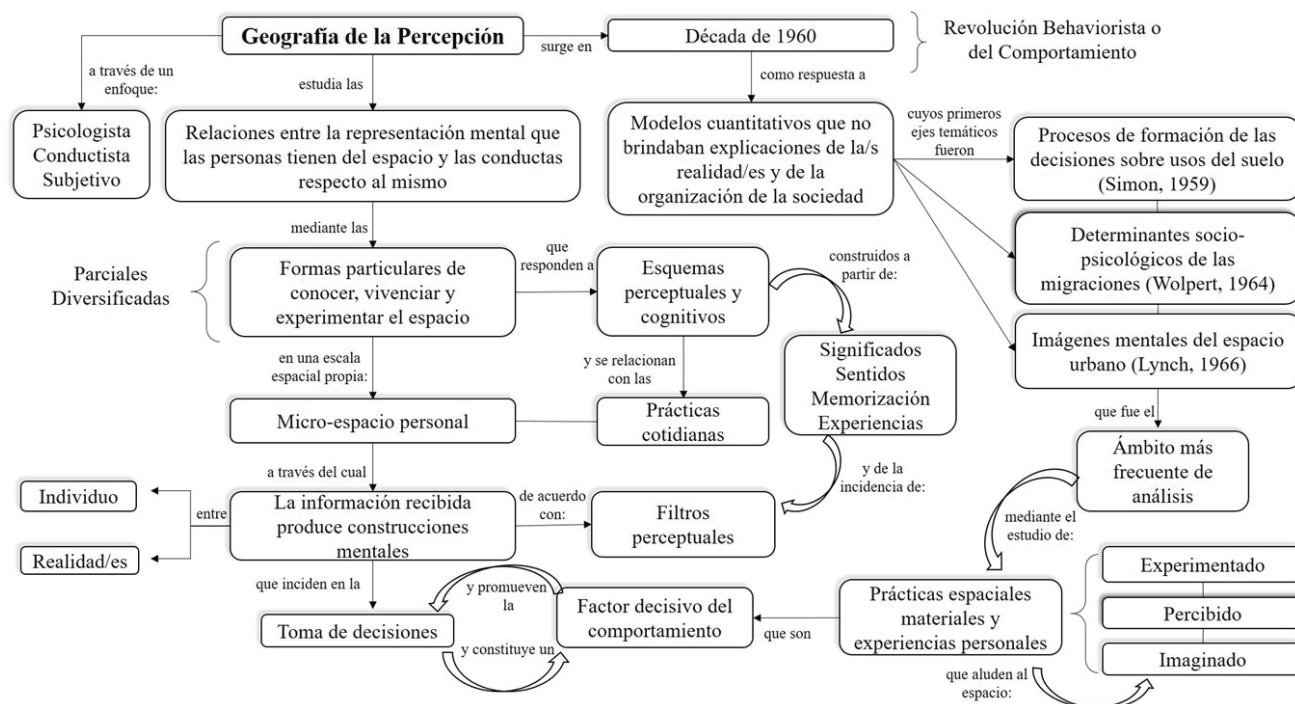
Con respecto a las prácticas espaciales materiales, Harvey (1998) sostiene que “... designan flujos, transferencias e interacciones físicas y materiales que ocurren en el espacio para asegurar la producción y reproducción social” (Harvey, 1998: 244). Así, los flujos de bienes, dinero, personas, fuerzas de trabajo e información generan una apropiación y uso del espacio que configuran una organización social específica. Es dable mencionar aquí, las características de los tipos de asentamiento, usos del suelo, formas de ocupación del espacio y particularidades de la edificación, de los equipamientos y de las actividades desarrolladas. Por ello, el espacio experimentado derivado de las prácticas espaciales materiales adquiere una configuración territorial y una dinámica social específica que debe tenerse en cuenta en la gestión del riesgo.

En relación con las representaciones del espacio, es decir, el espacio percibido Harvey (1998) sostiene que “... abarcan todos los signos y significaciones, códigos y saberes que permiten que esas prácticas materiales se comenten y se comprendan...” (Harvey, 1998: 244). De allí que, se consideran las medidas de distancia social, psicológicas y afectivas que las personas establecen con su entorno considerando las experiencias personales que inciden en la apropiación del espacio. Los vínculos que las personas entablan con el espacio en el que desarrollan sus prácticas responden al sentido y significación que le confieren al mismo a partir de la asistencia a lugares de encuentro que constituyen sitios de preferencia cargados de una especial valoración (Vara Muñoz, 2018). De este modo, los puntos de encuentro tales como, esquinas representativas, sociedades de fomento, paseos públicos, sedes de organizaciones vecinales y asociaciones deportivas, entre otras, adquieren un particular valor por las personas que lo transitan y vivencian.

Teniendo en cuenta los espacios de representación, es decir, el espacio imaginado, es posible afirmar que se trata de construcciones mentales definidas como: “... invenciones mentales (códigos, signos, discursos espaciales, proyectos utópicos, paisajes imaginarios y hasta construcciones materiales... que imaginan nuevos sentidos o nuevas posibilidades de las prácticas espaciales” (Harvey, 1998: 244). En tal sentido, cobran relevancia los sentimientos (atracción, rechazo, distancia, deseo) que las personas vivencian en la relación que establecen con el espacio.



**Figura 2.** Características, categorías y dimensiones de análisis de la Geografía de la Percepción<sup>6</sup>



Fuente: elaborado por A. Mastrandrea sobre la base de: Capel (1981), Gómez Mendoza *et al.* (1982), Bosque Maurel (1983), García Ballesteros (1986), Santos (1990), Harvey (1998), Santarelli y Campos (2002) y Vara Muños (2018).

### 3. La Teoría de las Representaciones Sociales en el reconocimiento de la percepción social del riesgo hídrico

El estudio de las representaciones sociales que las personas construyen a partir de los eventos que atraviesan en su vida diaria constituye una perspectiva de análisis enriquecedora para conocer los principales problemas que las afectan, las estrategias que adoptan para sobrellevarlos y las intervenciones que se producen en función de sus necesidades y posibilidades.

En primer lugar, se aborda la Teoría de las Representaciones Sociales a través de los aportes de Moscovici y Abric (1986), Jodelet (1986), Farr (2003), Araya Umaña (2002), Raiter (2002) y Kornblit (2004). En segundo lugar, dentro de esta teoría se selecciona el concepto de representaciones sociales del riesgo hídrico y las categorías de análisis de: conocimiento del sentido común, concepciones compartidas e imágenes significativas. Asimismo, se consideran las dimensiones: interacción social, asignación de sentidos y construcción de la realidad.

#### 3.1. Aproximación conceptual a la Teoría de las Representaciones Sociales

Con respecto a esta teoría, es importante considerar los diversos modos en los que se constituye el pensamiento social y la importancia otorgada a las construcciones de la realidad que realizan las personas. Moscovici *et al.* (1979) se propusieron comprender la naturaleza del pensamiento social y, para ello, rescataron las ideas provenientes del campo de la Sociología reelaborando el concepto de representaciones colectivas propuesto por Durkheim (2000). Los aportes de la sociología clásica

<sup>6</sup> Nota: La figura destaca el contexto en el cual se desarrolla el enfoque de la Geografía de la Percepción, desde el punto de vista epistemológico en lo que respecta a la evolución del pensamiento geográfico, el cual ha atravesado por un proceso complejo y discontinuo entre las tradiciones teórico-metodológicas positivista-antipositivista. Esta contraposición se evidencia en los modos de abordar el espacio, en las técnicas y métodos empleados y en las formas en las que se estudian los problemas de la realidad social.

durkheimiana constituyeron la base sólida sobre la cual Moscovici elaboró el concepto de representaciones sociales (Araya Umaña, 2002).

Según su teoría, la sociedad desempeña un rol activo en la construcción de las representaciones socialmente compartidas en un proceso de intercambios e interacciones donde la comunicación adquiere un protagonismo central y no es algo que se impone desde fuera del individuo como sostenía Durkheim (Araya Umaña, 2002). En tal sentido, las personas y grupos producen sus conocimientos en relación con otras representaciones ya construidas anteriormente. Así, Moscovici (1979) define el concepto de representaciones sociales como:

... una modalidad particular del conocimiento, cuya función es la elaboración de los comportamientos y la comunicación entre los individuos... La representación es un corpus organizado de conocimientos y una de las actividades psíquicas gracias a las cuales los hombres hacen inteligible la realidad física y social, se integran en un grupo o en una relación cotidiana de intercambios, liberan los poderes de su imaginación (Moscovici, 1979: 17-18).

Teniendo en cuenta dicha concepción Jodelet (1986) remarca que, en las Ciencias Sociales, Moscovici ha sido quien retomó el empleo de esta noción olvidada por Durkheim y especifica que las representaciones sociales hacen referencia a:

La designación de fenómenos múltiples que se observan y estudian a variados niveles de complejidad, individuales y colectivos, psicológicos y sociales... una nueva unidad de enfoque, fecunda para la psicología social, prometedora para otras ciencias sociales... designa una forma de conocimiento específico, el saber del sentido común... una forma de pensamiento social (Jodelet, 1986: 474).

Si bien el concepto aparece inicialmente en Sociología, la Teoría de las Representaciones Sociales va a ser elaborada por la Psicología Social debido a que los fenómenos que se estudian a distintos grados de complejidad (individuales, colectivos, psicológicos y sociales) requieren de una nueva unidad de enfoque. Esta cualidad es claramente expresada por Jodelet (1986) cuando sostiene que:

... las representaciones sociales se presentan bajo formas variadas, más o menos complejas. Imágenes que condensan un conjunto de significados; sistemas de referencia que nos permiten interpretar lo que nos sucede, e incluso, dar un sentido a lo inesperado; categorías que sirven para clasificar las circunstancias, los fenómenos y a los individuos con quienes tenemos algo que ver; teorías que permiten establecer hechos sobre ellos. Y a menudo, cuando se las comprende dentro de la realidad concreta de nuestra vida social, las representaciones sociales son todo ello junto (Jodelet, 1986: 472).

Van Dijk (1999) también ha utilizado el concepto de representaciones socio-mentales en relación con el análisis del discurso y destaca la dimensión cognitiva que hace posible los vínculos entre actores, acciones y los contextos. El autor remarca la importancia que adquiere la interacción social en la construcción individual y colectiva de las representaciones sociales ya que las mismas son compartidas por un grupo o cultura:

Las identidades de la gente en cuanto miembros de grupos sociales las forjan, se las atribuyen y las aprehenden los otros, y son por tanto no sólo sociales, sino también mentales. Los contextos son constructos mentales (modelos) porque representan lo que los usuarios del lenguaje construyen como relevante en la situación social (Van Dijk, 1999: 26).

Es importante mencionar que las concepciones se construyen a través de diferentes mecanismos de interacción social en función de: la información recibida, conceptos sobreimpuestos, conocimiento elaborado y reconstruido, aportes de los medios de comunicación, instituciones, organizaciones, grupos de pertenencia, vínculos con el lugar y experiencias de la vida cotidiana, entre otros aspectos. De allí que se concibe el sentido común -integrado por un conjunto de creencias, experiencias, vivencias y concepciones de la realidad- como un cuerpo de conocimientos compartidos por los miembros de grupos, instituciones y organizaciones sociales y gubernamentales.

Además del conocimiento del sentido común existen variables externas que influyen en los modos en los que se elaboran las imágenes significativas de las experiencias atravesadas por las personas como,

por ejemplo, la información recibida (Jodelet, 2008). Cada individuo conoce la realidad a través de diversos modos de interacción social y, por ello, es relevante analizar la función que desempeña el discurso de los actores sociales involucrados en la interpretación de los problemas que enfrentan.

Siguiendo a Moscovici *et al.* (1986) las representaciones sociales expresan la forma en la que las personas piensan y comprenden su mundo. En el pensamiento no normalizado abundan ideologías, imágenes, palabras y razonamientos influenciados por estereotipos o creencias anteriores y tradiciones culturales que han sido modificados por cada individuo en un proceso de interacción social. En dicha construcción de la realidad la interacción social que se manifiesta en la vida cotidiana de las personas juega un papel fundamental y su análisis contribuye a comprender las visiones individuales y grupales que los habitantes tienen del riesgo hídrico.

Un aporte especial es el de Raiter (2002) cuando define las representaciones sociales como las imágenes (mentales) que construyen los medios de difusión sobre los temas, eventos, acciones o procesos que conforman la agenda<sup>7</sup> pública. La actividad cognitiva de cada persona permite construir imágenes significativas que son representaciones de la realidad “... es decir realiza una operación mental sobre lo recibido y almacena el resultado de esa operación” (Raiter, 2002: 10). Es así que los diferentes actores construyen su propia representación del riesgo hídrico que depende de la interacción entre los procesos cognitivos y las imágenes ya existentes que han sido adquiridas y transmitidas a través de la interacción, motivo por el cual, dejan de ser individuales para convertirse en sociales.

Araya Umaña (2002) sostiene que existen diversas matrices socio-estructurales que influyen en las formas de percepción a través de las cuales se construyen las representaciones sociales. Las mismas son: el medio cultural en el que viven las personas, el lugar que ocupan en la estructura social, las experiencias cotidianas, los intereses, las informaciones, los conocimientos y los modelos de pensamiento recibidos por la tradición, la educación y la comunicación. Es así que los diferentes grupos sociales, en función de dichas matrices, construyen imágenes significativas y concepciones de las cosas, fenómenos o acciones. Es decir que las representaciones no son neutras sino elaboradas de acuerdo a ciertos objetivos, planificaciones y decisiones (Raiter, 2002).

Sin embargo, es preciso mencionar que las representaciones sociales se estructuran en un conjunto de elementos -más o menos estables- que responden a una organización en un núcleo central y elementos periféricos<sup>8</sup> (Moscovici, 1961 en: Petracci y Kornblit, 2007). El núcleo central contiene los elementos cognitivos responsables de la estabilidad y son innegociables entre los que se encuentran, por ejemplo, las creencias u opiniones del riesgo hídrico ligadas a la memoria colectiva y a la historia del grupo. Mientras que, el sistema periférico está compuesto por otros elementos de la representación que son responsables de la movilidad y desacuerdos entre los individuos, por ejemplo, las diferencias de opiniones respecto de las inundaciones o sequías que pueden estar influenciadas por el contexto.

Por lo mencionado en los párrafos precedentes, es posible afirmar que el estudio de las representaciones sociales del riesgo hídrico constituye una valiosa perspectiva para reconocer los modos en los que se elabora el conocimiento del sentido común y, de este modo, identificar las prácticas espaciales que desarrollan. Siguiendo a Araya Umaña (2002) nos aproxima a la visión del mundo que las personas o grupos tienen debido a que el conocimiento del sentido común es el que la gente utiliza para tomar posición ante los distintos problemas sociales. A continuación, se presenta un esquema en el que se sintetizan las características de la Teoría de las Representaciones Sociales, los conceptos, categorías y dimensiones de análisis consideradas (figura 3).

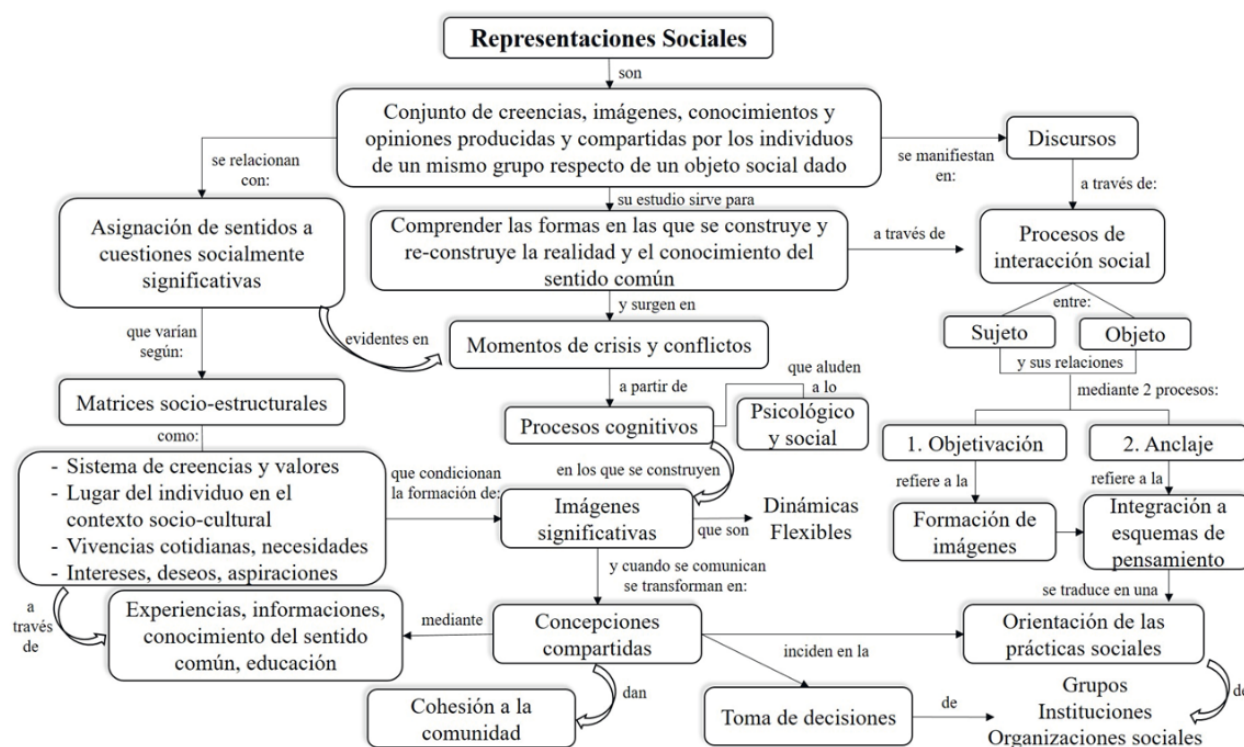
---

<sup>7</sup> Raiter (2002) denomina *agenda* a las representaciones activas en un momento dado.

<sup>8</sup> La Teoría del núcleo central fue propuesta originalmente por Abric (1976) derivada de las bases sentadas por Moscovici en 1961. Los postulados de dicha teoría fueron corroborados por diversos investigadores y su aporte radica en la posibilidad de realizar estudios comparativos de las representaciones de diferentes grupos sociales (Petracci y Kornblit, 2007).



**Figura 3.** Características, categorías y dimensiones de análisis de la Teoría de las Representaciones Sociales<sup>9</sup>



Fuente: elaborado por A. Mastrandrea sobre la base de: Jodelet (1986 y 2008), Moscovici *et al.* (1986, 1993), Raiter *et al.* (2002), Petracci *et al.* (2007) y De Souza Minayo (2009).

#### 4. Problemáticas y desafíos de la planificación y gestión del riesgo

La adopción de medidas tendientes a reducir los niveles de incertidumbre en relación con la probabilidad de ocurrencia de diferentes riesgos ambientales debiera constituir un eje prioritario en el diseño y elaboración de planes de gestión y ordenamiento territorial. La inclusión de estrategias que involucren una cultura preventiva y participativa se convierte en una necesidad fundamental en la gestión del riesgo hídrico. Por ello, es preciso avanzar en la definición de propuestas que consideren el conocimiento de la dinámica ambiental junto con la percepción y las representaciones sociales del riesgo hídrico que diferentes actores sociales construyen.

El concepto de cultura preventiva hace referencia a la capacidad de anticipación que tienen los diferentes actores sociales ante la ocurrencia de un evento potencialmente peligroso en una comunidad que es vulnerable a él (Medina y Vigo, 2020). Cada cultura tiene sus características particulares en cuanto a modos de interacción social y estrategias para enfrentar eventos que signifiquen una amenaza de acuerdo con un sistema ambiental históricamente construido. De allí que, es importante conocer el contexto socio-cultural para comprender las formas de actuación que intervienen tanto en la prevención como en la amplificación del riesgo hídrico.

Una de las primeras iniciativas que debiera considerarse para construir una cultura preventiva es la reorganización institucional a escala local y especialmente a una escala local en la que se reconozcan las peligrosidades y vulnerabilidades y se elaboren diagnósticos territoriales de carácter ambiental. Para llevar a cabo dichos diagnósticos es importante que todos los actores sociales se involucren en las tareas

<sup>9</sup> Nota: En el esquema se define el concepto de representaciones sociales y el principal propósito de su abordaje en el contexto de las Ciencias Sociales. Se destacan los componentes de dicha teoría, así como sus características y el modo en el que se construyen las concepciones que las personas elaboran en los procesos de interacción social.

a desempeñar a través de la participación individual o como miembros de grupos (barriales, religiosos, científico-técnicos, gubernamentales y no gubernamentales) con el fin de forjar vínculos de pertenencia, generar conciencia ambiental sobre los eventos que los afectan en su vida cotidiana y promover la responsabilidad individual y colectiva acerca de las intervenciones que, de modo consciente o inconsciente, realizan en su medio más próximo.

Así, dentro de la cultura participativa, es posible incluir la gobernanza del agua que según la Organización para la Cooperación y el Desarrollo Económico (OCDE, 2015) se refiere a la función de las instituciones y las relaciones entre las organizaciones y los grupos sociales implicados en la toma de decisiones, de forma transversal (entre sectores y espacios urbanos y rurales) y vertical (a diferentes escalas). La gobernanza incluye entonces las formas de organización de los múltiples actores sociales para gestionar y promover el desarrollo de sus territorios (Sili, 2019). En tal sentido, es fundamental generar una cultura de prevención que involucre una institucionalidad local y que considere la percepción y las representaciones sociales del riesgo hídrico de todos los actores sociales involucrados en el área de estudio.

La inclusión de una cultura preventiva y participativa se opone al tradicional modelo de gestión del riesgo hídrico proveniente del campo de las Ciencias Naturales (Natenzon *et al.*, 2015). Entre las características de este modelo se pueden mencionar que es reducido -en lo que respecta a la mera transmisión de conocimientos-; tecno-céntrico -centrado únicamente en el conocimiento científico-técnico sin considerar los saberes comunitarios-; unidireccional -parte de especialistas hacia la población sin intercambio de saberes-; comportamental -con recomendaciones de actuación y autoprotección-; factual -toma el riesgo como un hecho sin abordar las causas socio-históricas que reproducen situaciones de vulnerabilidad- y acrítico -trata el riesgo como una situación inevitable sin posibilidades de reflexionar sobre sus causas y modelos de desarrollo territorial- (Warner, en: Sulaiman y Jacobi, 2018). De este modo, la ausencia de una concepción del riesgo como constructo multidimensional (considerando la perspectiva espacial, temporal e interescalar en la que se produce) excluye del análisis determinadas cuestiones que caracterizan la relación sociedad-naturaleza.

Por lo tanto, es necesario implementar propuestas continuas de gestión del riesgo que se focalicen en la participación de todos los actores sociales con el propósito de compartir saberes y responsabilidades. En este marco, resulta primordial la conformación de espacios de diálogo y de aprendizaje colectivo para pensar en una concepción del riesgo hídrico centrada no sólo en la peligrosidad sino también en la vulnerabilidad. El intercambio social permite conformar espacios de reflexión, reconocer problemas prioritarios y necesidades, reforzar identidades colectivas, prevenir potenciales problemas ambientales, generar confianza en el proceso de toma de decisiones y promover una postura crítica y proactiva entre los actores sociales.

Teniendo en cuenta los aspectos mencionados anteriormente se destaca la riqueza de los planes de ordenamiento territorial que se focalicen en la gestión del riesgo. Es importante incluir en dichos instrumentos de planificación una serie de propuestas a desarrollar en el largo, mediano y corto plazo a partir de programas, proyectos y actividades correspondientes (Gómez Orea, 2007, Arzeno *et al.*, 2020). Así, la gestión del riesgo hídrico es un proceso continuo en el que las acciones a desarrollar se manifiestan tanto en la etapa previa a la ocurrencia del evento peligroso como en el momento concreto en el que se manifiesta y en el período posterior a él. Así, es responsabilidad de los actores sociales involucrados establecer las estrategias para la gestión integral del riesgo que abarque el antes, el durante y el después ante un evento peligroso, así como también, aquellas que deban aplicarse en forma simultánea.

## Reflexiones finales

La consideración de los conceptos, categorías y dimensiones de análisis del riesgo hídrico desde una perspectiva multidisciplinar implica tener en cuenta la complejidad que reviste conocer y analizar la construcción social de escenarios de creciente criticidad ambiental. El aporte de la Teoría Social del

Riesgo, de la Geografía de la Percepción y de la Teoría de las Representaciones Sociales permite reconocer el significado que diferentes actores sociales le otorgan a la ocurrencia de los problemas ambientales que vivencian y sus implicancias según las prácticas espaciales materiales desarrolladas.

Así, la Teoría Social del Riesgo colabora en el conocimiento de los procesos que configuran las peligrosidades y vulnerabilidades y las condiciones que caracterizan la exposición e incertidumbre ante eventos de exceso y déficit hídrico. En tal sentido, se considera la importancia de abordar la relación sociedad-naturaleza y la percepción de los problemas ambientales que los actores sociales conciben como significativos, conocimiento que permite analizar la construcción social escenarios de riesgo. Se reconoce entonces, la importancia de concebir el riesgo como una característica intrínseca de las sociedades que deriva de un proceso sociohistórico complejo construido a través del tiempo y que involucra, desde un enfoque espaciotemporal, tanto los cambios como las permanencias.

Por su parte, la Geografía de la Percepción y la Teoría de las Representaciones Sociales permiten identificar la percepción social del riesgo hídrico y los principales problemas considerados significativos por diferentes actores sociales, conocimiento fundamental que interviene en las prácticas espaciales que desarrollan. La consideración de las categorías de análisis propuestas por Lefebvre (1974 en: Harvey, 1998) del espacio experimentado, percibido e imaginado por los habitantes incorpora una dimensión psicológica esencial para reconocer sus motivaciones, decisiones y acciones. En este contexto, dichas teorías proporcionan los fundamentos teórico-metodológicos necesarios para reconocer los modos en los que las personas conciben el espacio en términos de vivencias, valoraciones y experiencias construidas individual y colectivamente. En tal sentido, analizar la configuración de escenarios de criticidad ambiental a través de la percepción que las personas tienen y de las representaciones sociales que construyen, especialmente de las inundaciones y sequías, permite reconocer los significados de las prácticas espaciales llevadas a cabo por estos junto con las necesidades y demandas de la población.

En síntesis, este abordaje teórico-conceptual constituye un enfoque enriquecedor para incrementar los niveles de conocimiento sobre el riesgo y orientar las acciones que, desde diferentes ámbitos (científico-técnicos, gubernamentales y no gubernamentales, organizaciones sociales), debieran implementarse en la gestión integral y participativa del riesgo hídrico. La consideración de diferentes variables, correspondientes tanto al medio natural como sociocultural, ofrece la posibilidad de realizar un análisis espacial mediante la triangulación de métodos, técnicas y fuentes de datos que integre diversas perspectivas en función de la articulación de los escenarios reales y los percibidos. Dichas perspectivas permiten diseñar, desde un enfoque multiescalar e interdisciplinar, propuestas tendientes a la planificación y gestión continua y participativa del riesgo hídrico.

## Referencias

- Araya Umaña, S. (2002). “Las representaciones sociales: Ejes teóricos para su discusión”. *Cuadernos de Ciencias Sociales*. 127. 1-79. Recuperado de: <http://www.flacso.org.uy/fileadmin/documentos/FLACSO/cuaderno127.pdf>
- Arzeno, M. B., Muñecas, L., y Zanotti, A. S. (2020). “Ordenamiento territorial en cuestión: orden y contraespacio en el norte de Misiones, Argentina”. *Cuadernos de Geografía: Revista Colombiana de Geografía*. 29(1) 51-68. Recuperado de: <https://revistas.unal.edu.co/index.php/rcg/article/view/73524>
- Arteaga, A. (2020). *Desarrollo metodológico para la evaluación de la gestión del riesgo hídrico. Ejercicio de aplicación sobre el caso del Arroyo del Gato en la región Gran La Plata, Argentina*. (Tesis inédita de doctorado). Facultad de Arquitectura y Urbanismo. Universidad Nacional de La Plata.
- Beck, U. (1993). “De la sociedad industrial a la sociedad del riesgo. Cuestiones de supervivencia, estructural social e ilustración ecológica”. *Revista de Occidente*. (150). 19-40. Recuperado de: <https://dialnet.unirioja.es/servlet/articulo?codigo=222621>

- Bertoncello, R. y Alesandri, C. (2003). *Procesos Territoriales en Argentina y Brasil*. Instituto de Geografía. Facultad de Filosofía y Letras. UBA. Buenos Aires.
- Bosque Maurel, J. (1983). "Geografía, Historia y Geografía Histórica". *Estudios Geográficos*. 44(172). 317-338. Recuperado de: <https://dialnet.unirioja.es/servlet/articulo?codigo=7115>
- Cabré, M. T. (2008). "El principio de poliedricidad: la articulación de lo discursivo, lo cognitivo y lo lingüístico en Terminología (I)". *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE)*. (16). 9-36. Recuperado de: <https://www.redalyc.org/pdf/2870/287024065003.pdf>
- Capel, H. (1981). *Filosofía y Ciencia en la Geografía contemporánea. Una introducción a la Geografía*. Barcanova. Barcelona.
- Carballo, C. T. (2001). "Inundación, degradación urbana y construcción social del riesgo". *Revista Geográfica*. 129. 95-110. Recuperado de: <https://search.proquest.com/openview/2c7ccf52fa1607d57a3a968e42b06b88/1?pqorigsite=gscholar&cbl=35027>
- Carballo, C. T. y Goldberg, S. (2014). *Comunidad e información ambiental del riesgo: las inundaciones y el río Luján*. Editorial Dunken. Luján.
- Cardona, O. D. (1993). "Evaluación de la amenaza, la vulnerabilidad y el riesgo. Elementos para el ordenamiento y la planeación del desarrollo". En: Maskrey, A. (Comp.). *Los desastres no son naturales* (45-65). La Red. Bogotá.
- De Souza Minayo, M. C. (2009). *La artesanía de la investigación cualitativa*. Lugar Editorial. Buenos Aires.
- Farr, R. M. (2003). "De las representaciones colectivas a las representaciones sociales: ida y vuelta". En J. A. Castorina (Comp.). *Representaciones sociales. Problemas teóricos y conocimientos infantiles* (153-175). Editorial Ge-disa. Barcelona.
- García Acosta, V. (2005). "El riesgo como construcción social y la construcción social de riesgos". *Desacatos*. (19). 11-24. Recuperado de: <https://www.redalyc.org/pdf/139/13901902.pdf>
- García Ballesteros, A. (1986). *Teoría y práctica de la Geografía*. Alhambra Universidad. Madrid.
- Gómez Mendoza, J.; Muños Giménez, J. y Ortega Canrero, N. (1982). *El pensamiento geográfico*. Alianza Universidad. Madrid.
- Gómez Orea, D. G. (2007). *Ordenación territorial*. Mundi-Prensa Libros.
- Harvey, D. (1998). *La condición de Posmodernidad*. Amorrortu. Buenos Aires.
- Herzer, H. M. (2011). "Construcción del riesgo, desastre y gestión ambiental urbana: Perspectivas en debate". *Revista virtual REDESMA*, 5(51) 51-61. Recuperado de: <http://www.revistasbolivianas.org.bo/pdf/rvr/v5n2/a08.pdf>
- Jodelet, D. (1986). "La representación social: fenómenos, concepto y teoría". En S. Moscovici y M. Hewstone (Comp.), *Psicología social II* (469-494). Paidós. Barcelona-Buenos Aires.
- Jodelet, D. (2008). "El movimiento de retorno al sujeto y el enfoque de las representaciones sociales". *Cultura y representaciones sociales*. 3(5). 32-63. Recuperado de: <http://www.scielo.org.mx/pdf/crs/v3n5/v3n5a2.pdf>
- Kornblit, A. L. (2004). *Metodologías cualitativas: modelos y procedimientos de análisis*. Biblos. Buenos Aires.



Lavell (1996). “Degradación ambiental, riesgo y desastre urbano. Problemas y conceptos: hacia la definición de una agenda de investigación”. En M.A. Fernández (Comp.). *Ciudades en riesgo* (12-42). Usaid. Colombia.

Lavell, A. (1999). “Un encuentro con la verdad: los desastres en América Latina durante 1998”. *Anuario Político y Social de América Latina*. 2. 1-19. Recuperado de: <https://www.desenredando.org/public/articulos/1999/ideal1998/LosDesastresEnAmericaLatinaDurante1998-1.0.0.pdf>

Lavell, A., Mansilla, E., Maskrey, A., y Ramírez, F. (2020). “La construcción social de la pandemia COVID-19: desastre, acumulación de riesgos y políticas públicas”. *La Red (Red de Estudios Sociales y Prevención de Desastres en América Latina)*. 1-16. Recuperado de: <https://www.gub.uy/sistema-nacional-emergencias/sites/sistema-nacional>

Lindón, A. y Hiernaux, D. (2006). *Tratado de Geografía Humana*. Universidad Autónoma Metropolitana: Antrophos. México.

Maskrey, A. (1993). *Los desastres no son naturales*. La Red. Colombia.

Mastrandrea, A. y Pérez, M. I. (2020). “Representaciones sociales del riesgo hídrico en el sector inferior de la cuenca del arroyo Napostá Grande: un abordaje histórico-ambiental (1828-2018)”. *Investigaciones Geográficas*. (74). 197-222. Recuperado de: [https://rua.ua.es/dspace/bitstream/10045/109944/6/Investigaciones\\_Geograficas\\_74\\_10.pdf](https://rua.ua.es/dspace/bitstream/10045/109944/6/Investigaciones_Geograficas_74_10.pdf)

Medina, S. S. y Vigo, M. (2020). “Evaluación ambiental estratégica y participación social en la gestión del riesgo hídrico”. *Margen: revista de trabajo social y Ciencias Sociales*. (97)4. 1-8. Recuperado de: <https://www.margen.org/suscri/margen97/Sosa-Vigo-97.pdf>

Moscovici, S. (1979). *Introducción a la psicología social*. Paidós. Barcelona.

Moscovici, S. y Abrieu, J. C. (1986). *Psicología social II. Pensamiento y vida social*. Paidós. Barcelona-Buenos Aires.

Natenzon, C. (1995). “Catástrofes naturales, riesgo e incertidumbre”. *FLACSO. Serie Documentos e Informes de Investigación*. (197). 1-19. Recuperado de: <http://pirna.com.ar/files/pirna/PUB-Natenzon-Catastrofes-naturales-riesgo-e-incertidumbre.pdf>

Natenzon, C., Marlenko, N., González, S., Ríos, D., Murgida, A., Meconi, G y Calvo, A. (2003). “Las dimensiones del riesgo en ámbitos urbanos. Catástrofes en el Área Metropolitana de Buenos Aires”. En R. Bertoncello (Comp.), *Procesos territoriales en Argentina y Brasil* (255-276). Instituto de Geografía, Facultad de Filosofía y Letras, UBA. Buenos Aires.

Natenzon, C. y Ríos, D. (2015). *Riesgos, catástrofes y vulnerabilidades. Aporte desde la Geografía y otras ciencias sociales para casos argentinos*. Ediciones Imago Mundi. Buenos Aires.

Organización para la Cooperación y el Desarrollo Económico (OCDE, 2015). *Marco de indicadores de la gobernanza del agua de la OCDE*. Recuperado de: <http://dx.doi.org/10.1787/9789264292659-en>

Petracci, M., y Kornblit, A. (2007). “Representaciones sociales: una teoría metodológica-mente pluralista”. En: A. Kornblit (Coord.). *Metodologías cualitativas en Ciencias Sociales* (91-112). Buenos Aires: Editorial Biblos.

Raiter, A. (2002). *Representaciones sociales*. Editorial Universitaria de Buenos Aires. Buenos Aires.

Rees, J. (2006). *Riesgos y gestión integrada de recursos hídricos*. Global Water Partnership. Estocolmo.

- Ríos, D. y Natenzon C. (2015). “Una revisión sobre catástrofes, riesgo y Ciencias Sociales”. En C. Natenzon y D. Ríos (Eds.), *Riesgos, catástrofes y vulnerabilidades. Aportes desde la Geografía y otras Ciencias Sociales para casos argentinos* (1-27). Imago Mundi. Buenos Aires.
- Santarelli, S. y Campos, M. (2002). *Corrientes epistemológicas, Metodología y Práctica en Geografía. Propuestas de estudio en el espacio local*. EdiUNS. Bahía Blanca.
- Santos, M. (1990). *Por una Geografía nueva*. Espasa- Calpe. Madrid.
- Santos, M. (1996). *Metamorfosis del espacio habitado*. Oikos-Tau. Barcelona.
- Saurí Pujol, D. (2006). “Nota necrológica. Gilbert F. White (Chicago, 1911-Boulder, Colorado, 2006)”. *Documents d'anàlisi geogràfica*. (48). 13-14. Recuperado de: <https://ddd.uab.cat/pub/dag/02121573n48/02121573n48p13.pdf>
- Sili, M. (2019). “Crisis y renovación de la planificación y la gobernanza de los territorios”. *Perspectives on Rural Development*. (2). 17-35. Recuperado de: <http://sibaese.unisalento.it/index.php/prd/article/view/20182/17167>
- Sulaiman, S. N. y Jacobi, P. (Coord.). (2018). *Melhor prevenir. Olhares e saberes para a redução de risco de desastre*. IEE-USP. São Paulo.
- Van Dijk, T. (1999). *El análisis crítico del discurso*. Anthropos. Barcelona.
- Vara Muñoz, J. L. (2018). “Transformación urbana del Área Metropolitana Este de Madrid: Ejemplos de su percepción en el municipio de Coslada”. *Revista Universitaria de Geografía*. 27(1). 49-72. Recuperado de: <http://bibliotecadigital.uns.edu.ar/pdf/reuge/v27n1/v27n1a04.pdf>
- Vegas Meléndez, H. (2017). “Políticas públicas y gobernanza: articulación para una gestión pública local autónoma”. *Revista Latinoamericana*. (48). 155-172. Recuperado de: <https://scielo.conicyt.cl/pdf/polis/v16n48/0718-6568-polis-16-48-00155.pdf>
- Warner, J. (2018). “Laços invisíveis”. En: S. N. Sulaiman y P. Jacobi (2018) (Coord.) (2018). *Melhor prevenir. Olhares e saberes para a redução de risco de desastre* (12-37). IEE-USP. São Paulo.
- Wilches-Chaux, G. (1993). “La vulnerabilidad global”. En: A. Maskrey (Comp.). (1993). *Los desastres no son naturales* (9-50). Red de Estudios Sociales en Prevención de Desastres en América Latina. Bogotá.

## DE LA POSIBLE RENOVACIÓN DEL PENSAMIENTO CRÍTICO ANTE LA FILOSOFÍA EMANCIPADORA DE J. RANCIÈRE

*ON THE POSSIBLE RENEWAL OF CRITICAL THINKING IN THE FACE OF THE EMANCIPATORY PHILOSOPHY OF J. RANCIÈRE*

*SOBRE A POSSÍVEL RENOVAÇÃO DO PENSAMENTO CRÍTICO FACE À FILOSOFIA EMANCIPATÓRIA DE J. RANCIÈRE*

**Karla Castillo Villapudua**

*(Universidad Autónoma de Baja California)*

*castillo.karla@uabc.edu.mx*

Recibido: 24/02/2021

Aprobado: 04/04/2021

### RESUMEN

El siguiente artículo pretende rastrear algunas de las principales objeciones en contra de aquello que Rancière llama teoría crítica tradicional. Esto con el objetivo de explorar la posibilidad de redefinir las coordenadas del pensamiento crítico. Pues, para Rancière, la forma de operar de aquel continúa realizándose a través de formas como la lógica de la denuncia, la concientización de clase y la insistencia en la oposición entre apariencia y realidad. Finalmente, reflexionamos sobre este otro pensamiento crítico enfocado en investigar las ideas que perfilarían y sostendrían un mundo igualitario más allá de la lógica de la denuncia.

Palabras clave: Rancière. crítica social. pensamiento crítico. emancipación. condiciones de posibilidad.

### ABSTRACT

The following article aims to trace some of the main objections against what Rancière calls traditional critical theory. This approach is to explore the possibility of redefining the coordinates of critical thinking. Well, for Rancière, this way of operating continues to be carried out through ways such as the logic of the complaint, class awareness, and insistence on the opposition between appearance and reality. Finally, we reflect on this other critical thinking focused on investigating the ideas that would shape and sustain an egalitarian world beyond the logic of the complaint.

Keywords: Rancière. social critical. critical thinking. emancipation. possibility conditions.

### RESUMO

O artigo a seguir tenta rastrear algumas das principais objeções contra o que Rancière chama de teoria crítica tradicional. Isso com o objetivo de explorar a possibilidade de redefinir as coordenadas do pensamento crítico. Pois bem, para Rancière, a forma como opera continua a ser realizada através de formas como a lógica da denúncia, a consciência de classe e a



insistência na oposição entre aparência e realidade. Finalmente, refletimos sobre este outro pensamento crítico focado em investigar as ideias que irão moldar e apoiar um mundo igualitário, mas também a lógica da reclamação.

Palavras-chave: Rancière. crítica social. pensamento crítico. emancipação. condições de possibilidade.

## Introducción

En años recientes, el pensamiento del filósofo Jacques Rancière ha ocupado un lugar importante en el panorama filosófico contemporáneo de habla francesa. Por poco más de tres décadas, sus ideas sobre estética y filosofía política han ejercido una notable influencia en el ámbito académico. Textos como *El espectador emancipado*, *El desacuerdo*, o *El maestro ignorante*, entre otros, han sido traducidos a numerosas lenguas, convirtiéndose así en referencia obligada a nivel internacional. Sin embargo, a diferencia de lo que ocurre en otros países, incluyendo aquellos de habla hispana como España, Argentina y Chile, en México apenas y se le ha prestado atención a su obra.

El proyecto filosófico de J. Rancière, es una de esas prácticas que se deslindan del lugar común de los regímenes epistémicos dominantes de una época, para aventurarse en una búsqueda intelectual genuina y propia. De ahí, su fascinación por los desacuerdos y su firmeza en mantener la mirada crítica, más allá de lo que pensaban sus maestros y colegas. Esto, quizá, porque dicha sospecha le permitió ampliar su pensamiento, no con el afán de ser diferente, sino con la ambición de lograr dar un paso más allá, y no conformarse con diagnósticos fatalistas. Por ello, podríamos pensar al joven Rancière como ese alumno incómodo que no se toma al pie de la letra las enseñanzas de su profesor- en este caso Louis Althusser- y se atreve a navegar por otros senderos. Y bien, ¿en qué consistió dicho desacuerdo? Pues bien, como se sabe durante la gran efervescencia política e intelectual del 68, las tesis de Althusser sobre el deber de los intelectuales orgánicos para concientizar a las masas embrutecidas de los mecanismos de poder gozaban de gran popularidad académica, por tal motivo, eran de las ideas más estudiadas y seguidas.

De manera mínima, podemos recordar que Althusser abogaba por la vía de la concientización para lograr la emancipación. Sin embargo, a raíz de los acontecimientos del mayo francés, el joven Rancière se decepciona y empieza a cuestionar duramente dichos planteamientos, pues detectaba ciertas jerarquías y distinciones entre aquellos que tienen el privilegio de la voz, la palabra y el conocimiento, y los que de manera definitiva no tienen acceso a ello por su condición de clase. Por ello, bajo esa sospecha el filósofo inicia su exploración con la revisión de los archivos obreros del siglo XIX, y descubre que en ocasiones los mismos trabajadores toman la voz sin la dependencia de una figura externa. A partir de esto, el autor no ha dejado de producir una gran cantidad de escritos para desarrollar sus ideas partiendo de la emancipación, razón por la cual, podemos asumirlo como un pensador que resiste ante los discursos desoladores que no presentan alternativa alguna.

Ahora bien, dentro del pensamiento político de Jacques Rancière, su conferencia *Las Desventuras del pensamiento Crítico* (2010) es una obra clave para acercarse a su objeción a la tradición crítica social<sup>1</sup>, ya que ahí, el filósofo francés rastrea una inversión de sus “verdaderos fines” y describe cómo fue territorializada por el capitalismo. De esta manera, el autor sugiere que es importante hacer una crítica de la crítica que implique otra forma de poner en práctica el pensamiento crítico.

Aunado a lo anterior, Rancière (2009) en su artículo *Sobre la importancia de la Teoría Crítica para los Movimientos Sociales Actuales*, plantea que lo que antes era considerado como un potente dispositivo para la transformación de la realidad social, ahora ha sufrido una inversión, que consiste entre otros

---

<sup>1</sup> En este punto, resulta oportuno aclarar que en este texto entenderemos por “teoría crítica tradicional” a la crítica social que apunta los aspectos negativos de los órdenes sociales existentes. (Corcuff, 2018: 63). Señalando que Rancière utiliza el término como “tradición de la crítica social” o “paradigma crítico”, sin mencionar particularmente a qué variante se refiere. Sin embargo, a lo largo de su obra podemos observar que Rancière está en desacuerdo con el marxismo estructuralista de su profesor Althusser.

asuntos, en mercantilizar toda protesta convirtiéndola en espectáculo y, además, en propagar sólo diagnósticos decadentes sin solución alguna.

De ahí que para Rancière (2010) esta forma de hacer crítica radicaba básicamente en tres ideas claves: la denuncia de las mitologías de la mercancía, de las ilusiones de la sociedad de consumo y del imperio del espectáculo. No obstante, es evidente que desde la lectura de nuestro autor, esas premisas sufrieron una modificación radical y en la actualidad se han transformado en un saber desesperanzado, que ha dejado de lado formular proyectos de emancipación.

Pero algo, es verdad, ha cambiado. Todavía ayer esos procedimientos se proponían suscitar formas de conciencia y energías encaminadas hacia un proceso de emancipación. Ahora están, ya sea enteramente desconectadas de ese horizonte de emancipación, o bien claramente vueltas contra su sueño (p.36).

De lo anterior, Rancière (2009) parece dar indicios de la pertinencia de renovar el pensamiento crítico, pues en textos como los anteriores expresa la necesidad de apostar por un pensamiento que no se fundamente en la culpabilidad e ignorancia y recupere en cambio, la creación de un proyecto de emancipación a través de la igualdad de las inteligencias.

Sin embargo, la posibilidad de una renovación del pensamiento crítico no es un aspecto totalmente desarrollado en la obra de Rancière. Por lo que el objetivo principal de este trabajo consiste en reconocer algunos elementos conceptuales, para posteriormente proponer si es posible llevar a cabo una renovación del pensamiento crítico en el siglo XXI.

El presente trabajo, entonces, surge de las siguientes preguntas: ¿Cómo renovar el pensamiento crítico, dominado por la lógica de la denuncia y el presunto desenmascaramiento, a partir de la filosofía de Rancière? ¿Cuáles serían las condiciones de posibilidad de este pensamiento crítico?

Así pues, el objetivo de esta investigación es proponer una posible renovación del pensamiento crítico a través de la filosofía de Rancière. En un primer eje revisamos la crítica de la teoría crítica tradicional en la conferencia *Las Desventuras del Pensamiento Crítico*. En un segundo momento presentamos algunos de los indicios que señala Rancière para renovar el pensamiento crítico desde las condiciones de posibilidad a través del concepto de emancipación. Por último, reflexionamos sobre las posibles consecuencias de practicar la renovación del pensamiento crítico.

## **La crítica de la crítica: el cambio de trayectoria**

En su conferencia titulada *Las Desventuras del Pensamiento Crítico*, publicada en el libro *El Espectador Emancipado* (2010) Rancière plantea que pese a la posición de algunos teóricos que celebran el supuesto agotamiento de la teoría crítica, los insumos conceptuales aún siguen vigentes, pero estos han sufrido una inversión:

Me gustaría mostrar, a la inversa, que los conceptos y procedimientos de la tradición crítica no están para nada en desuso. Todavía funcionan muy bien, incluso en el discurso de aquellos que declaran su caducidad. Pero su uso presente testimonia una total inversión de su orientación y sus supuestos fines (p.29).

Tomando en cuenta dicha reflexión podemos afirmar que el teórico francés no da por agotada la llamada crítica social, sino que, bajo una lúcida lectura acorde al tiempo actual, rastrea las modificaciones registradas en las últimas décadas del siglo XX. Pues definitivamente, la misma naturaleza dialéctica de la teoría enquistada en las diversas transformaciones históricas tiende a realizar mutaciones y no permanece inamovible.

Por tal motivo, Rancière (2010: 29) supone: “De modo que es necesario tomar en cuenta la persistencia de un modelo de interpretación y la inversión de su sentido si queremos emprender una verdadera crítica de la crítica”, A partir de esto, el filósofo francés emprende la tarea de elegir algunos ejemplos del campo

de las artes, la política y la teoría para mostrar este cambio de trayectoria. A continuación, abordaremos algunos ejemplos de manera breve, con la finalidad de desarrollar nuestra hipótesis principal de trabajo.

De inicio, Rancière (2010) nos presenta dos fotografías: “Sin título, 2005, de Josephine Meckseper” y “Bringing the War Home: Ballons, fotomontaje 1967-1972, de Martha Rosler”. Estas imágenes tienen como finalidad concientizar a los espectadores sobre la realidad oculta bajo la que funciona el sistema de dominación, que por un lado ofrece artículos y mercancías para el disfrute hedónico de algunos, a causa del terror y sufrimiento de otros. De este modo, Rancière (2010: 31) señala: “Se suponía que la conexión de las dos imágenes produjera un doble efecto. La conciencia del sistema de dominación que ligaba la felicidad doméstica norteamericana con la guerra imperialista, pero también un sistema de complicidad dentro de un sistema”. Frente a esto podemos observar cómo Rancière desmantela con minuciosa perspicacia el *modus operandi* de dichas fotografías, las cuales se encuentran influenciadas por la tradición crítica que busca tornar conscientes a los espectadores y a la vez producir un efecto de culpa.

Asimismo, Rancière (2010: 32) no duda en apuntar: “Por un lado, la imagen decía: ésta es la realidad oculta que ustedes no saben ver, deben tomar conciencia de ella y actuar de acuerdo con ese conocimiento”. De ahí que estas premisas sigan reproduciendo la existencia de un ocultamiento de la realidad, la cual está velada para el “ciudadano común” al no tener conocimiento de las estrategias operativas del sistema dominante.

Sin embargo, Rancière se deslinda tajantemente de aquella forma de ejercer la crítica, puesto que desde su particular enfoque, el querer educar a las “masas acríticas, a través de imágenes que representan atrocidades que ocurren en la esfera social, no es garantía de que vayan a propiciar una revolución de las formas de experimentar su horizonte existencial ni mucho menos transformarlo. De tal forma que Rancière (2010) afirma:

Pero no existe evidencia de que el conocimiento de una situación acarree el deseo de cambiarla...El dispositivo crítico apuntaba así a un efecto doble: una toma de conciencia de la realidad oculta y un sentimiento de culpabilidad en relación con la realidad negada (p.32).

De lo anterior, podemos argumentar que, para Rancière (2010), las premisas de la tradición crítica enfocadas en la concientización de los secretos del sistema dominante, y los afectos de culpa derivados de la ignorancia, carecen de validez, puesto que el conocimiento de un malestar no incide en la modificación o erradicación de tal evento. Así, en esta misma dirección, Rancière (2010: 47) agrega que el dispositivo de la crítica social repite la misma fórmula: “Pero se trata siempre de mostrar al espectador lo que no sabe ver y de avergonzarlo de lo que no quiere ver, a riesgo de que el dispositivo crítico se presente a su vez como una mercancía de lujo perteneciente a la lógica que él mismo denuncia”. Ante esta situación, no queda duda que en el mencionado mecanismo crítico subyacen ciertos afectos como la culpa y la vergüenza, los cuales bajo este modelo son producto de la ignorancia y la ceguera, y que, por tanto, han de erradicarse a través de la denuncia y concientización.

Llegado a este punto, hemos intentado mostrar algunos argumentos de los que se vale Rancière (2010) para señalar que la forma de operar de las artistas, anteriormente mencionadas, subyace bajo la influencia de la teoría crítica tradicional. Esto quiere decir que estas obras sólo reproducen el mecanismo de concientización a partir de las imágenes. Es decir, las fotografías tienen como objetivo hacer ver a los que no saben ver, y por ende, producir efectos de culpa ante su ignorancia. De igual forma, Rancière (2010) localiza este mismo modo de practicar la crítica en algunos pensadores contemporáneos, los cuales vamos a exponer de manera resumida a continuación.

De inicio, Rancière (2010) realiza una crítica al filósofo alemán Peter Sloterdijk, quien, a pesar de ser un personaje discreto dentro del mundo académico, su tesis sobre lo “antigravitorio” le ha dotado de cierta presencia en la reflexión teórica contemporánea. Ahora bien, una de las premisas centrales del pensador alemán radica en describir la época actual como una constante celebración ante los avances de la técnica y la conquista del espacio. Es decir, para el teórico, los viajes espaciales a través de aeronaves

sofisticadas, y el flujo mediático que anuncia incansablemente sus triunfos, sirven para maquillar la miseria de la realidad terrestre.

Por lo tanto, la popular frase de Marx “todo lo sólido se desvanece en el aire” ahora ha sido invertida bajo la supuesta conquista del espacio, que nos conduce a una pérdida de gravedad, razón por la cual enfocamos nuestra mirada hacia el cielo exterior. No obstante, como sugiere Rancière (2010: 36): “Por muy provocadoras que se pretendan, estas tesis no dejan de estar encerradas en la lógica de la tradición crítica... continúan denunciando una incapacidad de conocer y un deseo de ignorar”. Al respecto está claro que, desde la perspectiva de Rancière, estas formas de explicar cómo se ha transformado la realidad social siguen funcionando bajo el paradigma de la crítica social, puesto que se muestra una ausencia de conocimiento y también una espectacularización de la industria espacial.

Asimismo, Rancière (2010: 41) objeta que la crítica tradicional no se apoya en hechos verificables, dado que se limita a decir que las cosas no son lo que parecen ser, y que por tanto solo reproduce una mirada desencantada sobre un mundo en el que la interpretación crítica del sistema se ha convertido en un elemento más del sistema. Frente a esto, resulta claro que la concepción tradicional de crítica está gobernada por la intención de revelar cómo las cosas realmente son, más allá de los engaños del sistema dominante y, además, propagar cierto desencanto ante el panorama catastrófico de la realidad social.

Por la razón anterior, no es de extrañar que Rancière, con minuciosa atención, afirme que ese modo de hacer crítica y, por ende, de pensar críticamente, olvidó enfocar sus esfuerzos cognitivos para indagar cómo hacer posible la emancipación, pues sin lugar a dudas, observó un cambio producto de un desgaste sin esperanzas que lo volvió imposible. Al respecto, Rancière (2010: 36) sostiene: Todavía ayer esos procedimientos se proponían suscitar formas de conciencia y energías encaminadas hacia un proceso de emancipación. Ahora ya están enteramente desconectadas de ese horizonte de emancipación, o bien claramente vueltas contra su sueño.

A partir de lo anterior, es lógico pensar que desde la tesis de Rancière (2009), los quehaceres de la crítica tradicional se agotaron, y si antes se dedicaban a concientizar, ahora se han subsumido en un horizonte desencantado que supone que ha pasado el tiempo de la emancipación. De modo tal que no hay conexión alguna entre la crítica que opera bajo las ilusiones del mercado y el espectáculo, y la intención de crear una propuesta emancipadora. Sin embargo, Rancière no se detiene en esta sola observación de la inversión actual de la crítica, cuyo objetivo emancipador se habría convertido en una declaración de impotencia y una sumisión a lo existente, una inversión que también explica la transformación paradójica de la crítica en el discurso dominante.

Además, el pensador francés señala que esta característica estuvo desde el principio, es decir, desde las primeras formas de crítica social: “La desconexión actual entre la crítica del mercado y el espectáculo y cualquier objetivo emancipador es la forma última de una tensión que ha habitado el movimiento de la emancipación social desde su inicio” (Rancière, 2010: 33). Ante esto resulta prudente preguntarnos, ¿Qué significa esta tensión? Para tratar de responder rápidamente, la tensión es el conflicto desigual entre la posición ocupada por el crítico y la posición de aquellos a quienes se dirige y destina su discurso crítico. Esto quiere decir que, desde el principio, no solo habría una brecha entre estas dos posiciones, sino una verdadera jerarquía, y ahí radica el problema, puesto que el discurso de la crítica social es un discurso cuyo objetivo es la emancipación, pero cuyo dispositivo renueva el obstáculo para cualquier posible emancipación, es decir, la existencia de jerarquías (Rancière, 2010).

En este caso, la jerarquía entre el que sabe y el que no sabe, entre la crítica lúcida de los intelectuales y la falta de crítica de los embrutecidos que viven en la ilusión, pues están cegados y a pesar de ello se deleitan en su ceguera. Por esta razón, el filósofo presupone que en los cimientos de la crítica social existe una forma elitista de hacer crítica, por el hecho de excluir a los embrutecidos, es decir, a esos actores sociales que no tienen ninguna competencia crítica para embarcarse en la enorme tarea de desenmascarar las mentiras de la maquinaria opresora. A esto sigue la superioridad intelectual del crítico social, pues es el intelectual crítico quien mejor sabe qué hacer con los ignorantes, para liberarlos de las ilusiones y mistificaciones en las que están atrapados. Por ello, Rancière (2010) no dudará en relacionar

a esta figura de crítico social bajo los influjos de la herencia platónica, es decir, en aquellas premisas que nos condenan a ser prisioneros de una cueva tomando imágenes por realidades y potenciando nuestra ignorancia a partir de meras copias del conocimiento del mundo ideal.

Ahora bien, aunado a este engaño y pérdida de esperanza, Rancière (1996) también señala que esta forma de pensar la crítica también se fundamentaba en una desigualdad de principio. Donde a través de los postulados del marxismo cientificista, se afirmaba que sólo a partir de la concientización los individuos podrían liberarse de los dispositivos de dominación y mejorar sus condiciones de vida. Bajo esta idea, entonces, la emancipación aparecería cuando se tuviese el conocimiento de la ley de la dominación y, por tanto, sería necesario liberarse de las ilusiones y de la ignorancia de ese proceso:

A partir de allí, la emancipación ya no se concebía como la construcción de nuevas capacidades: era la promesa de la ciencia a aquellos cuyas capacidades ilusorias no podían sino ser la otra cara de su incapacidad real. Pero la lógica misma de la ciencia era la del aplazamiento indefinido de la promesa (p.47).

De lo anterior, podemos afirmar que dicha promesa no era otra cosa más que la responsabilidad otorgada a la ciencia para fungir como una estrategia transmisora de conocimiento y, por lo tanto, con el compromiso adquirido de sacar del letargo del espectáculo a los incapaces. No obstante, la forma de alfabetizar científicamente está mediada por una relación desigual que establece dos figuras: el que sabe y el que no sabe. De ahí que Rancière (2010) sostiene que esta estrategia de la ciencia crítica social reproduce las jerarquías entre el intelectualismo y el sentido común.

En consecuencia, esta crítica hace referencia a una ignorancia que despliega en la ceguera de la mirada, pero también hace alusión a una ausencia de reflexión sobre lo que se ve y, además, a una parálisis para transformar sus conocimientos en energía militante: “Los procedimientos de la crítica social, en efecto, tienen la finalidad de curar a los incapaces, a los que no saben ver, a los que no comprenden el sentido de lo que ven, a los que no saben transformar el saber adquirido en energía militante” (p.50). Visto de este modo, Rancière (2010) concibe la figura del intelectual como un médico que posee el deber moral de curar a esos ciegos incapaces de mirar correctamente el *modus operandi* del sistema dominante, y contradictoriamente, para que sigan curando la enfermedad, esta tiene que seguir existiendo.

En suma, la teoría crítica social desde el análisis de Rancière (2010) se concibe a sí misma como el aspecto intelectual de todo proceso histórico de emancipación y, por tanto, sus mecanismos de operación pierden todo propósito de eficacia puesto que desde sus fundamentos está diseñada desde la desigualdad que presupone ser o no ser crítico. Entonces, bajo esta lupa resulta necesario indagar en nuevas formas de emanciparse, sin las ataduras del capital intelectual y también, sin la figura clave del maestro explicador. Esto, a su vez, presupone una forma distinta de concebir el pensamiento crítico, es decir, un pensamiento que parte de la premisa de la capacidad de cualquiera y, que, además, no se estanque en el diagnóstico y la denuncia sino que, por el contrario, sea capaz de imaginar y encontrar nuevas alternativas de resistir y vivir en la sociedad.

Ahora bien, llegado a este punto hemos intentado exponer de manera mínima algunos argumentos de la lectura realizada por Rancière al paradigma de la crítica tradicional, bajo esta empresa, entonces, el autor se encamina a esbozar algunos indicios para formular otro pensamiento crítico lejos de la ignorancia y la denuncia. Esto supone que el pensamiento crítico no sería propiedad única y exclusiva de la clase inteligente y pensante, sino que considera que la inteligencia es colectiva y, por tanto, el pensamiento crítico puede ser puesto en acción por cualquiera, apelando, sobre todo, a sus condiciones de posibilidad.

En consecuencia, proponemos con Rancière retomar el concepto de crítica desde la perspectiva kantiana, esto es, la crítica como condiciones de posibilidad, donde sea posible reformular los mecanismos cognitivos y orientarlos hacia la investigación de otro horizonte existencial. A continuación, exponemos de qué manera el filósofo francés desarrolla estas ideas.

## **Las condiciones de posibilidad de otro pensamiento crítico**



Es importante señalar que el análisis de los insumos teóricos desarrollados por Rancière (2009) le permite sostener que los mecanismos de la teoría crítica tradicional<sup>2</sup> perdieron de vista proponer un proyecto de emancipación. En este sentido, el problema observado por Rancière (2003) consiste entre otros aspectos en afirmar que esta, en su afán por desenmascarar la realidad oculta y concientizar a los embrutecidos, olvidó o le prestó poca importancia a la creación de una política emancipadora. Por consiguiente, esta práctica fue incapaz de imaginar otro horizonte existencial, pues agotó toda su potencia en explicar y analizar las diversas formas operatorias de la maquinaria capitalista.

Una vez esbozado lo que se comprende por crítica y teoría crítica nos remitimos a las preguntas fundamentales de este artículo: ¿Cómo renovar el pensamiento crítico social<sup>3</sup> de la denuncia desde la filosofía emancipadora de Rancière? ¿Cuáles serían las condiciones de posibilidad de este pensamiento crítico? Tomando en cuenta estas interrogantes, a continuación, trataremos de proponer algunas salidas a partir de ciertas reflexiones de J. Rancière.

De acuerdo con Rancière (2009) cuando pensamos en crítica solemos asociar dicha práctica a la actividad de juzgar bajo el binomio de lo bueno o lo malo, es decir, realizamos juicios para evaluar si las ideas de los intelectuales, las obras de los artistas, o los movimientos sociales, alcanzan cierto grado de realidad o no. No obstante, el concepto de crítica que Rancière (2009) pretende resignificar, y a su vez reactivar, recupera y replantea la noción de crítica como condiciones de posibilidad del filósofo alemán Immanuel Kant<sup>4</sup>:

La crítica en general no es la actividad que juzga si las ideas, obras de arte o movimientos sociales son buenos o no. Por contra, es la actividad la que perfila el tipo de mundo que esas ideas, obras o movimientos proponen, o el tipo de trabajo dentro del cual toman consistencia. Podemos pensar aquí en la idea Kantiana de la crítica (p.6).

Tomando en referencia lo anterior, resulta plausible entender que, para el filósofo francés, la tarea de la crítica no consiste en emitir juicios de naturaleza maniqueísta y que, por tanto, no le compete evaluar las diversas expresiones de la ciencia, el arte, o los movimientos sociales bajo el calificativo de lo bueno o lo malo. Cabe destacar también que el pensador realiza una apropiación del concepto de crítica al margen de su significado trascendental. Esto implica investigar cuáles son los elementos necesarios para que ocurra la emancipación, es decir, dar un paso más allá de la denuncia y aventurarse a enfocarse en la búsqueda de los factores que fracturan los regímenes de opresión.

En este caso, podemos entender que la crítica para Rancière (2009) funciona como una práctica que investiga cómo construir otros mundos de vida y, por tanto, su atención se centra en pensar bajo qué características, rasgos, o elementos, un acontecimiento político y emancipatorio se vuelve posible o se sostiene. De ahí que no se interese por diagnósticos y denuncias, dado que su esfuerzo cognitivo se dirige hacia las ideas que perfilarían y sostendrían un mundo igualitario.

En este mismo orden de ideas, Rancière (2009: 8) supone: “La crítica es una investigación de las condiciones de posibilidad. Es un discurso que se refiere a las formas de posible conocimiento o juicio de sus condiciones de posibilidad”. De ahí que, bajo este supuesto, según entendemos, hacer crítica requiera investigar cuáles serían las condiciones mínimas para que puedan emerger otras formas de vivir, pensar y sentir. Pues, no se trata de conformarse con el mero conocimiento de las características

---

<sup>2</sup> Precisamente, como señala Rancière (2009): “Esa presunción de incapacidad está incrustada en el corazón de la tradición crítica, la idea de que la dominación se autoimpone ante la ignorancia e ilusiones de sus sujetos. En cierto momento la presunción iba acompañada de la promesa de que la ciencia los liberaría. Hoy en día la ciencia está satisfecha con explicar por qué tal liberación es imposible”.

<sup>3</sup> El pensamiento crítico derivado de la teoría crítica tradicional, opera bajo la consigna de que sólo los inteligentes son capaces de ejercerlo. Además, este pensamiento tiene como misión desenmascarar la realidad oculta de los regímenes de dominación. De esta forma, su práctica consiste en ver lo que los ignorantes no saben ver para después visibilizarlo, y hacer que los engañados se den cuenta. No obstante, de acuerdo con Rancière (2009) en el corazón mismo de esta crítica está la desigualdad.

<sup>4</sup> En este punto resulta prudente observar que Rancière en su conferencia *El Método de la Igualdad*, expresa a modo de comentario que su crítica está inspirada en la reflexión kantiana. Sin embargo, es importante mencionar que, dado el carácter fragmentario de la escritura rancieriana, el filósofo hasta el momento no ha desarrollado un texto o argumento completo donde especifique con claridad en qué punto se deslinda con exactitud del filósofo alemán, pues sólo lo menciona ocasionalmente, para enfocarse en las condiciones de posibilidad de otra crítica, y a su vez, de otro pensamiento crítico.

operativas de la realidad social del presente sino, por el contrario, en correr el riesgo de potenciar ideas y proyectos que tornen posible dibujar otra cartografía existencial.

Una vez que hemos visto el significado de la crítica recuperada por nuestro autor, resulta igualmente oportuno revisar algunos de los pequeños indicios que muestra respecto al pensamiento crítico bajo este esquema interpretativo. Inicialmente, Rancière en su texto *Sobre la importancia de la teoría crítica para los movimientos sociales actuales* (2009: 8) enfatiza: “El pensamiento crítico debería de tener como punto de comienzo la investigación acerca del poder de configurar mundos alternativos inherentes a esas formas de dominación”. Esta aseveración nos permite observar que, el verbo debería indica su inconformidad con el estado actual del pensamiento crítico, pues de cierta manera como hemos marcado en párrafos anteriores, el filósofo señala el desgaste de la lógica de la denuncia, acción tan popular en la crítica tradicional y, por ende, del pensamiento crítico bajo estas características. Con base en ello, puntualizamos que este pensamiento crítico por venir supondría inicialmente una acción investigativa que no sólo señale y elabore diagnósticos del malestar de este mundo.

Pues, como señala Rancière (2011: 262): “...un pensamiento crítico en sentido kantiano, es decir, una forma de pensar que hace posible las diferencias que instituye tal o cual dominio sensible, o, también, tal o cual dominio de inteligibilidad”. De allí que este pensar se caracterizaría por la suspensión de los afectos y pensamientos del orden dominante, lo cual implica una transformación de lo sensorial e inteligible. De ahí que, para posibilitar estas transformaciones, sea preciso seguir lo que el autor francés conceptualiza bajo el nombre de disenso, concepto clave<sup>5</sup> para tratar de comprender su proyecto filosófico.

En este orden de ideas, Rancière (2019: 14) señala: “Un pensamiento crítico es también una forma de pensar según la cual la institución de esos dominios se concibe como el producto de operaciones críticas o disensos”. Bajo este supuesto, entonces, se pone de manifiesto la radicalidad de un pensamiento crítico que no se conforma con la descripción de los órdenes existentes sino, por el contrario, tiene el potencial de propiciar otros modos existenciales desobedeciendo los mecanismos del orden policial.

Ahora bien, ¿qué significa que un pensamiento crítico pueda ser producto de operaciones críticas o disensuales? A primera vista, podemos argumentar que este pensar sería efecto del disenso y, por ende, es una de las consecuencias derivadas de asumir la interrupción de los regímenes dominantes. Básicamente se trata de apostar por experiencias disensuales como una de las condiciones que harán posible la ruptura de las jerarquías cognitivas. Esto quiere decir que este modo de pensar reestructura y suspende el orden instaurado abriendo paso a la construcción de nuevas capacidades.

Por lo tanto, argumentamos que las condiciones de posibilidad para este pensamiento crítico renovado es que al menos cuente con dos aspectos: igualdad de las inteligencias y disenso. Respecto al primer punto, resulta relevante recordar que para nuestro filósofo, la inteligencia es una y, por ende, no existen jerarquías cognitivas. En esa perspectiva, no existen seres embrutecidos esperando salir de la ignorancia e invocando la ayuda de la inteligencia erudita de un intelectual. Antes bien, esta estrategia conceptual le permite a Rancière (2007) sostener que la igualdad de las inteligencias es la capacidad que tiene cualquiera (*no importe qui*) de hablar y ocuparse de asuntos comunes.

En esta misma vía, Rancière (2010: 133) en su conferencia *¿Comunistas sin comunismo?* plantea: “la igualdad no es una meta a alcanzar, es un punto de partida, una presuposición que abre el camino para una posible verificación. En segundo lugar, la inteligencia es una. No hay una inteligencia del maestro y una inteligencia del alumno, una inteligencia del legislador, y otra del artesano, etc.” De acuerdo con ello, podemos pensar que el pensamiento crítico no es una habilidad exclusiva de los intelectuales, por el contrario, bajo el supuesto de la igualdad, cualquiera desde una condición horizontal puede aprender por sí mismo, y también aventurarse en el encuentro de otras formas de pensar. Al respecto, es indispensable dejar de

---

<sup>5</sup> “Es el trabajo que produce disenso, que cambia los modos de presentación sensible y las formas de enunciación al cambiar los marcos, las escalas o los ritmos, al construir relaciones nuevas entre la apariencia y la realidad, lo singular y lo común, lo visible y su significación” (Rancière, 2010: 67)



señalar a los “ignorantes” para comprender que no hay jerarquías y diferencias en torno a quienes poseen el conocimiento y quienes no.

Esto supone que bajo dicha renovación: “El pensamiento crítico investiga bajo qué condiciones y posibilidades podemos construir otro mundo y, por ende, abandona la obsesión por el desenmascaramiento de la ilusión, la sombra y el engaño” (Castillo, 2018: 12). Así pues, resulta factible decir que esta forma de renovar el pensamiento crítico, evita caer en teorizaciones descreídas y poco esperanzadoras, que anulan toda posibilidad de ruptura con el orden social.

En suma, las condiciones de posibilidad para renovar el pensamiento crítico, implicaría inicialmente la igualdad de las inteligencias y el disenso. Por supuesto, no se trata de replicar la denuncia y la concientización sino, por el contrario, no perder de vista las posibilidades de emancipación. De allí que este pensamiento no sea un análisis conformista de la realidad oculta, por el hecho de que traza las coordenadas para investigar otro mapa existencial. En definitiva, esta singular contribución de Rancière nos invita a seguir interrogando sobre las posibles formas de ponerlo en práctica, tanto en las acciones educativas, como en los procesos de investigación.

## Conclusiones

El objetivo de este artículo consistió en proponer una posible renovación del pensamiento crítico a través de la filosofía emancipatoria de J. Rancière. En razón de ello, revisamos algunas de las críticas que el autor realiza a la teoría crítica tradicional con el propósito de mostrar cómo desde su particular punto de vista, esta se queda arraigada en formas como las de la lógica de la denuncia, la concientización de clase, así como en la insistente oposición entre apariencia y realidad. Vimos cómo el filósofo subvierte estos supuestos a través de un cambio de trayectoria y señala cómo la crítica dejó de lado pensar en proyectos de emancipación, pues sólo se dedicó a realizar diagnósticos y radiografías de las diversas características adversas de la sociedad.

La tesis de Rancière (2010) sugiere que el pensamiento crítico tiene que ser replanteado, lo que implica entre otros asuntos rechazar la premisa que establece la dualidad entre apariencia-realidad. Con esto, el pensador supone que la acción de develar una realidad maquillada y oculta por los distintos dispositivos de poder perdió su eficacia, dado que en estos tiempos cínicos ya no hay ilusión cómplice, ni tampoco el deseo de ignorarla. De esto deriva, además, la objeción que realiza el filósofo a la vía de la concientización como el mecanismo que haría posible la liberación de los trabajadores, dado que en el corazón mismo de esta vía está infiltrada la desigualdad: unos saben y otros no saben.

Bajo estas premisas, observamos cómo Rancière declara tomar distancia de las formas operativas de la crítica tradicional, para reformular nuevos medios sin el afán de reproducir la ignorancia, la culpabilidad y la denuncia. Ello supone que una renovación del pensamiento crítico asume la igualdad de las inteligencias y, por lo tanto, no existen unos más inteligentes que otros. De ahí que los cualquiera –los ignorantes- también puedan ejercer la crítica como una potencia disensual, desmontando la hipótesis del marxismo cientificista que señalaba que solo algunos pueden pensar.

En contraste, este renovado pensamiento crítico ya no sería propiedad exclusiva de los intelectuales y científicos, por el hecho de que ya no se parte de la división de las inteligencias puesto que la inteligencia es una y, por tanto, se desvanece la idea de que solo una clase dentro de las jerarquías sociales puede hacer crítica. Por consiguiente, resulta factible la abolición del binomio entre los que saben y no saben, dejando de lado figuras como el intelectual orgánico y el maestro explicador.

Aunado a lo anterior, la renovación del pensamiento crítico recomendaría salir de la lógica de la denuncia, puesto que su propósito no coincide en señalar y reparar los desajustes de la sociedad para que funcione mejor, sino en apostar precisamente por otra esfera social al margen del orden dominante. Esto plantea, entre otras cosas, socializar este pensamiento crítico con el objetivo de ir aminorando la inercia de las denuncias amarillistas que sólo normalizan la desigualdad del mundo cotidiano, pues pocas veces

convocan a cambios radicales a través de la implementación transformadora de propuestas emancipadoras.

Hemos puntualizado que esta otra crítica deriva de recuperar la praxis del disenso. Ello con el objetivo de provocar desconexiones y nuevas formas de enfocar el pensamiento y los sentidos, pues como afirmó Febbro (2019) hay formas de pensar no diferentes sino en disidencia, de no repetir el agotador mensaje del diagnóstico del mal y, al mismo tiempo, de descubrir las posibilidades infinitas de la acción humana, lo que permitirá dislocarnos de nuestras distribuciones sensoriales y cognitivas habituales para crear e investigar otros horizontes vitales. A partir de ahí podemos imaginar que otro pensamiento crítico es posible, donde los dispositivos jerárquicos de alfabetización se desvanecen ante la evocación de una inteligencia horizontal al margen de la desigualdad.

Ciertamente, una renovada perspectiva del pensamiento crítico destacaría el poder de la política de la imaginación, cuyo más alto propósito no coincidiría en apuntar los desajustes de la sociedad, sino en apostar precisamente por imaginar otra esfera social.

Finalmente, la eficacia de este pensamiento crítico apostaría por desactivar los escenarios dominantes de la “realidad social” para transformar las dinámicas vitales, porque finalmente se caracterizaría por la igualdad de las inteligencias y el disenso. Al final podemos recordar la sentencia de Merleau-Ponty (1981) quien consideró que no se trata de reparar la naturaleza ni la cultura, sino más bien, en tener la fuerza imaginativa para pasar por encima de sus estructuras y crear otras.

## Referencias

- Althusser, L. (1988). *La filosofía como arma de la revolución*. Siglo XXI. Buenos Aires.
- Castillo, K. (2018). “Claves pedagógicas en J. Rancière”. *Ixtli, Revista Latinoamericana de Filosofía de la Educación*. 4 (2) 121-134. Recuperado de: <http://ixtli.org/revista/index.php/ixtli/article/view/94>
- Castillo, K. Miramontes, Ma. (2019). “Educar para el disenso: política y estética en J. Rancière”. *Revista Inclusiones* 6 (3) 127-139. Recuperado de: <http://www.revistainclusiones.com/gallery/8%20vol%206%20num%203%202019julsep19incl.pdf>
- Corcuff, Philippe. (2019). “De la posible renovación de la teoría crítica en Francia, entre desventuras académicas y tensiones Bourdieu/Rancière”. *Revista de Ciencias Sociales*. 32(44). 61-80. Epub 01 de junio de 2019. <https://dx.doi.org/10.26489/rvs.v32i44.3>
- Etchegaray, R. (2014). “La filosofía política de Jacques Rancière”. *Nuevo pensamiento, Revista de Filosofía*. 4 (4).
- Febbro, E. (2019). “No estamos frente al capitalismo, sino que vivimos en su mundo”. Recuperado de: <https://www.pagina12.com.ar/188829-no-estamos-frente-al-capitalismo-sino-que-vivimos-en-su-mundo>
- Martínez, J.J. (2016). “Entre la (filosofía) crítica y la (filosofía de la) emancipación: el problema del orden social en el pensamiento de Jacques Rancière”.
- Merleau-Ponty, M. (1981). *La estructura del comportamiento*. FCE. Ciudad de México.
- Rancière, J. (2003). *El maestro ignorante*. Laertes. Barcelona.
- Rancière, J. (2009). “Sobre la importancia de la teoría crítica para los movimientos sociales actuales”. *Estudios Visuales*, (7). 81-90. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=6824455>

Rancière, J. (2010). *El espectador emancipado*. Manantial. Buenos Aires.

Rancière, J. (2011). *El tiempo de la igualdad*. Herder. Barcelona.

Rancière, J. (1996). *Momentos políticos*. Manantial. Buenos Aires.

Rancière, J. (2019). *Disenso. Ensayos sobre estética y política*. FCE. México.

## LO NUMINOSO EN LA LEYENDA POPULAR

### IMAGEN, MITO Y RELATO EN EL PERSONAJE DE LA LLORONA

#### *THE NUMINOUS AND POPULAR LEGEND*

*Image, myth and narrative in the character of la Llorona*

#### *O NUMINOSO E LENDA POPULAR*

*Imagem, mito e narração na personagem de la Llorona*

**Federico Aguilar Tamayo**

*(Universidad Autónoma del Estado de Morelos)*

*federicoaguilart@gmail.com*

**Lorena Noyola Piña**

*(Universidad Autónoma del Estado de Morelos)*

*noyola@uaem.mx*

**Laura Silvia Iñigo Dehud**

*(Universidad Autónoma del Estado de Morelos)*

*lauartis@hotmail.com*

Recibido: 01/02/2021

Aprobado: 26/04/2021

### RESUMEN

El presente artículo reflexiona en la leyenda de La Llorona como un fenómeno santo cuya cualidad es doble: como relato y como creencia. Noción que exploramos a partir de los conceptos de Rudolf Otto, lo racional y lo irracional de la experiencia santa. Como campo de observación recurrimos al relato oral y a la imagen como insumos investigativos. Como resultado observamos cómo los conceptos racional e irracional se articulan en la leyenda de La Llorona tanto en imagen como en relato, lo que revela que aquella es un complejo mítico contemporáneo donde lo santo, persiste como relato e imagen.

Palabras clave: santo, mito, leyenda, imagen, oralidad.

### ABSTRACT

This paper ponders the legend of La Llorona as a sacred phenomenon whose quality is twofold: as a narrative and as a belief. We explore this circumstance from Rudolf Otto's concepts, rational and irrational, in sacred or religious experience. As an observation field, we resort to oral narrative and image as investigative inputs. As a result, we observe how the rational and irrational concepts are articulated in the legend of La Llorona as image and narrative, which reveals La Llorona is a contemporary mythical complex where the holy persists.

Keywords: Sacrum, myth, legend, image, orality.

## RESUMO

O artigo reflete sobre a lenda de *La Llorona* como um fenômeno sagrado cuja qualidade é dupla: como história e como crença. Exploramos, a partir dos conceitos de Rudolf Otto, o racional e o irracional da experiência sagrada. Como campo de observação, recorremos à história oral e à imagem como subsídios investigativos. Como resultado, observamos como os conceitos racionais e irracionais se articulam na lenda de *La Llorona* tanto na imagem quanto na história, o que revela que este é um complexo mítico contemporâneo no qual o sagrado persiste.

Palavras-chave: sacro, mito, lenda, imagem, oralidade.

## Introducción

El presente artículo reflexiona en torno a lo numinoso expresado en una de las formas populares del relato oral: la leyenda. En esta perspectiva consideramos que lo santo (de lo cual se deriva lo numinoso) persiste de maneras diversas que se manifiestan más allá de las instituciones religiosas; es el caso de cuentos, leyendas, consejas, fábulas, que dan cuenta de fenómenos cuya problemática abarca tanto su índole narrativa como mítica.

Como caso de estudio nos centramos en una leyenda popular de México como es La Llorona, su difusión desborda el territorio nacional de tal forma que sus variantes se encuentran además en Centroamérica<sup>1</sup>. En términos generales dicha leyenda trata de un espanto en forma de mujer que aparece por las noches, se lamenta por la muerte de sus hijos, que, según las versiones más comunes, ella misma habría asesinado ahogándolos en el río.

Los encuentros con La Llorona son -al día de hoy- relatados por cantidad de personas que aseguran haberla visto, por lo que se considera esta leyenda como una creencia viva<sup>2</sup>. Según diversos relatos el encuentro con La Llorona resulta en perjuicio físico y psicológico de la víctima, de ahí que se trata de una leyenda de corte macabro, deviene en desgracia y es causa de susto, espanto y estupor.

Por otro lado, este personaje se considera todo un fenómeno identitario, por ejemplo, en Xochimilco, una localidad de la Ciudad de México la leyenda se arraiga hasta llegar a la diosa prehispánica Cihuacóatl<sup>3</sup>.

En el binomio relato-suceso de la leyenda de La Llorona, tenemos un fenómeno donde la majestad del personaje emana en términos de Rudolf Otto (2016) del absoluto heterogéneo (lo no humano, ni mundano, que es santo), si empleamos el término de Mircea Eliade (1992) hablaríamos de una hierofanía<sup>4</sup>, en el caso de Rudolf Otto (2016), -autor que nos servirá de referente teórico- empleamos el término numinoso.

En el año 1917 cuando aparece el texto principal de Rudolf Otto *Das Heilige*, (Lo Santo), el marxismo impera en los círculos intelectuales de Europa, aun resuena implacable la proclama de Marx (1844): “Para Alemania, en resumen, la crítica de la religión está terminada y la crítica de la religión es la premisa de toda crítica”<sup>5</sup>. Como observa Manuel Fraijó: “Decir religión en Alemania de comienzos del siglo XX era

<sup>1</sup> Por ejemplo, la leyenda de la Ziguana, madre del duende Cipitío, y de quien se dice, acecha en los ríos y lagos de El Salvador. Un relato literario en torno a este personaje puede encontrarse en: Salazar Arrué Salvador. (1999) *Cuentos de barro*. p.43-44. CONICULTURA: San Salvador.

<sup>2</sup> En relación a las creencias nos apegamos a la explicación de Rudolf Otto: “... no es lo mismo tener idea de lo santo, que percibirlo y aun descubrirlo como algo operante, eficiente que se presenta operando en fenómenos... lo suprasensible puede aparecerse en ciertos acontecimientos, hechos y personas - comprobaciones efectivas de la autorrevelación - y que junto a la revelación interior nacida en el espíritu existe una revelación externa de lo divino”. (Otto, 2016: 261). De la anterior derivamos que: *creer* implica el reconocimiento de una efectividad de lo santo, que además de interiorizarse se expresa en fenómenos externos o mundanos, ya sean relatos o imágenes que operan como signo de su presencia. Quienes tienen una creencia no necesariamente la comprenden en términos racionales, sino que esta se experimenta y relaciona con el contexto.

<sup>3</sup> En Xochimilco la leyenda de La Llorona se sincretiza además de la referida diosa Cihuacóatl con la virgen de Dolores o virgen de Xaltocan, cuya imagen simboliza la madre dolorosa, el rostro de esta virgen se representa surcado por grandes lágrimas. Mayor referencia puede encontrarse en: Aguilar, F., Escobar, S., Villegas, L., Olivares, E. (2013) *Las aventuras de Cristóbal el Brujo*. Un recorrido por las leyendas de México, p. 259. Ciudad de México: Ensamble Comics A.C.

<sup>4</sup> Entendida la hierofanía como una manifestación terrenal de lo divino.

<sup>5</sup> En la introducción *Para la crítica de la filosofía del Derecho de Hegel* (1968).



nombrar un enfermo terminal, los grandes filósofos del siglo anterior- Feuerbach, Marx, Engels, Nietzsche- la habían tratado con una severidad crítica quizás nunca antes alcanzada” (Fraijó, 2016: 11).

No obstante, el contexto donde se publica la primera edición de “Lo Santo” de Rudolf Otto, éste no pasó desapercibido, por lo contrario desde su publicación en 1917 hasta 1936, el libro tenía un total de 25 ediciones, y para 1958 apareció la trigésima edición (Fraijó, 2016).

En *Lo santo, lo racional y lo irracional en la idea de Dios*, Rudolf Otto (2016) explora lo santo en su modo experiencial, el movimiento es audaz porque distingue el estudio de las religiones del plano estrictamente filosófico o antropológico hacia el fenomenológico, tras esto, desarrolla toda una conceptualización novedosa y estimulante que tiene en lo *numinoso* su concepto central.

Rudolf Otto observa que el fenómeno religioso se constituye por dos aspectos fundamentales: el primero corresponde a lo *racional*, este conlleva la simbología, narrativas, enseñanzas, explicaciones, terminología y conceptos empleados por las diversas religiones. Cuestión con la que Otto es particularmente crítico. Si bien valora la importancia de conceptos claros y precisos en toda religión (particularmente en la cristiana), también observa que las instituciones han racionalizado el sentimiento religioso en exceso, al punto, que éste queda socavado entre conceptos. En la contraparte *irracional* de lo santo Rudolf Otto centrará su estudio e interés, lo *irracional* apela al sentido emocional, psicológico, y primigenio de toda creencia. Por medio del concepto clave de lo *numinoso*, El autor reflexiona en aquello que es afín y origen de toda religión y creencia: su forma experiencial, anímica y sentimental.

Llamamos, desde luego, racional en la idea de lo divino aquella parte de él que entra en la clara comprensión de nuestra facultad conceptual, en la esfera de los conceptos corrientes y definibles. Afirmamos después que bajo de esa esfera de desnuda claridad yace una oscura profundidad, a la que no hallan paso nuestros conceptos y que es lo que denominamos irracional (Otto. 2016: 136).

Rudolf Otto se desmarca del contenido moral implícito en el término santo, (generalmente esquematizado en lo bueno), por lo que acuña el concepto *numinoso*, derivado de las palabras *numen* (lo santo, lo divino), y el sufijo *oso*, que determina una circunstancia y momento, donde “lo numinoso es vivido como presente” (Otto. 2016: 55).

El encuentro con lo numinoso no es necesariamente bondadoso y lleno de paz como pudiera imaginarse coloquialmente, por el contrario, se describe en Otto (2016) como tremendo y paralizante, ya que el numen en su completa majestad inunda en pavor a quien lo observa.

Por citar un ejemplo ilustrativo observamos la noción de nulidad total y horror paralizante en la literatura, particularmente en los relatos del afamado círculo de H.P. Lovecraft<sup>6</sup>, donde a menudo emplean el término *inenarrable* para describir horrores cósmicos y dioses primigenios ante los cuales el ser humano se pierde completamente de sí.

Tal estremecimiento radica en la cualidad santa de lo *numinoso*, ante el numen –que es absoluto y majestuoso–, lo humano se asume y concibe como mera creación, a decir, polvo y ceniza. Constatar la condición de no ser más que una criatura, es un resultado inminente de la confrontación directa con el numen, lo que conduce hacia un estado de estupor en tanto que opera el reconocimiento de que se *es nada*. A este estado concomitante a la presencia del numen Rudolf Otto (2016) lo denomina *sentimiento de criatura*.

Más allá de las implicaciones sentimentales y psicológicas del encuentro con lo *numinoso* descritas por Rudolf Otto con pluma estimulante, casi literaria, asumir la condición de ser creado, representa un eje común de toda creencia de carácter santo. En este sentido, el estudio de Rudolf Otto (2016) toca aristas antropológicas; el autor fue un viajero conocedor e incluso participante de liturgias en diversos lugares del

---

<sup>6</sup> Nos limitamos a citar el caso del autor Howard P. Lovecraft como un ejemplo explicativo, no obstante, un debate entre mito, literatura y horror numinoso puede revisarse en: Olavarría, María Eugenia (1989). “Lo numinoso: mito y terror”. *Alteridades. Anuario de Antropología*. UAM-México. 113-122.

mundo, su pasión no quedaba constreñida en el cristianismo, hecho que lo llevó a lugares lejanos con el fin de observar y experimentar de primera mano la religiosidad en otras culturas<sup>7</sup>.

Lo *numinoso* como categoría de análisis es desglosado por Rudolf Otto (2016) en varios mecanismos, (para el presente artículo echamos mano de aquellos que nos resultan fundamentales para el entendimiento de nuestro objeto de estudio), el primero lo hemos explicado líneas arriba como *sentimiento de criatura* y refiere la absoluta nulidad y estupor ante el numen cuando éste se presenta y manifiesta. Lo que se observa en leyendas populares donde diversos relatos describen un horror paralizante frente a la presencia de entes numinosos como La Llorona, el Charro Negro, el Nahual<sup>8</sup>. De manera particular en la leyenda que nos atañe, el terrible lamento de La Llorona es motivo de miedo y parálisis para quienes refieren haberlo escuchado.

Un segundo aspecto de lo *numinoso* refiere la potencia de un misterio implícito en el numen, misterio que, por ser del orden santo, atrae a la vez que repele. Aunque el numen pudiese ser siniestro, o diabólico, resulta fascinante en tanto que es un misterio, por lo que Rudolf Otto (2016) le asigna el apelativo de *mysterium tremendum*.

Siguiendo al autor el *mysterium tremendum*, (misterio tremendo) opera en el sentimiento de la fascinación contemplativa, hasta formas de atracción pavorosa: “El tremendo misterio puede estallar desúbito en el espíritu, entre embates y convulsiones. Puede llevar a la embriaguez, al arrobó, al éxtasis. Se presenta en formas feroces y demoniacas, puede hundir al alma en horrores y espantos casi brujescos”(Otto, 2016: 59).

De lo anterior, que los aspectos centrales que observamos en lo *numinoso* son: 1) el sentimiento de criatura y 2) el tremendo misterio, se comprobará que ambos conceptos están presentes en los relatos de leyenda que se analizan en este artículo, con lo que se aporta un marco analítico para comprender la leyenda como un fenómeno de carácter santo.

Quizás la gran dificultad de implementar la fenomenología de este autor en un trabajo de investigación social radica en que la experiencia religiosa es eminentemente subjetiva e intersubjetiva además de emocional. En un punto bastante álgido de su texto el mismo Rudolf Otto (2016) sostiene que todo aquel que no haya pasado por un momento de «fuerte conmoción religiosa» debería sin duda renunciar a la lectura de su libro (p. 52). No obstante, el tono exacerbado por el que atraviesan algunos pasajes en el estudio de *Lo santo...*, Rudolf Otto analiza la experiencia santa de una manera eficientemente objetiva.

En primera instancia refiere que un estudio sistemático de lo *numinoso* enfrenta el problema que tal concepto carece de una definición exacta, por lo que la comprensión del mismo se encuentra en las relaciones que guarda con el mundo sensible, observable y experiencial, de esto deriva que lo *numinoso* tiene dos formas de expresión, a decir: 1) expresión directa, (que es el encuentro frontal con el numen y grados sentimentales de la liturgia, himnos y prácticas religiosas); y 2) expresión indirecta que se encuentra en formas más cercanas al aspecto racional como son arquitectura, relatos, imágenes sacras, que, si bien no tienen la potencia total del numen, de menos conllevan su carácter.

En relación con nuestro estudio, la observación de las formas de expresión indirecta de lo *numinoso* son en realidad la pauta que Rudolf Otto brinda para una propuesta investigativa. Para el presente artículo se observaron de manera concreta relatos e imágenes obtenidos mediante la entrevista etnográfica y la ilustración, lo que se detalla en el siguiente apartado.

---

<sup>7</sup> Entre sus datos biográficos se sabe que Rudolf Otto viaja a Tenerife, África, Japón, China y la India.

<sup>8</sup> Para mayor referencia en torno a leyendas populares y criaturas fantásticas de la tradición oral mexicana, consultar *Las aventuras de Cristóbal el brujo, un recorrido por las leyendas de México* (2013).

## Metodología

Para observar lo *numinoso* en los relatos de leyenda contemplamos como sugiere Rudolf Otto (2016), el relato en *viva vox*, cuestión que se puede lograr por medio de la entrevista.

Los extractos de la leyenda de La Llorona y personajes afines que se trabajan en el presente artículo fueron obtenidos de una investigación mayor comprendida entre los años 2015-2017. En dicha investigación se recopilaban diversos relatos de leyenda en varias regiones de Chiapas, México (ver Tabla 1), el objetivo fue observar la dinámica cotidiana de la leyenda a través del relato oral para posteriormente analizar su simbología, relaciones semánticas, identificar variantes y ofrecer interpretaciones que dilucidaran los posibles significados de dichos relatos.

Una característica notable de la leyenda es su capacidad de narrarse como episodio, donde generalmente el relato se estructura en forma lineal: un comienzo, una trama y un desenlace, por lo que su empleo resultó beneficioso para el estudio de los relatos de leyenda.

La técnica de entrevista episódica, como refiere Uwe Flick (2004), encausa la charla hacia un enfoque temático.

El punto de partida para la entrevista episódica es el supuesto de que las experiencias de los sujetos en un cierto dominio se almacenan, y recuerdan las formas de conocimiento narrativo episódico y semántico... El conocimiento episódico se organiza más cerca de las experiencias y se asocia a situaciones y circunstancias concretas (Flick, 2004: 118).

Los episodios favorecen la evocación de circunstancias concretas, la persona entrevistada relata momentos precisos para expresarlos a través de una estructura narrativa. Las entrevistas que realizamos para el presente estudio tuvieron una duración comprendida entre 45 minutos a 2 horas. Todas fueron grabadas y transcritas en su totalidad, lo que el etnólogo J. Spradley (1979) apunta como registro *verbatim*. de las cuales se conformó un inventario jerarquizado según el tipo de leyendas y simbología expresada en los relatos, lo que permitió ubicar aquellos referentes a la leyenda en cuestión.

Como soporte tecnológico contamos con la pluma inteligente *Live Scribe*, dispositivo que graba las entrevistas a la par de la escritura de notas de campo, lo que es posible por medio de un cuaderno interactivo que registra tanto notas escritas como voces de manera simultánea. Esta tecnología permite revisar los apuntes de campo a la vez que se escucha la grabación, dinamizando el proceso de jerarquización e identificación de dominios<sup>9</sup>

---

<sup>9</sup> Más información de la implementación de este dispositivo se encuentra en: Aguilar Tamayo Manuel F. (2014), "Nuevas tecnologías en la formación: asesoría hipermedia con plumas digitales como apoyo para la elaboración de trabajos de investigación", en Santiago R. Acuña, María A. Gabino Campos y Consuelo P. Martínez Lozano (coord.), *Multiculturalidad, imagen y nuevas tecnologías*. Fragua. Madrid. 263-291.

Tabla 1. Tabla de datos de entrevistados durante el trabajo de campo

Nombre del entrevistado (a)	Ocupación y/o cargo del entrevistado (a)	Lugar y fecha de realización de la entrevista	Tema: leyenda o narración relatada
<b>Sra. Lupita</b>	Locataria del mercado de los Ancianos.	22/abril/2015 Mercado de los Ancianos, Tuxtla Gutiérrez Chiapas.	El carretón de la muerte, La tita. Oficio de esotérica, Su historia en el mercado.
<b>David Díaz</b>	Periodista del Blog de leyendas Chiapas Nostrum. Canal 10 de Chiapas.	2/octubre/2015 Edificio de Canal 10 de Chiapas.	Leyendas de Tapachula, Chiapa de Corzo.  Cuevas, Sombrerón, Bájate carne, San Simón, Enrique Verdi, Duendes.
<b>Amín Guillén</b>	Cronista, periodista, escritor.	13/octubre/2015 Centro de Comitán	La esquina blanca, el Sombrerón, El Colosode Chiapas, la academia.
<b>Ángel</b>	Promotor cultural.	13/octubre/2015) Centro de Comitán.	Arco del Carmen, San Cristóbal.
<b>David Díaz</b>	Blog de leyendas Chiapas Nostrum. Periodista Canal 10 de Chiapas.	14/octubre/2015 Chiapa de Corzo.	La Pochota, Batalla del 21 de octubre, Enrique Verdi, duendes.
<b>Festival Magia y Tradiciones: Alexis Griseln Esmeralda y Marijose Sr. Catarino Mendoza.</b>	Participantes del 7º Festival Magia y Tradiciones 2015.	31/octubre/2015 Chiapa de Corzo, parque central.	Sombrerón, carretón de la muerte, Tishanila.
<b>Sra. Mercedes</b>	Estilista	3/noviembre/2015 Barrio de San Ramón. San Cristóbal de las Casas	Nahual, cuevas, Cadejo, Llorona, Sombrerón, Yehualcihuatl, negrito cimarrón.
<b>Sra. Francisca</b>	Vendedora de verduras	9/noviembre/2015 Parque de San Ramón, San Cristóbal de las Casas	Llorona, Sombrerón, bájate carne.
<b>Sra. Francisca García (+)</b>	Comerciante	9/noviembre/2015 Barrio la Merced San Cristóbal de las Casas.	Cadejo, Llorona, pilico, bájate carne.
<b>Grupo de discusión alumnos de arqueología: Diana. Margarita. Josué. Oscar. Obed. Brian.</b>	Estudiantes	13/noviembre/2015 Escuela de arqueología UNICACH en Chiapa de Corzo. Los estudiantes son: Diana-Tuxtla. Margarita-Chiapa. Josué-Palenque. Oscar-Tapachula. Obed-Palenque. Brayán-San Cristóbal.	Llorona, Yehualcihuatl, Charro negro, Sombrerón, duendes, fantasmas, Vampiro de San Cristóbal, Sirena de San Cristóbal.
<b>Sr. Rosel Hernández</b>	Ramilletero y músico.	6/Septiembre/2016 Tuxtla Gutiérrez, parque de la marimba.	Pichito encantado, Mala Mujer Sombrerón, duende, Cocha enfrenada, San Pascualito
<b>Sra. Magdalena Morales</b>	Restaurantera Promotora Cultural	21/Septiembre/2016 Tapalapa Chiapas	Sombrerón, duende (monchosi).
<b>Sr. Celso Morales</b>	Campeño	21/Septiembre/2016 Tapalapa Chiapas	Sombrerón duende mal espíritu
<b>Sr. Santos</b>	Músico, curandero, referente en la localidad como conocedor de leyendas.	21/Septiembre/2016 Tapalapa Chiapas	Viejo de la montaña (Sombrerón), duende, aves de mal agüero, el rayo, nahual, susto

<b>Sr. Mariano Pérez</b>	Campesino	2 /Octubre /2016. Yabteclum (Yabjteclum) Chenalhó Chiapas	Yehualcíhuatl Sombrerón
<b>Sra. Juana</b>	Campeña	2 /Octubre /2016. Yabteclum Chenalhó Chiapas	Negro cimarrón nahual, Sombrerón, Mala Mujer, fuego resplandeciente (que indica riqueza), Burrito encantado, YalamBeket (Bájate carne)
<b>Sr. JosuéSra. Rosa</b>	Comerciantes	16/ Septiembre/ 2016Tuxtla Gutiérrez	Cocha enfrenada, Sombrerón, nahual, Tisigua, San Pascualito, duendes.

Fuente: Tabla elaborada por Federico Aguilar Tamayo.

Del universo de entrevistados se seleccionaron episodios que narran acerca de la leyenda de la Lloronay personajes afines, teniendo como delimitación las siguientes entrevistas:

- David Díaz, Chiapa de Corzo, Chiapas. México, 2015.
- Rosel Hernández, Tuxtla Gutiérrez, Chiapas. México, 2016.
- Amin Guillén, Comitán de Domínguez, Chiapas. México, 2016.
- Francisca (Chica) García+, San Cristóbal de las Casas, Chiapas. México, 2015.
- Mariano. Yaubteclum (Altos de Chiapas) 2016.
- Margarita, Chiapa de Corzo, 2015.
- Celso, (acerca de los rezos y protección contra espantos), Tapalapa, Chiapas 2016.

## Recopilación de datos visuales

Desde hace algunas décadas la investigación social cualitativa incorporó los métodos visuales como parte de sus herramientas de análisis. El uso de las imágenes ya existentes o elaboradas por quien investiga se ha convertido en una forma variada y valiosa de obtener los llamados “datos visuales”.

La antropología y la sociología visual son pioneras en el uso de metodologías visuales para la investigación social. La imagen por sí misma es una potente portadora de conocimientos, un dispositivo que encierra saberes por su carácter omnipresente en la sociedad y su carácter polivocal, es decir, nos permite obtener diversas miradas y lecturas sobre un mismo fenómeno (Banks, 2010).

En general, los métodos de investigación visual tienden hacia lo exploratorio más que hacia lo confirmatorio. Es decir, los métodos visuales no se emplean tanto en calidad de método para recoger datos de un tamaño y forma predeterminados que confirmarán o refutarán una hipótesis planteada previamente, como en calidad de método diseñado para llevar al investigador a dominios que podría no haber considerado y hacia hallazgos no anticipados previamente (Banks, 2010: 29).

Para el presente artículo realizamos una experimentación visual, en la cual se solicitó a diversos ilustradores (ver tabla 2) un dibujo con el tema de la leyenda de la Llorona, el objetivo es observar aquellos elementos son recurrentes y si estos son consistentes con la leyenda en su forma de relato.

Los artistas fueron convocados a través de sus redes sociales<sup>10</sup>, cuyos perfiles continuamente muestran sus trabajos, dando constancia de su actividad profesional constante. Los artistas están en un rango de edad de 38 a 46 años, cuatro hombres y una mujer. Todos han tenido formación en materia de ilustración ya sea en talleres o profesional universitaria.

La convocatoria fue por invitación directa y expresa para el presente artículo, se realizó a partir de su afinidad con el tema de los mitos y leyendas, así como la relación profesional con uno de los autores del

<sup>10</sup> Principalmente Facebook, Instagram, Deviant Art.



presente artículo. La solicitud fue una ilustración del personaje de La Llorona con el objetivo de obtener un relato visual que representara lo que el ilustrador conoce de la propia leyenda.

Resta aclarar que la convocatoria de ilustradores profesionales obedece al interés de analizar las imágenes de la Llorona sin que los límites técnicos representaran un factor adicional. Si bien la mayoría somos capaces de articular un relato oral no sucede de la misma manera con el dibujo, por lo que resulta beneficioso para nuestro artículo recurrir a quienes no tengan impedimentos técnicos para representar sus ideas.

Tabla 2. Ilustradores participantes de la investigación

Nombre	Breve reseña
<b>Raúl Villafuerte</b>	Egresado de la carrera de Diseño Gráfico en la Escuela Nacional de Artes plásticas de la Universidad Nacional Autónoma de México. Se desempeña profesionalmente como animador e ilustrador. Ha trabajado en producciones de animación como "La Familia Del Barrio" (MTV), "El Chapulín Colorado Animado"(anima estudios), La leyenda de la Nahualla (anima estudios). Ganador con Flame Animation Studio del "Programa de Apoyo a Productores Independientes" en su emisión 2009 otorgado por el canal 22 de México.
<b>Andrés Morán</b>	Se forma en los talleres impartidos por la editorial de historietas EGEA en México. Se destaca como uno de los principales alumnos del maestro Alberto León, de quien aprende a elaborar historietas de carácter popular destinadas a venderse en los kioscos de revistas. Ha colaborado en publicaciones como <i>Cristóbal el Brujo</i> , <i>Sensacional de Traileros</i> , la historieta de lucha libre <i>Místico</i> y <i>Rock Mortis</i> .
<b>Daniel García Garduño</b>	Egresado de la carrera de Diseño Gráfico en la Escuela Nacional de Artes plásticas de la Universidad Nacional Autónoma de México. Se desempeña profesionalmente como dibujante de cómics y animador. Ha trabajado en producciones de animación como "La Familia Del Barrio" (MTV) "El Chapulín Colorado Animado"(anima estudios) entre otras producciones de cine y tv. Produce el cómic independiente "Volt" que explora el género de superhéroes desde perspectiva crítica hacia la sociedad.
<b>Silvia Gaona Moreno</b>	"TENEBRIS" Maestra en Artes Visuales con orientación en Gráfica por la FAD-UNAM y Licenciada en Diseño y Comunicación Visual por la ENAP-UNAM. En 2013, junto al artista visual Alejandro Pérez Cruz, fundó el "Taller Gráfica Clandestina", que propone la clandestinidad para realizar eventos y propuestas plásticas que enriquezcan el sentido de la obra gráfica en México. Durante el periodo 2015-2016 fue becaria del FOCAEM dentro del programa de estímulos a la creación y desarrollo artístico del CONACULTA Estado de México, donde desarrolló el proyecto Nahuall: El Alma en el Autorretrato. Es miembro de la Asociación Mexicana de Grabadores de Investigación Plástica AC, donde desarrolla proyectos de carpetas de gráfica. Su trayectoria como artista gráfica ha sido reconocida con menciones honoríficas y con un 3er. Lugar en el 3er Festival de Arte Armadillo, (Armadillo de los Infantes, San Luis Potosí).
<b>Mario A. Vázquez Vega</b>	Licenciado en Artes Visuales de la Escuela Nacional de Artes plásticas de la Universidad Nacional Autónoma de México, se desempeña profesionalmente como docente en la misma, ilustrador y grabador. Ha incursionado en la historieta con la obra "Zapata El Hijo del Maíz" publicado en 2008 por (Ensamble Comics A.C. y el Instituto Nacional de la Juventud), ilustrador de la Novela "El destino equivocado" del David López (Círculo Rojo editorial).

De esta manera, con elementos tanto orales como visuales nos aproximamos a la complejidad del personaje de La Llorona, no obstante, antes de entrar en materia concreta, es conveniente reflexionar algunos conceptos necesarios tales como mito, leyenda y la manera en cómo esto se articula con la propuesta teórica de nuestro autor Rudolf Otto.

## La leyenda y el mito

En la leyenda se anida el carácter primigenio del mito adaptado al tiempo actual, a ello se han referido diversos autores como Claude Levi Strauss (1987) cuando apunta que el mito posee una estructura permanente que opera de manera sincrónica y diacrónica: "Ella se refiere simultáneamente al pasado, al

presente y al futuro” (p. 232). Por su parte Mircea Eliade (1968) emplea el término *mito vivo* para referirse a creencias que persisten contemporáneamente.

De lo anterior observamos que la leyenda generalmente ha sido comprendida en asociación al mito, en este sentido Gerard van der Leeuw (1964) sostiene: "El mito está eternamente presente, mientras que la leyenda se relaciona íntegramente con el pasado" (p. 401). Este autor sugiere que, a través de la leyenda, el mito se incorpora a la temporalidad humana. "La leyenda es un mito que se ha quedado colgado en algún lugar o en algún hecho histórico" (p. 401).

En efecto, leyenda y mito se emparentan, siendo su divergencia (cuando la hay) de orden temporal. Si el mito se avoca a relatos de héroes del tiempo primigenio, la leyenda acontece en lo cotidiano, ocurre en la temporalidad humana. Acerca de la leyenda Alberto de Cuenca (2009: 107) enumera cuatro características fundamentales que marcan distancia respecto al mito:

1. Refiere hechos dotados de un trasfondo histórico, pero siempre extraordinarios, fuera de lo común. Se hace referencia también a hechos de situaciones próximas, es decir, las leyendas y sus personajes son también creencias vivas (la magia de la realidad).
2. Se desarrolla en un lugar preciso.
3. Los personajes humanos son personas concretas.
4. En su recitado se opera reteniendo las ideas capitales, no aprendiendo el texto de memoria. Esto deviene en múltiples variantes de una sola leyenda.

A lo anterior hay que añadir que la leyenda, se cuenta entre generaciones, se adapta con el tiempo dependiendo el contexto y el lugar. La leyenda, por tanto, adquiere un sentido colectivo porque a través del relato se recrea –por decirlo en términos coloquiales– «en boca de todos».

La leyenda se encuentra en la oralidad, pero también se difunde en medios electrónicos e impresos, no se requiere de especialización alguna para apropiarse de ella. Su peculiar plasticidad le permite contarse como ficción y como suceso, ocurre en la cotidianidad: «en esa calle espantan», «en ese cerro hay una bruja», «en ese árbol se aparece la Llorona», a menudo recurre a personajes cercanos: «le paso a un tío» «le pasó a mi abuelo», de ahí su marcado carácter identitario.

Las narraciones populares se transmiten mediante mecanismos orales que no necesariamente se ciñen a rigurosidades históricas ni a versiones unidireccionales. En la narrativa oral se observa que las variantes de una leyenda no son unánimes, a menudo se contradicen, se entremezclan y se recrean de tal forma que las diversas variantes se congregan. La amplitud de la narración oral expande las fronteras del relato hacia contextos sociales donde se modifican y actualizan conforme al repertorio cultural de los propios narradores.

En el caso particular de leyenda de espantos, se ha avanzado anteriormente en proponer que ésta es: “Una narración y creencia de corte sobrenatural que produce miedo, susto, involucra personajes de corte siniestro y en que el héroe mítico es tergiversado hacia una condición de víctima” (Aguilar, 2019: 47).

De ahí que en la leyenda de espantos se refieran personajes asociados a la muerte y el orden nocturno: fantasmas, bestias antiguas, maldiciones. Este es el caso de La Llorona cuyas apariciones según diversos relatos ocurren durante la noche.

Pensar la leyenda como mecanismo de transmisión oral que opera dentro del contexto actual, nos sitúa en lo que Walter Ong (1988) comprende como oralidad secundaria, surgida tras el advenimiento de dos procesos comunicacionales fundamentales: la imprenta y el desarrollo de medios electrónicos de comunicación, procesos que impactan en toda comunicación oral.

En un escenario en que leyenda y mito son del mismo orden, pero divergentes temporalmente, surge un segundo problema acerca de los márgenes donde opera la leyenda respecto a la religión, considerando en todo caso, que nuestro autor de referencia (Rudolf Otto) se avoca primordialmente al estudio de la religión. Si partimos de la noción que las prácticas y representaciones de La Llorona exceden el orden de las

instituciones religiosas, tendremos que reflexionar en las posibilidades y límites de la teoría de Rudolf Otto (2016) para comprender dicho planteamiento.

La leyenda tiene la ubicuidad de operar en los límites del pensamiento religioso institucional. Mientras que la doctrina religiosa generalmente reniega la leyenda al asociarla a superchería o reducirla a tradicionalismo, la leyenda responde como un mecanismo de asimilación mucho más complejo al tener al Dios rector como fuerza protectora ante lo diabólico. Esto que se observa en algunos relatos donde el rezo o la oración funcionan como protección contra espectros de la leyenda popular.

La leyenda se integra a lo social mediante ritos de protección al alcance de todos. Con esto queremos observar que en la leyenda opera un orden que engloba a la propia institución religiosa, porque admite que, a la par de un Dios padre, su madre sagrada, y consorte de santos, existen seres, milagros y fantasmagorías. De ahí que en la leyenda se entremezclan creencias, saberes, mitos, ritos y costumbres que ocurren alternos a la religión oficializada.

Que Dios me quite ese mal de aquí, que me libre de ese mal, de esa tentación porque sabemos que Dios tiene más poder que el Satanás. Sí y ese Dios tiene más poder y claro cuando uno se recomienda a Dios pues Dios también actúa en ese momento para que nos libre de los malos espíritus. Y si ya al momento con dos, tres palabras ya nos suelta, y si no pues de veras que uno ahí lo maltrata. Cuando uno no hace oración lo maltrata (Sr Celso, Tapalapa, 2016).

Otras formas de protección consisten en marcar el signo de la cruz en el camino, o bien persignarse haciendo la señal de la cruz.

Le cuento la historia de lo que vi y me dice: ¡Es que no te persinaste! cuando uno no se persina sueña uno feo, tiene pesadillas, que no te acuestas con Dios pues, te estás acostando con el diablo, por eso te vino a visitar. Ese es el Sombrerón me dice (Sr. Rosel Hernández, Chiapa de Corzo, 2016).

El estudio de Carlos Navarrete (2014), *Oraciones a la Cruz y al Diablo*, compila oraciones transmitidas por rezadores, cantores y cuenteros tradicionales del estado de Chiapas. En dicha investigación se evidencian relaciones entre conceptos en apariencia opuestos como la Cruz y el Diablo pero que en realidad operan como una dualidad complementaria.

Estas fuerzas se manifiestan a través de dos poderes personificados en Dios y el Demonio, en cuyo desdoblamiento se encausa la contradicción que rige el mundo. Son enemigos irreconciliables en apariencia, no en esencia; la contradicción entre opuestos positivo-negativo no es violenta y a menudo se entrecruzan. La naturaleza es el todo, «como el centro de dos espejos enfrentados (Navarrete, 2014: 27).

En relación con el advenimiento de la religión como practica espiritual hegemónica, la leyenda (mediante el relato) establece un tiempo mítico ancestral donde además de no haber religión, no había luz eléctrica, ni caminos pavimentados y la presencia de seres legendarios era mayor (aunque al día de hoy sigan presentes). Esta dicotomía que se simboliza en luz/oscuridad otorga a la religión cierto carácter civilizatorio, sin embargo, no borra a la leyenda, más bien la sublima al dotarle un carácter preexistente.

Pues ahorita ya están pavimentadas las calles, pero hace años no estaban pavimentadas, y entonces, por eso es que se escuchaba el golpeteo de la llanta de la carretilla [de San Pascualito, personaje de leyenda] (Rosel Hernández, Tuxtla Gutiérrez, 2016).

En Tapachula cuando yo era chamaco no había televisión, todo era radio, no había mucho alumbrado público, había nomás los focos de la casa, algunos focos en la calle, así que era el espacio ideal para las charlas (David Díaz, Chiapa de Corzo, 2015).

En el 55, 60 más o menos. No había ni luz, con qué trabajo entraba un carro. (Josué, Chiapa de Corzo, 2016).

Es que antes no había muchas casas, habitantes, no había nada, de aquí a Chenalhó, vas a Chenalhó caminando, no hay casas, vas a encontrar casas, pero hasta allá otra casa, hasta allá, ahorita ya puras casitas en el pueblo, ahorita ya no, ya no lo han visto, [al Sombrerón] ya no. Antes como de eso ya hace como 40 años, como 50 años lo contaban (Sr. Mariano, Yabteclum, 2016).

Para autores de corte evolucionista como el propio Rudolf Otto (2016), lo anterior (mitos, leyendas, consejas) constituye un origen *primitivo* de toda gran religión (cristianismo, islam, budismo). En la perspectiva del presente artículo ya sea por convicción ética, como por enfoque teórico, tomamos distancia de tal noción al considerarla más que nada producto de su época, en todo caso nos alumbra el hecho de que el propio autor reconoce afinidades entre mito, leyenda y religión al proponer el concepto de *numinoso* como eje común de estos estadios. Otto explica que los dos aspectos centrales de lo *numinoso*: el misterio tremendo y el sentimiento de criatura se observan en relatos míticos al igual que religiosos, en lo que Rudolf Otto (2016) comprende mediante el concepto de *pavor demoniaco*.

El *pavor demoniaco* echa mano de personajes fantásticos, milagros y toda una suerte de misterios siniestros e inquietantes que conservan □ aun siendo considerados por Rudolf Otto como fantasías □ el carácter potente del numen:

De este sentimiento y de sus primeras explosiones en el ánimo del hombre primitivo ha salido toda la evolución histórica de la religión. En él echan sus raíces lo mismo los demonios que los dioses y todas las demás creaciones de la «apercepción mitológica» y de la fantasía que materializa y da cuerpo a estos entes (Otto, 2016: 62).

Del *pavor demoniaco* y sus relatos como es el caso de las «historias de fantasmas», «cuentos de duendes y de espectros», Otto refiere que hay en éstos algo de terror fantasmal que por su naturaleza *numinosa* resulta fascinante:

...el encanto singular del espectro es que es un *mirum*, y como tal opera por sí mismo sobre la imaginación y despierta en ella el interés de un incentivo poco común y una intensa curiosidad. Él mismo, esa singular y extraña atrae a la fantasía... por ser un prodigio, un «absurdo», una cosa como, en realidad no hay otra, por ser «absolutamente heterogéneo» con el hombre; en suma, por no pertenece al círculo de nuestra realidad, sino a otro distinto, que provoca en nuestro ánimo un interés irrefrenable (Otto, 2016: 83).

Aún sin darle el peso específico que en nuestra perspectiva la leyenda requiere, a menudo encontramos a Rudolf Otto (2016) casi fascinado por leyendas populares, y cuentos de fantasmas a los que en diversos pasajes de *Lo santo...* describe con pluma magistral: “Es este un terror de íntimo espanto, que nada de lo creado, ni aun lo más amenazador y prepotente, puede inspirar. Palpita en él algo de terror por los fantasmas” (Otto, 2016: 61).

De los espectros que nos competen como son los personajes de las leyendas, Rudolf Otto admite que como *mirum* posibilitan la comprensión y análisis del origen de lo *numinoso*, en nuestro caso particular nos posibilita comprender tales leyendas desde una perspectiva teórica de la cual obtendremos una mirada novedosa que complementa la aproximación de carácter etnográfico y posibilita el encuentro con otros aspectos de la leyenda comprendidos en su imagen.

## La Llorona

La leyenda de La Llorona se difunde por el territorio que abarca México y se expande hasta Centroamérica, desde nuestra perspectiva esta amplitud es posible gracias a la gran cantidad de variantes por medio de las cuales se adapta a circunstancias locales. En términos generales, la versión que tiene más difusión encuentra su origen en los estados del centro de México, la leyenda describe una mujer entiempos de la Nueva España que por despecho amoroso asesinó a sus hijos ahogándolos en el río, desde entonces su alma en pena se aparece por las noches en callejones, ríos, lagos, puentes y espacios liminales de carácter siniestro, su lamento descrito como aterrador es característico en este personaje. Al grito de ¡Ay mis hijos! La Llorona lamenta eternamente su crimen.

Podemos ubicar la leyenda de La Llorona en tres temporalidades distintas; por un lado, un origen prehispánico asociado a divinidades como Coatlicue y Cihuacóatl, esto se ha estudiado por diversos autores, entre los que podemos destacar se encuentra el trabajo de los cronistas Artemio del Valle Arizpe (1936) y Luis Gonzáles Obregón [1937], (1988), el propio filósofo y lingüista Miguel de León Portilla (1999) menciona al respecto relacionándola con el *sexto agüero* de los presagios funestos de la caída del imperio

Mexica relatados por Sahagún: “Muchas veces se oía: una mujer lloraba, iba gritando por la noche, andaba dando grandes gritos: - ¡Hijitos míos, pues ya tenemos que irnos lejos! Y a veces decía:

-Hijitos míos, ¿a dónde los llevaré?” De este presagio, León-Portilla refiere en una nota al pie de página de su obra *La visión de los vencidos* (1999): “El texto parece referirse a *Cihuacóatl* que gritaba y lloraba por la noche. Es este uno de los antecedentes de la célebre “Llorona” (León, 1999: 4).

Una segunda temporalidad es en el periodo colonial donde la leyenda de La Llorona adquiere forma de alma en pena a la vez que se urbaniza al tener por escenarios los callejones de las ciudades novohispanas. En lo que identificamos como un tercer momento, esta leyenda se revitaliza en relatos contemporáneos donde los escenarios son lugares cotidianos, ocurre a miembros de la familia, amigos y conocidos e incluso se cuentan relatos en primera persona.

Diversos autores han recopilado versiones de la leyenda de La Llorona a lo largo de toda su complejidad histórica, es el caso de Rogelio Álvarez (2005), así como Alberto Círigo Villagómez y Marisol González Olivo (2009). Una relación completa de relatos en torno a La Llorona se encuentra en “La Llorona. Nuestra madre” de Marcia Trejo Silvia (2013).

En la dimensión que nos atañe, tenemos que la leyenda de La Llorona se recrea cotidianamente, La Llorona se entremezcla con otras leyendas de tal manera que se asocia a diversas denominaciones locales, para darse una idea clara de esto, se observa que tan solo en nuestros estudios de campo se encontraron las siguientes denominaciones: Tisigua, Tishanila, Matlancihua, Yehualcíhuatl, Llorona, Mala Mujer, Malora, Xpak'intè, Pajkintaj. En revisión bibliográfica aparecen otras variantes nominales, por ejemplo, en el caso de Centroamérica se habla de la Ziguanaba. Estas variantes coinciden en describir una aparición en forma de mujer de extraordinaria belleza, que posteriormente revelará su carácter demoníaco como *Misterio Tremendo*, en ocasiones su rostro no es visible al estar cubierto por un largocabello, un velo de novia, o la oscuridad de la noche, generalmente porta un vestido blanco entallado y cabello suelto, durante la noche La Llorona, Tisigua, Tichanila o Mala mujer provoca a algún varón que bajo efectos etílicos ha perdido el rumbo, éste la sigue con la promesa de placer sexual, el hombre es engañado como resultado de tal encuentro. Se dice que el hombre muere ahogado en un río o recibe alguna clase de mal, así lo refieren testimonios recogidos en el trabajo de campo en el estado de Chiapas, México:

La Tisigua es la famosa Mala Mujer pues, que cometió un pecado, o se murió mala, mató a alguien o le hicieron algo. Y tú la encuentras, y tú la encuentras a medianoche convertida en una hermosa mujer, tú la encuentras a medianoche, y te da jalón la chava ¿no? te dice: ¿pues qué? y tú la empiezas a seguir, en aquellas calles oscuras ¿no?, a seguir, hasta que se metía al río, empezaba a nadar o a media agua ¿ves? está hondo, entonces se metía pues al río, cuando venían a ver se volteaba en una gran clavera... y ¡oonde de ahí te llevaba ahí al río! Esa es la Tisigua. (David Díaz, Tuxtla Gutiérrez, 2015).

La Mala Mujer se viste siempre de blanco, muy elegante, pero los que lo han visto aseguran que sí es guapa, es alta, bien vestida, pero las patas las tiene al revés, y no los tiene como gente sino los tiene como un gallo pero las patas volteadas, pero eso sí, por eso cuando se anda de noche y si se encuentra una mujer que esté vestida de blanco y que esté guapa y caminando solita hay que tener cuidado de verle las patas, porque si las tiene volteadas seguro que es el diablo (Rosel Hernández, Tuxtla Gutiérrez, 2016).

En nuestra perspectiva consideramos que dichas versiones son integrales a un solo cuerpo mítico donde La Llorona es el personaje central, al tomar en cuenta que los propios entrevistados comúnmente se refieren a estos númenes de manera indistinta, si bien, en ocasiones emplean el nombre de La Llorona, a veces refieren para el mismo personaje otras denominaciones como la Tisigua, Mala Mujer, etcétera.

Esta postura, contrasta con miradas que preponderan el carácter local de la leyenda o defienden su recorrido histórico situando las variantes en lugares y tiempo específicos. Consideramos que la eficacia de la leyenda es precisamente su capacidad de compenetración y aglutinación lo que le brinda un carácter simbólico por encima de su fijación histórica o geográfica. Tal como postula Levi-Strauss, “el mito es el conjunto total de todas sus variantes” (Lévi-Strauss, 1987: 240), es decir, que éste no obedece a una sola versión privilegiada.



Las diversas variantes locales tienden a desdibujar el carácter histórico de La Llorona porque integran bajo su amplitud otros númenes que, si bien, en términos históricos no se corresponden con la leyenda en cuestión, dentro de la oralidad se fusionan de tal manera que las versiones concurren como una sumatoria, más que como diferenciación, con ello alcanzamos una complejidad que no puede ser observada solamente en la historicidad, sino en correlación con la oralidad y su contexto, cuestión fundamental para observar todo fenómeno santo en su contemporaneidad.

### El carácter *racional* de La Llorona

Comprendemos el carácter racional de la leyenda de La Llorona como aquella instancia que reflexiona en las características y los porqués de la leyenda, es decir, que no ahonda hacia el sentido emocional o pavoroso del encuentro con lo numinoso, sino que más bien dibuja cierta taxonomía del personaje, tal como se analiza en el siguiente ejemplo:

Es una representación de La Llorona que yo la confundo con la Tisigua igual. Que es la que llama a los hombres, los pierde, y pues ya cuando sale el sol, y quedan un poco traumatados (Margarita, Chiapa de Corzo, 2015).

En el ejemplo anterior observamos una explicación *racional* de la leyenda, por un lado confronta de manera reflexiva dos variantes similares, una de ellas La Llorona y la otra conocida localmente como la Tisigua, explica el poder fascinante de lo *numinoso* observable en la oración «llama a los hombres», el efecto del *sentimiento de criatura* se observa en las frases: «los pierde», «quedan poco traumatados»; por otro lado, lo siniestro asociado al orden nocturno se deduce por oposición de la oración «pues ya cuando sale el sol».

De manera similar vemos la forma en que otro participante de manera *racional* refiere:

Lo que tiene es que esta mujer cuando se le va acercando, ella posiblemente atrae a los borrachines o se va acercando al hombre y se va viendo que igual tiene los pies al revés, pero tiene una cara de caballo o algo así (Brayan, Chiapa de Corzo, 2015).

En el párrafo anterior se introduce un elemento central: *la víctima*, caracterizada como un *borrachín* o ebrio. El elemento *numinoso* se constituye de imágenes potentes que exhiben el *tremendo misterio* como son: la revelación de los pies al revés, y una divergente cara de caballo. Estas imágenes explican la alteridad del numen en donde una vez roto el misterio, la belleza se torna en monstruosidad.

En otro relato recogido en Comitán de Domínguez, Chiapas, el cronista Amín Guillén refiere una variante denominada “La esquina blanca”, en referencia a una esquina de un barrio de la ciudad donde se cuenta que aparece la Llorona:

Aparecía la reina de blanco y empezaba a llamar a los transeúntes borrachos, trasnochados que pasabanahí, no faltó pues quien iba a abrazarla porque los invitaba al sexo, al abrazo, los llamaba y llegaban, abrazaban a esta mujer y aquí está lo importante de esta leyenda, de que se dormían en sus brazos de su amante y amanecían en un cerro abrazados de un nopal y con espinas, y ahí amanecían con la borrachera del otro día ...como que ahí está muy incisivo el castigo, la maldad, lo malo de este personaje que es La Llorona (Amin Guillén, Comitán de Domínguez, 2016).

En este relato se explica otra característica de la leyenda de La Llorona como es la maldad del personaje. Aunado al engaño, La Llorona castiga de forma cruel a quien ha caído en su encanto.

En el objetivo de entender la leyenda como fenómeno santo, observamos que ésta se adhiere a diversos relatos, es decir, que requiere de mecanismos y estructuras *racionales* para preservarse y transmitirse, Rudolf Otto (2016) no es ajeno a esta idea, de hecho, considera que, entre lo *racional* e *irracional* de lo santo debe, en todo caso, primar una “*sana y bella armonía*” (p.259).

Mientras que –como todo mito– La Llorona recurre a mecanismos narrativos, esta leyenda no pierde su virtud como creencia en tanto que se mantiene en la emocionalidad de quienes cuentan el relato. Estanoción

se aborda en el apartado siguiente y alude a lo que Rudolf Otto (2016) considera la contraparte *irracional* de la creencia sacra.

## El carácter *irracional* de la leyenda de La Llorona

En este apartado reflexionamos en el carácter *irracional* de la leyenda de La Llorona, observamos expresiones de pavor *demoniaco* en diversos relatos donde el *sentimiento de criatura*, referido anteriormente, se observa como el miedo paralizante ante la presencia del espanto, la incapacidad de reacción para librarse del embrujo, la caída en pecado, revelado en frases coloquiales como: «¡se lo ganó!», «¡se lo llevó!», «¡lo perdió!».

En un relato resultado de una charla casual en torno a las leyendas, se habla del característico llanto de La Llorona, observamos que ese lamento no es un clamor cualquiera, por sus características numinosas se impone horrorosamente en el ánimo de quien lo escucha:

Y sí, una vez estaba lavando mi mamá y lo oímos que gritó ahí en el caracol, cuando entra mi mamá ¡perohasta con el pelo parado! y el otro grito ya lo oímos hasta allá por la primavera. Y ahí se fue.

—¿Y qué gritaba?

¡Aaayy mis hijos!

(pregunta) ¿Y qué sentiste?

¡Ay Jesús!, hijo, ¡horrible! Nos abrazamos con la Terita (diminutivo de Teresa) porque yo dormía yo con ella, la Rosita dormía allá en su cuarto de mi mamá con mi papá, era la criatura pué'. Y la Zoila quedormía sola. ¡Jesús!, estaba que pegaba de gritos. Se abrazó de mi mamá y todo, pero no es nada, es el aire, no ves gente, no ves nada, porque Alfonso abrió la ventanita y no miró nada, nada... ¡nada! (Francisca (Chica) García+, San Cristóbal de las Casas, 2015).

En diversas variantes de La Llorona los relatos describen algún tipo de revelación de su misterio y tras lo cual exhibe algún tipo de monstruosidad, revirtiendo su condición de belleza hacia una alteridad monstruosa, por ejemplo, se muestra como una bestia con cara de caballo, se torna en una enorme serpiente, o bien exhibe que debajo de su largo cabello hay un rostro cadavérico (Trejo S. 2013), eventotétrico que invade de estupor a la víctima quien pasa súbitamente del deseo sexual a la parálisis horrorizante, lo *monstruoso* —explica Otto— es “justamente lo misterioso en grado superlativo” (Otto, 2016, p. 168).

Como *misterio tremendo*, La Llorona —aun en su rasgo de monstruosidad— posee un carácter fascinante, esto lo hemos reflexionado con Rudolf Otto (2016) cuando refiere que la imagen de los cuentos de fantasmas es como un *mirum*, si bien el *misterio tremendo* es de *absoluto inalcanzable*, las imágenes pueden referir, aun de manera indirecta, algo del misterio del numen. En esto, el autor —incluso atribuyendo ingenuidad a las imágenes míticas—, no deja de referirlas como la presencia de santo, o lo *numinoso*:

...así como decíamos antes que lo *tremendum* pasa ser estímulo de la imaginación, de modo que ésta escogelo terrible por medio expresivo o inventa formas originales terribles para representarlo, así también lo misterioso se transforma en el incitante más fuerte de la ingenua fantasía que aguarda el milagro, lo inventa, losiente, lo refiere. Lo misterioso es también acicate nunca fatigado para la inagotable invención de cuentos, mitos, consejos y leyendas... (Otto, 2016: 146).

## La representación del *misterio tremendo*, en La Llorona. Modos indirectos de lo numinoso

Entrando en materia de la representación del *misterio tremendo*, a continuación, presentamos una colección de cinco imágenes referentes a La Llorona, estas fueron elaboradas *exprofeso* para la presente investigación, el ejercicio consistió en solicitar dibujos o ilustraciones a cinco artistas profesionales, el ejercicio fue crear una imagen del personaje de leyenda. Con esto se busca una suerte de relato o testimonio dibujado donde se observan desde la perspectiva visual los aspectos *racional* e *irracional* en la leyenda de La Llorona, sirvan

estos también como un fascinante *mirum* donde podemos comparar las relaciones entre oralidad y visualidad en lo santo.

Como antecedente proporcionado por el mismo Rudolf Otto (2016) tenemos que reconocer ciertas formas indirectas de expresión de lo numinoso que remite hacia el universo de las imágenes e íconos, menciona imágenes de dioses antiguos: "el aspecto terrible y espantoso de las imágenes o descripciones primitivas de los dioses"(p. 143).

A continuación, se observa la colección de cinco imágenes con las que reflexionaremos en la representación del misterio tremendo e imagen de la leyenda de La Llorona.



Figura 1

"La Llorona"  
Raúl Villafuerte 2017



La figura 1 elaborada por el ilustrador y animador Raúl Villafuerte representa La Llorona con su característico vestido blanco, la habilidad artífice del autor logra con economía de líneas que el personaje asemeje flotamiento ya que no se distingue piso ni pies y el ropaje se muestra en movimiento.

El rostro tiene en negro las cuencas de los ojos signo que remite a la muerte o la siniestralidad. Destacamos un elemento simbólico primordial de La Llorona como es el cabello negro y largo, que también está presente en otras imágenes del corpus visual de esta investigación.



Figura 2

“La Llorona”  
Andrés Morán Muñoz 2017

La figura 2 elaborada por el historietista Andrés Morán representa a La Llorona en un estado desesperado, cierto dramatismo se logra formalmente mediante el encuadre al que el dibujante recurre al representar al personaje desde una perspectiva elevada, lo que la hace imponente. Morán ubica a La Llorona en un

escenario conformado por tres planos lo que resulta en una ilustración de carácter realista por lo que sugiere un encuentro casi frontal con el personaje.

La imagen es tensa porque bastaría con que La Llorona bajase la mirada para descubrir al espectador. Mientras tanto, se puede contemplar con cierto estupor y fascinación el detalle de la mortandad de su rostro, asimismo se observa el dolor del personaje tanto por su gesto facial surcado por lágrimas, como por sus manos tensas. Anímicamente Morán nos presenta una Llorona al borde del colapso, quizás a punto de soltar su lamentable grito. En esta ilustración el *tremendo misterio* se revela de manera casi filmica, observamos sin ser observados, más no por ello la imagen —en tanto referencia a lo *numinoso*— deja de ser inquietante. El orden nocturno es representado mediante la luna.



Figura 3

"La Llorona"  
Daniel García Garduño 2016

La figura 3 elaborada por el ilustrador, animador y creador de comics Daniel García Garduño, es la imagen de la colección que probablemente representa de manera más directa y frontal el carácter monstruoso de La Llorona, el trazo es mediante líneas, logra un retrato con mensaje potente: el rostro de La Llorona. Al igual que las otras imágenes, ésta asocia a La Llorona con la muerte por medio de los ojos y la dentadura cadavérica. Aquí, el *tremendo misterio* de La Llorona se muestra casi como un muerto que acaba de salir



del ataúd, evidenciados por sus ropajes desgarrados, incluso sus ojos siguen cerrados, al igual que las otras imágenes, se presenta el signo del cabello largo y negro.



Figura 4

"La Llorona"  
Silvia Gaona 2017

La figura 4 elaborada por la artista Silvia Gaona incorpora elementos que profundizan en el aspecto emocional de la leyenda. Formalmente emplea signos como la luna que remite al orden nocturno, al fondo de la imagen un escenario natural es representado mediante montañas, las aves negras posadas en cráneos humanos son, en esencia, un signo siniestro de carácter gótico.

En su aspecto anímico y emocional observamos el gesto de dolor en el rostro de La Llorona reforzado por las lágrimas que surcan su cara afligida; un gran corazón en forma de cresta corona al personaje, mediante este símbolo se observa que esta leyenda tiene un fuerte elemento emocional como es el dolor sentimental.

El mecanismo visual-sentimental de la imagen de Silvia Gaona funciona formalmente al contraponer elementos como: La Llorona quien porta su particular vestido blanco, la muerte representada en las costillas visibles del personaje y en los cráneos del primer plano, otros elementos se suman como las aves negras (mal presagio, mal agüero), un corazón representado como órgano que alude a un dolor emocional correspondido por un gesto de dolor en el personaje.

En la imagen lo elementos se constituyen como un tremendo misterio, que además toca aristas que hasta el momento no se habían vislumbrado en el presente artículo: el carácter santo del dolor. Un dolor numinoso.

En la figura 5, última de la serie, se observa a La Llorona elaborada por el artista visual Mario A. Vázquez



Figura 5

"La Llorona"  
Mario A. Vázquez Vega 2020

Vega. Se trata de una evocación de las leyendas del México colonial. Como escenario se tiene un callejón, lugar predilecto para esta aparición en los relatos referidos por cronistas como Artemio del Valle Arizpe (1936) y Luis Gonzáles Obregón (1988). El vestido es elegante, de otro tiempo, va cubierto con un velo que de alguna manera enmascara su misterio, pero deja entrever un rostro fantasmal. Al igual que en otras imágenes los ojos negros operan como mirum del carácter numinoso del personaje.

El gesto corporal de La Llorona es de dolor, se toma el pecho en señal de despecho mientras con dificultad casi enfermiza se apoya en las paredes del callejón para caminar. Si no fuera por su fantasmal rostro pareciera una dama de la época en situación de fragilidad, sin embargo, otros signos como la noche cuya insignia es la luna llena, y la calavera adosada a uno de los muros como representación de la muerte, confieren a la imagen su carácter siniestro y sobrenatural.

## Resultados

Si acaso se puede ubicar nuestra colección de imágenes en un género específico propondríamos decir que se enclavan en el género del terror, evocan símbolos como la muerte. La noche es el marco natural del personaje de La Llorona. Aquí hemos de observar que los símbolos (noche), adquieren un carácter siniestro en correlación con otros (la muerte), en dicha correlación el creador yuxtapone significados para dar origen al *mirum*, a la vez que relata la leyenda mediante el mecanismo narrativo de la ilustración.

Las imágenes se construyen mediante un mecanismo *racional*, casi descriptivo que apela a características generales de La Llorona (cabello largo, ropajes de otra época, gesto de dolor, ojos negros referentes a la muerte). Las distintas representaciones se ajustan a formas donde el espectro, el fantasma o el numen se encarna de manera arquetípica, de tal manera que su imagen remite esencialmente a un ente *numinoso*, o fantasmal aun si no se tuviera referencia concreta del personaje. Otro elemento a destacar es el detalle de los ojos, representado de manera similar en las cinco imágenes, donde operan como símbolo de la muerte.

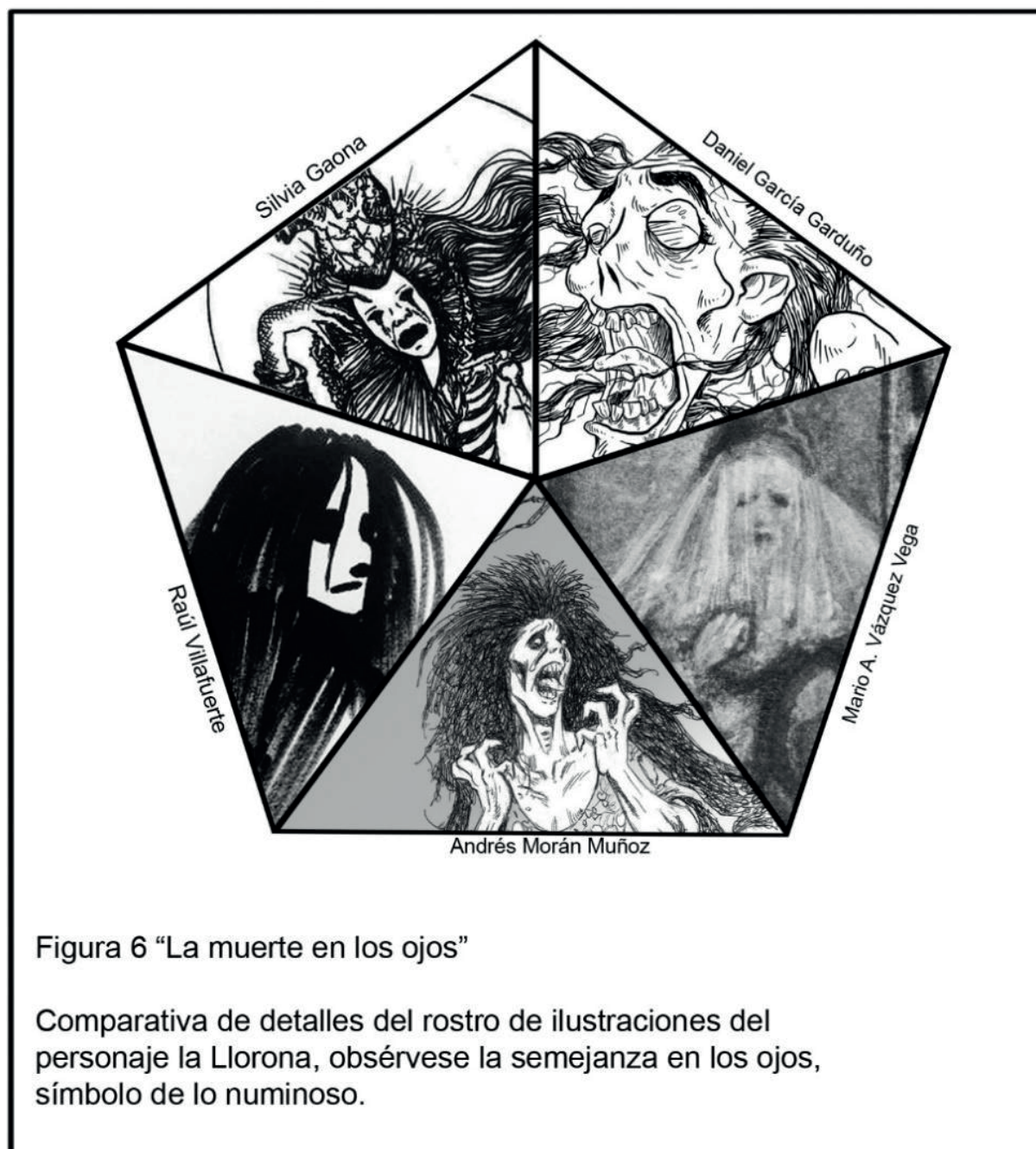
Hay ciertas correlaciones con los relatos que confirman la imagen de la Llorona como un numen de aspecto tenebroso cuyo misterio revelado es de carácter monstruoso o fatal, lo que principalmente se observa en la figura 3 y en la 2. No obstante se encontraron más semejanzas en lo simbólico (muerte, noche, cabello largo) que en lo narrativo. Lo anterior se observa en las ilustraciones centradas en la imagen del personaje, mientras que las narraciones orales enfatizan además en la víctima y los lugares donde ocurren las apariciones.

Se puede observar en el conjunto de imágenes que aquellas que muestran la vestimenta desgarrada (Figuras 2,3 y 4) apelan a la imagen de La Llorona en su forma monstruosa, ya sea cuando su misterio se revela y es causa de susto, espanto o estupor, o bien cuando ésta profiere su espantoso lamento. Cuando se ilustra al personaje en su carácter fantasmal (figura 1 y 5) la vestimenta se representó sin roturas.

Al observar la colección de imágenes en conjunto (ver figura 6), es posible encontrar relaciones con númenes de otras culturas. Cuenta el mito que Medusa petrifica con su sola mirada a todo aquel que le mira directo a los ojos, circunstancia que Jean Pierre Vernant (1986) describe metafóricamente como *la muerte en los ojos*. Más allá de la petrificación que resulta mortal, ante la mirada de Medusa, la víctima se enfrenta al *pavor demoniaco* que queda plasmado por siempre en la gestualidad de la víctima así convertida en estatua del horror. De esto que la mirada tanto en Medusa como en La Llorona tengan implícito el *tremendo misterio* del numen cuya revelación resulta fatal.

La *muerte en los ojos* de la que nos cuenta Vernant (1986), es resultado del encuentro horizontal y sin intermediación alguna con lo divino, resulta fatal porque se contraponen directamente dos órdenes cuya asignación se encuentra normada por los dioses, a decir, una oposición entre lo civilizado comprendido como el espacio social, y lo salvaje, entendido como el páramo donde las potencias divinas habitan: lagos, montañas, callejones, cruces de caminos, lugares numinosos donde los espantos aguardan.





De todo lo anterior que la complejidad de la leyenda de espantos la configure como un fenómeno santo que continuamente se revitaliza sin perder su hondura histórica, son relatos que persisten como creencia al grado que operan de manera liminar a las creencias religiosas institucionalizadas.

## Conclusión

En este artículo hemos revisado la leyenda de La Llorona a través de un mecanismo fenomenológico que reflexiona en lo santo como experiencia, con esto nos aproximamos al estudio de las manifestaciones populares considerándoles en igualdad de importancia con otro tipo de expresiones santas o religiosas, particularmente el caso de la leyenda cuyo tratamiento investigativo ha sido generalmente compilatorio, o etnográfico.

Como propuesta teórica empleamos los mecanismos racional e irracional en lo *numinoso* para comprender un fenómeno como la leyenda de espantos, aportando una perspectiva que permite comprender aspectos en la leyenda en su virtud de creencia y en su cualidad de narración e imagen.

Por otro lado, propusimos una metodología que contempla la imagen como insumo investigativo y no meramente como ilustración de un texto determinado. Con esto nos abrimos a la intersección de disciplinas donde la diversidad en las fuentes investigativas expande los campos de conocimiento para contemplar objetos de estudio multidimensionales, en los que la realidad compleja apunte hacia una comprensión holística, humana.

Creemos que bien vale el esfuerzo de expandir horizontes disciplinares si acaso el objeto a estudiar así lo requiere, en el caso particular del presente artículo, nos fue necesario integrar conceptos provenientes del estudio de las religiones, antropología, arte, diseño y comunicación.

El tono de nuestro autor (Rudolf Otto) nos llevó a lugares profundos de la psicología religiosa, donde discernimos y enfrentamos miedos y anhelos expresados de formas misteriosas y fantásticas. Observamos junto con el autor, la manera en que el *misterio tremendo* reverbera en las narraciones orales e imágenes de La Llorona, las formas de horror y *pavor demoniaco* expresadas de manera indirecta en los relatos e imágenes.

En nuestro recorrido por el personaje de La Llorona comprendimos que, hay una persistencia vigente en prácticas, creencias y relatos. Éstos nos enseñan, nos advierten, nos reflejan. Sería un error concebir las leyendas como creencias finiquitadas e inamovibles, enclaustradas en el mero tradicionalismo.

Los relatos e imágenes de lo *numinoso* se conforman de creencias, historia, identidad nacional, tales circunstancias no son del todo observables mediante la propuesta teórica de Rudolf Otto (2016) y de hecho reconocerla representa cierto deslinde de su posicionamiento ético y teórico. Con influencia de Kant, Otto fundamenta que lo santo es un sentimiento a priori e interno en todo ser humano, con esto deja de lado el aspecto cultural, no obstante, si ampliamos la mirada se comprenderá que la leyenda se ancla y actualiza su repertorio materializando el mito en la memoria colectiva.

De lo anterior, concluimos que en la leyenda se reconoce un carácter identitario que potencia al fenómeno santo más allá de la contención ejercida por la religión institucionalizada. Por lo que la leyenda además de operar en un plano tradicional, se proyecta según su contexto como forma política y resistencia cultural, que al preponderar narrativas locales participa de la identidad de quienes la cuentan.

La leyenda es capaz de incorporar elementos tanto locales como globales para resignificarlos. Lo anterior puede observar en las imágenes de la Llorona elaboradas por los artistas invitados, que constantemente apelan a representaciones arquetípicas y reconocibles en diversos contextos, a decir, el gran símbolo de “la muerte en los ojos” y su carácter espectral, fantasmagórico.

La complejidad de la leyenda se explica reconociendo el contexto de una oralidad secundaria propuesto por Walter Ong (1988) permeada por tecnologías como la imprenta y los medios de comunicación masivos. De esta manera, las imágenes de La Llorona retoman –además de la leyenda misma–, un conjunto de formas y arquetipos racionales que hacen posible su representación y transmisión.

Con lo anterior se abren preguntas para futuras investigaciones, particularmente en el modo en que la narrativa popular plantea su poder performativo al articularse a un complejo de construcciones identitarias más amplias, que además exceden el marco de las creencias religiosas, donde incluso resuenan mitos de origen, saberes locales, y donde el relato resulta indispensable.



## Referencias

- Aguilar M. (2014). “Nuevas tecnologías en la formación: asesoría hipermedia con plumas digitales como apoyo para la elaboración de trabajos de investigación”. En Santiago R. Acuña, María A. Gabino Campos y Consuelo P. Martínez Lozano (coord.). *Multiculturalidad, imagen y nuevas tecnologías*. Fragua. Madrid. 263-291.
- Aguilar, F., Escobar, S., Villegas, L., Olivares, E. (2013). *Las aventuras de Cristóbal el Brujo. Un recorrido por las leyendas de México*. Ensamble Comics A.C. Ciudad de México.
- Álvarez, J. (2005). *Leyendas mexicanas. Antología*. Everest. Madrid.
- Banks, M. (2010). *Los datos visuales en investigación cualitativa*. Morata S.L. Madrid.
- Brandomín, J. (1990). *Leyendas y Tradiciones Oaxaqueñas*. Edición del Autor. Oaxaca.
- Cirigo, A. y González, M. (2009). *Leyendas mexicanas de todos los tiempos (prehispánicas, coloniales y modernas)*. Editores Mexicanos Unidos. México.
- De Gutiérrez, A. (1993). *Comp. Tradición Oral de El Salvador*. CONACULTA. El Salvador/México.
- Del Valle, A. (1936). *Historias de vivos y muertos: leyendas, tradiciones y sucesos del México virreinal*. Biblioteca Nueva. México.
- Eliade, M. (1992). *Lo Sagrado y lo Profano*. Labor. Barcelona.
- \_\_\_\_\_(1968). *Mito y Realidad*. Ediciones Guadarrama. Madrid.
- Flick, U. (2004). *Introducción a la investigación cualitativa*. Morata. Madrid.
- Fraijó, M. (2016). “Estudio introductorio” en Otto, Rudolph. *Lo Santo. Lo racional y lo irracional en la idea de Dios*. Alianza. Madrid.
- González, L. (1988). *Las calles de México: Leyendas y sucesos, vida y costumbres de otros tiempos*. Porrúa. México.
- León, M. (1999). *La visión de los vencidos Relaciones indígenas de la Conquista*. Universidad Nacional Autónoma de México. Ciudad de México.
- Lévi-Strauss, C. (1987). *Antropología estructural: mito, sociedad, humanidades*. Siglo XXI. México.
- Marx, K. (1968). “Introducción para la crítica de la Filosofía del derecho de Hegel” en Hegel, G.W.F. *Filosofía del derecho*. Editorial Claridad. Buenos Aires.
- Olavarria, M. (1989). “Lo numinoso: mito y terror”. *Alteridades*. Anuario de Antropología. UAM México. 113-122.
- Ong, W. J. (1988). *Oralidad y escritura. Tecnologías de la palabra*. Fondo de Cultura Económica. México.
- Otto, R. (2016). *Lo Santo. Lo racional y lo irracional en la idea de Dios*. Alianza. Madrid.
- Pineda del Valle (2011). En Cruz Coutiño, Antonio. *Mitología maya contemporánea*. UNACH-CAEME. Tuxtla Gutiérrez.

Salazar, S. (1999). *Cuentos de barro*. CONCULTURA. San Salvador.

Trejo, S. (2013). *La Llorona, nuestra madre*. Trillas. México.

Vernant, Jean-Pierre. (1968). *La muerte en los ojos, figuras del Otro en la antigua Grecia*. Gedisa. Barcelona.

### **Entrevistas**

Celso, Tapalapa. Chiapas 2016.

David Díaz, Chiapa de Corzo. Chiapas. México, 2015.

Rosel Hernández, Tuxtla Gutiérrez. Chiapas. México, 2016.

Amin Guillén, Comitán de Domínguez. Chiapas. México, 2016.

Francisca (Chica) García+. San Cristóbal de las Casas. Chiapas. México, 2015.

## ¿ES POSIBLE MITIGAR SESGOS COGNITIVOS<sup>1</sup>?

UN ANÁLISIS CRÍTICO DE DIVERSAS PROPUESTAS PARA REDUCIR EL SESGO A MI FAVOR

*IS IT POSSIBLE TO MITIGATE COGNITIVE BIASES?*

*A critical analysis of different proposals for reducing myside bias*

*É POSSÍVEL MITIGAR VIESES COGNITIVOS?*

*Uma análise crítica de diferentes propostas para reduzir o viés a meu favor*

**Alba Massolo**

(Universidad Católica de Córdoba)  
albamassolo@gmail.com

**Mariel Traversi**

(Universidad Católica de Córdoba)  
mariel.traversi@gmail.com

Recibido: 19/11/2020

Aprobado: 01/06/2021

### RESUMEN

En este artículo, realizamos una evaluación crítica de diferentes estrategias empleadas para reducir el sesgo a mi favor. Este sesgo cognitivo ha sido caracterizado como la tendencia a producir y evaluar argumentos de manera influenciada por las propias opiniones e ideologías. Proponemos una división en dos grandes conjuntos de estrategias: un primer conjunto, vinculado a las teorías de los procesos cognitivos duales, que apunta a mejorar las habilidades del razonamiento en solitario y un segundo conjunto, vinculado a las teorías evolucionistas sobre el pensamiento humano, que se basa en el intercambio de puntos de vista en contextos dialógicos de argumentación. Mostramos que estos dos conjuntos de estrategias están enfocados en activar el mismo tipo de procesos cognitivos, sólo que mientras que el primero intenta hacerlo mediante instrucciones directas para activar estos procesos en el razonamiento en solitario, el segundo lo hace mediante el intercambio argumentativo con pares. Asimismo, argumentamos que las estrategias pertenecientes al segundo conjunto pueden resultar más promisorias puesto que esta manera de razonar con pares podría replicarse con posterioridad en otras situaciones y contextos de producción y evaluación de argumentos. Finalmente, destacamos la importancia de seguir explorando estrategias destinadas a estimular las prácticas metacognitivas.

Palabras clave: racionalidad. razonamiento. sesgos cognitivos. argumentación. mitigar sesgos.

---

<sup>1</sup> Este artículo fue elaborado en el marco del Proyecto de Investigación: “Lógica y razonamiento humano: un estudio sobre sesgos de razonamiento” radicado en la Unidad asociada a Conicet – Área de Ciencias Sociales y Humanidades de la Facultad de Filosofía y Humanidades, Universidad Católica de Córdoba. Financiado por la Universidad Católica de Córdoba y el Conicet.

## ABSTRACT

In this paper, we critically evaluate different strategies used to reduce myside bias. This cognitive bias has been characterized as the tendency to produce and evaluate arguments in a way that is influenced by one's own opinions and ideologies. We propose a division into two major sets of strategies: a first set, linked to dual processes theories, which aims to improve reasoning skills individually, and a second set, linked to evolutionary theories of human reasoning, which is based on the exchange of points of views in dialogical contexts of argumentation. We show that these two sets of strategies focus on activating the same cognitive processes, but while the first set tries to do it through direct instructions in solo reasoning, the second set does it through argumentative exchange with peers. Likewise, we argue that the strategies belonging to the second set may be more promising since this way of reasoning with peers could be replicated later in other situations and contexts of argument production and evaluation. Finally, we highlight the importance of continuing to explore strategies aimed at stimulating metacognitive practices.

Keywords: rationality. reasoning. cognitive biases. argumentation. bias mitigation.

## RESUMO

Neste artigo, avaliamos criticamente diferentes estratégias usadas para reduzir o viés de confirmação. Este viés cognitivo tem sido caracterizado como a tendência de produzir e avaliar argumentos influenciados pelas próprias opiniões e ideologias. Propomos uma divisão em dois grandes conjuntos de estratégias: um primeiro conjunto, vinculado às teorias dos processos cognitivos duais, que visa melhorar as habilidades de raciocínio individual, e um segundo conjunto, vinculado às teorias evolucionistas sobre o pensamento humano, que se baseia na troca de pontos de vista em contextos dialógicos de argumentação. Mostramos que estes dois conjuntos de estratégias estão focados na ativação do mesmo tipo de processos cognitivos, apenas que, enquanto o primeiro tenta fazê-lo por meio de instruções diretas para ativar somente estes processos no raciocínio, o segundo o faz por meio de trocas argumentativas com pares. Da mesma forma, argumentamos que as estratégias pertencentes ao segundo conjunto podem ser mais promissoras, uma vez que esta forma de raciocínio com pares poderia ser replicada posteriormente em outras situações e contextos de produção e avaliação de argumentos. Finalmente, destacamos a importância de continuar explorando estratégias destinadas a estimular práticas metacognitivas.

Palavras-chave: racionalidade. raciocínio. vieses cognitivos. argumentação. vieses atenuantes.

## Introducción

Los sesgos cognitivos constituyen una robusta y persistente tendencia de la cognición humana que está presente en una amplia variedad de situaciones y contextos. La discusión en torno a los sesgos cognitivos posee un lugar central en el debate actual acerca de la racionalidad. Durante los últimos cincuenta años, diversos estudios en el ámbito de la psicología cognitiva y la economía han dado cuenta de la presencia de este fenómeno en los procesos de toma de decisiones. Asimismo, se ha advertido acerca de los efectos perniciosos que tienen estos sesgos en ciertas decisiones de fundamental importancia, ya sea a nivel individual o social. Es por esto que una parte considerable de la investigación ha estado dirigida a estudiar de qué manera se pueden mitigar los sesgos cognitivos.

El denominado sesgo a mi favor<sup>2</sup> (*myside bias*), de aquí en más SMF, ha sido caracterizado como la tendencia a considerar y ponderar de manera positiva aquellos argumentos consistentes con las propias opiniones, al mismo tiempo que se ignoran y se evalúan de manera negativa aquellos argumentos que contradicen estas opiniones y creencias previas (Perkins, 2019). Desde algunas líneas de investigación, se han destacado las virtudes de este sesgo. En particular, se ha argumentado que el SMF constituye una ventaja adaptativa cuya función consiste en dividir el trabajo cognitivo, evitando de esta manera la sobrecarga de tareas para cada razonador individual (Mercier & Sperber, 2017). Este último argumento podría poner en duda la necesidad de mitigar este sesgo cognitivo, ya que si el SMF constituye una ventaja adaptativa, entonces su reducción podría resultar contraproducente. Sin embargo, como argumentaremos en este artículo, existen determinadas situaciones que poseen consecuencias de fundamental importancia, como por ejemplo la participación en la política democrática, donde resulta imprescindible que las decisiones que se tomen estén basadas en procesos cognitivos realizados con esfuerzo y meticulosidad.

En este artículo, proponemos realizar una evaluación crítica de diferentes estudios empíricos que intentaron mitigar el SMF. Nuestro objetivo central consiste en realizar una revisión de diferentes estrategias que han sido empleadas a fin de reducir este sesgo cognitivo. Para cumplir este objetivo, es menester analizar además las diferentes perspectivas teóricas subyacentes a cada una de las estrategias que serán revisadas. Por esto, el artículo seguirá la siguiente estructura, comenzamos ofreciendo una caracterización del SMF y establecemos algunas distinciones con otros fenómenos cognitivos similares. En la tercera sección, presentamos dos líneas de investigación a partir de las cuales se ofrecieron explicaciones sobre el SMF, a saber, las teorías de los procesos cognitivos duales y las teorías evolucionistas sobre el pensamiento humano. Asimismo, argumentamos por qué resulta importante reducir el SMF en ciertas situaciones argumentativas, más allá de la línea de investigación que se adopte. En la cuarta sección, nos enfocamos en la evaluación crítica de las diferentes propuestas para reducir el SMF. Dividimos estas estrategias en dos conjuntos: el primer conjunto, vinculado a las teorías de los procesos duales, apunta a mejorar las habilidades del razonamiento en solitario y el segundo conjunto, vinculado a las teorías evolucionistas sobre el pensamiento, se basa en el intercambio dialógico en contextos grupales de argumentación. En la quinta sección, realizamos una discusión crítica y argumentamos que las dos estrategias empleadas para reducir el SMF están enfocadas en activar el mismo tipo de procesos cognitivos, sólo que mientras que la primera estrategia intenta hacerlo mediante instrucciones directas para activar estos procesos en el razonamiento en solitario, la segunda estrategia lo hace a partir de incentivar el intercambio argumentativo con pares. Mostramos, además, que la segunda estrategia puede resultar más promisorio puesto que esta manera de razonar con pares podría replicarse con posterioridad en otras situaciones de evaluación y producción de argumentos. A modo de conclusión, destacamos la importancia de seguir explorando estrategias destinadas a estimular las prácticas metacognitivas.

## 1. El sesgo a mi favor y otros fenómenos cognitivos afines

El SMF ha sido caracterizado como la tendencia a generar y evaluar argumentos de manera influenciada por las propias creencias y opiniones (Stanovich, West & Toplak, 2013). Este fenómeno cognitivo está vinculado tanto con la producción como con la evaluación de argumentos. Con respecto a la producción de argumentos, existe una marcada tendencia a generar argumentos que contemplan un solo lado de la discusión, a saber, el lado con el que acuerda la persona que elabora el argumento (Baron, 1995; Wolfe & Britt, 2008). Con respecto a la evaluación de argumentos, existe una manifiesta tendencia a considerar mejores y más fuertes aquellos argumentos que respaldan la propia postura, de manera independiente de

---

<sup>2</sup> En español, no existe una traducción estandarizada de la expresión “myside bias”. En trabajos anteriores, hemos traducido el nombre de este sesgo cognitivo como “sesgo a mi favor” a partir de adecuar a la gramática del español la expresión “sesgo en favor mío” empleada por Mercier & Sperber (2019). Otros autores, sin embargo, han preferido traducir esta expresión como “sesgo de mi lado”. Ante la falta de una traducción estándar, vamos a mantener el criterio de trabajos anteriores y emplearemos la expresión “sesgo a mi favor” en este artículo.



la calidad o validez lógica que posean los argumentos en cuestión (Baron, 1995; McCruden & Barnes, 2016; Čavojová, Šrol & Adamus, 2018; Kessler, Braasch & Kardash, 2019).

El SMF está estrechamente vinculado con el sesgo de confirmación. Este último se define como la tendencia a buscar y considerar información para confirmar las creencias previas, sin dejar lugar a los contraejemplos o a la información que podría falsar esas creencias (Nickerson, 1998). Si bien ambos sesgos cognitivos poseen una fuerte similitud y a veces son tratados como si fuesen fenómenos idénticos, consideramos necesario enfatizar una diferencia entre ellos. Mientras que el sesgo de confirmación caracteriza la tendencia confirmatoria general del razonamiento humano, el SMF caracteriza la tendencia a confirmar la información compatible con las propias opiniones y a cuestionar, o incluso a falsar, la información incompatible con estas opiniones y creencias (Mercier, 2017).

Existe también una estrecha relación entre el SMF y el razonamiento motivado. El razonamiento motivado se caracteriza como la tendencia de las personas a ajustar la evaluación de la información de acuerdo con algún objetivo o fin particular (Kahan, 2013). Las motivaciones que dirigen el razonamiento pueden ser dos: obtener una conclusión adecuada o exacta y obtener una conclusión deseada (Kunda, 1990). Mientras que la primera categoría puede vincularse con un mayor esfuerzo cognitivo, la segunda se vincula con procesos cognitivos sesgados y está directamente relacionada con el SMF. Sin embargo, el razonamiento motivado involucra otras ilusiones cognitivas como es el caso de la búsqueda sesgada de información en la memoria.

Finalmente, es posible establecer vínculos entre el SMF y dos fenómenos de la cognición social: la necesidad de cognición y la necesidad de cierre. Por un lado, la necesidad de cognición (Cacioppo & Petty, 1982) refiere a la disposición del pensamiento humano a involucrarse y disfrutar de actividades relacionadas al pensamiento. Si bien en un principio se especuló que esta disposición del pensamiento podría correlacionarse negativamente con el SMF, los resultados empíricos no encontraron correlaciones significativas entre la escala que mide la necesidad de cognición y el SMF (Macpherson & Stanovich, 2007). Por otro lado, la necesidad de cierre (Kruglanski, 2004) es un estilo de regulación epistémica que refiere a la motivación de buscar y mantener una respuesta definitiva frente a un problema, por ejemplo asegurar una opinión con respecto a un tema y así evitar la incertidumbre y la ambigüedad. Se ha mostrado que existen correlaciones positivas entre el SMF y este estilo de regulación epistémica (Sá, West & Stanovich, 1999).

## **2. Dos explicaciones sobre el sesgo a mi favor**

En esta sección, presentamos dos teorías explicativas sobre el SMF. La primera corresponde a las teorías de los procesos cognitivos duales y se vincula estrechamente al programa de heurísticas y sesgos iniciado por Kahneman y Tversky (1974), mientras que la segunda está vinculada con las teorías evolucionistas sobre el razonamiento humano y con el programa de la racionalidad ecológica propuesto por Gigerenzer (2008).

### **2.1 Teorías de los procesos duales**

Las teorías de los procesos duales postulan la existencia de dos tipos de procesamiento en el razonamiento y en las funciones cognitivas superiores que son cualitativamente diferentes, a saber, los procesos tipo 1 y los procesos tipo 2. La característica distintiva de los procesos tipo 1 es que son automáticos y que no requieren de la memoria de trabajo. Por su parte, los procesos tipo 2 requieren de una gran demanda de la memoria de trabajo y están correlacionados con los procesos de desconexión cognitiva y de simulación mental (Evans & Stanovich, 2013). Tanto los procesos tipo 1 como los procesos tipo 2 involucran una multiplicidad de sistemas cognitivos subyacentes. Frente a un problema de razonamiento o de toma de decisiones, los procesos tipo 1 producen una respuesta intuitiva por defecto, posteriormente un procesamiento tipo 2 puede intervenir (o no) ofreciendo una respuesta alternativa (Evans, 2019).

La génesis filosófica de las teorías de los procesos duales se remonta a los escritos de Platón y su planteo de la división del alma humana en tres partes: el alma racional, el alma irascible y el alma apetitiva. Es posible trazar ciertas analogías entre el alma racional y los procesos tipo 2, en tanto ambos tendrían la capacidad de controlar los juicios superficiales y los deseos originados en las otras dos partes del alma (Frankish & Evans, 2009). Asimismo, esta idea de la mente humana dividida puede rastrearse en la discusión moderna acerca de la mente animal en filósofos como Descartes, Leibniz y Locke. Sin embargo, es probable que el antecedente más reconocido sobre la dualidad de la mente humana sean las investigaciones de Freud vinculadas al concepto de inconsciente (Ibid.). En un sentido más acotado, las teorías de los procesos duales son producto de los últimos cuarenta años de investigación en ciencia cognitiva. Es posible distinguir al menos cuatro corrientes de investigación independientes que sin estar influenciadas entre sí dieron origen a los primeros estudios sobre los procesos duales: el trabajo experimental de Reber sobre el aprendizaje implícito (Frankish & Evans, 2009); los experimentos del programa de heurísticas y sesgos (Kahneman, 2012); los estudios empíricos del ámbito de la psicología del razonamiento (Evans, 2010); y, finalmente, los desarrollos sobre cognición social (Smith & Collins, 2009). El punto de convergencia de todos estos desarrollos consiste en postular la existencia de dos mecanismos de procesamiento diferentes, a saber, los procesos tipo 1 y tipo 2. Estos procesos emplean diferentes procedimientos y arrojan resultados diferentes, muchas veces conflictivos entre sí. Uno de estos mecanismos de procesamiento, es rápido, inconsciente, automático, está contextualizado y no requiere de esfuerzo ni de grandes demandas de la memoria de trabajo. El otro mecanismo es lento, controlado, consciente, flexible, está descontextualizado y requiere de esfuerzo y demanda de la memoria de trabajo (Frankish & Evans, 2009).

Según una versión en extremo simplificada de esta teoría, los procesos tipo 1 son responsables de los sesgos cognitivos, mientras que los procesos tipo 2 son garantía de respuestas correctas. No obstante, se ha observado que en ciertas circunstancias los procesos tipo 1 pueden conducir a respuestas correctas al tiempo que los procesos tipo 2 pueden ser responsables de respuestas sesgadas. Sobre este asunto, resulta necesario establecer una distinción entre ambientes benévolos y ambientes hostiles (Stanovich, West & Toplak, 2016). Por un lado, un ambiente benévolo contiene pistas o indicios que resultan de utilidad para la aplicación de heurísticas o procesos tipo 1. En estos ambientes, la práctica o la propia historia evolutiva facilitan la posibilidad de explotar estos indicios. La interacción del ser humano con la naturaleza constituye un ejemplo de un ambiente de este tipo. Por otro lado, un ambiente hostil no contiene estas pistas de utilidad para la aplicación de heurísticas o mecanismos tipo 1. En estos ambientes, existen otros agentes capaces de ajustar estas pistas en su propio beneficio de manera tal que las heurísticas pueden resultar perjudiciales para las personas que las emplean. De esta manera, desde las teorías de los procesos duales se ha enfatizado que en los ambientes hostiles la correlación entre los procesos tipo 1 y las respuestas correctas es muy baja, siendo necesaria la intervención de una respuesta basada en procesamientos tipo 2 para evitar una respuesta sesgada.

El SMF pone de manifiesto la dificultad para realizar una desconexión cognitiva con respecto a las creencias previas a la hora de producir y evaluar argumentos. Asimismo, este sesgo está vinculado con el procesamiento tacaño de la información, es decir, con la tendencia a procesar la información a partir de mecanismos que requieren una baja demanda computacional (Stanovich, 2018). De esta manera, las respuestas por defecto obtenidas a partir de los procesos cognitivos tipo 1 estarían vinculadas con la aparición del SMF. Sin embargo, los procesos cognitivos tipo 2 parecen ser condición necesaria aunque no suficiente para evitar este sesgo. Los procesos tipo 2 sólo van a resultar efectivos si se aplican de manera cuidadosa, se posee el conocimiento sobre las reglas de razonamiento adecuadas y además se tiene la capacidad cognitiva de aplicar estas reglas sin errores (Evans, 2019). Estas observaciones son consistentes con las correlaciones encontradas entre la baja incidencia del SMF y las puntuaciones altas en las diferentes escalas propuestas para medir el denominado pensamiento racional, como es el caso del CART (evaluación exhaustiva de pensamiento racional) (Stanovich, West y Toplak, 2016) o el AOT (pensamiento activo de mente abierta) (Baron, 2019).

En el marco de las teorías de los procesos duales, la hipótesis de la cognición de protección de identidad (Kahan, 2017) cuestiona este vínculo entre pensamiento racional y SMF. Esta hipótesis esgrime que las

personas realizan un esfuerzo cognitivo para adecuar la información a la posición defendida en los propios grupos de pertenencia (Kahan, 2013). De esta manera, el SMF estaría mayormente correlacionado con los procesos esforzados tipo 2, puesto que las personas utilizarían todas sus habilidades cognitivas a fin de proteger los puntos de vista que se sostienen en sus grupos de identificación. No obstante, se podría cuestionar si esta correlación positiva entre el SMF y los procesos tipo 2 se vincula necesariamente con el pensamiento racional, puesto que, como ha sido advertido, el esfuerzo cognitivo no implica necesariamente pensamiento racional.

## **2.2 Teorías evolucionistas sobre el pensamiento humano**

Las teorías evolucionistas sobre el razonamiento humano proponen una reconsideración del rol que desempeñaron los sesgos cognitivos en nuestra historia evolutiva. Si bien no vamos a detallar las diferencias internas de estas teorías, resulta necesario señalar que mientras que algunas de estas teorías resaltan la evolución genética, otras teorías ponen el acento en la evolución cultural. Desde la racionalidad ecológica, se cuestionó el programa de heurísticas y sesgos por centrar la discusión en torno a la racionalidad en la adecuación del razonamiento humano con las leyes de la lógica y de la probabilidad. La teoría de la racionalidad ecológica define la racionalidad a partir de la coherencia entre las conductas cognitivas humanas y el medio ambiente, es decir, según el éxito que una determinada cognición tenga en el mundo, independientemente de si esta conducta se ajusta o no a las reglas de la lógica (Gigerenzer, 2008). Los procesos cognitivos esforzados no siempre llevan a mejores respuestas: muchas veces las respuestas basadas en heurísticas constituyen la mejor respuesta posible.

En sintonía con esta línea de investigación, los estudios recientes sobre el pensamiento humano encuentran en la interacción con el medio ambiente y con los otros miembros de la especie una explicación sobre el desarrollo de las funciones cognitivas superiores. Según la hipótesis de la intencionalidad compartida (Tomasello, 2014), el pensamiento humano evolucionó a partir de la intencionalidad individual que era característica de la cognición esencialmente competitiva de los grandes simios. En este contexto, resultó indispensable para la supervivencia de la especie alguna forma de adaptación para poder lidiar con los problemas de coordinación social, ya sea para cooperar o para comunicarse con los otros congéneres. Así, las grandes demandas de cooperación de la especie, indispensables para la supervivencia, pudieron haber impulsado la configuración de formas de cognición más complejas. En este escenario evolutivo, la función del razonamiento parece estar vinculada con sobrescribir nuestras creencias y deseos de manera tal que la alineación de intenciones con los otros miembros de la especie resulte más sencilla, facilitando de esta manera una decisión conjunta que puede ser mutuamente ventajosa (Norman, 2016). De aquí que el SMF pueda resultar de ayuda para alinear intenciones, puesto que formar y mantener una perspectiva común con los miembros de un grupo facilita una colaboración mutualista. De esta manera, se argumenta que el SMF constituye una ventaja adaptativa.

Partiendo de presupuestos similares, el enfoque interaccionista de la razón (Mercier & Sperber, 2017) también ofrece un argumento a favor de las ventajas adaptativas del SMF. Según Mercier y Sperber, la razón desempeña una función eminentemente social, a saber, intercambiar justificaciones y argumentos. Los seres humanos, a diferencia de otras especies animales, desarrollaron un sistema de cooperación y comunicación extremadamente complejo. La historia evolutiva de la razón se ve reflejada en dos características del razonamiento humano: los sesgos y la pereza cognitiva. Estas condiciones de vida ultra sociales llevaron por un lado, a que las personas analizaran con detenimiento en qué congéneres podían confiar y, por otro lado, a mostrarse a sí mismos como individuos dignos de confianza. Como puede advertirse, ambas tareas involucran un intercambio argumentativo. Por un lado, a fin de aceptar razones y justificaciones para confiar en otros individuos es necesario realizar previamente un escrutinio exhaustivo. Por otro lado, a fin de ser considerado un cooperador confiable resulta necesario ofrecer buenas razones y justificaciones para expresar nuestras propias ideas. Este sistema de vigilancia epistémica permite ir ajustando la confianza entre los miembros de una comunidad. De esta manera, como la razón ejerce funciones diferentes al producir y al evaluar argumentos, los rasgos que exhibe al

realizar estas dos actividades también son diferentes: la razón es exigente a la hora de evaluar argumentos ajenos pero es perezosa a la hora de elaborar argumentos propios. Esta descripción de los rasgos de la razón coincide con la descripción del SMF. Puede verse entonces, que para el enfoque interaccionista de la razón el SMF es una ventaja adaptativa. Su función en un contexto social de intercambio argumentativo consiste en dividir la labor cognitiva garantizando una economía de recursos cognitivos.

Es necesario resaltar una diferencia entre dos niveles en la discusión acerca de las ventajas adaptativas de los sesgos cognitivos: el nivel individual y el nivel grupal. El enfoque interaccionista de la razón se centra en las ventajas adaptativas del SMF a nivel individual. En este sentido, la división de la labor cognitiva propiciada a partir del SMF resulta una ventaja para el individuo en tanto supone que cada individuo economice recursos cognitivos. No obstante, otros enfoques evolucionistas han resaltado las ventajas adaptativas de los sesgos cognitivos a nivel grupal (Norman, 2016).

Frente a estas explicaciones evolucionistas sobre el SMF, resulta necesario preguntarse si los intentos por mitigar este sesgo no serían contraproducentes, puesto que si el SMF constituye una ventaja adaptativa, ya sea a nivel individual o grupal, su disminución podría resultar perjudicial. A continuación ofrecemos dos argumentos de por qué incluso desde una perspectiva evolucionista resulta necesario reducir la incidencia del SMF en ciertos contextos de toma de decisiones.

En primer lugar, es menester enfatizar que las ventajas adaptativas a nivel grupal no constituyen necesariamente ventajas a nivel individual (Boudry, Vlerick & McKay, 2015). Consideremos por ejemplo el caso de la abeja que al picar a un potencial atacante de su colmena pierde su aguijón y muere. De manera similar, podría argumentarse que el SMF es una ventaja adaptativa a nivel grupal, puesto que facilita la distribución de la tarea cognitiva dentro de una comunidad. Sin embargo, a nivel individual es difícil argumentar a favor de las ventajas adaptativas de este sesgo, puesto que puede plantear serias dificultades a las personas que intentan dirimir por sí mismas sobre una cuestión controvertida (Maynes, 2015). Mercier y Sperber (2017) reconocen que ciertas instituciones sociales, como la ciencia y la justicia, poseen mecanismos para evitar la influencia del SMF en la toma de decisiones individuales. Este es el caso del sistema de evaluación por pares en el ámbito de la ciencia y de los procedimientos de acusación, investigación y defensa y de la posterior decisión de un tribunal en el ámbito de la justicia. En este sentido, en muchos contextos argumentativos, adelantarse a una posible objeción y poder elaborar una respuesta con anticipación parece ser más ventajoso a nivel individual que esperar a que se señale una crítica y ofrecer una réplica imprevista. De esta manera, parece bastante plausible que estar en sintonía con el contexto argumentativo podría ser una característica más adaptativa que mantener un dogmatismo obstinado (Dutilh-Novaes, 2018).

En segundo lugar, atendiendo a la distinción anteriormente señalada entre ambientes benévolos y hostiles (Stanovich, West & Toplak, 2016), podría argumentarse que si bien el SMF puede constituir una ventaja adaptativa (tanto a nivel individual como grupal) en ambientes benévolos, la situación se modifica radicalmente en ambientes hostiles. Esta distinción entre ambientes benévolos y hostiles fue elaborada en el marco de la teoría de los procesos duales para señalar algunas dificultades inherentes al programa de la racionalidad ecológica. Se ha señalado que los ambientes naturales que dieron origen al desarrollo evolutivo de la especie humana eran esencialmente ambientes benévolos, donde las heurísticas y los sesgos pudieron haber funcionado como ventajas adaptativas tanto para el individuo como para la especie. Sin embargo, la configuración de las sociedades humanas actuales, mediadas por la tecnología y con la existencia de diversos grupos sociales con intereses encontrados, plantea ciertos recaudos. Así, en estos ambientes hostiles, el SMF puede ser utilizado de manera intencional por determinados actores sociales de manera tal que la evaluación o producción de argumentos y, en consecuencia, la toma de decisiones resulte en detrimento del bienestar de un individuo o grupo social. Es por esto que resulta difícil encontrar ventajas adaptativas a nivel individual o grupal del SMF en ambientes hostiles.

En conclusión, si bien el SMF pudo haber sido una ventaja a nivel grupal en nuestra historia evolutiva y puede seguir siendo ventajoso, incluso a nivel individual, en contextos benévolos de toma de decisiones,



nuestras sociedades actuales poseen una complejidad tal que resulta necesario en ciertas circunstancias realizar un esfuerzo cognitivo para tomar una decisión. Fenómenos actuales tales como la polarización política y los discursos de odio, que claramente afectan el bienestar individual y social, están en gran medida vinculados a procesos sesgados de razonamientos como el SMF. En este sentido, reducir la incidencia del SMF en contextos de relevancia, como es el caso de la democracia participativa o de las finanzas individuales, resulta fundamental en pos del bienestar tanto social como individual.

### **3. Estrategias para reducir el sesgo a mi favor**

En esta sección, analizamos diversas estrategias para mitigar el SMF. Para esto realizamos una búsqueda bibliográfica exhaustiva en la base de datos EBSCO. Además, efectuamos una búsqueda manual a partir de las referencias de los estudios seleccionados. Las palabras utilizadas en la búsqueda fueron “myside bias”, “mitigating bias” y “reducing bias”. En un principio, se seleccionaron únicamente estudios vinculados al SMF. Sin embargo, como la cantidad de artículos que satisfacía este criterio fue algo escasa, decidimos incluir también artículos vinculados al sesgo de confirmación, ya que como se mencionó en la sección 2, ambos sesgos se encuentran estrechamente vinculados. Finalmente, luego de la lectura y el análisis, seleccionamos 16 de 38 artículos preseleccionados. El criterio de exclusión estuvo relacionado a cuestiones de pertinencia al tema de revisión.

Proponemos una división en dos grandes conjuntos de las estrategias para mitigar el SMF: (i) estrategias enfocadas en activar habilidades de razonamiento en solitario y (ii) estrategias basadas en incentivar el intercambio argumentativo en grupos. Las estrategias del conjunto (i) están vinculadas teóricamente con las teorías de los procesamientos duales. Estas estrategias se enfocan en la activación de diferentes habilidades analíticas del pensamiento, en particular, enfatizan la desconexión cognitiva. Las estrategias del conjunto (ii) se relacionan con las teorías evolutivas sobre el pensamiento humano, en tanto están enfocadas en incentivar los intercambios dialógicos grupales a partir de los cuales habría evolucionado el razonamiento.

#### **3.1 Estrategias enfocadas en activar habilidades de razonamiento en solitario**

En la primera sección de la tabla 1, se pueden observar los estudios empíricos que proponen estrategias para activar habilidades de razonamiento en solitario. Estas estrategias apuntan a estimular la desconexión cognitiva apelando a: adoptar una perspectiva opuesta a la propia, inducir la atención plena, incentivar las habilidades de pensamiento científico, estimular la simulación mental contrafáctica y apuntalar la metacognición.

En primer lugar, presentamos tres propuestas basadas en una estrategia dirigida a considerar una perspectiva opuesta o diferente a la propia. Por un lado, Wolfe, Britt y Butler (2009) analizan los efectos de considerar una perspectiva opuesta a partir de fomentar la inclusión de objeciones, contraargumentos y refutaciones en una tarea de argumentación escrita. Este estudio empleó un diseño entre-sujetos donde se ofreció a una parte de la muestra un tutorial que enfatizaba la importancia de precisar la conclusión de un argumento, de proveer razones elaboradas para respaldar la conclusión y de presentar y refutar posibles objeciones. Como resultado, los participantes que usaron este tutorial incluyeron un 42% más de contraargumentos y objeciones en sus ensayos que los participantes del grupo control. El estudio sugiere que un tutorial que explique la importancia de incorporar información opuesta puede ayudar a reducir el SMF en los ensayos argumentativos escritos. No obstante, se trata de una estrategia que emplea una instrucción directa sobre cómo resolver la tarea que se evalúa. Por otro lado, una estrategia similar fue propuesta por van Brussel, Timmermans, Verkoeijen y Paas (2020) para reducir el sesgo de confirmación. Este estudio de diseño pre-test y post-test entre-sujetos solicitó a un grupo de participantes que resolviera una serie de tareas de testeo de hipótesis. Una parte de la muestra observó un video donde se explicaba la importancia de reducir sesgos cognitivos y además se instruía a los participantes a tomar una perspectiva opuesta antes de resolver las tareas del post-test. Los participantes que recibieron esta instrucción ofrecieron menos respuestas sesgadas que los participantes que no recibieron esta



instrucción. Sin embargo, la diferencia encontrada entre ambos grupos fue sólo moderada. Un punto a favor de este estudio es que no dieron instrucciones directas sobre cómo se debía resolver la tarea a evaluar. Siguiendo una estrategia similar, basada en la idea de que tomar otra perspectiva puede evocar procesos analíticos de pensamiento que permiten a los individuos distanciarse de su propia perspectiva y así lograr una desconexión cognitiva, McCrudden, Barnes, McTigue, Welch y MacDonald (2017) propusieron una tarea de evaluación de argumentos, tanto fuertes y débiles como consistentes e inconsistentes con las creencias previas, sobre el cambio climático. En este estudio, el tipo de argumento y la fuerza del argumento fueron explicitadas como variables intra-sujeto, mientras que la toma de perspectiva se admitió como variable entre-sujeto. La fuerza del argumento se determinó mediante el período de tiempo en el cual se recogió la evidencia utilizada: los argumentos fuertes utilizaban evidencia de un período de cien años y los débiles utilizaban pruebas de un período de dos años. Se pidió a los participantes que realizaran esta tarea de evaluación de argumentos desde la propia perspectiva y desde la perspectiva de un científico especialista en cambio climático. La conclusión principal de este estudio fue que tomar otra perspectiva reduce de manera significativa las respuestas sesgadas para el caso de los argumentos débiles pero no para los argumentos fuertes. Específicamente, cuando se les pide a los participantes que consideren un argumento sobre cambio climático desde la perspectiva de un científico, las creencias previas casi no afectan la evaluación de argumentos débiles aunque siguen afectando la evaluación de argumentos fuertes. Como señalamos respecto del estudio anterior, un punto a favor es que no se dieron instrucciones directas sobre cómo se debía resolver la tarea a evaluar. Aunque los efectos concretos para mitigar el SMF no fueron los esperados para los argumentos fuertes.

Jurkovič (2016) propuso para mitigar el SMF una estrategia de inducción de atención plena (*mindfulness*) previamente a la resolución de una serie de tareas de razonamiento así como lecturas sobre sesgos cognitivos. El estudio empleó un diseño entre-sujetos y partió del supuesto de que durante una meditación de atención plena el practicante entrena sus funciones ejecutivas y aprende a inhibir las reacciones habituales de la mente y, de esta forma, toma distancia de sus propias creencias. Esto podría resultar efectivo para reducir sesgos como el SMF. Sin embargo, los resultados que se obtuvieron no fueron significativos y mostraron que ni la inducción de atención plena a corto plazo ni las lecturas sobre sesgos fueron suficientes para reducir el SMF.

Algunas propuestas se enfocaron en activar habilidades de pensamiento científico. Por un lado, Drummond y Fischhoff (2019) propusieron una tarea de evaluación de evidencia científica sobre la ley conocida como Obamacare<sup>3</sup> y analizaron los efectos de administrar previamente una escala de razonamiento científico, una prueba de aritmética y dar una instrucción explícita para usar habilidades de pensamiento científico en la resolución de la tarea asignada. El estudio empleó un diseño entre-sujetos. Los resultados mostraron que si bien el solo hecho de tomar una escala de razonamiento científico o de aritmética no fue suficiente para reducir el SMF, las instrucciones específicas de emplear estas habilidades de pensamiento científico en la tarea de evaluación de evidencia resultaron efectivas para reducir en una leve medida este sesgo cognitivo. Los participantes que resolvieron la tarea bajo la instrucción de emplear habilidades de pensamiento científico evaluaron de manera más crítica la evidencia y aumentaron la capacidad para identificar limitaciones en los estudios. No obstante, enfatizamos que la diferencia encontrada entre los dos grupos fue débil. Por otro lado, en un diseño intra-sujetos, Čavojová, Šrol y Jurkovič (2019) administraron un conjunto de pruebas con el objetivo de analizar la interacción entre la capacidad de razonamiento científico, las habilidades cognitivas de pensamiento y las disposiciones analíticas de pensamiento con la mentalidad cerrada, la susceptibilidad a adoptar creencias epistémicamente sospechosas<sup>4</sup> y los sesgos cognitivos. Luego de realizar análisis correlacionales y regresiones, los resultados permitieron concluir que el pensamiento científico se correlaciona positivamente con las disposiciones hacia el pensamiento analítico ( $r=.203$ ,  $p=.001$ ) y la habilidad cognitiva ( $r=.366$ ,  $p=.001$ ) y negativamente con el dogmatismo ( $r=-.261$ ,  $p=.001$ ), las creencias epistémicamente sospechosas ( $r=-.209$ ,  $p=.001$ ) y la susceptibilidad a los sesgos cognitivos ( $r=-.379$ ,  $p=.001$ ). Además, tanto la inteligencia como la capacidad de razonamiento científico surgieron

<sup>3</sup> La Ley del Cuidado de Salud a Bajo Precio (en inglés, Patient Protection and Affordable Care Act), conocida como Obamacare, tiene como objetivo expandir la cobertura de seguro médico a más habitantes estadounidenses.

<sup>4</sup> Se consideran creencias epistémicamente sospechosas las creencias en fenómenos paranormales, pseudociencias y teorías conspirativas.

como predictores independientes significativos de la susceptibilidad a creencias epistémicamente sospechosas; aunque esto no se aplica a las creencias infundadas alineadas con la afiliación política o religiosa. Si bien la correlación encontrada entre el pensamiento científico y los sesgos cognitivos es negativa, este grupo de investigadores destaca que se debería seguir trabajando en el desarrollo de estrategias específicas que permitan aplicar estas habilidades de pensamiento científico en la reducción de sesgos cognitivos.

Otra propuesta consistió en activar la simulación mental en un escenario contrafáctico. Strachanová y Valus (2019) presentaron a los participantes de su estudio distintas situaciones-problema de personas hipotéticas, a saber, una persona tratando de determinar si un compañero era introvertido, una persona tratando de explicar un comportamiento extraño de un nuevo colega de trabajo y una persona reconsiderando su actitud hacia las vacunas luego de obtener nueva información sobre el tema. Posteriormente, los participantes respondieron una serie de preguntas críticas y resolvieron tareas de razonamiento en relación con estas situaciones hipotéticas. Estas tareas de razonamiento analizaron: la búsqueda selectiva de información para confirmar una hipótesis, el rechazo de explicaciones alternativas y la falta de predisposición para reconsiderar una opción por defecto frente a la obtención de nueva información. Los resultados demostraron que la simulación mental en un escenario contrafáctico ayuda a ignorar las respuestas intuitivas de manera significativa en comparación con lo que sucede en un escenario neutral. De esta manera, la preparación contrafáctica facilitó sustancialmente la búsqueda de evidencia desconfirmadora y la consideración de explicaciones alternativas. Estos patrones sugieren que la simulación mental podría ayudar a mitigar el SMF.

Finalmente, presentamos tres propuestas enfocadas en la metacognición. En un estudio de diseño mixto realizado por Bryzc, Wyszomirska-Góra, Bar-Tal y Wiśniewski (2014) basado en el sesgo de confirmación, se analizó la influencia de la metacognición en la formación de impresiones sobre los demás. Se presentó a los participantes una serie de afirmaciones sobre distintas personas a quienes debían juzgar de acuerdo a dos tipos de características: características comunitarias (es una persona honesta /deshonesta) y características competitivas (es una persona inteligente /poco inteligente). Los resultados evidenciaron que la estrategia de confirmación estuvo presente con mayor frecuencia en las impresiones elaboradas por los individuos con altos niveles de metacognición, sin encontrar diferencias entre los juicios realizados con respecto a los dos tipos de características. De esta manera, el estudio concluyó que ser consciente de los propios sesgos y conductas irracionales no parece suficiente para reducirlos. Resultados similares fueron obtenidos por Perkins (2019) en su estudio sobre el impacto del repertorio metacognitivo de las personas en la evaluación de argumentos. En este estudio, se ofrecieron a los participantes programas de instrucción sobre pensamiento crítico y guías generales de razonamiento informal. Se observó que las personas pueden mejorar sus capacidades de razonamiento a partir de la instrucción. Sin embargo, el hallazgo más significativo indica que aquellos participantes con mejores habilidades metacognitivas fueron también los más sesgados. A diferencia de estas conclusiones, en el estudio llevado a cabo por McCrudden y Barnes (2016) se obtuvieron resultados promisorios en relación al vínculo entre metacognición y reducción del SMF. Se observó que los participantes menos sesgados utilizaron las mismas estrategias para evaluar tanto argumentos consistentes como inconsistentes con sus propias opiniones e ideologías, mientras que los participantes que mostraron mayor nivel de SMF utilizaron estrategias menos reflexivas y menos deliberativas a la hora de evaluar argumentos inconsistentes con sus propias opiniones. Este estudio mostró que existen diferencias entre la metacognición conceptual y procedimental. Por esto, a fin de reducir el SMF no sólo resulta necesario apuntalar la metacognición conceptual, es decir, el conocimiento sobre los propios sesgos, sino también resulta imprescindible trabajar la metacognición procedimental, es decir, una vez que se reconocen estos sesgos saber qué hacer con ellos.

### **3.2 Estrategias basadas en el intercambio argumentativo dialógico**

En la segunda sección de la tabla 1, se pueden observar los estudios empíricos con sus respectivas estrategias basadas en el intercambio argumentativo en contextos dialógicos.

En primer lugar, la propuesta de Felton Crowell y Liu (2015) fue diseñada para reducir los efectos del SMF en la argumentación escrita. El objetivo de esta investigación consistió en examinar si existen diferencias entre el diálogo de búsqueda de persuasión y el diálogo de búsqueda de consenso en la reducción del SMF. Para esto, los participantes de este estudio fueron asignados a dos condiciones diferentes: diálogo de persuasión y diálogo de consenso. Los participantes de la primera condición tenían como objetivo persuadir a los otros participantes para que adoptaran su posición respecto de la pena de muerte, mientras que los participantes de la segunda condición tenían como objetivo llegar a un consenso con sus pares sobre la pena de muerte. Posteriormente, cada participante escribió un ensayo argumentativo sobre ese tema. Los resultados mostraron que el 72% de los participantes de la condición diálogo de búsqueda de consenso incluyó contraargumentos en sus escritos frente a un 54% de los participantes de la condición diálogo de persuasión. A su vez, los participantes de la primera condición desarrollaron argumentos integradores<sup>5</sup> en mayor medida que los participantes de la condición diálogo de persuasión (58% frente a 2%). Además, un 8% de los estudiantes en la condición diálogo de persuasión cambió su posición respecto de la pena de muerte, mientras que en la otra condición el porcentaje de cambio de postura ascendió a un 18%. De esta manera, los resultados de este estudio mostraron que los objetivos del discurso previo a la escritura tienen un impacto en el SMF en los ensayos escritos. Si bien los autores reportaron que los participantes de ambos grupos mostraron un leve efecto de polarización, los participantes del grupo de búsqueda de consenso moderaron sus argumentos usando estrategias integradoras. Adoptando un objetivo similar, el estudio llevado a cabo por Christensen-Branum, Strong y D'on Jones (2018) destaca la importancia de la enseñanza de la argumentación balanceada, donde se contemplen las diferentes posturas relativas a un tema en discusión, para mitigar el SMF. Los autores de este estudio proponen reconsiderar la inclusión en clases de debates entre posturas a favor y posturas en contra, así como generar y responder preguntas críticas sobre argumentos, y sostienen que estas prácticas podrían exacerbar el fenómeno del SMF a partir de hacer foco en la persuasión y la competencia en lugar de enfatizar el consenso y la colaboración.

En esta misma línea, Mercier, Boudry, Paglieri y Trouche (2017) presentaron una serie de sugerencias para mejorar las habilidades de razonamiento y disminuir los sesgos cognitivos basadas teóricamente en el enfoque interaccionista de la razón. Estos autores aseguran que los contextos que favorecen la discusión grupal permiten a los estudiantes mejorar tanto sus habilidades como argumentadores así como sus habilidades para razonar de manera individual. Con respecto a la reducción del SMF, estos investigadores afirman que deberían propiciarse contextos de discusión grupales donde se presenten perspectivas alternativas y donde cada participante se sienta libre de expresar sus opiniones y de cambiar su punto de vista si fuera necesario. En una línea similar, se ha observado que ir más allá del énfasis en el razonamiento individual y adoptar entornos dialógicos de intercambio argumentativo debería ser el objetivo central de la educación del pensamiento crítico, puesto que el diálogo crítico grupal constituye una posible clave para mitigar el SMF (Dacey, 2020). Un estudio empírico que parece abonar estas propuestas es la tesis doctoral de J. Morris (2017). En esta investigación, el autor compara los efectos de dos intervenciones para reducir el SMF en la argumentación escrita: la instrucción directa, donde un docente ofrece información sobre las diferentes aristas de un tema y el trabajo colaborativo en grupos, donde los estudiantes realizan un intercambio argumentativo y de información para aprender sobre un tema. Los resultados de este estudio muestran que no existen diferencias significativas en el desempeño de los participantes con respecto a las producciones escritas en estas dos condiciones experimentales. Sin embargo, el SMF es notablemente superior en la condición control, integrada por participantes que no realizaron ningún trabajo previo a la tarea de argumentación escrita. La probabilidad de que un estudiante en cualquiera de las dos condiciones experimentales considere ambos lados de un problema fue alrededor de cuatro veces mayor que la probabilidad de que un estudiante en el grupo control considere ambos lados. Un resultado muy interesante de este estudio muestra que en tareas de argumentación escrita posteriores, es decir, pasados tres meses luego de las dos intervenciones, comienza a advertirse una diferencia significativa en el desempeño de los dos grupos experimentales. Mientras que un 30% de los participantes de la condición trabajo colaborativo en grupo logró integrar

---

<sup>5</sup> Un argumento integrador es un argumento que contempla las dos partes de una discusión. Asimismo, en un argumento integrador además de considerarse contraargumentos, se los responde y refuta.

argumentos de ambos lados en sus escritos, sólo un 10% de los participantes de la condición instrucción directa logró hacerlo. Asimismo, en la condición trabajo colaborativo en grupos, el SMF se ve reducido incluso en tareas argumentativas con considerable distancia conceptual con respecto a la tarea resuelta en un principio. En conclusión, puede advertirse una clara ventaja de los efectos del diálogo colaborativo con pares en la reducción del SMF.

Finalmente, en un estudio orientado a resaltar la importancia del respeto entre grupos en la reducción de sesgos cognitivos, se ha afirmado que el trato respetuoso proveniente de un grupo de opinión opuesta sirve para reducir (o incluso eliminar) la evaluación sesgada de argumentos (Eschert & Simon, 2019). En este estudio, que consistió en evaluar una serie de argumentos sobre la necesidad de controlar el uso de las redes sociales, se introdujo experimentalmente una opinión respetuosa y una opinión irrespetuosa por parte de un supuesto grupo que mantenía la posición opuesta a la que adhería el participante. Así, antes de realizar una tarea de evaluación de argumentos, se informó a un grupo de participantes que personas que pensaban diferente estaban interesadas en conocer su opinión con respecto a la regulación del uso de redes sociales (muestra de respeto), mientras que a otro grupo de participantes se les informó que personas que pensaban diferente no tenían ningún interés en saber su opinión (muestra de poco respeto). Como consecuencia, los participantes que recibieron una muestra de respeto sintieron que su grupo fue reconocido como igual y la evaluación de argumentos resultó mucho menos sesgada. Esto fue radicalmente diferente en el caso de los participantes que recibieron una muestra de poco respeto por parte del supuesto grupo opositor. En ese caso, la evaluación sesgada de argumentos se incrementó considerablemente.

## **Discusión general**

Tanto en el primer conjunto de estrategias vinculadas a mejorar las habilidades de razonamiento en solitario, incentivando la desconexión cognitiva, como en el segundo conjunto de estrategias destinadas a fomentar el intercambio argumentativo en contextos dialógicos existen diversas propuestas que resultaron efectivas a la hora de reducir el SMF. En particular, creemos que la estrategia diseñada por Strachanová y Valuš (2019) de incentivar la simulación mental en un escenario contrafáctico obtuvo uno de los mejores resultados. Asimismo, la propuesta desarrollada por Wolfe, Britt y Butler (2009) de instar a considerar la perspectiva opuesta a partir de un tutorial con instrucciones específicas también obtuvo resultados interesantes. Sin embargo, como señalamos anteriormente, esta estrategia ofrece instrucciones directas sobre cómo resolver la tarea que se va a evaluar. Además, la propuesta de McCrudden y Barnes (2016) para activar mecanismos de metacognición en la toma de decisiones resulta una vía muy interesante para seguir explorando. Con respecto al segundo conjunto de estrategias, la propuesta de Felton, Crowell y Liu (2015) que consiste en incentivar el intercambio argumentativo dialógico de búsqueda de consenso también arrojó resultados muy promisorios. Esto permite mostrar que los dos conjuntos de estrategias logran sacar partido de sus desarrollos teóricos de base. Mientras que las estrategias del primer conjunto apuntan a la desconexión cognitiva incentivando el uso de procesamientos tipo 2, las estrategias del segundo conjunto procuran replicar los contextos dialógicos de intercambio argumentativo a partir de los cuales se habría configurado la razón a lo largo de nuestra historia evolutiva.

Si bien toman como punto de partida bases teóricas diferentes, es posible advertir que las propuestas para reducir el SMF diseñadas desde cada uno de estos dos conjuntos de estrategias están destinadas a activar procesos cognitivos análogos. La diferencia radica en el modo que cada conjunto de estrategias elige para activar estos procesos. Por un lado, las estrategias que pertenecen al primer conjunto apuntan a activar estos procesos cognitivos intentando direccionar el razonamiento en solitario, es decir, procuran incentivar la aplicación de procesos cognitivos esforzados durante la resolución de una tarea de razonamiento que se debe responder de manera individual. Estas estrategias entienden que frente a una tarea de razonamiento, es posible reducir la incidencia del SMF a partir de estimular la desconexión cognitiva activando mecanismos tales como el pensamiento contrafáctico o la toma de una perspectiva opuesta. Este tipo de activación, vinculada a la desconexión cognitiva, favorece una evaluación y



producción de argumentos menos sesgada. Por otro lado, desde el segundo conjunto de estrategias la apuesta se dirige a activar estos procesos cognitivos esforzados a partir del intercambio de razones y puntos de vista en un contexto dialógico de argumentación. En este caso, la misma dinámica del intercambio argumentativo va a llevar a que se consideren las diversas perspectivas de la cuestión bajo discusión. De esta manera, los participantes de este diálogo a partir de encontrar objeciones y contraargumentos y de reconocer la postura opuesta a la propia, deberán realizar un esfuerzo cognitivo a fin de reforzar sus propios argumentos. Así, tanto la producción como la evaluación de argumentos van a resultar menos sesgadas. Resulta necesario remarcar que hablar de esfuerzo cognitivo en el marco de las teorías evolucionistas sobre el razonamiento humano no implica adherir a una modalidad de procesamiento dual. De hecho, Mercier & Sperber (2017) cuestionan la fuerza explicativa de las teorías de los procesos duales al cuestionar la vaguedad de la distinción entre procesos tipo 1 y tipo 2 y señalar que estas formas de procesamiento responden más bien a una diferencia de grado y no tanto a una dicotomía de tipos. En este enfoque teórico, se reconoce la importancia del esfuerzo cognitivo en determinadas circunstancias, aunque estos procesamientos esforzados no se consideren procesos tipo 2. Es importante señalar, además, que para que este tipo de estrategias basadas en el intercambio argumentativo con pares sean efectivas resulta necesario garantizar la presencia de diversos puntos de vista y que se trate de un diálogo colaborativo y respetuoso de búsqueda de consenso. Como puede advertirse, desde ambas estrategias el énfasis consiste en tratar de que se considere una perspectiva opuesta o diferente aunque para esto se empleen diferentes metodologías.

¿Es posible dirimir acerca de qué conjunto de estrategias resulta más efectivo? A fin de abordar esta cuestión, vamos a retomar los resultados obtenidos por J. Morris (2017) en su disertación doctoral. Recordemos que en esta investigación Morris compara los efectos de dos intervenciones para reducir el SMF en la argumentación escrita, a saber, la instrucción directa y el trabajo colaborativo en grupos. Los resultados de este estudio muestran que el SMF es superior en la condición control y en la condición instrucción directa. En la condición de trabajo colaborativo en grupos, el SMF se ve reducido incluso en tareas argumentativas con considerable distancia conceptual con respecto a la tarea resuelta en un principio. Esta diferencia podría explicarse a partir de la posibilidad que otorga la modalidad del trabajo colaborativo en grupos de transferir las habilidades adquiridas durante el intercambio argumentativo a otras tareas de argumentación. Este resultado empírico, aporta elementos para resaltar la potencialidad de las estrategias que estimulan el intercambio argumentativo en contextos dialógicos como forma de reducir el SMF. De esta manera, la posibilidad de internalizar una práctica argumentativa grupal y de aplicarla en otras situaciones argumentativas que se resuelven de manera individual constituyen uno de los puntos más fuertes del segundo conjunto de estrategias. Queda por explorar la pregunta acerca de si estas nuevas habilidades internalizadas permanecen a lo largo del tiempo.

Como fue señalado al principio de esta sección, una hipótesis promisorio para seguir explorando está vinculada con la posibilidad de reducir el SMF a partir de estimular las prácticas metacognitivas. Se ha argumentado que motivar la exploración minuciosa del repertorio disponible de tácticas de razonamiento puede reducir el SMF (Perkins, 2019). En este sentido, la metacognición conceptual y la metacognición procedimental es decir, tener conocimiento sobre la propia cognición y saber cómo controlarla, pueden ser una herramienta de gran ayuda (McCrudden & Barnes, 2016). Así, resulta necesario no sólo saber que las creencias previas pueden sesgar los procesos de producción y evaluación de argumentos, sino también controlar los procesos cognitivos que se emplean al realizar estas actividades argumentativas. Uno de los aspectos centrales que ha sido discutido a lo largo de todo este artículo es cómo determinar en qué situaciones resulta necesario evitar respuestas intuitivas y ofrecer, en cambio, respuestas más elaboradas y en qué situaciones es preferible ofrecer una respuesta rápida. Creemos que la metacognición puede contribuir a dirimir esta cuestión. Frente a un escenario de toma de decisiones, hay dos cuestiones que parecen centrales a fin de ofrecer una respuesta óptima: por un lado, examinar el contexto y advertir las consecuencias implicadas por la respuesta que se ofrezca y, por otro lado, reconocer la batería de recursos cognitivos disponibles para elaborar una respuesta. Estimular las prácticas metacognitivas puede resultar de ayuda para reconocer la importancia de la situación a resolver y, a partir de esto, determinar qué recursos cognitivos se deberían aplicar en la resolución de esa situación. Para el caso del SMF, si el diagnóstico de la situación lo amerita, se podría advertir que es menester tomar distancia de



las propias opiniones para evaluar o producir un argumento, activando de esta manera la desconexión cognitiva. En este sentido, una propuesta que podría trabajarse en el marco de los intentos por reducir el SMF es la impulsada por la investigadora brasileña Selma Leitão quien analiza la interacción de los procesos de debate crítico y de los procesos metacognitivos en las prácticas argumentativas que tienen lugar en los contextos áulicos (Leitão, de Chiaro & Cano, 2016).

Tabla 1 - *Estrategias para mitigar sesgos según conjunto de estrategias*

	Título	Autores y año	Estrategia
ESTRATEGIAS PARA ACTIVAR HABILIDADES DE RAZONAMIENTO EN SOLITARIO	Argumentation Schema and the Myside Bias in Written Argumentation	Wolfe, Britt & Butler (2009)	Tomar una perspectiva opuesta a la propia
	The effect of metacognitive self on confirmation bias revealed in relation to community and competence	Brycz, Wyszomirska-Gorá, Yoram & Winiewski (2014)	Metacognición
	The effect of perspective-taking on reasoning about strong and weak belief-relevant arguments	McCrudden, Barnes, McTigue, Welch & MacDonald (2017)	Tomar una perspectiva opuesta a la propia
	Effect of Short-term Mindfulness Induction on Myside Bias and Miserly Processing: A preliminary Study	Jurkovic (2016)	Inducción de atención plena ( <i>mindfulness</i> )
	Differences in student reasoning about belief-relevant arguments: A mixed methods study	McCrudden & Barnes (2016)	Metacognición
	Why should we try to think like scientists?	Cavojová, Šrol & Jurkovic (2019)	Estimular el pensamiento científico
	Does putting on your thinking cap reduce myside bias in evaluation of scientific evidence?	Drummond & Fischhoff (2019)	Estimular el pensamiento científico
	Learning to reason: The influence of instruction, prompts and scaffolding, metacognitive knowledge, and general intelligence on informal reasoning about everyday social and political issues	Perkis (2019)	Metacognición
	Mental Simulation as a Remedy for Biased Reasoning	Strachanová & Valus (2019)	Simulación mental contrafáctica
	‘Consider the Opposite’ – Effects of elaborative feedback and correct answer feedback on reducing confirmation bias – A pre-registered study	van Brussel, Timmermans, Verkoeijen & Paas (2020)	Considerar una perspectiva opuesta a la propia
ESTRATEGIAS DE INTERCAMBIO ARGUMENTATIVO DIALÓGICO	Arguing to Agree: Mitigating My-Side Bias Through Consensus-Seeking Dialogue	Felton, Crowell & Liu (2015)	Diálogo de búsqueda de consenso
	Natural-Born Arguers: Teaching How to Make the Best of Our Reasoning Abilities	Mercier, Boudry, Paglieri & Trouche (2017)	Diálogo colaborativo
	Effect of collaborative learning and direct instruction on myside bias	Morris (2017)	Comparación entre instrucción directa y diálogo argumentativo
	Mitigating myside bias in argumentation	Christensen-Branum, Strong & D’On Jones (2018)	Diálogo de búsqueda de consenso
	Respect and political disagreement: Can intergroup respect reduce the biased evaluation of outgroup arguments?	Eschert & Simon (2019)	Respeto entre grupos
	Come Now, Let Us Reason Together: Cognitive Bias, Individualism, and Interactionism in Critical Thinking Education	Dacey (2020)	Diálogo crítico colaborativo

## Referencias

- Baron, J. (1995). "Myside bias in thinking about abortion". *Thinking & Reasoning*. 1(3). 221-235.
- Baron, J. (2019). "Actively open-minded thinking in politics". *Cognition*. 188. 8-18.
- Boudry, M., Vlerick, M., & McKay, R. (2015). "Can evolution get us off the hook? Evaluating the ecological defence of human rationality". *Consciousness and cognition*. 33. 524-535.
- Brycz, H., Wyszomirska-Góra, M., Bar-Tal, Y., & Wiśniewski, P. (2014). "The effect of metacognitive self on confirmation bias revealed in relation to community and competence". *Polish Psychological Bulletin*. 45(3). 306-31.
- Cacioppo, J. & Petty, R. (1982). "The Need for Cognition". *Journal of Personality and Social Psychology*. 42(1). 116-131.
- Čavojová, V., Šrol, J., & Adamus, M. (2018). "My point is valid, yours is not: myside bias in reasoning about abortion". *Journal of Cognitive Psychology*. 30(7). 656-669.
- Čavojová, V., Šrol, J., & Jurkovič, M. (2020). "Why should we try to think like scientists? Scientific reasoning and susceptibility to epistemically suspect beliefs and cognitive biases". *Applied Cognitive Psychology*. 34(1). 85-95.
- Christensen-Branum, L., Strong, A., & Jones, C. D. O. (2019). "Mitigating Myside Bias in Argumentation". *Journal of Adolescent & Adult Literacy*. 62(4). 435-445.
- Dacey, A. (2020). "Come Now, Let Us Reason Together: Cognitive Bias, Individualism, and Interactionism in Critical Thinking Education". *Informal Logic*. 40(1). 47-76.
- Drummond, C., & Fischhoff, B. (2019). "Does "putting on your thinking cap" reduce myside bias in evaluation of scientific evidence?" *Thinking & Reasoning*. 25(4). 477-505.
- Dutilh Novaes, C. (2018). "The enduring enigma of reason". *Mind & Language*. 33(5). 513-524.
- Eschert, S., & Simon, B. (2019). "Respect and political disagreement: Can intergroup respect reduce the biased evaluation of outgroup arguments?" *PloS one*. 14(3). e0211556.
- Evans, J. S. B. (2010). "Intuition and reasoning: A dual-process perspective". *Psychological Inquiry*. 21(4). 313-326.
- Evans, J. S. B. (2019). "Reflections on reflection: the nature and function of type 2 processes in dual-process theories of reasoning". *Thinking & Reasoning*. 25(4). 383-415.
- Evans, J. S. B., & Stanovich, K. E. (2013). "Dual-process theories of higher cognition: Advancing the debate". *Perspectives on Psychological Science*. 8(3). 223-241.
- Felton, M., Crowell, A., & Liu, T. (2015). "Arguing to agree: Mitigating my-side bias through consensus-seeking dialogue". *Written Communication*. 32(3). 317-331.
- Frankish, K. & Evans, J. S. B. (2009). "The duality of mind: an historical perspective". En Evans, J. S. B. & Frankish, K. (Eds.). *In Two Minds: Dual Processes and Beyond*. Oxford University Press. New York.
- Gigerenzer, G. (2008). "Why heuristics work". *Perspectives on Psychological Science*. 3(1). 20-29.

- Jurkovič, M. (2016). "Effect of short-term mindfulness induction on myside bias and miserly processing: A preliminary study". *Studia Psychologica*. 58(3). 231.
- Kahan, D. M. (2013). "Ideology, motivated reasoning, and cognitive reflection". *Judgment and Decision Making*. 8. 407–424.
- Kahan, D. M. (2015). "The expressive rationality of inaccurate perceptions". *Behavioral & Brain Sciences*. 40. 26-28.
- Kahneman, D. (2012). *Pensar rápido, pensar despacio*. Editorial Debate. Barcelona.
- Kahneman, D., & Tversky, A. (1974). "Judgment under Uncertainty: Heuristics and Biases". *Science* 185 (4157). 1124-1131.
- Kessler, E. D., Braasch, J. L., & Kardash, C. M. (2019). "Individual Differences in Revising (and Maintaining) Accurate and Inaccurate Beliefs About Childhood Vaccinations". *Discourse Processes*. 56(5-6). 415-428.
- Kruglanski, A. (2004) *The Psychology of Closed Mindedness*. Psychology Press. New York.
- Kunda, Z. (1990). "The case for motivated reasoning". *Psychological Bulletin*. 108(3). 480.
- Leitão, S., de Chiaro, S., Cano, M. (2016). "El debate crítico. Un recurso de construcción de conocimiento en el aula". *Textos de didáctica de la lengua y la literatura*. 73. 26-33.
- Macpherson, R. & Stanovich, K. (2007). "Cognitive ability, thinking dispositions, and instructional sets as predictors of critical thinking". *Learning and Individual Differences*. 17(2). 115-127.
- McCrudden, M. T., & Barnes, A. (2016). "Differences in student reasoning about belief-relevant arguments: A mixed methods study". *Metacognition and Learning*. 11(3). 275-303.
- McCrudden, M. T., Barnes, A., McTigue, E. M., Welch, C., & MacDonald, E. (2017). "The effect of perspective-taking on reasoning about strong and weak belief-relevant arguments". *Thinking & Reasoning*. 23(2). 115-133.
- Maynes, J. (2015). "Critical thinking and cognitive bias". *Informal Logic*. 35(2). 183-203.
- Mercier, H. (2017). "Confirmation bias-myside bias". En R. F. Pohl (Ed.) *Cognitive Illusions: Intriguing Phenomena in Thinking, Judgment and Memory*. Routledge. London.
- Mercier, H., Boudry, M., Paglieri, F., & Trouche, E. (2017). "Natural-born arguers: Teaching how to make the best of our reasoning abilities". *Educational Psychologist*. 52(1). 1-16.
- Mercier, H., & Sperber, D. (2017). *The Enigma of Reason*. Harvard University Press. Cambridge.
- Mercier, H., & Sperber, D. (2019) "Resumen de The Enigma of Reason". *Teorema. Revista Internacional de Filosofía*. 38(1). 59-67.
- Morris, J. A. (2017). *Effect of Collaborative Learning and Direct Instruction on Myside Bias* (Doctoral dissertation, University of Illinois at Urbana-Champaign).
- Nickerson, R. S. (1998). "Confirmation bias: A ubiquitous phenomenon in many guises". *Review of General Psychology*. 2(2). 175-220.
- Norman, A. (2016). "Why we reason: intention-alignment and the genesis of human rationality". *Biology & Philosophy*. 31(5). 685-704.

- Perkins, D. (2019). "Learning to reason: The influence of instruction, prompts and scaffolding, metacognitive knowledge, and general intelligence on informal reasoning about everyday social and political issues". *Judgment and Decision Making*. 14(6). 624.
- Sá, W., West, R., & Stanovich, K. (1999). "The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill". *Journal of Educational Psychology*. 91(3). 497–510.
- Smith, E., & Collins, E. (2009). "Dual-process models: A social psychological perspective". En Evans, J. S. B. & Frankish, K. (Eds.). *In Two Minds: Dual Processes and Beyond*. Oxford University Press. New York.
- Stanovich, K. E. (2018). "Miserliness in human cognition: The interaction of detection, override and mindware". *Thinking & Reasoning*. 24(4). 423-444.
- Stanovich, K. E., West, R. F., & Toplak, M. E. (2013). "Myside bias, rational thinking, and intelligence". *Current Directions in Psychological Science*. 22(4). 259-264.
- Stanovich, K. E., West, R. F., & Toplak, M. E. (2016). *The Rationality Quotient: Toward a Test of Rational Thinking*. MIT Press. Cambridge.
- Strachanová, D., & Valuš, L. (2019). "Mental simulation as a remedy for biased reasoning". *Studia Psychologica*. 61(2). 99-109.
- Tomasello, M. (2014). *A Natural History of Thinking*. Harvard University Press. Cambridge.
- van Brussel, S., Timmermans, M., Verkoeijen, P., & Paas, F. (2020). "'Consider the Opposite'—Effects of elaborative feedback and correct answer feedback on reducing confirmation bias—A pre-registered study". *Contemporary Educational Psychology*. 60. 101844.
- Wolfe, C. R., & Britt, M. A. (2008). "The locus of the myside bias in written argumentation". *Thinking & Reasoning*. 14(1). 1-27.
- Wolfe, C. R., Britt, M. A., & Butler, J. A. (2009). "Argumentation schema and the myside bias in written argumentation". *Written Communication*. 26(2). 183-209.

## LA HISTORIA COMO CIENCIA AUTÓNOMA DE LOS ASUNTOS HUMANOS

### COLLINGWOOD Y LA EXPERIENCIA COMO EXPERIENCIA HISTÓRICA

#### *HISTORY AS THE AUTONOMOUS SCIENCE OF HUMAN AFFAIRS*

*Collingwood and experience as historical experience*

#### *A HISTÓRIA COMO CIÊNCIA AUTÔNOMA DOS ASSUNTOS HUMANOS*

*Collingwood e a experiência como experiência histórica*

**Mariano Pettarin**

*(Universidad de Buenos Aires*

*Consejo Nacional de Investigaciones Científicas y Técnicas)*

*pettarinmariano@gmail.com*

Recibido: 10/11/2020

Aprobado: 16/06/2021

### RESUMEN

Este trabajo se propone elucidar la noción de experiencia propia de la filosofía collingwoodiana de la historia. Se argumentará que la experiencia resulta el elemento operante central e implícito desde el cual Collingwood construye los dos conceptos centrales de su crítica a la doctrina epistemológica realista según la cual conocer no implica diferencia alguna sobre lo que es conocido: la lógica de las preguntas y respuestas y la teoría de las presuposiciones absolutas. Primero, daremos cuenta como ambos conceptos surgen de la propia práctica de investigación del autor como alternativa al ejercicio de prácticas pseudo-históricas anidadas en el realismo y perimidas, como la historia de tijeras y engrudo y la historia crítica. Luego, apelaremos a las críticas de Quine y Davidson a los dos dogmas del empirismo demostrando su convergencia con las discusiones sostenidas por Collingwood contra sus adversarios realistas y la comparación entre los dos conceptos collingwoodianos y la lógica de investigación planteada por los dos autores, para así dar cuenta el pleno funcionamiento de un concepto collingwoodiano de experiencia capaz tanto de sortear eficazmente las críticas de idealismo y relativismo histórico, como de explicar eficazmente una práctica científica auténticamente histórica y abierta a la cooperación entre las diferentes formas de conocimiento o modos de experiencia.

Palabras clave: autonomía. experiencia. idealismo. presuposiciones. realismo.

### SUMMARY

This work aims to elucidate the original notion of experience present in Collingwood's philosophy of history. It presents how experience is the main implicit operating element behind the two concepts that undermine the realist epistemological doctrine. Knowing makes no difference to what is known: the logic of question and answer and the theory of absolute presuppositions. First of all, it will be explained how both concepts emerge from the author's historical practice as an alternative to entrenched, obsolete pseudo-historical practices such as the scissors and paste history and critical history. Furthermore, Quine and Davidson's



critique of the two dogmas of empiricism will prove helpful for showing the convergence between these two authors and Collingwood's quarrel with realism and comparing the two concepts as mentioned earlier and the logic of investigation as proposed by the two other authors. This approach will show a Collingwoodian concept of experience capable of withstanding possible idealist or relativistic labels and effectively explain a truly authentic scientific historical practice, open to cooperation between the different forms of knowledge or modes of experience.

Keywords: autonomy. experience. idealisms. presuppositions. realism.

## RESUMO

Este artigo visa elucidar a noção de experiência própria da filosofia da história collingwoodiana. Argumentaremos que a experiência é o elemento operacional central e implícito a partir do qual Collingwood constrói os dois conceitos centrais de sua crítica à doutrina epistemológica realista, segundo a qual saber não implica nenhuma diferença do conhecido: a lógica das perguntas e respostas e a teoria das pressuposições absolutas. Primeiramente, mostraremos como ambos os conceitos surgem da própria prática de pesquisa do autor como alternativa ao exercício de práticas pseudo-históricas aninhadas no realismo e desatualizadas, como o método histórico de “cola e tesoura” e a história crítica. Em seguida, apelaremos às críticas de Quine e Davidson a dogmas do empirismo, demonstrando sua convergência com as discussões mantidas por Collingwood contra seus oponentes realistas e a comparação entre os conceitos collingwoodianos e a lógica de pesquisa semeada pelos autores, para relatar o pleno funcionamento de um conceito collingwoodiano de experiência capaz tanto de classificar efetivamente as críticas do idealismo e do relativismo histórico quanto de explicar efetivamente uma prática científica autenticamente histórica aberta à cooperação entre diferentes formas de conhecimento e modos de experiência.

Palavras-chave: autonomia. experiência. idealismo. premissas. realismo.

## Introducción

Uno de los grandes objetivos de la empresa intelectual de Collingwood fue la defensa de la historia como ciencia autónoma de los asuntos humanos. Una ciencia dedicada al estudio de la *res gestae* o la actividad libre, las acciones de los hombres en su calidad de seres racionales e inscriptos en contextos sociales y culturales específicos (Collingwood, 2011 [1946]: 406).

Dicha autonomía, esbozada en *An Autobiography* y *Idea of History*, fue para este autor una expresión del valor práctico asignado a la historia en la comprensión de nuestra realidad presente. Lo que el historiador profesional provee es un *ojo entrenado* sobre cómo experimentamos esta realidad. De hecho, Collingwood insiste en que existe una relación estrecha entre los problemas propios de la investigación histórica y los problemas de la vida cotidiana, pues los primeros emergen de los segundos (Collingwood, 1982 [1939]: 114).

La riqueza de la noción de historia en Collingwood reside entonces en el hecho de que el rol fundamental de la historia en la constitución de identidades colectivas e individuales (su valor práctico) se vincula con cierta forma de experiencia humana en particular (Belvedresi, 2001: 136). Dicha vinculación ofrece pistas en relación con un aspecto central pero raramente explicitado de su filosofía de la historia. En otras palabras, el concepto de experiencia en Collingwood constituye un elemento central sobre el cual se configura la defensa de la ciencia histórica como actividad intelectual fundamental en nuestros tiempos modernos, ya que toda experiencia es experiencia histórica.

Sin embargo, Collingwood no se ha caracterizado por ser un pensador que explicita los conceptos con los cuales opera, flexibilizando su vocabulario para apelar de forma más eficaz a los problemas directrices de cada una de sus obras (D'Oro, 2005: lxx)<sup>1</sup>. Ello dificulta precisar la noción de experiencia que promueve y su articulación con otros conceptos fundamentales de su filosofía de la historia. Este escrito propone como problema de investigación la elucidación del concepto de experiencia operante en la filosofía collingwoodiana de la historia.

Se plantea como hipótesis que dicho concepto representa no sólo una elaboración original por parte del autor, sino que constituye el elemento nodal desde el cual se puede explicar el resto del andamiaje conceptual sobre el cual se construye toda su filosofía de la historia. Una consideración del concepto collingwoodiano en estos términos permitirá, a su vez, aprehender de forma más precisa las implicancias epistemológicas, potencialidades prácticas y posibles vinculaciones teóricas de su noción de historia.

Con un enfoque en sus obras principales sobre filosofía en general y filosofía de la historia en particular, procederemos a un análisis de dichos textos collingwoodianos poniéndolos en diálogo con concepciones alternativas sobre la experiencia. Ello nos permitirá ir precisando los contornos conceptuales de la experiencia tal como es concebida por Collingwood. Asimismo, dicha puesta en relación obedece a una secuencia lógica específica y consecuente con la evolución intelectual del autor.

En primer lugar, se establece una reposición del contexto de discusión del Collingwood, en el cual el autor rechaza ciertas formas de concebir la experiencia en tanto las considera anti-históricas y *realistas*, específicamente a partir de su crítica de la historia de *tijeras y engrudo* (scissors and paste) y la *historia crítica* (critical history)<sup>2</sup>.

Luego, reconstruiremos cómo Collingwood articula una concepción propia de la experiencia, operante de modo implícito en su filosofía de la historia pero que refleja en determinadas prácticas de investigación auténticamente históricas. Se analizará el movimiento collingwoodiano de alejamiento del *realismo* para la formulación de una posición propia anidada su teoría de las *presuposiciones absolutas* y su *lógica de preguntas y respuestas*.

Finalmente, Nuestra propuesta consistirá en hacer inteligibles las reflexiones de Collingwood sobre la experiencia histórica a partir del tratamiento que Quine y Davidson efectuaron en relación con la experiencia y la noción de interpretación. Dicha puesta en diálogo permitirá considerar cómo la experiencia entendida en términos collingwoodianos no sólo implica un distanciamiento efectivo de la antinomia entre *realismo* e *idealismo*<sup>3</sup> imperante el clima académico de Oxford, sino que también representa una formulación original comprensible en términos de que toda experiencia es experiencia histórica. Dicho planteamiento resulta funcional a su proyecto de defensa de la autonomía de la historia sin derivar tampoco en un relativismo histórico.

## 1. Collingwood y el rechazo a las formas “realistas” de hacer historia

En su escrito autobiográfico, Collingwood da cuenta del derrotero intelectual desde el cual irá cristalizando posiciones filosóficas opuestas a las enarboladas por lo que era la tendencia dominante en

---

<sup>1</sup> Por ejemplo, para el caso de *The Idea of History*, Collingwood busca clarificar los presupuestos subyacentes a la historia científica y marcar la utilidad de la historia en la vida pública. Si bien en *An Autobiography* Collingwood da cuenta de dichos presupuestos, su punto de partida es la crítica específica a lo que él consideraba como el *establishment* filosófico de Cambridge, Cook Wilson y sus realistas. El último de sus escritos editado en vida, *New Leviathan*, constituyó un alegato político contra el fascismo explicitando los procesos históricos de su surgimiento y consolidación. De esta forma, no resulta extraño ver como este autor modifica su terminología de acuerdo a los objetivos primarios de sus escritos.

<sup>2</sup> Tal como estos términos aparecen expresados en *The Idea of History* y *An Autobiography*.

<sup>3</sup> Utilizaremos el término de idealismo en el sentido de inmaterialismo, el compromiso con la tesis ontológica de la sola existencia de entidades mentales. En relación al historicismo radical, tomaremos la definición según la cual, “En términos generales, historicismo implica el punto de vista de que la perspectiva histórica es la única manera de comprender el mundo humano” (Requate, 1995: 94).

el ambiente académico filosófico de Oxford, el *realismo* y sus *minute philosophers*<sup>4</sup> quienes tenían en John Cook Wilson al máximo de sus referentes.

A partir del trazado de las consecuencias epistemológicas propias de la doctrina central y positiva de los realistas, “conocer no implica diferencia alguna sobre lo que es conocido” (Collingwood. 1982 [1939]: 44), Collingwood es capaz de desplegar una crítica certera contra aquellas posiciones. Como bien lo ha explicado D’ Oro (2006), la oposición de Collingwood hacia el *realismo* no implica un compromiso con el inmaterialismo, sino su objeción frente a la afirmación de la posibilidad del acceso a la realidad sin mediación epistémica, una idea

Promovida por John Cook Wilson en Oxford, Ernst Mach en Vienna y G.E.Moore y Bertrand Russell en Cambridge –como sus contrapartes en Estados Unidos- revivieron las tempranas pre-kantianas tradiciones del empirismo Británico. Collingwood fue uno de los primeros filósofos en darse cuenta que ello no era posible (Toulmin, 1982: xiii).

Como consecuencia de sostener que “conocer no implica diferencia sobre lo conocido” se adopta una consideración “simplificada” y “aprehensiva” del conocimiento, lo que supone un divorcio entre la teoría y práctica, haciendo de ambas actividades fútiles (Collingwood, 1982 [1939]: 25). Pero Collingwood no se detuvo aquí, en tanto buscó explicitar los compromisos epistemológicos subyacentes a dicha pasividad. En este sentido, lo que el historiador-filósofo británico cuestiona son los presupuestos empiristas-positivistas de los realistas, la concepción de la mente humana como una *tabula rasa*, sustentada en la convicción de que todas nuestras actividades mentales se derivan en última instancia de una percepción sensible de objetos externos, lo que en términos de esta doctrina puede caracterizarse como experiencia (Collingwood, 1982 [1939]: 25-26).

La experiencia concebida realísticamente no sólo opera sobre un receptor pasivo, sino que también habilita su potencial verificación gracias a las capacidades de traducibilidad en un lenguaje idealmente desprovisto de ambigüedades en el que existe “una correspondencia uno a uno entre las proposiciones y las oraciones indicativas, expresando cada oración indicativa una proposición, siendo una proposición definida como la unidad de pensamiento, o aquello que es verdadero o falso” (Collingwood, 1982 [1939]: 36). El proyecto de construir un lenguaje cristalino enfatiza el carácter reduccionista de la concepción *realista* de la experiencia, la correspondencia uno a uno entre un segmento de la realidad que se nos presenta desde la percepción (dejando su registro en la *tabula rasa*) y un enunciado expresando una unidad mínima de pensamiento. Desde este punto de vista, los enunciados son susceptibles de ser verificados o refutados de forma aislada: “que una proposición sea o verdadera o falsa por sí misma, siendo la verdad o falsedad cualidades de esas proposiciones... denominar –una proposición- como verdadera o falsa implica asignar una relación de correspondencia o no correspondencia entre ella y algo que no es una proposición, un ‘estado de cosas’ o hecho (Collingwood, 1982 [1939]: 36). De esta forma, Collingwood critica con especial fuerza la versión atomista de *realismo*.

El interés de Collingwood por explicitar los compromisos de la teoría *realista* no responde a una mera ambición de refutación *per-se*, sino por las consecuencias que ello acarrea para la investigación histórica: Los filósofos *realistas* ejercitaban una filosofía encapsulada en el academicismo, directa expresión de su central doctrina sobre que el conocer no tiene consecuencia alguna sobre lo que es conocido y un reflejo de la convicción de que el acceso a la realidad no precisa mediación epistémica alguna. Según Collingwood, los mismos compromisos ontológicos subyacen a ciertas prácticas presentes en la historiografía profesional, prácticas que Collingwood supo condensar en la historia de *tijeras y engrudo* y su versión un tanto más compleja pero de la misma naturaleza, la *historia crítica*.

Una de las formas de investigación pseudo-históricas que Collingwood (2011 [1946]) rechaza fuertemente es la historia de tijeras y engrudo, la cual “como ya he dicho, no es realmente historia, pero no tenemos otro nombre que darle” (341). El único procedimiento de la historia de *tijeras y engrudo* es

---

<sup>4</sup>En este caso, “minute” implica la consideración de estos filósofos como filósofos casuales o improvisados.

la recopilación y condensación de testimonios de un tema determinado (con especial predilección por los testigos directos).

Frente a los testimonios de las denominadas autoridades, el historiador oficia como un mero compilador y reproductor, un medio para que ellas puedan desplegarse (Collingwood 2011 [1946]: 341). Para Collingwood un historiador de estas características es un historiador-realista, un pseudo-historiador, una *tabula rasa* que reedita la doctrina central realista sobre que el conocer no implica diferencia alguna sobre lo que es conocido.

Pero a Collingwood le resulta mucho más interesante desarrollar cómo la *historia crítica*, emergente de la práctica de las tijeras y el engrudo, reafirma los compromisos ontológicos del *realismo*. La *historia crítica* surge de la propia comparación de las autoridades y del hecho que la presencia de testimonios conflictivos representa un escollo que no puede resolverse en la mera recopilación. Ante la presencia de evidencia conflictiva, el historiador crítico debe resolver el problema por la incorporación o la exclusión de un testimonio (Collingwood, 2011 [1946]: 343).

Si bien Collingwood destacó que la *historia crítica* es más sofisticada que la de *tijeras y engrudo*, ella no problematiza la concepción aprehensiva de la experiencia. Es decir, el criterio con el cual se incorporan o descartan las autoridades -devenidas ahora en fuentes- es desde la verosimilitud de sus enunciados. El británico ofreció el caso de sus colegas que estudiaban la Grecia Antigua, a quienes acusaba de enquistarse en una ortodoxia según la cual “el último evento significativo en los estudios históricos sobre Grecia clásica... fue el de comparar las dos versiones de la Revolución Ateniense provistas por Tucídides y Aristóteles, decidiendo punto por punto cual era la más verosímil” (Collingwood, 1982 [1939]: 83).

Siguiendo a Collingwood, la historia crítica alcanzó su zenit en historiadores como Mommsen o Maitland, quienes promovieron un gran avance del conocimiento de los hechos gracias a la elevación sin precedentes del examen exacto y crítico de las pruebas históricas, siendo cautelosos a la hora de aplicar sin más los principios de generalización provenientes de las ciencias naturales (Collingwood, 2011 [1946]: 200).

Collingwood se interesó especialmente en Mommsen porque fue su obra la que encarnaba de forma más cristalina los principios metodológicos de la *historia crítica*, con los logros y las limitaciones que ellos conllevan: 1) La consideración de un hecho como una unidad comprobable mediante un acto cognoscitivo o un proceso de investigación separado, cuyo correlato es la atomización del campo de lo históricamente cognoscible en una infinidad de hechos minúsculos y 2) La consideración de los hechos como unidades independientes exige la eliminación de los elementos subjetivos por parte del investigador.

De esta forma, para Collingwood los logros de Mommsen en su *Historia de Roma* se limitaron a la resolución de problemas de pequeña escala, como el efectivamente detectar donde se reclutaban las legiones (en diferentes períodos temporales claramente especificados) gracias a la apelación a una minuciosa recopilación de un corpus de derecho romano. A la hora de analizar los procesos de transformación en la Roma Antigua (lo que Collingwood define como problemas de gran escala), el intento de Mommsen se quiebra (Collingwood, 2011 [1946]: 204).

Pero las objeciones a la historia de *tijeras y engrudo* e *historia crítica* trascienden lo metodológico. Más bien, las limitaciones metodológicas expresan los compromisos ontológicos subyacentes de quienes la practican. Las críticas de este pensador hacia los realistas en su *Autobiography* convergen con las refutaciones a las formas pseudo-históricas presentes en *The Idea of History* en tanto ambas reproducen una concepción pasiva de la experiencia. Ello aparece con mayor claridad en *Principles of History*:

Puede decirse, en respuesta a esta contención, que todo lo que el historiador moderno puede hacer es ordenar un mosaico de hechos, siendo estos hechos los conocidos por los propios contemporáneos, sin incorporar ningún hecho que estos desconocían. Esto es como la contención humana de que toda idea debe derivar de

una impresión y que todo lo que la imaginación hace es redistribuir los materiales derivados de las sensaciones (Collingwood, 1991: 150).

Esta concepción pasiva de la experiencia como compromiso ontológico fundamental se reflejaba, siguiendo a Collingwood, en el cercenamiento de la autonomía de la historia y su servilismo hacia los principios de las ciencias naturales, específicamente la necesidad de considerar los hechos históricos como unidades atómicas verificables y expresables en enunciados singulares.

El resultado de estas prácticas no puede ser otro que una (pseudo) historia sumamente limitada a la resolución de problemas académicos a pequeña escala y desconectada de los problemas contemporáneos. En el caso de los pseudo-historiadores con ambiciones de generalización, que buscan encontrar en la historia el cumplimiento de leyes generales (o algún tipo de regularidad) al estilo de las ciencias naturales, la frustración resulta mayor al observar la persistencia de lo contingente<sup>5</sup>. En otras palabras, ambas versiones constituyen expresiones de *realismo* y configuran un polo opuesto al proyecto collingwoodiano de la defensa de la historia como la ciencia autónoma de los asuntos humanos.

Si, siguiendo a Collingwood, es en la práctica (pseudo) histórica donde se puede constatar el principio fundamental sobre la concepción pasiva de la experiencia, sólo el análisis de la investigación auténticamente histórica puede arrojar pistas sobre una concepción correctamente histórica de la experiencia o, en otras palabras, del compromiso normativo ineludible para el ejercicio y la defensa de una ciencia autónoma preocupada por la resolución de las problemáticas contemporáneas.

Collingwood hace uso de dos conceptos centrales para dar cuenta la investigación auténticamente histórica. El primero de ellos, la *lógica de preguntas y respuestas*, explica cómo se desarrolla la práctica histórica científica propiamente dicha. Luego, las *presuposiciones absolutas* apuntan hacia los compromisos subyacentes a cualquier forma de conocimiento.

## 2. Las presuposiciones absolutas subyacentes a la lógica de las preguntas y respuestas.

Frente a la convicción realista de que los enunciados pueden ser verificados aisladamente, Collingwood argumentó que los statements no deben ser desagregados, sino que forman parte de un complejo de preguntas y respuestas en el cual cada enunciado tiene sentido en función de su contribución a un todo significativo (Collingwood, 1982 [1939]: 37).

Entonces para Collingwood la posibilidad de significar (verificar) escapa a los enunciados particulares, un movimiento argumental que el autor expresa en la prioridad de la palabra correcto sobre verdadero a la hora de juzgar la pertinencia de un complejo de enunciados: “Por ‘correcto’ no quiero decir ‘verdadero’. La respuesta ‘correcta’ a una pregunta es aquella que posibilita continuar con el proceso de preguntas y respuestas” (Collingwood, 1982 [1939]: 37). Por ejemplo, en el caso de la historia, la comparación, enunciado a enunciado, de diferentes explicaciones, es desplazada por la coherencia de la narrativa histórica como un todo, un complejo articulado de preguntas y respuestas. Además, la refutación de Collingwood admite una posición más activa del investigador antes que la aprehensión como punto de partida. Serán calificadas como más fructíferas no aquellas narrativas capaces de comprimir la mayor cantidad de enunciados veritativos, sino las que habiliten la integración de nuevas preguntas dentro de un complejo de preguntas y respuestas crecientemente sofisticado.

Al describir su derrotero intelectual, el autor comentó cómo sus conclusiones en relación al proceso de preguntas y respuestas emergieron de su actividad como arqueólogo e historiador. Esto es específicamente ilustrado por el autor al criticar con dureza la aplicación de principios de investigación obsoletos a la excavación arqueológica. Para Collingwood, el *blind digging*, o la mera excavación dentro

---

<sup>5</sup> Collingwood hace especial referencia a un historiador de corte Spencerista, de nombre Robert Mackenzie, quien publicó en 1880 un libro titulado *The Nineteenth Century. A History*, “donde caracterizaba el siglo XIX como una época de progreso a partir de un estado de barbarie, ignorancia y bestialidad, que apenas puede exagerarse, para llegar a un reino de ciencia, ilustración y democracia” (Collingwood, 2011 [1946]: 219).



de un espacio físico esperando recuperar la mayor cantidad de artefactos, pero sin ninguna problemática en mente, no era otra cosa que una expresión del *realismo* (Collingwood, 1982 [1939]: 30)<sup>6</sup>. Pero Collingwood veía como en la arqueología se habían comenzado a desarrollar investigaciones dentro del marco de complejos de preguntas y respuestas, atendiendo a problemáticas específicas antes de iniciar el proceso de excavación, como la búsqueda del propósito de determinadas lógicas edilicias imperantes en la Britania Romana, o el alcance del proceso de romanización cultural de las élites británicas considerando la evidencia arqueológica perteneciente también a los períodos posteriores a la caída del Imperio (Collingwood, 1982 [1939]: 127-130).

Por fuera de la arqueología, el ejemplo que más le interesaba a nuestro autor era el de John Bagnell Bury, un historiador anglo-irlandés de fines del Siglo XIX e inicios del XX. Collingwood calificó a Bury como un positivista crecientemente perplejo. Inicialmente, Bury defendía la postura que la historia consistía en un compuesto de hechos aislados, cada uno de ellos comprobable sin referencia a los otros. En su *A History of the Roman Empire From its Foundation to the Death of Marcus Aurelius* (1893), Bury replica a rajatabla los métodos del positivismo importados de las ciencias naturales, al tratar de agrupar hechos aislados en tipos de hechos, para luego intentar ver si obedecen al funcionamiento de determinadas leyes. Pero luego Bury, en una serie de conferencias entre las cuales Collingwood destaca *The Place of Modern History in the Perspective of Knowledge* (1902), comienza a cuestionar el método que le había guiado hasta ese entonces, argumentado que el pensamiento histórico debía ser necesariamente algo diferente y relativamente nuevo a las ciencias naturales, aunque sin ofrecer demasiadas precisiones. Profundiza este problema en *Cleopatra's Nose* (1916), donde reflexiona sobre la contingencia en la historia para concluir que su verdadera función no es distinguir entre lo accidental y lo necesario, sino explicar y desarrollar lo individual. Sin embargo, se lamenta Collingwood, Bury termina retrocediendo, para explicar como lo individual deviene en un elemento irracional, esperando que con el avance de la ciencia ese elemento pueda desaparecer en los estudios históricos (Collingwood, 2011 [1946]: 221-224).

Lo importante del caso de Bury para Collingwood es que ilustra cómo no puede haber una transformación hacia prácticas auténticamente históricas sin una crítica profunda de los propios presupuestos subyacentes a dichas prácticas, y que el propio Bury es seducido hacia, pero finalmente no efectúa una transición. En otras palabras, Bury manifiesta un *pensamiento filosófico* pero de carácter limitado. Detecta que existen ciertos compromisos epistémicos necesarios para el ejercicio de una disciplina (en su caso, el apego a que una explicación auténticamente científica responde al sometimiento de los fenómenos a regularidades legaliformes) y que esos compromisos se encuentran en conflicto con su práctica como historiador, pero no llega a criticarlos ni comprender como funcionan de forma articulada con la fundamental doctrina *realista* y la noción de experiencia anidada a ella. El rechazo definitivo del intuicionismo *realista* converge entonces con la necesidad de desarrollo de un concepto capaz de explicar los compromisos epistémicos habilitantes (y aquellos que cercenan) del despliegue de los procesos de preguntas y respuestas.

En *An Essay on Metaphysics*, Collingwood (2014 [1939]) desarrolló como un complejo de preguntas y respuestas habilita que en su interior existan enunciados capaces de ser verdaderos o falsos, a los cuales denominó proposiciones (25). Pero lo que le interesó a Collingwood no es que ciertas proposiciones sean verdaderas o falsas, sino la razón por la cual ellas son enunciadas en primer lugar. Aquello cuya eficacia lógica permite que ciertas proposiciones sean verificadas son las presuposiciones absolutas a las que debemos comprometernos. De esta forma, las presuposiciones no constituyen nunca respuestas a preguntas, y por ello no pueden estar sometidas a los criterios de verdad o falsedad. En otras palabras, postular que una presuposición absoluta es verdadera o falsa carece de sentido (Collingwood, 2014[1939]: 32-33).

El concepto de presuposiciones absolutas reformula la objeción central de Collingwood frente a sus adversarios realistas, quienes sostienen una postura anticientífica que no asume las presuposiciones absolutas involucradas en cualquier forma de conocimiento con pretensiones de ser sistemático. En el

---

<sup>6</sup> Cabe resaltar que Collingwood no hacía una distinción tajante entre la historia y la arqueología. De hecho consideraba a la arqueología como una forma de historia que trataba con tipos de fuentes específicas. (Collingwood, 1982 [1939]: 131).

caso específico de los realistas, es su ceguera frente al hecho que sostener que el conocimiento las cosas se dan por aprehensión resulta una convicción, antes que el producto de la mera intuición. Collingwood explicó cómo una de las razones por las cuales dichas presuposiciones parecen no estar operando en nuestro acceso al conocimiento es por el hecho de que su precedencia respecto al complejo de preguntas y respuestas no es temporal sino lógica o, en términos más precisos de Collingwood, eficazmente lógica.

Una de las características más importantes de este tipo de presuposiciones es que no configuran categorías universales. Las presuposiciones absolutas se agrupan en constelaciones históricas y son, por lo tanto, inevitablemente transitorias. De hecho, *The Idea of History* (su primera sección), *The Principles of Art* y *The Idea of Nature* representan el desarrollo de las presuposiciones absolutas en el arte, las ciencias físicas y la ciencia histórica desde la antigüedad hasta la física, biología e historia modernas. Por ejemplo, para el caso de la historia, el barómetro desde el cual medir los cánones de la historia científica no es ahistórico, sino es gracias a prácticas de investigación y pensamiento histórico instituidas (en realidad en vías de institución y que activan aquellos procesos resistencia -por ejemplo por parte de los *realistas*- criticados duramente por el británico) que se pueden juzgar ciertos anquilosamientos académicos como perimidos o pseudo-históricos. De hecho, uno de los problemas más acuciantes es que la persistencia de determinadas prácticas obsoletas en ciencias sociales se debe a que ciertas presuposiciones absolutas que rigen el estudio de los fenómenos naturales son trasladadas sin mayores resquemores a la comprensión de las acciones de los hombres. Es decir, se asumen los compromisos ontológicos de los *realistas*, la equivalencia entre experiencia y percepción. Por contraste, la tesis sobre las presuposiciones absolutas en su relación de eficacia lógica con la *lógica de preguntas y respuestas* anida en una concepción mucho más sofisticada, aunque no explicitada de experiencia que implica, a su vez, una superación de la noción de experiencia realista. En otras palabras, para que el despliegue de un proceso de preguntas y respuestas sea posible, es necesaria la operación de una determinada constelación de presuposiciones absolutas históricamente configurada, siendo una noción específica de experiencia la parte central, el mortero de dicha constelación. Pero, al ser una concepción que funciona de manera implícita en Collingwood, apelaremos a desarrollos de autores posteriores pero afines en determinados aspectos al historiador-arqueólogo británico para elucidar dicho concepto.

### 3. La experiencia más allá del *realismo* e *idealismo* (y del relativismo).

Una de las objeciones recurrentes al pensamiento de Collingwood por parte de los interlocutores es que su tesis son una expresión de *idealismo*. Idealista en el ambiente filosófico de esos tiempos no implicaba tan sólo el rechazo de una postura filosófica, sino que también acarreaba una recusación de tipo académico-política: Significaba asociar a cualquiera que sea depreciado con dicha denominación la pertenencia a la facción desplazada de los discípulos de Green y Bradley (Toulmin, 1982: 19). Para los realistas, idealistas eran aquellos que sostenían el inmaterialismo, o el compromiso con la tesis ontológica de la sola existencia de entidades mentales. Pero, como hemos visto y cómo apunta claramente D'Oro, Collingwood no se interesó principalmente en criticar de forma directa el compromiso ontológico realista respecto la existencia de objetos por fuera de nuestra mente, sino de forma indirecta, refiriendo a los procedimientos por los cuales los realistas verifican (como verdaderos o falsos) las propiedades de dichos elementos: La doctrina de que el conocer no implica diferencia alguna sobre lo que es conocido. Collingwood le interesaba la crítica como punto de partida para la construcción, rescatando los aportes positivos de sus adversarios intelectuales, como el tratamiento sistemático y riguroso de la evidencia legada de la *historia crítica*. Es por ello que la acusación de idealista, como una muletilla descalificatoria y herramienta de cuestionamiento de todo un edificio teórico sin haberse adentrado rigurosamente en él, carecía de validez para nuestro autor.

En su lectura de Collingwood, Louis Mink explica cómo Collingwood fue siempre reacio a posicionarse frente a contraposición entre *realismo* o *idealismo* siendo esta entendida en términos de un dualismo insalvable, un débil compromiso y no una síntesis iluminadora (Mink, 1969: 112).<sup>7</sup>

Esta síntesis o, en palabras de Mink, teoría comprensiva (de la cual las presuposiciones absolutas y la lógica de las preguntas y respuestas son una expresión) implica la construcción de conceptos y la convicción de que dichos conceptos son parte de una constelación de presuposiciones absolutas robusta, en la cual su elemento central, la experiencia, no debe ser necesariamente puesto en superficie. Demostraré cómo esa concepción puede ser explicitada desde las formulaciones de Quine y Davidson.

En *Desde un punto de vista lógico*, Quine critica los compromisos epistemológicos (afines a los sostenidos por los realistas de Collingwood) a partir de la puesta en relieve de los dos dogmas del empirismo. El primero de ellos siendo la distinción entre lo analítico y lo sintético, o la diferencia fundamental entre verdades basadas en significaciones independientes de las consideraciones fácticas y aquellas basadas en los hechos. El segundo dogma es el del reduccionismo, la creencia de que lo que otorga sentido a un enunciado es la referencia a la experiencia inmediata (Quine, 2002 [1980]: 62).

En términos collingwoodianos, se podría argumentar que los dos dogmas del empirismo configuran la doctrina central del *realismo* que conocer no hace diferencia alguna sobre lo que es conocido en tanto implican percepción y aprehensión.

La teoría de la verificación, la cual sostiene que el sentido o significación de un enunciado es el método de confirmación o refutación empírica, colapsaría junto a la distinción analítico-sintética. De hecho, un enunciado analítico representa un caso límite quedando confirmado en cualquier supuesto (Quine, 2002 [1980]: 81). Este desarrollo es afín a la renuencia por parte de Collingwood de otorgar significado a los enunciados aislados, por fuera de los complejos de preguntas y respuestas a los cuales pertenecen.

Siguiendo a Quine, la relación de la ciencia con la experiencia es mucho más compleja que la planteada por el reduccionismo, ya sea sólo por el simple hecho de que el lenguaje que empleamos para significar lo extra-lingüístico es opaco y no habilita establecer relaciones de correspondencia uno-a-uno entre enunciados y datos de la experiencia. La réplica realista frente a los escollos del lenguaje ordinario podría ser la de articular un lenguaje artificial, más cristalino y desde el cual se pueda definir con claridad cómo los dos dogmas del empirismo pueden ser directamente verificados por la experiencia. Pero esta empresa culmina en un sinsentido al postular que la noción de analiticidad, presupuesto sobre el cual se asienta el reduccionismo empirista, no se puede definir más que auto-referencialmente. Es decir, postular la existencia de enunciados analíticos es en sí mismo un enunciado analítico (Quine, 2002 [1980]: 79-80).

Quine aplica esta crítica a lo que el propio Collingwood denominó como dos vertientes de una misma postura realista: El “imposible empirismo de Locke y Hume” y la tesis de Frege, con la cual “el enunciado llegó a ser reconocido, en vez del término, como la unidad relevante para una crítica empirista” (Quine, 2002 [1980]: 82).

Frente al reduccionismo, Quine expone cómo nuestros enunciados acerca del mundo externo se someten como cuerpo total al tribunal de la experiencia sensible y no individualmente, lo que no es otra forma de decir que nuestro corpus lingüístico configura un todo interconectado. Nuestro conocimiento, sea en el campo de la física, matemática, la historia, es una fábrica construida por el hombre que no está en contacto con la experiencia más que lo largo de sus lados. “O, con otro símil, el todo de la ciencia es como un campo de que no está en contacto con la experiencia más que a lo largo de sus lados” (Quine,

---

<sup>7</sup> La razón principal desde la cual Collingwood explica las limitaciones de la historia crítica de Bradley es por la imposibilidad de su idealismo para trascender la doctrina de que conocer no implica diferencia alguna sobre lo que es conocido: “De esta suerte, Bradley ha legado un dilema a sus sucesores. O la realidad es un flujo de la vida subjetiva, en cuyo caso es subjetiva pero no objetiva, se disfruta pero no es posible conocerla, o es aquello que conocemos, en cuyo caso es objetiva pero y no subjetiva y es un mundo de cosas reales externas a la vida subjetiva de nuestra mente y externas las unas respecto a las otras... El mismo Bradley agarró el primer cuerno del dilema; pero agarrar cualquiera de los dos es caer en el error fundamental de concebir la vida de la mente como un flujo inmediato de percepciones y sensaciones, vacío de toda reflexión y autoconocimiento” (Collingwood, 2011 [1946]: 215).

2002 [1980]: 87). Lo extra-lingüístico determina sólo tangencialmente la configuración total de un sistema. Resulta entonces absurdo buscar una divisoria entre enunciados sintéticos, que valen contingentemente por experiencia, y analíticos, que valen en cualquier caso. La correspondencia entre enunciados-experiencia representa tan sólo la periferia de una malla interconectada de creencias, y los enunciados periféricos suelen ser los primeros en ser sacrificados ante un flujo de experiencia conflictivo. Es decir, mientras los enunciados centrales de la red presentan poca conexión preferencial con algún elemento de la experiencia, los enunciados periféricos funcionan como *fusibles* dada “nuestra natural tendencia a evitar perturbar el sistema en su conjunto” (Quine, 2002 [1980]: 88).

Desde este punto de vista, la propuesta de Collingwood puede ser interpretada como convergente con la de Quine. Los enunciados inscriptos en procesos de preguntas y respuestas no pueden verificarse aisladamente, como *units of thought*. Además, la relación entre las presuposiciones absolutas que habilitan la emergencia (o que gracias a ellas podemos formular determinadas preguntas) de los complejos de preguntas y respuestas y la continuidad de la actividad inquisitiva, es afín a la distinción de Quine entre los elementos centrales de nuestra malla de creencias y los elementos periféricos. Es decir, reformulando la tesis de Quine en términos de Collingwood, resulta mucho más conveniente para la continuidad de la investigación admitir la incorrección en el derrotero de cierto complejo que comprometer nuestras presuposiciones absolutas.

Si bien Quine y Collingwood confrontan el empirismo reduccionista desde posiciones que pueden ser leídas como afines, la tesis de Quine tiene derivas dualistas con las que Collingwood no se compromete. Nos referimos específicamente al concepto de esquema conceptual, del que Quine hace uso para delimitar las diferentes mallas de creencias (expresadas en enunciados) con las que damos cuenta del flujo de la experiencia: “La ontología de cada cual es básica para el esquema conceptual mediante el cual interpreta todas las experiencias, incluso las más tópicas” (Quine, 2002 [1980]: 49). Desde este punto de vista, una afirmación ontológica que es axiomática en un esquema conceptual puede ser juzgada como trivial en el marco de otro. El peligro que acarrea el hacer uso de la noción de esquema conceptual como contraposición a la distinción entre los enunciados analíticos es que su reemplazo da lugar a otro dualismo, el de esquema-contenido.

Davidson da cuenta de los componentes necesarios para la construcción de una propuesta de tales características, “el lenguaje como fuerza organizadora, sin que se haga una distinción clara con respecto de la ciencia... (y) lo que se organiza, la corriente de experiencia sensible” (Davidson, 2011 [1984]: 9). Entonces, el supuesto de la experiencia neutra (la pasividad del sujeto frente a la experiencia o, en términos realistas el hecho de que “conocer no implica diferencia alguna sobre lo que es conocido”) es una condición necesaria para la existencia de un esquema conceptual, alimentando la metáfora “de un espacio único dentro del cual cada esquema tiene una posición y proporciona un punto de vista” (15). Es decir, si son el conjunto de los enunciados dentro de un esquema conceptual los que hacen frente al tribunal de la experiencia, siendo la totalidad de evidencia existente lo único que puede confirmar o refutar las entidades postuladas en los enunciados, la caracterización de lenguajes o esquemas conceptuales en términos de adecuación a alguna entidad se reduce a la idea de que algo es un esquema conceptual aceptable si es verdadero. Dentro de la teoría de Quine, se puede admitir que un esquema conceptual sea verdadero pero no traducible (Davidson, 2011 [1984]: 14). De esta forma, es preciso dar el paso que propone Davidson, abandonar la metáfora de un espacio único donde los diferentes esquemas conceptuales conviven y refieren a la misma realidad neutral para así habilitar la posibilidad de traducción e interpretación. Es tomando este paso que se puede comprender con mayor precisión el concepto de experiencia anidado en un proyecto collingwoodiano, que critica ciertas formas de pensar el pasado al ser ellas expresión de presupuestos realistas e importados de las ciencias naturales.

#### **4. La experiencia como experiencia histórica.**

La defensa de la autonomía de la historia depende de la capacidad de que diferentes formas de pensar el pasado, constelaciones diferentes de presuposiciones absolutas puedan ser traducidas, revisadas y criticadas desde nuestro punto de vista. Es decir, desde una posición que puede ser entendida en términos



holistas (aunque el autor jamás empleo ese término), a Collingwood no le interesa verificar la validez de enunciados individuales, más bien problematizar los complejos de preguntas y respuestas. Desde esta consideración, los complejos de preguntas y respuestas no presentan las derivas problemáticas del dualismo esquema contenido apuntando por Davidson, en tanto configuran un proceso histórico y que es a su vez experiencia. De esta forma es posible comprender cómo el realismo no hace sino expresar el desarrollo dentro de un complejo preguntas y respuestas antes que sostener una relación ontológica entre entidades y un sujeto para quien el conocer no hace diferencia alguna sobre lo que es conocido.

Collingwood acordaría con Davidson cuando el segundo afirma que no se puede apelar a la evidencia para argumentar que las creencias de otros no son verdaderas cuando estas difieren o parecen tener grandes diferencias con las nuestras (Davidson, 2011 [1984]: 19). El criterio que usa Collingwood es aquí el mismo, ya que el historiador realista no puede hacer otra cosa que fracasar en comprender el pasado porque es incapaz de aprehender la naturaleza de la experiencia humana en tanto histórica y procesual. En su lectura de Collingwood, es el propio Davidson quien marca esta diferencia:

El éxito aparente de la ciencia en la explicación de los fenómenos físicos pone sobre relieve la cuestión sobre si la misma metodología es aplicable a las acciones humanas y a los estados mentales y eventos a los cuales apelamos para explicar dichas acciones. El trabajo de R.G. Collingwood hizo del estudio de historia algo central en tanto ahonda en la explicación y comprensión de la acción. Según Collingwood, la metodología de la historia (o cualquiera de las ciencias sociales que tratan del comportamiento humano individual) difiere marcadamente de la metodología de las ciencias naturales (Davidson, 2005: 282).

Al argumentar que la experiencia es un proceso histórico, Collingwood rechaza la separación tanto de esquema y contenido (aunque su interlocutor directo no haya sido Quine y Collingwood no lo haya planteado en estos términos) como de sujeto y objeto. De hecho, para este autor tanto el *idealismo* como el *realismo* reproducen dichos dualismos al otorgar prioridad epistémica a alguno de estos polos.

De esta forma, la propuesta de Collingwood rechaza tanto la existencia de un mundo externo como de un mundo interno independiente para transformar al proceso como principal referencia de análisis. De allí su insistencia a afirmar que el *idealismo* no es otra cosa que una redición del *realismo* ya que, aunque tengan nombres diferentes, ambos arrojan las mismas consecuencias epistemológicas anti-procesuales (Collingwood, 2015 [1933]: 246).

Dentro de dicho proceso se transforman las constelaciones de presuposiciones absolutas con las cuales construimos nuestras teorías sobre un mundo externo que, como es siempre importante reiterar, es inaprensible por fuera de dichas presuposiciones. Dichos procesos de transformación son necesariamente históricos. Ello implica que no existe una línea divisoria pre-establecida desde la cual delimitar procesos, ya que las propias nociones de principio y fin atentan contra la misma idea de proceso (Collingwood, 1982 [1939]: 98).

El término específico con el que Collingwood denomina a esta dinámica es el complejo de preguntas y respuestas, en los cuales (insistimos) los criterios de verdad y falsedad de los enunciados dependen de la corrección del complejo como un todo, o sea de la capacidad de que las respuestas habiliten preguntas nuevas y viceversa. Cuando Collingwood afirma que tanto el estudio de los fenómenos naturales como los fenómenos humanos se resuelven en la historia, lo que pretende no es trasladar la hegemonía del *realismo* al relativismo histórico, sino tan sólo remarcar cómo “La ciencia natural como una forma de pensamiento existe y siempre ha existido en el contexto de la historia” (Collingwood, 2014 [1945]: 177). Collingwood no busca impugnar por completo el *realismo*, sino más bien marcar la transitoriedad de sus compromisos ontológicos (sus presuposiciones) al integrarlos dentro de un proceso histórico. Sobre este punto, Davidson también acuerda al remarcar que la propia noción de objetividad, la externalidad del objeto de investigación respecto al sujeto (antes que el carácter construido del mismo), es una concepción filosófica históricamente constituida (Davidson, 2003: 297).

Siguiendo a Collingwood, la única constante de la experiencia es su procesualidad, su historicidad, lo que hace a cualquier concepción de experiencia revisable. Esto resulta un elemento fundamental a la



hora de entender cómo la defensa collingwoodiana de la autonomía de la historia es una expresión de las diferentes, legítimas e inter-traducibles *formas de experiencia* (modes of experience).

El problema de los *modos de experiencia* es planteado desde un punto de vista filosófico en *An Essay in Philosophical Method*. En este texto, Collingwood (1933) se pregunta sobre el enigma que surge al comparar cómo diferentes formas de investigación ofrecen explicaciones diferentes sobre una misma instanciación: Lo que puede ser descrito como un suceso histórico en términos de las acciones de los hombres (como los hombres proyectan y desarrollan sus intenciones en contextos adversos) también admite una explicación en términos de evento, como expresión de regularidades (por ejemplo en el caso de una epidemia). Estas formas de instanciarse sobre el mismo sustrato de la realidad encarnan *modos de experiencia*, que expresan el funcionamiento de diferentes conceptos filosóficos (término que será reemplazado por el de presuposiciones absolutas en *An Essay of Philosophical Method*).

De hecho, para Collingwood la distinción principal no reside en el hecho de que la historia estudia el pasado, pues las ciencias naturales también lo hacen. Lo importante es cómo las diferentes presuposiciones absolutas operan en la construcción de un pasado natural y un pasado histórico. El pasado natural se basa en la uniformidad de la naturaleza, y busca en fenómenos ya acaecidos la manifestación de leyes empíricas. Por el contrario, el pasado histórico se interesa por las acciones de los hombres, impredecibles y cambiantes (Collingwood, 2011 [1946]: 221-224). El meollo de la cuestión es que en las explicaciones de ciencias naturales el concepto principal que construye ese *modo de experiencia* es el de evento, que habilita el establecimiento de regularidades y predicciones, mientras que en la historia (y las ciencias sociales) es la mente -o la idea de acción intencional- la que permite expresar los fenómenos como el producto de acciones de hombres en determinados contextos culturales en constante transformación. Ello explica el fracaso del naturalismo histórico, en su fallido intento de explicar el comportamiento de los hombres a partir de la uniformización de normas socioculturales.

Pero la defensa de la autonomía de la historia no implica la inconmensurabilidad entre la historia y las ciencias naturales, más bien todo lo contrario. Al pensar la relación entre ambos *modos de experiencia* en términos dualistas y esquemáticos, resulta de gran utilidad presentarla desde el punto de vista de Davidson (2011 [1984]), aplicando el principio de caridad, según el cual establecemos que cuando una persona sostiene creencias diferentes a las nuestras, debemos tomar que esas creencias son verdaderas (16-17). Para el caso del estudio de los fenómenos del pasado, la doctrina de la re-actualización, que no es otra forma de expresar que una explicación propiamente histórica debe comprender las acciones en tanto producto de creencias legítimas -racionales-, expresa la caridad davidsoniana<sup>8</sup>.

Luego, es gracias al establecimiento del común acuerdo que podemos construir las bases del desacuerdo significativo. Esta es la propuesta que articula Collingwood ya desde *Speculum Mentis*, donde traza como los diferentes *modos de experiencia*, “la religión, el arte, la ciencia, la historia y la filosofía”, se encuentran interconectados, aunque en ciertas actividades una forma “predomine” sobre las otras: Nos referimos, por ejemplo, al uso que hace la historia de las ciencias auxiliares como la geología o la biología. La preocupación principal en *Speculum Mentis* es justamente tratar de demostrar cómo la inconmensurabilidad entre las diferentes formas de experiencia se ha configurado en lo que el autor denomina como el principal escollo de la vida moderna, “(o la falta de) interpenetración entre las varias actividades de la mente” (Collingwood, 2014 [1924]: 27): El *realismo*, ferozmente atacado en *An Autobiography*, no es otra cosa que la expresión de la imposición de un *modo de experiencia* desde el franco rechazo hacia los otros y la invisibilización de los conceptos o presuposiciones gracias a las cuales esto es posible. Pero, para Collingwood, es gracias a que toda experiencia es procesual, transitoria e histórica que siempre debe haber y habrá diferentes *modos de experiencia* inextricablemente interconectados dado que, aunque existen desacuerdos significativos en relación con presuposiciones absolutas divergentes, siempre formarán parte de la misma constelación histórica.

---

<sup>8</sup>La doctrina de re-actualización sea quizás uno de los aspectos más controversiales de la filosofía de la historia de Collingwood, cuya lectura simplista ha sido utilizada para impugnar el proyecto collingwoodiano por ser psicologista. Para ver las principales posiciones sobre la tesis de reactualización, consultar Van derDussen (2016: 15-20) y Retz (2017).

## A modo de conclusión: la experiencia como punto de partida

Hemos visto a lo largo de este escrito como el propio desarrollo intelectual de Collingwood, orientado hacia la defensa de la autonomía de la historia, la exigió la construcción de un arsenal teórico original sustentado en una concepción operante (aunque implícita) de la experiencia según la cual toda experiencia es proceso y, por ende, histórica.

El primer paso hacia una construcción alternativa de experiencia implicó un rechazo de la concepción que hasta ese entonces dominaba en el ámbito académico británico de inicios del S XX. Nos referimos específicamente a la consideración de la experiencia como equivalente a la percepción. La traducción de este compromiso ontológico en los términos doctrinarios del *realismo* significaba que la convicción de que conocer no implica diferencia alguna sobre lo que es conocido. A nivel de las prácticas históricas, ello se traducía en el intento por la verificación de hechos como unidades atómicas y la ambición de reducir dichos hechos a la manifestación de regularidades en consonancia con los cánones de las ciencias naturales.

La propia práctica del autor en su investigación histórico-arqueológica junto con los desarrollos de otros pensadores (como Mommsen, Bradley o Bury) le sugirieron a Collingwood que en la historia auténtica el conocimiento no es mera verificación, más bien es construido y condensado en complejos de preguntas y respuestas, cuya corrección e incorrección bien habilita o cercena la continuidad de la actividad intelectual. De hecho, Collingwood, antes que rechazar plenamente los avances de los historiadores previos, les integraba dentro de un proceso del cual el formaba parte, ofreciendo él mismo las respuestas a los interrogantes previamente abiertos. El caso de Bury, quien le marca el camino, sea quizás el más emblemático. De esta forma, su tesis sobre las presuposiciones absolutas avanzada en *An Essay on Philosophical Method* y sistematizada en *An Essay on Metaphysics* surge de la necesidad de dar cuenta los compromisos subyacentes a cualquier forma de conocimiento, para así advertir contra la importación de ciertas presuposiciones absolutas presentes en las ciencias naturales a el estudio de los fenómenos del hombre.

El ejercicio académico auténticamente histórico (en palabras de Collingwood, científico), reposa en una serie de presuposiciones absolutas que configuran constelaciones no ontológicas sino transitorias. Ello le permite a Collingwood no rechazar los avances hechos por el realismo, sobre todo en relación al tratamiento crítico y sistemático de la evidencia, sino a valorar su contribución dentro de los compromisos normativos adquiridos y exigidos en la investigación histórica. Para nuestro autor es tan cierto que la ciencia histórica no debe tener como horizonte la búsqueda de leyes o regularidades al estilo de las ciencias naturales como que no se puede ignorar la necesidad de verificar los hechos históricos que integran los complejos narrativos de preguntas y respuestas.

Pero si bien Collingwood explicita el compromiso ontológico central de sus adversarios *realistas* a partir de su crítica certera a una experiencia concebida en términos intuitivos, el núcleo central de la constelación de presuposiciones absolutas habilitantes de la práctica genuinamente histórica, su propia concepción de experiencia, opera de forma implícita en su filosofía de la historia. Las críticas de Quine y Davidson a los dogmas del empirismo cumplen aquí un rol de extrema importancia a la hora de elucidar la noción de experiencia operante en la filosofía collingwoodiana al relevar sus rasgos centrales: a- La experiencia en Collingwood evita en la dicotomía entre *realismo* e *idealismo* sin caer en el relativismo, el temor a la inconmensurabilidad entre diferentes constelaciones de presuposiciones absolutas, b- Como bien lo ha indicado Quine, no es necesario individualizar (o explicitar) cada uno de nuestros compromisos para que ellos funcionen. La lógica de investigación sugiere más bien que estamos dispuestos a sacrificar un elemento que entra en conflicto con nuestra experiencia para defender un bloque (o, en términos collingwoodianos, una constelación) de presuposiciones, c- Una vez abandonada definitivamente la concepción intuitiva de la experiencia, Collingwood opera bajo la presuposición absoluta que la experiencia es algo procesual, que nuestra estrategia a la hora de interpelar el mundo que nos rodea consiste en el despliegue de complejos de preguntas y respuestas históricamente situados. Es

por ello que la correcta manera de explicitar la concepción collingwoodiana de la experiencia es que toda experiencia es experiencia histórica.

Se abre de esta forma una relación de cooperación antes que enfrentamiento entre la historia y las ciencias naturales. Al resaltar el carácter histórico y construido de cada una de las ciencias, Collingwood dentro de un grupo de constelación de presuposiciones emergen diferentes *modos de experiencia* permeables entre sí. Desde este punto de vista, la autonomía de la historia, sustentada en una concepción específica y original de experiencia, no es una reacción especular al imperialismo de las ciencias sociales, es una reafirmación de la identidad de una forma de interpelar al mundo en vistas al diálogo con otras.

Luego, la consideración de la experiencia como experiencia histórica habilita además a una lectura mucho más enriquecedora de un concepto un tanto controvertido en Collingwood, el de re-actualización. Bajo estos términos, la re-actualización puede ser vista no como la continuidad de un proyecto idealista, sino aquello sin lo cual la comprensión de las acciones del sujeto pasado no es posible. La afinidad entre la re-actualización y el principio de caridad de Davidson constituye tan sólo el inicio de una aproximación mucho más promisorio y enriquecedora de dicha doctrina.

## Referencias

Belvedresi, R. (2001). "Historia y sentido: los aportes de R.G. Collingwood al reconocimiento del valor práctico de la historia". En E. Adamovsky (Ed.). *Historia y sentido: exploraciones en teoría historiográfica*. El cielo por asalto. Buenos Aires. 135-156.

Collingwood, R. [1924] (1956). *Speculum Mentis or The Map of Knowledge*. Clarendon. Oxford.

Collingwood, R. [1933] (2005). *An Essay on Philosophical Method*. Introducción de Connelly, J y D'Oro, G. Oxford University Press. Oxford.

Collingwood, R. [1939] (1982). *An Autobiography*. Introducción de Stephen Toulmin. Clarendon. Oxford.

Collingwood, R. [1939] (2014). *An Essay on Metaphysics*. Clarendon. Oxford.

Collingwood, R. [1945] (2014). *The Idea of Nature*. Martino Publishing. Mansfield.

Collingwood, R. [1946] (2011). *Idea de la Historia*. Traducción de Edmundo O' Gorman. FCE. México.

Collingwood, R. (1999). *Principles of History*. Introducción de William Dray y Jan van der Dussen. Oxford University Press. Oxford.

Davidson, D. [1984] (2011). *Sobre la idea misma de un esquema conceptual*. Traducción de Manuel Pinedo y Jesús Palomo. Recuperado de [http://www.ugr.es/~pinedo/papers/sobre\\_la\\_idea\\_misma.htm](http://www.ugr.es/~pinedo/papers/sobre_la_idea_misma.htm).

Davidson, D. (2003). *Subjetivo, Intersubjetivo, Objetivo*. Traducción de Olga Fernández Prat. Ediciones Cátedra. Madrid.

Davidson, D. (2005). *Truth, Language and History*. Clarendon. Oxford.

Dewey, J. (1916). *Essays on experimental logic*. Dover. New York.

Dewey, J. (1938). *Logic: The Theory of Inquiry*. Southern Illinois University Press. Carbondale.

D' Oro, G. (2006). Robin George Collingwood. *Stanford Encyclopedia of Philosophy*. Recuperado de <https://plato.stanford.edu/entries/collingwood/>

- D' Oro, G. (2018). "The touch of King Midas: Collingwood in why actions are not events". *Philosophical Explorations*. 21(1). 160-169. doi: 10.1080/13869795.2017.1421697.
- Mead, G. H [1932] (2008). "La filosofía del presente. Traducción de Ignacio Sánchez de la Ycera". Centro de Investigaciones Sociológicas. Madrid.
- Mink, L. (1969). "Mind, History and Dialectic: The Philosophy of R.G" *Collingwood*. Indiana University Press. Bloomington.
- Quine, W. [1980] (2002). *Desde un punto de vista lógico*. Traducción de Manuel Sacristán. Paidós. Barcelona.
- Retz, T. (2017). Why Re-enactment is not Empathy, Once and for All. *Journal of the Philosophy of History*. 11 (3). 306-323. doi: 10.1163/18722636-12341376.
- Van der Dussen, J. (2017). *Studies on Collingwood, History and Civilization*. doi: 10.1007/978-3-319-20672-1.

## REVISITANDO CRÍTICAMENTE NOCIONES DE SIMILITUD EN LA TEORÍA DE JUEGOS COMPORTAMENTALES<sup>1</sup>

*REVISITING CRITICALLY SIMILITUDE NOTIONS IN BEHAVIORAL GAME THEORY*

*REVENDO CRITICAMENTE NOÇÕES DE SIMILARIDADE NA TEORIA DOS JOGOS COMPORTAMENTAIS*

**Carlos Maximiliano Senci**

*(Instituto de Investigaciones Económicas y Sociales del Sur  
Consejo Nacional de Investigaciones Científicas y Técnicas  
Universidad Nacional del Sur  
Universidad Salesiana - Bahía Blanca)  
maxisency@msn.com*

Recibido: 22/10/2020

Aprobado: 12/07/2021

### RESUMEN

El objetivo de este trabajo consiste en revisar críticamente diferentes concepciones de similitud en la teoría de juegos que se encuentran presupuestas en diversos conceptos de solución. Pasamos revista a cuatro concepciones: 1) similitud empírica, 2) la doctrina Harsanyi-Aumann, 3) Superracionalidad, y 4) pensamiento mágico. Evaluamos críticamente las diferentes nociones en función de la similitud que presuponen, la viabilidad para resolver problemas de coordinación y/o cooperación basados en ellas, y su fundamentación racional.

Palabras clave: teoría de juegos. similitud. superracionalidad. razonamiento en equipo. pensamiento mágico.

### ABSTRACT

This work aims to critically revise different conceptions of similarity in game theory that are presupposed in various solution concepts. We review four conceptions: 1) empirical similarity, 2) the Harsanyi-Aumann doctrine, 3) Superrationality, and 4) magical thinking. We critically evaluate the different notions based on the similarity they presuppose, the viability of solving coordination and cooperation problems based on those notions, and their rational foundation.

Keywords: game theory. similarity. superrationality. team reasoning. magical thinking.

### RESUMO

O objetivo deste trabalho é revisar criticamente diferentes concepções de similaridade na teoria dos jogos que são pressupostas em vários conceitos de solução. Revisamos quatro concepções: 1) similaridade empírica, 2) a doutrina Harsanyi-Aumann, 3) Superracionalidade e 4) pensamento mágico. Avaliamos criticamente as diferentes noções

---

<sup>1</sup> Agradezco los comentarios y sugerencias de Fernando Tohmé a una versión previa de este trabajo.



com base na semelhança que pressupõem, na viabilidade de resolução dos problemas de coordenação e cooperação a partir delas e no seu fundamento racional.

Palavras-chave: teoria dos jogos. similaridade. superracionalidade. raciocínio de equipe. pensamento mágico.

## Introducción

Un ámbito de aplicación crecientemente extendido de la teoría de juegos (TJ a partir de ahora) consiste en el estudio de las propiedades de la interacción social de agentes intencionales. Pero hay que tener en cuenta que la TJ no es una teoría social, sino matemática. Como dice Ross (2008: 60):

La teoría de juegos *per se* no es una rama de las ciencias sociales, sino una rama de la matemática: es una teoría de la maximización simultánea de ciertas clases de funciones en objetos abstractos llamados “campos de preferencia” cuando las funciones en cuestión deben tomar recíprocamente los outputs de los demás entre sus inputs. El “agente racional” aquí es un tipo técnico, no un tipo empíricamente determinado.

Como herramienta formal la TJ es útil para modelar la interacción de dos o más agentes, y es muy útil por la capacidad de ofrecer predicciones precisas acerca del comportamiento de dichos agentes, en el sentido de que las decisiones simultáneas de los jugadores unívocamente determina un resultado o un conjunto de resultados.<sup>2</sup> Las predicciones pueden ser precisas justamente porque se basan únicamente en la estructura de pagos de los juegos, especificada por las funciones de utilidad de los agentes. Como consecuencia de que los agentes que modela la TJ son altamente idealizados, las predicciones pueden no ser descriptivamente adecuadas. En este sentido la TJ es normativa, ya que no está interesada en los procesos de toma de decisiones tal como en realidad se llevan a cabo por seres humanos de carne y hueso. Las predicciones son capturadas por distintos conceptos de solución, esto es, reglas que prescriben lo que deben hacer los agentes para que los resultados que pretenden conseguir cumplan ciertos criterios, entre ellos criterios de racionalidad, de eficiencia, entre otros. Los conceptos de solución asignan una combinación de perfiles de estrategia, i. e., las soluciones, a cada juego estratégico finito.

El concepto de solución más extensamente usado en la TJ no-cooperativa es el concepto de equilibrio de Nash (EqN a partir de ahora; Nash, 1950; Myerson 1991: 105). Informalmente, un EqN es una tupla de estrategias, tales que ninguno de los jugadores tiene incentivos en modificar su estrategia, dadas las estrategias de los otros jugadores. Dicho de otra manera, cada jugador tiene conjeturas correctas acerca de las acciones de los otros, y actúa racionalmente, esto es, adopta su mejor estrategia (dadas esas conjeturas). Desde los trabajos pioneros de Aumann sabemos que las condiciones epistémicas que requiere el EqN son sumamente demandantes. En particular, en el caso de dos jugadores la condición epistémica de equilibrio demanda que haya conocimiento mutuo (cada uno sabe lo que sabe el otro) de la racionalidad de los jugadores (i. e., los jugadores son similares en cuanto a su racionalidad), de los pagos, y de las conjeturas (Aumann y Brandenburger, 1995). En el caso de equilibrios en estrategias puras esto significa que hay conocimiento mutuo acerca de qué acciones van a elegir los otros.

En muchas situaciones ubicuas el EqN no provee de soluciones satisfactorias en relación con al menos algún criterio. El caso típico es el Dilema del Prisionero, en el que no pueden satisfacerse al mismo tiempo un criterio de racionalidad individual (entendida como maximización individual de los pagos) y un criterio de eficiencia colectivo. El EqN tampoco parece brindar soluciones aceptables en otra clase de interacciones, e. g., juegos de coordinación como la Caza del Ciervo (*Stag Hunt*) o el Hi-Lo. En estos juegos hay equilibrios que son de recompensa dominante<sup>3</sup> (*payoff-dominant*) y sin embargo el EqN no brinda las herramientas para discriminar entre estos equilibrios y aquellos que no lo son. Estos juegos son instancias de una amplia clase de juegos en los que existen múltiples EqN. En estos casos el EqN no

<sup>2</sup> Para ser más precisos, estoy excluyendo juegos en los que puede introducirse un factor exógeno de incertidumbre en los que los resultados pueden depender de procesos aleatorios (Myerson 1991, capítulo 2).

<sup>3</sup> Un resultado es de recompensa dominante si todos los jugadores obtienen una recompensa mayor que la que obtendrían en cualquier otro resultado.

provee de predicciones precisas, a lo sumo puede ofrecer clases de soluciones. Esta insatisfacción respecto de las soluciones ofrecidas ha generado un programa de refinamientos al EqN (Harsanyi y Selten, 1988). La TJ comportamental también puede verse como una forma de refinamiento del abordaje estándar en la TJ, aunque no necesariamente de sus conceptos de solución (Ross, 2019), cuyo objetivo es proponer conceptos de solución no-ortodoxos, i. e., que modifican o adicionan supuestos y restricciones impuestos por el EqN. En este artículo me enfoco en un conjunto de nociones alternativas que van más allá de la visión clásica apartándose del EqN.

Una intuición empírica que se puede adicionar a la estructura formal de los juegos consiste en la similitud entre los jugadores. A veces la similitud puede referirse a aspectos normativos o formales, tales como la racionalidad. En otras ocasiones los aspectos relevantes pueden ser meramente empíricos y no estar referidos a aspectos normativos. Como señala Morton (1994: 22):

...los agentes humanos se reconocen claramente entre sí como similares: en la familia, en el entorno social o simplemente como especie. Y este reconocimiento juega un papel en sus interpretaciones de la acción de los demás y sus propensiones a la cooperación. (p. 22; mi traducción; todas las traducciones son propias a menos que se indique lo contrario).

También ciertas nociones de similitud han permeado en la literatura comportamental, en general como reelaboraciones de nociones previas que pueden encontrarse en la psicología social. La intuición básica consiste en que tiene sentido para un jugador reducir la incertidumbre que experimenta cuando enfrenta a otros jugadores considerando la similitud con los demás. Este fenómeno conductual ha sido documentado en una serie de trabajos recientes, que estudiaron el efecto de formas diferentes de *priming* de similitud en juegos de coordinación y/o cooperación (ej., Fischer, 2009; Di Guida y Devetag, 2013; Mussweiler y Ockenfels, 2013; Chierchia y Coricelli, 2015; Rubinstein y Salant, 2016).

Hay diferentes formas por medio de las cuales los teóricos de juegos han intentado capturar diversos sentidos a partir de los cuales pueden pensarse relaciones de similitud entre jugadores. En lo que sigue voy a esbozar una taxonomía, que no pretende ser excluyente, acerca de diferentes formas de pensar la similitud entre jugadores. En primer lugar, lo que llamaré la *concepción inductiva o empírica de la similitud*, en la que discutiré dos abordajes principales: la teoría de puntos focales y el pensamiento en equipo (*team reasoning*).

En segundo lugar, abordaré lo que se ha llamado la *doctrina Harsanyi*. En tercer y cuarto lugar me referiré a soluciones que no han tenido tanta difusión en la literatura económica, pero que especialmente la segunda de ellas ha tenido relevancia en el ámbito filosófico: me refiero a la noción de *superracionalidad* y a la de *pensamiento mágico*.

## **Similitud empírica en el razonamiento entre los seres humanos**

En su reciente libro *Understanding Institutions*, Francesco Guala (2016) profundiza en lo que él llama el problema de la lectura mental (*mind reading*) en los juegos de coordinación. La coordinación depende de que las personas tengan una “teoría de la mente”, es decir, la capacidad de atribuir estados mentales a los demás. Aunque su forma de enmarcar el problema es de naturaleza lewisiana, Guala rechaza la conocida solución de Lewis, y aplica otros dos conceptos de solución al problema de la lectura mental: el enfoque del pensamiento de solución (*solution thinking*) de Morton (Morton 1994), así como el enfoque de razonamiento en equipo (RE a partir de ahora) desarrollado por Bacharach, Sugden y Gold (Bacharach 1999, 2006; Gold, 2012; Gold y Sugden 2007; Sugden, 2000, 2003). A pesar de sus similitudes, el RE y el pensamiento de solución poseen un enfoque distinto del de la TJ ortodoxa. Este enfoque se basa en la TJ psicológica (Geanakoplos, Pearce y Stacchetti, 1989), que modifica “la teoría ortodoxa mediante la introducción de principios formales de razonamiento que pueden ayudar a explicar las observaciones empíricas y las intuiciones ampliamente compartidas que la teoría ortodoxa deja sin explicar” (ver Colman, 2003: 149-150), mientras que la solución de Lewis y el enfoque del punto focal se basan en los equilibrios de Nash derivados de creencias mutuamente consistentes. Lo que tienen estos

enfoques en común es que proponen una forma de razonamiento que modifica la forma de razonamiento estándar e introducen principios formales cuyo objetivo es explicar regularidades empíricas que en muchos casos se presentan como anomalías para la TJ estándar. Es importante tener en cuenta que en última instancia la justificación acerca de las soluciones que pueden brindar a diferentes problemas como el de coordinación, o cooperación, es una justificación empírico-inductiva que apela a la experiencia individual y a la psicología.

En las teorías de coordinación de Schelling, Lewis, Bacharach y Sugden se hacen varias suposiciones sobre las características del razonamiento que son comunes a (y conocimiento común dentro de) poblaciones relevantes de potenciales jugadores. Estas son *regularidades empíricas*, basadas en características comunes de la psicología humana y la experiencia compartida de los jugadores en la vida humana. Por lo tanto, puede haber situaciones en las que sea razonable que un jugador asuma que el razonamiento de los demás será similar al suyo.

Schelling abordó el problema de coordinación tratándolo como un problema de predicción. La posibilidad de la coordinación requiere que primero sea posible coordinar las predicciones, esto es, “leer el mismo mensaje en una situación común, identificar el (único) curso de acción en el que las expectativas mutuas puedan converger” (Schelling, 1981: 64). Cualquier inferencia que funcione deberá contener varias creencias de orden superior. Para evitar una regresión al infinito, el proceso de razonamiento debe terminar en alguna etapa intermedia. En muchas ocasiones existen señales que coordinan las expectativas de todas las partes. De acuerdo con Schelling, los elementos que pueden funcionar como señales de ese tipo pueden ser muy diversos, siempre que sean lo suficientemente salientes (*salient*) para los jugadores, es decir, que tengan el tipo adecuado de prominencia o notoriedad. Nótese que las señales son irrelevantes desde el punto de vista de las recompensas y pueden “depender de analogía, precedente, disposición accidental, simetría, configuración estética o geométrica, razonamiento casuístico, o de quiénes son las partes y qué saben unos de otros” (Schelling, 1981, 57). Basado en esta idea, Schelling elaboró su solución de *punto focal* para el problema de coordinación.

Para cumplir su tarea los puntos focales deben satisfacer una propiedad de unicidad. De nada nos sirven puntos focales que no nos permitan discriminar entre diferentes posibles soluciones. Una de las formas de satisfacer este requisito consiste en que los puntos focales se destaquen *para todos los individuos por igual*, es decir, tienen que ser salientes para todos los jugadores que participan de la interacción. Para ello es evidente, a su vez, que los individuos tienen que ser lo *suficientemente similares* como para extraer la misma información de las señales: “los sujetos deberán ser similares en la importancia relativa que otorgan a diferentes aspectos de comparación; pero a menudo son suficientemente similares para resolver el problema” (Lewis, 1969: 35). Sea cual sea la forma de explicar de qué modo los puntos focales permiten la coordinación, lo cierto es que a la base de estas nociones hallamos el supuesto de similitud de los jugadores.

La teoría del RE de Bacharach y Sugden ha sido propuesta como una solución alternativa a ciertos problemas de coordinación. Según Bacharach, el RE es un modo de razonamiento práctico (volveremos más adelante sobre esto) motivado por una clase de juegos que tienen margen para una recompensa común (*scope for common gain*), entendida como la posibilidad de una mejora de Pareto. En esta clase de juegos encontramos puzzles de cooperación, como la paradoja de Hi-Lo (Schelling, 1981, p. 291) y el ya mencionado Dilema del Prisionero, para los cuales la teoría de la decisión estándar ofrece soluciones que son cuestionables en términos de su racionalidad.

En palabras de Bacharach: “alguien ‘razona en equipo’ si encuentra la mejor combinación de acciones posible para todos los miembros de su equipo, y entonces hace su parte en ello” (Bacharach 2006: 121). El RE equivale a pensar en uno mismo como miembro de un grupo, y derivar las implicancias prácticas de ello. Al permitir que se consideren como agentes unidades colectivas, el RE tuerce la sintaxis, por así decirlo, de los juegos estratégicos y permite a los jugadores pensar en sí mismos como miembros de un equipo y adoptar las preferencias del equipo como propias. Por lo tanto, la clave para entender el RE es reemplazar al agente individual por un equipo de razonadores. En este sentido, el RE es una forma de

enmarcar (*frame*) la interacción con otros agentes. Bacharach llama a esto “transformación agencial” (Bacharach, 2006). La idea básica es que al identificarnos con un grupo enmarcamos nuestras propias preferencias de una manera grupal. Como consecuencia de esta transformación, un individuo que se identifica con un grupo “se considera intercambiable con otros miembros del mismo grupo. Percibe esta equivalencia aunque sabe que es ‘numéricamente distinto’” (Bacharach 2006: 73). Una traducción directa de esta propiedad es la condición de simetría. Los razonadores de equipo deberían satisfacer una propiedad de invariancia respecto de la simetría del juego: esto es, al rotar la matriz sobre la diagonal los jugadores deberían obtener los mismos resultados en un juego en forma normal.

Gold y Sugden (2007) han propuesto analizar las intenciones del grupo como resultado de un modo específico de razonamiento práctico, instanciado por el RE. El elemento básico es el concepto de *esquema de razonamiento práctico*, inspirado en el silogismo práctico de Aristóteles. Se define la validez de un esquema como una propiedad que asegura la consecución del éxito, es decir, maximiza los pagos (entendidos como índices de utilidad).

Veamos cómo funciona el esquema aplicado al juego Hi-Lo (Tabla 1):

### Ejemplo 1: Hi-Lo

Tabla 1. Matriz de pagos del juego Hi-Lo

		Jugador 2	
		H	L
Jugador 1	H	2, 2	0, 0
	L	0, 0	1, 1

El juego Hi-Lo plantea un problema de coordinación ya que los jugadores prefieren coincidir en sus estrategias a no coincidir. Sin embargo, la teoría ortodoxa es estéril para proporcionar un resultado determinado, ya que en realidad hay dos Equilibrios de Nash, en los que ambos jugadores eligen la misma estrategia. Sin embargo, uno de los equilibrios (H, H) es de recompensa dominante, y por lo tanto parece trivial que los jugadores deben coordinarse en ese equilibrio. Esta no es, sin embargo, la manera en que la teoría ortodoxa resuelve este problema. De hecho, dado que el jugador 1 elige L, la mejor respuesta del otro jugador consistiría en elegir la misma estrategia. Y por lo tanto, (L, L) se convierte en un equilibrio de Nash. Parece insatisfactorio desde un punto de vista normativo y tampoco está a la altura de proporcionar un análisis descriptivo adecuado capaz de explicar cómo los individuos se coordinan exitosamente en juegos como este, como lo demuestran los resultados experimentales (ver Colman *et al.* 2008; Bardsley *et al.* 2010).

Consideremos el esquema práctico de racionalidad colectiva de Gold y Sugden:

1. Debemos elegir entre (H, H), (H, L), (L, H), o (L, L)
2. Si elegimos (H, H) el resultado será R1
3. Si elegimos (H, L) el resultado será R2
4. Si elegimos (L, H) el resultado será R3
5. Si elegimos (L, L) el resultado será R4
6. Queremos conseguir R1 en lugar de R2, R3 o R4.
7. Debemos elegir (H, H).

El esquema ofrece una conclusión práctica: teniendo en cuenta las premisas, que incluyen proposiciones sobre los objetivos del grupo, debemos inferir como conclusión lo que debemos hacer para garantizar nuestro objetivo. En la TJ ortodoxa los individuos maximizan una función de utilidad individual. Por el

contrario, en el RE la unidad de agencia es colectiva, es un “nosotros”, cuyo objetivo es maximizar una función de utilidad grupal. Un individuo no tiene razones para favorecer una u otra de las posibles estrategias en el juego si razona su camino hacia la conclusión aplicando preferencias estándar. Pero el objetivo práctico del agente colectivo debería consistir en maximizar el bienestar total del grupo.

La teoría del RE, tal como la proponen Bacharach y Sugden, no brinda una respuesta acerca de cuándo la racionalidad requiere de un jugador que adopte el RE (esta crítica ha sido formulada entre otros por Browne, 2018). Eso es correcto, pero la idea de que este es un problema presupone algo así como la doctrina Harsanyi (ver próxima sección), es decir, que cada juego tiene una solución racional única, que es trabajo del teórico encontrar o anunciar. Bacharach y Sugden argumentan explícitamente que esta doctrina está fundamentalmente injustificada. En relación con el RE, lo máximo que puede hacer una teoría de la racionalidad (instrumental) es decirles a los agentes qué hacer, una vez establecidos sus objetivos. No puede decir qué agentes (individuos o equipos) deberían existir racionalmente. Ser racional, desde el punto de vista de la TJ, no brinda un acceso epistémico a conocer o saber en qué situación se debería cambiar la forma de razonamiento y adoptar una preferencia de grupo. Bacharach y Sugden tratan la cuestión acerca del modo de razonamiento de los jugadores como una cuestión empírica, y proponen varias formas de responder esa pregunta en términos de la psicología y la experiencia de los jugadores (sobre esto ver Bacharach, 2006). En este punto las teorías de ambos difieren en un punto sustancial. Para Bacharach es racional decidir en favor del equipo aun cuando no lo sea individualmente. Para Sugden, en cambio, es racional jugar en equipo siempre y cuando se cumpla una condición de garantía (*assurance*) que brinde seguridad respecto de que los demás jugadores también «razonan en equipo». Es elocuente el ejemplo del dilema del prisionero. Mientras que la cooperación mutua es la mejor solución en términos del beneficio colectivo, desde el punto de vista estrictamente individual sigue siendo más conveniente no cooperar, independientemente de la elección del otro jugador. La elección de no-cooperar protege al jugador también en el caso de que el otro jugador elija no-cooperar. No hay ninguna razón instrumental válida que pueda persuadir a un jugador racional de que la estrategia que debe seguir es la cooperación. Por eso el RE se presenta a sí mismo como una teoría alternativa acerca de la racionalidad.

Bacharach introduce un parámetro  $w$  ( $0 \leq w \leq 1$ ), que representa la probabilidad de que un individuo se identifique con un grupo, o en otras palabras, razone en equipo. En otras versiones de la teoría de Bacharach (Gold, 2012),  $w$  representa la probabilidad de que el RE tenga poder motivacional para un individuo. Lo distintivo de la teoría de Bacharach es que admite situaciones en las que  $w < 1$ , esto es, en las que un individuo no tiene garantías de que los demás razonen en equipo. Eso lleva a que la teoría de Bacharach admita que los perfiles de estrategia fuera de la diagonal (C, D) y (D, C) sean ponderados mejor que el equilibrio de Nash (D, D) (ver ejemplo 2). Por lo tanto, si consideramos que la recompensa de equipo es la suma de las recompensas individuales, un jugador de equipo podría elegir cooperar aun cuando supiera con certeza que los otros jugadores no cooperarían ( $w = 0$ ), ya que  $(C, D) = (D, C) = 7 > 2 = (D, D)$ .

### Ejemplo 2: Dilema del Prisionero

Tabla 2. Matriz en forma normal del Dilema del Prisionero.

		Jugador 2	
		C	D
Jugador 1	C	4, 4	0, 7
	D	7, 0	1, 1

Para Sugden, esta solución no es aceptable. En efecto, en trabajos recientes Sugden (2011, 2015) se aparta de la versión de Bacharach del RE. En lugar de comparar el resultado del RE con el resultado no-



cooperativo, para Sugden el punto de referencia es el mejor pago que podría obtener el individuo unilateralmente. Por eso Sugden no puede aceptar que los perfiles fuera de la diagonal principal puedan ser válidos para el RE, a menos que los jugadores cuenten con la suficiente garantía de que los demás también jugarán la opción cooperativa. De esta manera Sugden no se aparta drásticamente del concepto de racionalidad estándar.

Lo anterior pareciera implicar que la teoría de RE de Bacharach no presupone una noción de similitud, en la medida en que admite que la identificación con el grupo puede ser unilateral, incluso *contrario sensu* de lo que prescribiría la racionalidad instrumental. Muy por el contrario, las últimas versiones de RE, que debemos a Sugden, parecen reintroducir la idea subyacente de que los jugadores tienen que ser, en un sentido sustancial, similares: tienen que tener la garantía (o la creencia) de que los demás jugarán la misma estrategia con un alto grado de probabilidad (o tener la garantía de que jugarán la misma estrategia). ¿Cuán alta debe ser esa probabilidad para que un jugador coopere en un dilema del prisionero a sabiendas de que cooperar no es su mejor elección *independientemente* de lo que haga el otro jugador?

Este es un tema que excede, desde luego, los límites de este trabajo, aunque de manera provisoria podría decirse que no parece haber una respuesta genuina a este interrogante (para intentos de solución, e interrogantes ver Gold y Colman, 2018; Gold, 2018).

3. La «Doctrina Harsanyi» debe su nombre a Aumann (1976), quien la bautizó en referencia a los supuestos que incluye el desarrollo teórico de Harsanyi acerca de los tipos. También suele aparecer en la literatura como la doctrina Harsanyi-Aumann. Según esta doctrina existen principios normativos universales de racionalidad que pueden identificar una solución única (posiblemente probabilística) para cada problema de decisión descrito en forma completa. Por lo tanto, en un mundo en el que hay conocimiento común de racionalidad, los jugadores racionales obtendrán las mismas soluciones (es decir, serán similares de un único modo racional).

Como es bien conocido, Harsanyi (1968) introdujo su teoría de tipos como una forma de reducir los juegos de información incompleta a juegos de información completa. La idea es que en la noción de tipo está incluida la información de un agente (sus creencias) acerca de su propio tipo y el tipo de los demás. Esto es, la incertidumbre respecto de la función de utilidad de los demás queda condensada en la noción de tipo. Para Harsanyi las únicas diferencias posibles entre jugadores se derivan de diferencias de información, y las diferencias respecto de las creencias individuales solamente pueden ser atribuidas a diferencias de información. Esto se traduce en dos condiciones que deben cumplir los jugadores: primero, los tipos son conocimiento común, y segundo, los jugadores deben tener una distribución de probabilidad *a priori* común acerca del conjunto de los estados del mundo. Esta es una restricción normativa al comportamiento de los jugadores en un juego que se puede traducir fácilmente en la simetría del juego. Dicho de otra manera, el supuesto de distribución *a priori* común de las creencias implica que los tipos de los jugadores no difieren respecto de las expectativas que tienen acerca de los eventos del mundo si poseen la misma información. Desde luego, podría ser el caso de que los jugadores estuvieran en posesión de diferentes conjuntos de información, y entonces las diferencias en cuanto a la forma de jugar el juego podrían atribuirse a diferencias de información. Esto es, los jugadores sólo diferirán en función de nueva información que adquieran.

La extensión de Aumann (1976) de la doctrina Harsanyi elimina las diferencias de información al argumentar que si los individuos tienen la misma distribución común *a priori* respecto del evento, i.e. respecto del juego, entonces sus distribuciones de probabilidad *a posteriori* con respecto a ese evento tienen que estar de acuerdo porque los individuos racionales revisan sus creencias a través de la actualización bayesiana cuando se enfrentan a diferencias informativas. En el mundo real hay especuladores que compran y otros que venden activos financieros, unos creyendo que van a subir de precio y los otros, que van a perder valor. Obviamente alguno está equivocado. Pero la transacción se lleva a cabo porque no tienen conocimiento común de qué creencias tienen. En el momento en que descubren que las creencias que tienen son inconsistentes, si son racionales tienen que revisar sus creencias hasta que converjan. Como lo expresó Aumann, “los jugadores no pueden estar de acuerdo en

estar en desacuerdo” (*agreeing to disagree*). A este resultado le subyace una noción de similitud entre jugadores.

Estos supuestos en conjunto implican lo que Hargreaves-Heap y Varoufakis (2004) llaman *alineación consistente de creencias*, y que está a la base del equilibrio de Nash. Esta noción implica que ningún jugador racional desde el punto de vista instrumental puede esperar de otro jugador similar (i. e., en cuanto a su racionalidad) que juegue de manera diferente<sup>4</sup>. Si ambos poseen la misma información, y son igual de racionales, entonces razonarán de la misma manera, y por ende, terminarán concluyendo de manera independiente lo mismo. Y el resultado tendría que ser necesariamente un EqN. Para que el resultado no fuese un EqN, tendría que ser falso o que los jugadores sean racionales, o que no tuvieran conocimiento común de racionalidad. Como señalan Hargreaves-Heap y Varoufakis (1995: 60), no podemos esperar que los jugadores tengan siempre sus creencias alineadas, a menos que hubiera una única forma racional de jugar el juego para cada jugador. Pero una cosa es suponer que de existir una única solución racional a un juego, esta solución sería la que resultaría de la interacción de jugadores racionales, bajo el supuesto de conocimiento común de racionalidad. Y otra distinta es suponer que siempre existe una única forma racional de jugar un juego. Con buen tino, Hargreaves-Heap y Varoufakis le llaman a este supuesto Principio de determinación racional (p. 60). Sigo en la discusión de este principio a Hargreaves-Heap y Varoufakis. Este principio está sustentado por la doctrina Harsanyi-Aumann en base a dos argumentos: que si los jugadores racionales tienen la misma información entonces sacarán las mismas conclusiones; y que los jugadores tienen la misma información respecto de las reglas del juego. Estos dos elementos conforman lo que podríamos llamar la *similitud racional* de los jugadores.

Hargreaves y Varoufakis presentan dos argumentos contra esta versión de la *similitud racional*: por un lado, la doctrina Harsanyi-Aumann parece requerir de los jugadores que estos sean similares en cuanto a su capacidad de procesar información. Si lo pensamos como un programa de computadora, frente a los mismos inputs los jugadores tiene que obtener los mismos outputs. Pero esto presupone que cada situación tiene una resolución algorítmica, esto es, que para cada problema existe un procedimiento mecánico para resolverlo respecto del cual los jugadores coincidirán. Esta visión algorítmica de la razón implicaría que sería siempre posible identificar un conjunto exhaustivo de reglas, cuya aplicación cada jugador estaría en condiciones de determinar. Sin embargo, como señalan Hargreaves y Varoufakis en su crítica a este aspecto de la doctrina Harsanyi-Aumann:

ningún conjunto de reglas puede ser exhaustivo: Ningún conjunto de reglas puede contener reglas para su propia aplicación. Cualquier configuración nueva donde las reglas puedan aplicarse, siempre puede ser individualizada de tal manera que no esté cubierta por las reglas existentes. En estas circunstancias, o bien es necesario crear nuevas reglas para cubrir cada nueva aplicación (y este proceso amenaza con una regresión al infinito), o bien las personas deben interpretar creativamente de qué forma las reglas existentes se aplican en las nuevas configuraciones. Pero una vez que se reconoce una interpretación de este tipo, no hay razón para suponer que todos los individuos serán creativos, en este sentido, de la misma manera (p. 64).<sup>5</sup>

Por otro lado, la segunda forma de similitud es mucho más fuerte, ya que supone que además de tener las mismas capacidades predictivas, los jugadores deben tener la misma información. El teorema de Aumann consiste en lo siguiente: los desacuerdos respecto de cómo jugar el juego provocan revisiones (bayesianas) de las creencias o probabilidades asociadas a las alternativas del juego, y en el largo plazo, esto es, después de las revisiones necesarias los jugadores terminarán con idéntica información. Pero aplicado a juegos en forma normal el teorema de Aumann parece introducir un proceso de convergencia que no estaría incluido en la descripción del juego. ¿Cómo y en qué momento tendría lugar ese proceso de revisión de creencias?

<sup>4</sup> Por supuesto que Harsanyi asume que cada jugador conoce su propio tipo, es decir, un jugador no puede no saber que es racional.

<sup>5</sup> Aquí Hargreaves y Varoufakis parecen presuponer que el conjunto de principios a aplicar para resolver un juego cualquiera es «abierto», es decir que siempre van a existir casos no cubiertos. Pero la teoría de juegos ha tratado este tema extensamente (ver por ejemplo Lipman, 1991) y no necesariamente la respuesta es negativa. La autorreferencia *per se* no es un problema.

Según Hargreaves y Varoufakis esta forma de similitud introduce un elemento difícil de justificar en juegos *one-shot* que se refiere al proceso de convergencia en una distribución *a posteriori* común:

...cuando las creencias se refieren a cómo jugar el juego, y las creencias divergentes solo se revelan en el juego, es más que un poco difícil ver cómo se aplica el argumento a las creencias que los agentes sostienen antes de jugar el juego. Naturalmente, cuando se repite el juego, la idea tiene mucho más sentido (p. 66).

## Superracionalidad

Un concepto de solución que hace de la similitud entre jugadores su fundamento es la noción de *Superracionalidad*. Esta noción fue propuesta por Hofstadter (1983, 1985) para brindar un fundamento a la solución cooperativa en el Dilema del Prisionero (DP) (Flood y Dresher, 1952). La noción de superracionalidad de Hofstadter comparte un parecido de familia con una serie de nociones que toman la simetría de los jugadores como argumento central para justificar la cooperación en el DP (Campbell, 1989). Estas nociones incluyen un supuesto adicional a la descripción del DP según el cual los jugadores son *similares* en cierto sentido, similitud que se evidenciaría en una facultad epistémica de los jugadores para reconocer que se enfrentan a una misma situación y que poseen las mismas estrategias (Hofstadter, 1985).

La noción de Superracionalidad describe una propiedad de los perfiles en la diagonal de un juego simétrico. Hofstadter parece suponer que los jugadores superracionales tienen conocimiento mutuo de que son racionales. Eso parece deducirse de la idea de que los “pensadores superracionales ...incluyen en sus cálculos el hecho de que están en un grupo de pensadores superracionales” (Hofstadter, 1985: 748). La noción de superracionalidad se aplica naturalmente a juegos simétricos. De la aceptación del principio de determinación racional, se sigue para los pensadores superracionales que hay una única solución racional al juego. Dado que el juego es simétrico, la estrategia racional debe ser la misma para todos los jugadores, y debe encontrarse en la diagonal principal del juego. Y dado que estos no pueden diferir en cuanto a sus poderes de razonamiento (i. e., son similares) en un meta-nivel los jugadores pueden darse cuenta que de las estrategias disponibles deben elegir aquella que sea óptima.

La única solución racional estaría determinada conjuntamente por dos meta-principios: el Principio de Razón Insuficiente (PRI) (Sinn, 1980) y el principio de coordinación (Bacharach, 1999). Según el primero, ninguna estrategia debería tener mayor probabilidad si son *a priori* indistinguibles. Es evidente que si los jugadores son racionales, y si sus preferencias son indistinguibles en términos de su racionalidad (i. e., como son racionales deben razonar de la misma manera), entonces sus estrategias deberían ser indistinguibles. De acuerdo con este principio, en el ejemplo 2, si los jugadores son superracionales, la aplicación del PRI nos llevaría a considerar como posibles sólo los perfiles de estrategias sobre la diagonal principal del juego (ver tabla 2), ninguno de los cuales debería tener primacía sobre otro. Es decir, si sólo aplicáramos el PRI tanto (C, C) como (D, D) serían igual de probables. Para brindar una solución precisa, superracionalidad requiere de otro principio que pueda arbitrar una solución. La noción de superracionalidad adiciona a la doctrina Harsanyi-Aumann la idea de que la solución racional a cualquier juego debe ser una solución Pareto eficiente o de recompensa dominante (Harsanyi y Selten, 1988). Esta idea está capturada por un principio que vemos asumido en las distintas formulaciones de la teoría del RE, tal como en el principio de coordinación de Bacharach (Bacharach, 1999), o en el principio individual de racionalidad en equipo de Janssen (2001). En ambos casos lo que requiere el principio es que cada jugador haga su parte en una combinación de estrategias Pareto eficiente.

Otra vez, en el ejemplo 2, es evidente que, de los perfiles de estrategias (C, C) y (D, D), la solución identificada por el principio de coordinación es la de recompensa dominante (C, C).

Ahora bien ¿Cuál es la justificación para dar ese paso? Es evidente que puede haber condiciones particulares de la situación o de los jugadores que hagan uno u otro supuesto más plausible. Pueden existir situaciones en las cuales la no-cooperación, o la elección de la opción no-paretiana (pensemos en

cuestiones como la deserción durante una guerra), estuviera penada con la pena capital (y esta fuera prácticamente inexorable), quizás la opción alternativa ni siquiera sería pensable para los involucrados. Sin embargo, este supuesto parece injustificado, incluso en términos racionales.

## Pensamiento mágico

El pensamiento mágico es otra noción que ha aparecido bajo diferentes nombres, y que esencialmente se reduce a la idea de que existe alguna manera de saber *a priori* la elección de otro jugador. Pensemos en el contexto de un juego *one-shot* en forma normal (i.e. una interacción que tiene lugar una sola vez sin posibilidad de comunicación), en el que la única información que poseen los jugadores consiste en la estructura estratégica del juego (las estrategias posibles, las recompensas, etc.). Los jugadores no tienen ningún acceso epistémico privilegiado a los estados de conciencia de los demás, por lo tanto la posibilidad de «conocer» la disposición a elegir una u otra estrategia no es en absoluto trivial. El mecanismo preciso que presentan las distintas soluciones que agrupo bajo el concepto de pensamiento mágico pueden variar.

La idea de pensamiento mágico consiste en la creencia de que la propia acción de alguna manera tiene una influencia causal sobre la acción de otro jugador (Shafir y Tversky, 1992; Daley y Sadowsky, 2016). En consecuencia, las creencias respecto de las acciones de los demás varían de acuerdo a las propias acciones. Esto es, si voy a cooperar en el DP, es más probable que crea que el otro jugador también cooperará. Y además, si voy a cooperar en el DP le atribuyo una mayor probabilidad a que el otro jugador también coopere.

Una noción familiar es la de translucidez (*translucency*), que fuera propuesta por Gauthier como fundamento de su teoría del contrato social. La noción de translucidez captura la idea de que los jugadores pueden, a través de algún medio (cognitivo), establecer si otro jugador puede estar dispuesto a cooperar (Gauthier, 1986; Frank, Gilovich, y Regan, 1993; Spiekermann, 2007; Capraro y Halpern, 2015). En forma general, los modelos de pensamiento mágico capturan la creencia de que, debido a las similitudes existentes entre los seres humanos, un jugador puede elegir un modo de razonamiento a sabiendas de que *cualquiera que elija* será también elegido por el otro jugador (si suponemos un entorno de dos jugadores). En contraste, la doctrina de Harsanyi y la superracionalidad implican solo que si el jugador usa el modo de razonamiento que realmente se requiere normativamente, este será el mismo modo de razonamiento que usaría otro jugador completamente racional.

Es claro que el pensamiento mágico requiere alguna justificación filosófica para ser considerado un concepto de solución válido. En general, se ha justificado el pensamiento mágico por medio de teorías psicológicas. Posiblemente la más conocida sea la teoría de la proyección social (Alicke, Dunning y Krueger, 2005; Acevedo y Krueger 2005). Según esta teoría la mayoría de las personas espera que los otros se comporten como ellos lo hacen. Desde luego, la solución que promueven de los dilemas sociales, o los problemas de coordinación es sencilla: si elijo cooperar, y mis co-jugadores se me parecen, existe una probabilidad mayor de que elijan cooperar a que no lo hagan. Desde luego, la proyección social provee una solución psicológica, esto es, empírica, a los dilemas sociales y los problemas de coordinación, pero no una solución que pueda considerarse racionalmente fundamentada.

Hay otras versiones de la teoría de la proyección social que tienen un parecido de familia entre sí. La percepción de la similitud es particularmente importante en la investigación en psicología social, y ha sido propuesta para explicar la cooperación en juegos *one-shot* (Aksoy y Weesie, 2012). El comportamiento de los actores generalmente depende de las creencias sobre los motivos sociales y las estrategias de los demás (e incluso creencias de orden superior, como las creencias sobre las creencias de los demás acerca de las propias motivaciones sociales). Este punto ha sido enfatizado en la literatura de psicología social como se ejemplifica en la hipótesis del triángulo (Kelley y Stahelski, 1970), el sesgo de similitud estructural (SSE) (Kuhlman, Brown y Teta, 1992) y el modelo de cono de Iedema (1993). Estas tres teorías plantean diferentes relaciones entre las preferencias y las creencias acerca de las preferencias de los demás. La hipótesis del triángulo clasifica a los individuos en competidores y



cooperadores, y establece que los competidores esperan menos variación en la población que los cooperadores. Según SSE, los actores clasificados en una categoría de orientación social esperan que otros pertenezcan de manera predominante a su propia categoría. En coincidencia con el SSE, el modelo de cono de Iedema también predice que los actores de todas las categorías sociales esperan una mayor frecuencia de actores de su propia categoría. Esto está en línea con el efecto de falso consenso (Ross, Greene y House, 1977), que consiste en que en ausencia de mayor información, las personas esperan que las demás personas sean similares a ellos en términos de gustos, preferencias y otros atributos socialmente relevantes.

Uno de los problemas de estas nociones consiste en que el supuesto de independencia causal parece ser violado (sobre esta noción ver Bicchieri y Green, 1997), lo que descarta la elección de resultados fuera de la diagonal. Basta con que un jugador se forme la intención de elegir cierta estrategia para asegurarse que el otro jugador hará lo mismo. Mientras que la dependencia causal de las decisiones podría considerarse producto de una forma de razonamiento inválida, la noción de dependencia causal puede volver a entrar en escena en la forma de la correlación intrínseca de creencias (o considerando jerarquías de creencias). Entonces, en cierto modo, incluso si se respeta el principio de independencia causal de que las elecciones de un jugador no dependen causalmente de las de otro, las elecciones podrían estar correlacionados al nivel de las creencias (de unos sobre otros). Como señalaron Brandenburger y Friedenberg (2008), esta es una adaptación a la TJ de la idea del principio de causa común o correlación (Reichenbach, 1956).

Hay otras nociones que parecen seguir la misma línea, en el sentido de que no respetan el principio de independencia causal. Una de esas nociones es la propuesta por Morton (1994). La idea de Morton consiste en que cuando dos individuos tienen un objetivo en común pueden predecir sus acciones recíprocamente e interpretar las creencias de cada uno a través de un procedimiento que llama pensamiento de solución (*solution thinking*) ¿En qué consiste la noción propuesta por Morton? Es un intento de solución para el problema de la lectura mental en los juegos de coordinación. Al igual que el RE y la teoría de puntos focales, la solución propuesta por Morton pretende racionalizar la posibilidad de la coordinación para obtener beneficios mutuos en una amplia variedad de juegos, para los cuales el EqN no ofrece soluciones satisfactorias. El problema, según el autor, parece provenir de los supuestos que incorpora la teoría de la acción racional. Por eso es necesario ir más allá del alcance de la teoría pura, e incorporar intuiciones empíricas (Morton: 22). Para ello Morton recurre explícitamente a un concepto de similitud:

...el simple hecho de la similitud tiene muchas de las consecuencias que tendría el conocimiento mutuo perfecto. La similitud, utilizada correctamente, logra lo que parece necesitar creencias sobre creencias, creencias sobre creencias sobre creencias, etc. (p. 22)

El método de simulación defendido por Morton es una forma de simulación desde el punto de vista de la primera persona. Consiste en que un agente puede imaginar lo que haría si estuviera en el lugar del otro: lo que una persona haría en la situación de interés, le sirve para predecir lo que harían los demás. Es decir, para entender lo que harían los demás necesitamos ponernos a nosotros mismos en la situación en la que están ellos, y el resultado de la simulación en nosotros mismos podemos atribuirla a los otros.

A mi modo de ver la «*solution thinking*» de Morton es un esquema (necesariamente) inválido de razonamiento, una forma de razonamiento mágico. Para entender por qué sería instructivo ver su representación en el esquema de Gold y Sugden. Seguimos para ello la reconstrucción de Guala de la solución de Morton que tiene la virtud de enfatizar los aspectos de similitud entre jugadores (Guala 2016: 97). Guala discute la solución de Morton en relación con la teoría de puntos focales, es decir, como una noción que se encuadraría dentro de las nociones empíricas de similitud en nuestra taxonomía. Sin embargo, como quedará claro en la reconstrucción que realizo, el lugar adecuado para la solución de Morton es como una forma de pensamiento mágico.

Veamos el siguiente esquema de racionalidad individual para la solución de Morton:



1. 'Yo' identifico el punto focal obvio
  2. Dado que es exactamente como yo, el otro jugador también identifica el mismo punto focal
  3. Yo derivo mis acciones y las del otro jugador por medio de un simple razonamiento instrumental
  4. Predigo lo que hará y lo que el otro jugador cree que haré
  5. Quiero lograr (H, H)
- 
- Debo elegir H

Nótese que el pensamiento de solución se realiza en el modo de agencia individual, contrario al RE. De acuerdo con Guala, la simulación ocurre entre los pasos 2 y 4. Una vez que se identifica el punto focal cabe hacerse la pregunta acerca de qué fundamentos tenemos para pensar que el otro jugador podrá reconocer el mismo punto focal que yo, y además si lo elegirá. De acuerdo con los pasos (1) y (2) «Yo» identifico el punto focal, y como el otro jugador es exactamente como yo, entonces él/ella también lo identificará. Hasta aquí, todo lo que se requiere es un principio de simetría (que desde luego habría que justificar también). El problema surge en los pasos (3) y (4), que permanecen sin explicación. A partir de la premisa «simple razonamiento instrumental» todo lo que puedo concluir es que sería bueno para mí, y también para el otro jugador, que él/ella y yo eligiéramos H. Si eso fuera suficiente para inferir «Debo elegir H», entonces el paso en el razonamiento según el cual el otro jugador razona de manera *similar* a como razono yo, me permitiría inferir “Él/Ella elegirá H”. Pero ¿de dónde pueden provenir estas inferencias si no de alguna forma de pensamiento mágico? Guala no parece proveer un argumento genuino que permita justificar esta inferencia.

Adviértase que si hubiera un argumento que justificara la solución de Morton, ese mismo argumento sería atendible para sostener un argumento en favor de la noción de superracionalidad.

## Conclusión

En este artículo se ha pasado revista a una serie de enfoques sobre nociones de solución en teoría comportamental de juegos que presuponen diferentes nociones de similitud. Además se ha ofrecido una revisión crítica de dichas soluciones. Bajo la noción de similitud empírica se han agrupado las teorías de punto focal de Schelling y Lewis y el RE. Se han desarrollado diferentes maneras en que el RE se aparta de la noción de racionalidad estándar en la teoría de juegos ortodoxa, y hemos analizado las diferentes soluciones que ofrecen al dilema del prisionero.

También se ha discutido la llamada doctrina Harsanyi-Aumann, para subrayar las dos formas de similitud entre jugadores que sustenta el EqN: por un lado, la idea de que si los jugadores tienen la misma información, entonces obtendrán las mismas inferencias, por lo que los jugadores deberían ser similares en cuanto a su capacidad cognitiva; por otro lado, la similitud planteada por Aumann respecto de que los jugadores no pueden tener diferencias de información, porque frente a diferencias de información revisarían sus creencias hasta converger. Un problema que se plantea a esta visión de la similitud consiste en que este proceso de convergencia no puede ocurrir en juegos *one-shot*, por lo que parece más apropiado en el caso de juegos repetidos.

Por último, se ha ejemplificado el *pensamiento mágico* a partir de una noción de solución propuesta por Morton y fundamentada en formas de simulación, para mostrar que incurre en una forma de razonamiento inválido.

Diferentes nociones de similitud también han comenzado recientemente a permear en las ciencias del comportamiento. La similitud es un mecanismo cognitivo que tienen las personas para reducir la complejidad de las situaciones. Creemos que proporcionar una representación formal clara es una de las tareas que los filósofos analíticos pueden emprender, y que puede resultar en un trabajo propedéutico para la elaboración de modelos formales.

## Referencias

- AKSOY, O.; WEESIE, J. (2012). "Beliefs about the social orientations of others: A parametric test of the triangle, false consensus, and cone hypotheses". *Journal of Experimental Social Psychology*. 48. 45-54.
- ALICKE, M. D.; DUNNING, D. A.; KRUEGER, J. (2005). *The self in social judgment*. Psychology Press.
- AUMANN, R. (1976). "Agreeing to disagree". *The Annals of Statistics*. 4 (6). 1236-1239.
- AUMANN, R.; BRANDENBURGER, A. (1995). "Epistemic conditions for Nash equilibrium". *Econometrica*. 63(5). 1161-1180.
- BACHARACH, M. (1999). "Interactive team reasoning: A contribution to the theory of cooperation". *Research in Economics*. 53: 117-157.
- BACHARACH, M. (2006). *Beyond Individual Choice: Teams and Frames in Game Theory*. Editado por N. Gold y R. Sugden. Princeton. Princeton University Press. NJ.
- BARDSLEY, N.; MEHTA, J.; STARMER, C.; SUGDEN, R. (2010). "Explaining focal points: cognitive hierarchy theory versus team reasoning". *The Economic Journal*. 120 (543). 40-79. <https://doi.org/10.1111/j.1468-0297.2009.02304.x>
- BICCHIERI, C.; GREEN, M.S. (1997). "Symmetry arguments for cooperation in the prisoner's dilemma". En Bicchieri, C., Jeffrey, R., y Skyrms, B. (Editores). *The Logic of Strategy*. Kluwer Academic Publisher/Oxford University Press. New York. 229-249.
- BRANDENBURGER, A.; FRIEDENBERG, A. (2008). "Intrinsic correlation in games". *Journal of Economic Theory*. 141(1). 28-67. <https://doi.org/10.1016/j.jet.2007.09.012>
- BROWNE, K. (2018). "Why should we team reason". *Economics and Philosophy* 34. 185-198. <https://doi.org/10.1017/s0266267117000347>
- CAMPBELL, R. K. (1989). "The Prisoner's Dilemma and the Symmetry Argument for Cooperation". *Analysis*. 49 (2). 60-65.
- CAPRARO, V.; HALPERN, J. (2019). "Translucent players: Explaining cooperative behavior in social dilemmas". *Rationality and Society*. 31(4). 371-408.
- CHIERCHIA, G.; CORICELLI G. (2015). "The impact of perceived similarity on tacit coordination: propensity for matching and aversion to decoupling choices". *Frontiers in Behavioral Neuroscience*. 9. 202. <https://doi.org/10.3389/fnbeh.2015.00202>
- COLMAN, A. M. (2003). "Cooperation, psychological game theory, and limitations of rationality in social interaction". *Behavioral and Brain Sciences*. 26. 139-198. <https://doi.org/10.1017/s0140525x03000050>
- COLMAN, A. M.; PULFORD, B. D.; ROSE, J. (2008). "Collective rationality in interactive decisions: Evidence for team reasoning". *Acta Psychologica*. 128(2). 387-397. <https://doi.org/10.1016/j.actpsy.2007.08.003>
- DALEY, B.; SADOWSKY, P. (2017). "Magical Thinking: A Representation Result". *Theoretical Economics*. 12. 909-956. <https://doi.org/10.3982/te2099>
- GAUTHIER, D. (1987). *Morals by Agreement*. Oxford University Press.

- GEANAKOPOLOS, J.; PEARCE, D.; STACHETTI, E. (1989). "Psychological games and sequential rationality". *Games and Economic Behavior*. 1(1). 60-79.
- GOLD, N., Ed. (2005). *Teamwork: Multi-Disciplinary Perspectives*. Palgrave Macmillan. New York.
- GOLD, N. (2012). "Team reasoning, framing and cooperation". En Okasha, S. y Binmore, K. (Eds.) Publisher. Cambridge University Press. 185-212. <https://doi.org/10.1017/CBO9780511792601.010>
- GOLD, N. (2018). "Team reasoning: controversies and open research questions". En Jankovic, M. y Ludwig, K. (Eds.) *Handbook of Collective Intentionality*. Routledge. New York. 221-232.
- GOLD, N.; COLMAN, A. (2018). "Team Reasoning and the Rational Choice of Payoff-Dominant Outcomes in Games". *Topoi*. 39(2). 305-316. <https://doi.org/10.1007/s11245-018-9575-z>
- GOLD, N.; SUGDEN, R. (2007). "Collective intentions and team agency". *The Journal of Philosophy*. 104 (3). 109-137.
- DI GUIDA, S.; DEVETAG G. (2013). "Feature-Based Choice and Similarity Perception in Normal-Form Games: An Experimental Study". *Games*. 4(4). 776-794.
- FLOOD, M.; DRESHER, M. (1952). "Some experimental games". *Research memorandum RM-789*. Rand. Santa Monica, CA.
- FRANK, R. H.; GILOVICH, T.; REGAN, D. T. (1993). "The evolution of one-shot cooperation: An experiment". *Ethology and Sociobiology*. 14: 247-256.
- GAUTHIER, D. (1986). *Morals by Agreement*. Oxford University Press. Oxford.
- GUALA, F. (2016). *Understanding Institutions: The Philosophy and Science of Living Together*. Princeton University Press. Princeton.
- HARGREAVES, S.; VAROUFAKIS, Y. (1995). *Game Theory. A critical introduction*. Routledge. London.
- HARSANYI, J. C. (1968). "Games with incomplete information played by Bayesian players". Parts I, II, III. *Management Science*. 14. 159-182, 320-334, 486-502.
- HARSANYI, J. C.; R. SELTEN, R. (1988). *A General Theory of Equilibrium Selection in Games*. MIT Press. Cambridge. MA and London.
- HOFSTADTER, D. R. (1983). "Dilemmas for superrational thinkers, leading up to a luring lottery". *Scientific American*. 248(6).
- HOFSTADTER, D. R. (1985). *Metamagical themas: Questing for the essence of mind and pattern*. Basic Books. New York.
- HURLEY, S. (2005). "Rational agency, cooperation and mind-reading". En Gold, N. (ed). *Teamwork: multi-disciplinary perspectives*. 9. Palgrave MacMillan. Basingstoke. 200-215. [https://doi.org/10.1057/9780230523203\\_10](https://doi.org/10.1057/9780230523203_10)
- IEDEMA, J.; POPPE, M. (1994). "Causal attribution and self-justification as explanations for the consensus expectations of one's social value orientation". *European Journal of Personality*. 8. 395-408.
- JANSSEN, T. (2001). "Rationalizing focal points". *Theory and Decision*. 50. 119-148.

- KELLEY, H. H.; STAHELSKI, A. J. (1970). "Social interaction basis of 'cooperators and competitors' beliefs about others". *Journal of Personality and Social Psychology*. 16(1). 66-91. <https://doi.org/10.1037/h0029849>
- KRUEGER, J. I.; ACEVEDO, M. (2005). "Social Projection and the Psychology of Choice". En M. D. Alicke, D. A. Dunning, y J. I. Krueger (Eds.). *Studies in self and identity. The Self in Social Judgment*. Psychology Press. 17-41.
- KRUEGER, J.I.; DIDONATO, T.E.; FREESTONE, D. (2012). "Social projection can solve social dilemmas". *Psychol Inq*. 23. 1-27. <https://doi.org/10.1080/1047840X.2012.641167>
- KUHLMAN, D. M., BROWN, C.; TETA, P. (1992). "Judgements of cooperation and defection in social dilemmas: The moderating role of judge's social orientation". En W. LEWIS, D. K. 1969. *Convention: A Philosophical Theory*. Harvard University Press. Cambridge.
- LIPMAN, B. (1991). "How to Decide How to Decide How to...: Modeling Limited Rationality". *Econometrica*. 59(4). 1105-1125.
- MORTON, A. (1994). "Game Theory and Knowledge by Simulation". *Ratio*. 7(1). 14-25. <https://doi.org/10.1111/j.1467-9329.1994.tb00150.x>
- MUSSWEILER, T.; OCKENFELS, A. (2013). "Similarity increases altruistic punishment in humans". *PNAS*. 110(48). 19318-19323.
- MYERSON, R. B. (1991). *Game Theory: Analysis of Conflict*. Harvard University Press. Cambridge.
- NASH, J. F. (1950). "Equilibrium points in n-person games". *PNAS*. 36 (1). 48-49. <https://doi.org/10.1073/pnas.36.1.48>
- QUATTRONE G.A.; TVERSKY, A. (1984). "Causal versus diagnostic contingencies: on self-deception and on the voter's illusion". *Journal of Personality and Social Psychology*. 46(2). 237-248. <https://doi.org/10.1037/0022-3514.46.2.237>
- RAPOPORT, A. (1960). *Fights, games, and debates*. University of Michigan. Ann Arbor.
- REICHENBACH, H. (1956) *The Direction of time*. University of Los Angeles Press. Berkeley.
- ROSS, D. (2016). *Philosophy of Economics*. Palgrave. New York: <https://doi.org/10.1057/9781137318756>
- ROSS, L.; GREENE, D.; HOUSE, P. (1977). "The false consensus effect: An egocentric bias in social perception and attribution processes". *Journal of Experimental Social Psychology*. 13(3). 279-301. [https://doi.org/10.1016/0022-1031\(77\)90049-x](https://doi.org/10.1016/0022-1031(77)90049-x)
- RUBINSTEIN, A.; SALANT, Y. (2016). "'Isn't everyone like me?': On the presence of self similarity in strategic interactions". *Judgment and Decision Making*. 11 (2). 168-173.
- SCHELLING, T. (1981). *The Strategy of Conflict*. Harvard University Press. Cambridge.
- SHAFIR, E.; TVERSKY, A. (1992). "Thinking Through Uncertainty: Nonconsequential Reasoning and Choice". *Cognitive Psychology*. 24(4). 449-474. [https://doi.org/10.1016/0010-0285\(92\)90015-t](https://doi.org/10.1016/0010-0285(92)90015-t)
- SINN, H. W. (1980). "A rehabilitation of the principle of insufficient reason". *Quarterly Journal of Economics*. 94(3). 493-504. <https://doi.org/10.2307/1884581>

SPIEKERMANN, K. (2007). “Translucency, assortment, and information pooling: how groups solve social dilemmas”. *Politics, Philosophy and Economics*. 6. 285-306.

SUGDEN, R. (2000). “Team preferences”. *Economics and Philosophy*. 16(2). 175-204

SUGDEN, R. (2003). “The logic of team reasoning”. *Philosophical Explorations*. 6(3). 165-181.  
<https://doi.org/10.1080/10002003098538748>

SUGDEN, R. (2011). “Mutual advantage, conventions and team reasoning”. *Int Rev Econ*. 58(1). 9–20.  
<https://doi.org/10.1007/s12232-011-0114-0>

SUGDEN, R. (2015). “Team reasoning and intentional cooperation for mutual benefit”. *Journal of Social Ontology*. 1. 143–166. <https://doi.org/10.1515/jso-2014-0006>



## ALEGATO POR EL CONTROL DEL RAZONAMIENTO

*A PLEA FOR THE CONTROL OF REASONING*

*UM APELO A FAVOR DO CONTROLE DO RACIOCÍNIO*

**Enric Trillas**

(Universidad de Oviedo)  
*etrillasetrillas@gmail.com*

**Alejandro Sobrino Cerdeiría**

(Universidad de Santiago de Compostela)  
*alejandro.sobrino@usc.es*

Recibido: 18/08/2020

Aprobado: 05/02/2021

### RESUMEN

El objetivo de este trabajo es mostrar que la falsación popperiana es un tipo extremo de control del razonamiento. Para ello se muestra el ‘esqueleto del razonamiento ordinario’, que parte de la consideración de que un enunciado  $q$  puede ser una consecuencia de otro  $p$  si cabe afirmar el enunciado condicional ‘Si  $p$ , entonces  $q$ ’, abreviadamente,  $p < q$ , denotando  $<$  la relación de inferencia, una relación que, al no ser entendido el enunciado condicional de la misma forma en todos los contextos que abarca el lenguaje, se supondrá primitiva y que, en cada caso, deberá explicitarse en la forma que corresponda. Ello permitirá examinar el falsacionismo desde los conceptos formales de refutación y especulación, permitiendo una nueva mirada sobre lo que pretende ser un criterio de demarcación entre lo que es científico y lo que no lo es, así como arrojar dudas sobre la identificación entre ‘conjetura’ e ‘hipótesis’, al mostrar que hay varios tipos de conjeturas.

Palabras clave: falsacionismo. conjetura. hipótesis. especulación.

### ABSTRACT

The objective of this work is to show that Popperian falsification is an extreme type of reasoning control. The ‘skeleton of ordinary reasoning’ is shown to approach that purpose, starting from the consideration that a statement  $q$  can be a consequence of another  $p$  ‘If  $p$ , then  $q$ ’, abbreviated,  $p < q$ , denoting  $<$  the inference relation, a relation that. Since the conditional statement is not understood in the same way in all the contexts covered by the language, it will be assumed primitive and, therefore, must be approached correspondingly in each case. This standpoint will allow examining falsifiability from the formal concepts of refutation and speculation, allowing a new look at the demarcation criterion and cast doubt on the identification between ‘conjecture’ and ‘hypothesis’ by showing various kinds of conjecture.

Keywords: falsationism. conjecture. hypothesis. speculation.

## RESUMO

O objetivo deste trabalho é mostrar que a falseabilidade popperiana é um tipo extremo de controle do raciocínio. Para isso, é mostrado o 'esqueleto do raciocínio ordinário', que parte da consideração de que um enunciado  $q$  pode ser consequência de outro  $p$  se o enunciado condicional 'Se  $p$ , então  $q$ ', abreviadamente,  $p \rightarrow q$ , denotando  $\rightarrow$  a relação de inferência, relação que, uma vez que o enunciado condicional não é entendido da mesma forma em todos os contextos que a linguagem abrange, será assumida como primitiva e que, em cada caso, deve ser explicitada da forma adequada. Isso nos permitirá examinar o falsificacionismo a partir dos conceitos formais de refutação e especulação, permitindo um novo olhar sobre o que se pretende ser um critério de demarcação entre o que é científico e o que não é, bem como lançar dúvidas sobre a identificação entre 'conjectura' e 'hipótese', por mostrar que existem vários tipos de conjecturas.

Palavras-chave: falsificacionismo. conjectura. hipótese. especulação.

## Introducción

Este artículo no trata directa, específica y únicamente del 'racionalismo crítico', como Karl R. Popper denominó (Popper, 1934, 1963) y su seguidor David Miller aún denomina (Miller, 1994) la doctrina filosófica también conocida por *falsacionismo* o *refutacionismo*. No se intentará más que ofrecer una visión previa de cuánto y en relación con esa doctrina se refiere a la refutación, desde el punto de vista del modelo matemático que el primer autor ha llamado el *esqueleto formal* del razonamiento ordinario, de sentido común o de cada día y que se efectúa en y con un lenguaje natural que domine cada pensador (Trillas, 2017, 2018, 2020, en prensa).

Se trata de un modelo que, siendo un marco formal y muy general de ese razonamiento, permite establecer resultados válidos en cualquier otro tipo de razonamiento que no acepte ni como premisas ni como conclusiones enunciados auto-contradictorios. Tal esqueleto formal soporta, se encuentra, en muchos de los modelos que desarrolla la Lógica la cual y desde siempre, se ha entendido como el estudio de las condiciones que hacen correcto al razonamiento deductivo; un claro antecedente de control del razonamiento. El 'esqueleto' conduce a una interpretación de las palabras de Albert Einstein, *La ciencia proviene de refinar el pensamiento ordinario*.

Contemplar el falsacionismo desde los conceptos formales de refutación y especulación podrá, tal vez, echar algo de luz acerca de una visión filosófica que nada menos pretende ser un doble criterio de demarcación entre lo que es científico y lo que no lo es, así como entre la inducción y lo que no lo es y negando, a la vez, que la inducción sirva para alcanzar el conocimiento científico, para lo cual sólo aceptan la deducción. Repensar la visión de Popper desde aquel esqueleto significa hacerlo con una perspectiva que, no coincidente con lo que aportan sus escritos y los de Miller, ofrece, sin embargo, un punto de vista nuevo que permite, a la vez, considerarla de manera 'naïve' y más amplia, al no afectar únicamente al razonamiento científico. Tanto Popper como Miller, por ejemplo, identifican 'conjetura' con 'hipótesis' pero, como se verá, hay otros tipos de conjeturas.

Originalmente, se trató de un tema íntimamente relacionado con los debates sobre el método científico que, básicamente mantenidos en una parte de la Filosofía de la Ciencia tienen, no obstante, y razonablemente, escaso eco en los ámbitos científicos.

En todo caso, este artículo no pretende ser conclusivo, es únicamente una reflexión en favor del 'control del razonamiento' y considerando pernicioso su descontrol; es bien sabido que algo descontrolado ofrece peligro y una de sus pretensiones reside en mostrar que la falsación popperiana es un tipo extremo de control del razonamiento. Es gracias al razonamiento cómo, finalmente y de la manera que sea, se obtiene el conocimiento científico.

## El esqueleto del razonamiento ordinario

Se debe, ante todo, exponer el modelo matemático llamado ‘esqueleto formal’ del razonamiento ordinario (Trillas, 2020) que se ha citado. Tal modelo parte de observar, en el lenguaje, que un enunciado  $q$  puede ser una consecuencia de otro  $p$ , si cabe afirmar el enunciado condicional ‘Si  $p$ , entonces  $q$ ’, abreviadamente,  $p < q$ , y llamando a la relación binaria  $<$  así definida la relación de inferencia.

Una relación que, al no ser entendido el enunciado condicional de la misma forma en todos los contextos que abarca el lenguaje, se supondrá primitiva y que, en cada caso, deberá explicitarse en la forma que corresponda; por ejemplo y entre otras, se entiende como el enunciado incondicional ‘no  $p$  o  $q$ ’, abreviado por  $p' + q$ , o como ‘no  $p$  o ( $p$  y  $q$ )’, abreviado por  $p' + p \cdot q$ , e incluso, cuando carece de sentido considerar no- $p$ , por ‘ $p$  y  $q$ ’, abreviado por  $p \cdot q$ .

El razonamiento ordinario es contextual, cuanto maneja se refiere a situaciones concretas; por ello, los enunciados deben ser tratados mediante su significado local, especificándolos por lo que indican en la concreta situación del discurso y del contexto que lo circunda (Trillas, 2017, 2018).

2.1. Con ello, una *consecuencia* de  $p$  es toda  $q$  tal que  $p < q$ , una *hipótesis* o explicación para  $p$  es toda  $h$  tal que  $h < p$ , en tanto que una *especulación* desde  $p$ , es toda  $e$  que ni sea consecuencia ni hipótesis y bien verifique  $e' < p$ , o bien  $e'$  y  $p$  no sean comparables entre sí por la relación de inferencia, lo cual se abreviará  $p \diamond e'$ . En el primer caso,  $p \diamond e$  y  $e' < p$ , se dice que  $e$  es una especulación *débil* y en el segundo,  $p \diamond e$  y  $p \diamond e'$ , que es una especulación *fuerte*.

En cuanto a la relación  $<$ , sólo se le supondrán dos propiedades universales:

- 1) La ley reflexiva,  $p < p$ , para toda  $p$ ;
- 2) La regla del Modus Ponens,  $(p \cdot (p < q)) < q$ .

La primera garantiza que  $<$  no es vacía y la segunda (o regla MP) que  $<$  es efectiva, es decir, que dados  $p$  y  $p < q$ ,  $q$  es alcanzado. Sin ellas, la relación  $<$  podría no existir o ser inútil. Nótese que si (1) garantiza que el conjunto de las consecuencias de  $p$ ,  $p\uparrow = \{q; p < q\}$ , no es vacío ya que por lo menos es  $p \in p\uparrow$ , (2) garantiza que los demás elementos de  $p\uparrow$  puedan ser alcanzados. Nótese que si la fórmula (2) del Modus Ponens, puede ser resuelta como una inecuación, entonces cabe encontrar todas las expresiones para la relación  $<$ ; ello es el caso en las álgebras de Boole (Birkhoff, 1948), puesto que la inecuación  $p \cdot (p < q) < q$  es, en ellas e identificando  $<$  con su orden reticular, equivalente a  $(p < q) < p' + q$ ; es decir, en las álgebras de Boole cualquier expresión menor o coincidente con  $p' + q$  (no  $p$  o  $q$ ), verifica MP y por tanto puede representar  $p < q$ . Es el caso, por ejemplo, de  $p \cdot q$  ( $p$  y  $q$ ) puesto que, en esas álgebras, es  $p \cdot q < q < p' + q$ .

En cuanto a la negación ‘no’, sólo se supondrá que satisface la ley universal llamada de *contraposición*:

Si  $p < q$ , entonces  $q' < p'$ . Abreviadamente,  $(p < q) < (q' < p')$ .

A la vez, se excluirá a cualquier enunciado  $p$  tal que verifique  $p < p'$ , que sea *auto-contradictorio*. Tal denominación proviene de definir que  $r$  contradice o refuta a  $p$ , si es  $p < r'$  y con lo cual  $p < p'$  indica que  $p$  se refuta o contradice a sí mismo, se auto-refuta o auto-contradice; es formalmente imposible.

Nótese que si es  $p < p'$ , entonces y en virtud de la ley de contraposición, es  $(p'')' < p'$ . Cuando un enunciado  $s$  es tal que verifica  $s'' = (s'')' < s'$ , también verifica  $s'' < s'''$ ,  $s^{iv} < s'''$ , etc., y se dice que es una *tautología*. Por tanto, la negación de un enunciado auto-contradictorio es una tautología, es tautológico.

La negación se clasifica según su comportamiento en cada enunciado  $q$  respecto de  $<$ ; de hecho, ello es una propiedad *local* y no universal como lo es la reversibilidad que, en realidad, es la definición de una

negación como aplicación que asigna a cada enunciado  $q$  el  $\text{no-}q = q'$ . Algo que, además, debe comprobarse en cada caso; es decir, comprobar, dado  $q$ , la existencia real de un  $q'$  sólo sometido a cumplir la ley de contraposición. Pues bien, la negación se llama:

a) *Intuicionista* en  $q$ , si es  $q'' < q$ .

b) *Débil* en  $q$ , si es  $q < q''$ ;

obviamente se trata de dos definiciones no incompatibles y, por ello, la negación también se llama:

c) *Fuerte* en  $q$  de ser  $q'' < q$  y  $q < q''$ , abreviadamente,  $q \sim q''$ ,

d) *Salvaje* en  $q$  cuando ni es  $q'' < q$ , ni  $q < q''$ , es decir, cuando  $q$  y  $q''$  no son comparables por la relación de inferencia, abreviadamente  $q \diamond q''$ .

Fuerte significa a la vez intuicionista y débil en  $q$ , en tanto que salvaje significa ni intuicionista ni débil en  $q$ .

Con independencia del carácter de la negación, la ley de contraposición permite probar que la relación  $<$  de inferencia verifica la llamada Regla Modus Tollens, o regla MT:

Si dado  $p < q$  se conoce  $q'$ , entonces también se conoce  $p'$ .

En efecto,  $[q' \cdot (p < q)]$  implica  $[q' \cdot (q' < p')]$  que, por la regla MP, lleva a  $p'$ ; en particular,  $[q \in p \uparrow \Rightarrow p' \in q \uparrow]$ .

Debe notarse que ambas reglas MP y MT sólo son efectivamente aplicables si las respectivas premisas,  $p \cdot (p < q)$  y  $q' \cdot (p < q)$  no son auto-contradictorias. Por ejemplo, en un álgebra de Boole (Birkhoff, 1948) e interpretando el enunciado condicional  $p < q$  por el enunciado incondicional  $p \cdot q$ , no puede aplicarse la regla MT:  $q' \cdot (p < q) = q' \cdot (p \cdot q) = p \cdot (q' \cdot q) = p \cdot 0 = 0$ , lo cual asegura, como se verá a continuación, que la premisa  $q' \cdot (p < q)$  es auto-contradictoria. No sucede lo mismo de interpretar  $p < q$  como  $p' + q$ , ya que:  $p \cdot (p' + q) = p \cdot p' + p \cdot q = 0 + p \cdot q = p \cdot q < q$ . Nótese que, de prescindir del último paso anterior, la ecuación  $p \cdot (p < q) = p \cdot q$  puede permitir encontrar más fácilmente algunas expresiones para  $p < q$ ; así, como de ella se sigue

$$p' + p \cdot (p < q) = p' + p \cdot q \Rightarrow (p' + p) \cdot (p' + (p < q)) = (p' + p) \cdot (p' + q) \Rightarrow p' + (p < q) = p' + q,$$

cualquier expresión de  $p < q$  que la satisfaga será una solución, como es el caso, por ejemplo, de  $q + p' \cdot q'$ , ya que es inmediato comprobar  $p' + q + p' \cdot q' = p' + q$ .

Los nombres de ambas reglas provienen, respectivamente, de abreviar las expresiones latinas *Modus Ponendo Ponens*, el modo de ‘poner’ la conclusión tras poner la premisa, y *Modus Tollendo Tollens*, el modo de ‘quitar’ la negación de la premisa tras quitar la negación de la conclusión.

2.2. Definida la contradicción, el nombre auto-contradictorio no requiere explicación adicional, pero lo requiere el de tautología que, originalmente, se refirió a aquellos enunciados que se verifican para todos los elementos a los que se apliquen; por ejemplo, el enunciado ‘ $n$  es par’ es una tautología si  $n$  es cualquier número par y no lo es si  $n$  es cualquier número natural aunque lo es el enunciado ‘ $n = n$ ’, todo número natural es igual a sí mismo.

En las álgebras de Boole (Birkhoff, 1948), donde la relación  $<$  se considera igual al orden reticular subyacente:  $p < q \Leftrightarrow p + q = q \Leftrightarrow p \cdot q = p$ , la fórmula  $p < q'$  implica la  $p \cdot q < q' \cdot q = 0$  que, a su vez implica  $p \cdot q = 0$ ; es decir, si  $q$  contradice  $p$ , ambos son incompatibles. Recíprocamente, si  $p \cdot q = 0$ , en virtud de la ley booleana del reparto perfecto,  $p = p \cdot q + p \cdot q'$ , es  $p = 0 + p \cdot q' = p \cdot q' \Leftrightarrow p < q'$ . Por

lo tanto, en las álgebras de Boole es  $p < q' \Leftrightarrow p \cdot q = 0$ , sus muchas leyes no permiten distinguir entre contradicción e incompatibilidad.

En esas álgebras, la negación es fuerte en todos sus elementos; el carácter fuerte de la negación es universal en ellas. Además, la anterior relación  $\sim$  es la igualdad, ya que  $p < q$  y  $q < p$  significa, en ellas,  $p = q$ , puesto que  $p < q \Leftrightarrow p = p \cdot q$ ,  $q < p \Leftrightarrow q = q \cdot p$  y, al ser la conjunción ( $\cdot$ ) conmutativa, sigue  $p = p \cdot q = q \cdot p = q$ .

Con ello, es  $p < p' \Leftrightarrow p \cdot p = 0 \Leftrightarrow p = 0$ ; el único elemento auto-contradictorio es el mínimo 0. Nótese que,  $p' < p \Leftrightarrow p' \cdot p = 0 \Leftrightarrow p' = 0 \Leftrightarrow p = 1$ , es decir,  $p$  es el máximo y en el cual ‘están contenidos’ todos los elementos del álgebra, todos la verifican: es la tautología de la misma. Por ejemplo, en el álgebra de Boole  $2^4$  con cuatro átomos  $a, b, c, d$ , y 16 elementos obtenidos con todas las disyunciones posibles de los átomos más el mínimo  $0 = a \cdot b = a \cdot c = \dots = c \cdot d$ , el elemento máximo es  $a + b + c + d = 1$  y, obviamente, es:  $a < 1, b < 1, c < 1, d < 1, a + b < 1$ , etc. De darse 1 se dan, automáticamente, todos los elementos del álgebra de Boole. De ahí que, pese a las muchísimas leyes universales de que gozan esas álgebras, se haya extrapolado el nombre tautología para designar a los  $p$  tales que  $p' < p$ .

Debe observarse que, en el lenguaje en y con el cual se desarrolla el razonamiento ordinario, la existencia de enunciados mínimo o máximo, 0 y 1, es decir que para cualquier otro enunciado  $q$  sea  $0 < q$  o  $q < 1$ , no es evidente en todos los contextos. Su existencia no puede considerarse más que una propiedad local; sin embargo, en poco se retomará el asunto.

2.3. De hecho, dos enunciados  $u$  y  $v$  se llaman *equivalentes por inferencia*, cuando valen simultáneamente  $u < v$  y  $v < u$ , lo cual, y como se dijo, se abrevia  $u \sim v$ . Una relación que es, evidentemente, reflexiva y simétrica, pero que, en general, no es transitiva; en general, no es una relación de equivalencia algebraica (Trillas, 2020), (Birkhoff, 1948). Sin embargo, una condición suficiente para ello es que  $<$  sea por lo menos localmente transitiva, es decir, que verifique  $[p < q, q < r \Rightarrow p < r]$ . En efecto, de ser  $<$  transitiva localmente en la terna  $(p, q, r)$ , entonces:

$$p \sim q \ \& \ q \sim r \Leftrightarrow ((p < q) \ \& \ (q < p)) \ \& \ ((q < r) \ \& \ (r < q)) \Leftrightarrow ((p < q) \ \& \ (q < r)) \ \& \ ((r < q) \ \& \ (q < p)) \Rightarrow (p < r) \ \& \ (r < p) \Leftrightarrow p \sim r.$$

Cuando la transitividad de  $<$  es universal, como es el caso de las álgebras de Boole, la relación  $\sim$  es una equivalencia algebraica o igualdad; como se mostró, en esas álgebras es  $p \sim q \Leftrightarrow p = q$ .

En cualquier fórmula involucrando la relación de inferencia  $<$ , puede substituirse todo enunciado  $q$  por otro  $r$  equivalente a él y obteniéndose un resultado también equivalente. Así, en la fórmula booleana  $p + q \cdot q'$  puede substituirse  $q \cdot q'$  por 0 y obtenerse  $p + q \cdot q' = p + 0 = p$ ; los enunciados  $p + q \cdot q'$  y  $p$  son equivalentes en las álgebras de Boole y aunque en general sólo verifiquen  $p < p + q \cdot q'$ .

En el lenguaje no siempre puede asegurarse que una terna sea transitiva (Trillas, 2020) y el fallo de la transitividad provoca que, por ejemplo, no pueda asegurarse que, como se probará de inmediato, una consecuencia o una hipótesis sean conjeturas.

Si una refutación de  $p$  es un enunciado  $r$  tal que  $p < r'$ , una *conjetura*  $q$  es un enunciado que no refuta a  $p$ , es decir, tal que  $p \not< q'$ , no es  $p < q'$ ; obviamente, dado  $p$  no hay más que refutaciones y conjeturas. En particular, las especulaciones son conjeturas ya que tanto  $e' < p$  como  $e' \diamond p$  implican  $p \not< e'$ . Con todo ello y en presencia de transitividad local:

1) Si es  $p < q$ , con  $p \not< p'$  ( $p$  no auto-contradictorio), también es  $p \not< q'$ .

En efecto, si fuese  $p < q'$ , como sea que  $p < q \Rightarrow q' < p'$ , la transitividad asegura que vale el absurdo  $p < p'$ ; por lo tanto, es  $p \not< q'$ . Es decir, bajo transitividad local (en la terna  $(p, q', p')$ ) por lo menos, toda consecuencia de una premisa no auto-contradictoria es una conjetura.



2) Si es  $h < p$  y  $h < h'$ , también es  $h < p'$ .

En efecto, de ser  $h < p'$ , a causa de que  $h < p \Rightarrow p' < h'$ , seguiría el absurdo  $h < h'$ ; por lo tanto, es  $h < p'$ . Es decir, bajo transitividad local (en la terna  $(h, p', h')$  por lo menos), toda hipótesis no auto-contradictoria es una conjetura.

Nótese que si es  $p < q$ , no puede ser  $q < q'$ . De ser así, la transitividad local en la terna  $(p, q, q')$  llevaría a  $p < q'$ , que es absurdo por (1). Por lo tanto, bajo transitividad local no hay consecuencias auto-contradictorias de una premisa no auto-contradictoria. Es más, dos consecuencias cualesquiera no pueden ser contradictorias: De ser  $p < q$  y  $p < u$ , si fuese  $u < q'$ , al ser también  $q' < p'$ , la transitividad local implicaría el absurdo  $p < p'$ .

Sin embargo, esos resultados no pueden deducirse en el caso de las hipótesis, en el cual la experiencia enseña que, muchas veces, se presentan hipótesis contradictorias de las cuales, por tanto, sólo una de ellas podrá aceptarse. Nada garantiza, a la vez, que una hipótesis deba ser no auto-contradictoria; por ello, una tal hipótesis se excluye y, realmente, las hipótesis que se consideran son aquellos enunciados  $h$  tales que  $h < h'$ ,  $h < p$  pero sin que sea  $p < h$ , es decir, sin que sea  $h \sim p$ , que equivaldría a explicar  $p$  con  $p$ .

Así pues, bajo transitividad local en las ternas que corresponda y dada una premisa  $p$  no auto-contradictoria, no hay más que refutaciones, consecuencias, hipótesis y especulaciones. Bajo transitividad, las conjeturas se clasifican perfectamente en consecuencias, hipótesis y especulaciones débiles y fuertes.

2.4. Por todo lo dicho, el razonamiento ordinario sólo consiste en buscar como conclusiones, refutaciones, consecuencias, hipótesis y especulaciones. Además, bajo condiciones de transitividad local adecuadas, las tres últimas conclusiones son las únicas conjeturas posibles. En resumen, razonar es bien *refutar*, o bien *deducir* o *abducir*, o bien *especular*; abducir no es más que deducir hacia atrás y se considera que la Lógica es el estudio de la deducción correcta. Ciertamente, razonar equivale a conjeturar y refutar.

Al modelo 'esqueleto' le faltan aún las leyes universales de la conjunción ( $\cdot$ ) y la disyunción ( $+$ ). Son las siguientes,

a) *Conjunción*:  $p \cdot q < p, p \cdot q < q$

b) *Disyunción*:  $p < p + q, q < p + q$ .

De ellas se sigue que también vale:

$q \cdot p < p, q \cdot p < q, q < q + p, p < q + p,$

sin suponer que se trate de operaciones conmutativas.

El hecho que tanto  $p$  como  $q$  estén por encima de  $p \cdot q$  y por debajo de  $p + q$ , respecto de la relación  $<$  de inferencia, tampoco significa que  $p \cdot q$  sea el menor de todos los enunciados que están por encima de ambos  $p$  y  $q$ , ni que  $p + q$  sea el mayor de cuantos están por debajo, como sucede en los retículos (Birkhoff, 1948); enunciados que, por otra parte, sería muy difícil describir en forma lingüística explícita. Nótese que, si las anteriores leyes no llevan a un retículo; sin embargo, no son incompatibles con esa estructura tan rígida; la toleran.

Cuanto cabe afirmar es que tanto  $p$  como  $q$  están 'entre'  $p \cdot q$  y  $p + q$ :  $p \cdot q < p$  y  $p < p + q$ , indican que es  $p \cdot q < p < p + q$ , así como  $p \cdot q < q < p + q$ , fórmulas que, bajo transitividad, muestran  $p \cdot q < p + q$ , así como que de ser  $p \cdot q$  el menor de los de encima y  $p + q$  el mayor de los de debajo (es decir, de generar un retículo), entonces  $p$  y  $q$  son los únicos enunciados entre su conjunción y su disyunción que,

además y necesariamente, son ambas conmutativas y asociativas aunque sin, también necesariamente, contar con máximo y mínimo.

Es posible, en cambio y aunque tampoco sea necesario, que existan minimales o maximales, es decir, enunciados  $u$  tales que no exista ningún  $z$  verificando  $z < u$ , y enunciados  $v$  tales que no exista ningún  $z$  verificando  $v < z$ , respectivamente. Obviamente, si hay un único minimal es el mínimo y si hay un único maximal es el máximo y, en tal caso, ambos son únicos en el sentido de ser equivalentes por inferencia.

En efecto, de haber dos mínimos  $0_1$  y  $0_2$ , se verificarían simultáneamente  $0_1 < 0_2$  y  $0_2 < 0_1$ , con lo cual sería  $0_1 \sim 0_2$ . Y análogamente de existir dos máximos:  $1_1 < 1_2$  y  $1_2 < 1_1 \Rightarrow 1_1 \sim 1_2$ .

Cuando la relación  $<$  es transitiva y, por ello, también lo es la relación  $\sim$ , entonces las clases  $[u] = \{v; v \sim u\}$  son clases de equivalencia algebraica y, por tanto, son una clasificación perfecta o *partición* del conjunto de todos los enunciados involucrados. Tales clases reciben el nombre de *proposiciones*; sin embargo, el lenguaje no maneja siempre proposiciones sino enunciados. Además, y en el mismo, aunque esas clases existan, pueden no ser disjuntas entre sí.

La transitividad, por lo menos local, hace más seguro el razonamiento ordinario; por ello, cuando la transitividad falla, se dice que el razonamiento ordinario es irregular y, en él, pueden existir consecuencias o hipótesis que no sean conjeturas, así como pares de consecuencias contradictorias y proposiciones con elementos comunes. El irregular es un tipo de razonamiento ordinario muy poco seguro; por ello es conveniente asegurarse de que haya transitividad en las ternas de enunciados que corresponda, que el razonamiento sea regular.

2.5. En el esqueleto regular se verifican los dos teoremas siguientes,

1) *No Contradicción*. Cualquiera que sea  $p$ , es  $p \cdot p'$  auto-contradictorio, imposible.

En efecto,  $p \cdot p' < p \Rightarrow p' < (p \cdot p')'$  y como también es  $p \cdot p' < p'$ , la transitividad en la terna  $(p \cdot p', p', (p \cdot p')')$  conduce a  $p \cdot p' < (p \cdot p')'$ .

Se trata de un resultado que garantiza el 'principio' de No Contradicción enunciado por Aristóteles como tal al ser, según él, no susceptible de prueba.

2) *Tercero Excluido*. Cualquiera que sea  $p$ ,  $p + p'$  es una tautología.

En efecto,  $p < p + p' \Rightarrow (p + p')' < p'$  y como también es  $p' < p + p'$ , la transitividad en la terna  $((p + p')', p', p + p')$  conduce a  $p + p' < (p + p')'$ .

Se trata de un resultado que Aristóteles también enunció como un 'principio', si bien lo hizo de una manera menos clara que el de No-Contradicción.

La prueba de estos antiguos principios aristotélicos, ahora teoremas y siempre vistos a lo largo de la historia como una garantía de razonamiento correcto, representa un soporte intelectual al modelo 'esqueleto'. Es por ello que los seis principios en los que se basa, dos para la relación de inferencia, uno para la negación, los dos para la conjunción y la disyunción, más la transitividad local, pueden llamarse 'principios formales del razonamiento ordinario'. Unos principios que, formalmente, son anteriores a los dos principios de Aristóteles.

2.6. Con cuanto se ha expuesto surge una 'regimentación' del razonamiento ordinario que, básicamente, lo clasifica en dos categorías distintas, es más, disjuntas: Conjeturar y refutar. Con ello, deductivamente y sea hacia adelante o hacia atrás, surge la pregunta, ¿Qué representan las conclusiones respectivas?

Por lo que respecta a las conjeturas,  $p < c'$ , negar la relación  $<$  obliga a considerar las distintas posibilidades que contiene. Si  $c$  es una consecuencia,  $p < c$ , la regla MP indica que conocida la premisa, también se conoce la consecuencia por inferencia

Si  $c$  es una hipótesis,  $c < p$ , y si es aplicable la regla MT, conocido  $p'$  también se conoce  $c'$  y, negándola a  $(c')'$  si la negación no es salvaje en  $c$ , bien de  $c < (c')'$  se abduce  $c$ , bien de  $(c')' < c$  se deduce  $c$ . Es decir, si la negación no es salvaje, conocida la premisa también se conoce la hipótesis por inferencia.

Si  $c$  es una especulación,  $p \diamond c$ , nada puede asegurarse deductivamente excepto si se trata de una especulación débil cuando, como se mostrará en el siguiente apartado, se obtiene deductivamente y resulta efectivamente alcanzable. La gran diferencia con las especulaciones fuertes es su alcance que, no obstante, y como se verá, puede lograrse por medio de aplicar sucesivamente la regla MP.

En cuanto a las refutaciones,  $p < r'$  implica  $(r')' < p'$  y entonces, supuesta la transitividad de  $<$  y que la negación es débil o intuicionista en  $r$ ,  $r < (r')'$ , sigue  $r < p'$ , se llega a  $\neg p$  deductivamente. La refutación lleva a la negación de la premisa.

Debe notarse que, en varios de los casos, se acaba en un enunciado negativo que coexiste con su original, con lo cual se llega al absurdo de su conjunción; así es en el de las refutaciones donde se llega a  $p \cdot p'$ . En el caso de las consecuencias,  $p < c \Rightarrow c' < p'$  y la regla MP muestra que se dispone tanto de  $p$  como de  $p'$ . En el caso de las hipótesis,  $c < p \Rightarrow p' < c'$  y coexisten  $c$  y  $c'$ . En el siguiente apartado se verá qué tiene todo ello que ver con la presunta ‘verdad’ tanto de la premisa como de la conclusión.

## Especular, abducir y refutar

3.1. El razonamiento ordinario se produce, real y efectivamente, mediante procesos secuenciales que alternan deducción hacia adelante y hacia atrás o abducción (Trillas, 2020, en prensa). La deducción lleva a consecuencias, la abducción a hipótesis y, cuando el camino mezcla deducción y abducción, cuando es mixto, puede llevar a especulaciones.

Por ejemplo, en un álgebra de Boole finita  $2^4$  y cuyos cuatro átomos sean  $a, b, c, d$ , de la premisa  $p = a + b$  se llega, añadiendo otros átomos, a todas sus consecuencias  $\{a + b + c, a + b + d, a + b + c + d = 1\}$ , en tanto que quitando átomos se llega a todas sus hipótesis  $\{a, b\}$ . Si se quitan y añaden simultáneamente átomos, se llega a elementos ortogonales a  $p$ , que pueden ser especulaciones:

- El camino mixto:  $p = a + b > a < a + c = e$ , lleva a  $e$  que ni es una consecuencia ni una hipótesis de  $p$ , le es ortogonal y al ser  $e' = b + d$  también ortogonal a  $p$ , es una especulación fuerte.

Por su parte,

- El camino mixto  $p > a < a + c + d = e$  lleva a una especulación débil puesto que valen  $p \diamond e$  y  $e' = b < p$ .

Nótese que el camino mixto  $p = a + b > a < a + b + c$  lleva, sin embargo, a una consecuencia y el  $p = a + b < a + b + d > b$  a una hipótesis. Cabe, por tanto, llegar a consecuencias, hipótesis y especulaciones por caminos mixtos. Si es obvio que en las álgebras de Boole finitas todas las conclusiones pueden encontrarse así, no lo es en cuanto la estructura es distinta.

Sin embargo, en el caso de una especulación débil,  $p \diamond e$  y  $e' < p$ , el modelo esqueleto permite observar lo que sigue: 1)  $e'$  se obtiene desde  $p$  mediante abducción y, negando  $e'$ , se obtiene  $e''$ ; 2) Si la negación es débil en  $e$ ,  $e < e''$ , entonces y de nuevo por abducción se llega a  $e$ ; 3) Si la negación es intuicionista en  $e$ ,  $e'' < e$ , entonces por deducción se llega a  $e$ . En el primer caso, se trata del camino:  $p > e' \rightarrow e'' > e$ , que es abductivo. En el segundo, el camino es mixto:  $p > e' \rightarrow e'' < e$ . Obviamente, si la negación es fuerte, el proceso se abrevia así:  $p > e' \rightarrow e'' \sim e$ ; pero, si la negación es salvaje en  $e$ , nada cabe concluir.

En el caso de una especulación fuerte lo anterior no es reproducible puesto que también es  $p \diamond e$ ; sólo resta la posibilidad de llegar desde  $p$  a  $e$  por otro camino mixto, como se ha mostrado en la anterior álgebra de Boole atómica.

Es, por tanto, razonable ver el razonamiento como la realización de esos caminos mixtos, una especie de *zigzag* browniano a partir de la premisa. Unos caminos que se siguen hasta llegar a una conclusión que se considere suficiente para aquello que se busca y que, muchas veces, obliga a parar y volver a empezar. Se trata de procesos de prueba y error hasta encontrar una solución satisfactoria y por más que ésta no reste garantizada al efectuarlos.

Tan propio del razonamiento ordinario es deducir y abducir, como lo es especular y por más que quepa considerar a la deducción como la principal operación intelectual, al permitir probar las conclusiones con la mayor seguridad posible en el razonamiento.

3.2. Una vez en posesión de una especulación  $e$  desde  $p$ , entonces (Trillas, 2018, 2019, en prensa).

1)  $p + e$ , es una consecuencia de  $p$ ; 2)  $p \cdot e$  es una hipótesis de  $p$ , supuesto que no es auto-contradictoria ni equivalente a  $p$ ; 3) Si  $<$  es transitiva localmente donde convenga y la negación es intuicionista o fuerte en  $p$ ,  $p' \cdot e$  es una refutación de  $p$ .

En efecto: La primera sigue de  $p < p + e$ : la segunda de  $p \cdot e < p$ , y la tercera de  $p' \cdot e < p' \Rightarrow p'' < (p' \cdot e)'$  que, al ser  $p < p''$  y por la transitividad en la terna  $(p, p'', (p' \cdot e)')$ , lleva hasta  $p < (p' \cdot e)'$ .

Por lo tanto, la especulación ayuda a deducir, abducir y refutar. Especular es parte importante en la reflexión sobre  $p$ , sobre repensar  $p$  y cuanto de  $p$  se conozca. Especular permite llegar a prever y explicar, hallar consecuencias e hipótesis; reflexionar sobre algo comporta, en buena medida, especular.

Como sea, además, que las especulaciones son las conjeturas que no son consecuencias, ni hipótesis, cabe identificar la especulación con la inducción que, entonces y de probarse universal la manera de llegar a ellas mediante un *zigzag* por caminos mixtos de inferencia, dejaría de ser el 'misterio' con el cual se la ha visto siempre.

De momento, sabemos que si la información sobre la premisa  $p$ , objeto de la reflexión, puede organizarse como un álgebra de Boole finita, especular no es sino tal *zigzag*. Aunque ello puede dejar de lado a muchas especulaciones, tampoco parece que cuando alguien reflexiona efectúe todas las inducciones posibles; de ordinario, sólo se llega a algunas de ellas, las que resulten de combinar por negación, conjunción y disyunción cuanto se recuerde. En cuanto a quienes se niegan a aceptar que la inducción exista realmente, lo que niegan es la especulación y cabe preguntarles si, con independencia de su verdad o falsedad, niegan la existencia de enunciados ortogonales a uno dado. En todo caso, es una visión con la cual inducir es tan parte del razonamiento como lo es deducir; el razonamiento no sería posible sin ambas operaciones intelectuales.

3.3. En cuanto a los grados de verdad, realmente sobre cuán verdadero es un enunciado, se definen como las aplicaciones que a cada enunciado  $u$  asignan un número entre 0 y 1,  $t(u) \in [0, 1]$ , de manera que la verdad no retroceda por deducción y cambie simétricamente por negación, es decir, sometido a verificar los dos axiomas:

1)  $p < q \Rightarrow t(p) \leq t(q)$ , y 2)  $t(p') = 1 - t(p)$ .

Está claro que una consecuencia nunca será menos verdadera que la premisa y que, de ser ésta completamente verdadera,  $t(p) = 1$ , de  $1 \leq t(q)$  seguirá  $t(q) = 1$ : Toda consecuencia de una premisa completamente verdadera es completamente verdadera. En cuanto a las hipótesis,  $h < p$ , de  $t(h) \leq t(p)$  sigue que nunca serán más verdaderas que la premisa y que, de ser ésta completamente verdadera, las hipótesis no tienen por qué serlo, pero de ser una hipótesis completamente verdadera también lo será la premisa y, de ser ésta completamente falsa, también lo será la hipótesis.

La definición aportada por los dos axiomas anteriores está, además, en consonancia con la ley de contraposición de la negación  $[p < q \Rightarrow q' < p']$ , puesto que si vale la segunda sigue  $t(q') = 1 - t(q) \leq t(p') = 1 - t(p) \Leftrightarrow t(p) \leq t(q)$ , como corresponde a la primera.

Nótese que siendo

$$p \cdot p' < (p \cdot p')', \text{ es } t(p \cdot p') \leq 1 - t(p \cdot p') \Leftrightarrow t(p \cdot p') \leq \frac{1}{2},$$

si  $t$  sólo toma los valores 0 y 1 y, en el caso de únicamente considerar enunciados completamente verdaderos o completamente falsos, necesariamente es  $t(p \cdot p') = 0$ : Los enunciados  $p \cdot p'$  son completamente falsos. Análogamente, de  $(p + p')' < p + p'$ , sigue  $\frac{1}{2} \leq t(p + p')$  que implica  $t(p + p') = 1$ : Los enunciados  $p + p'$  son completamente verdaderos. Los grados de verdad permiten un 'control numérico' de la validez de las conclusiones a las que se llega razonando.

3.4. En cuanto al grado de verdad de las especulaciones, nada cabe decir de las fuertes, pero las débiles y a causa de

$$e' < p \Rightarrow 1 - t(e) \leq t(p) \Leftrightarrow 1 - t(p) \leq t(e),$$

son verdaderas si la premisa es falsa y, de ser ésta verdadera, tanto pueden ser verdaderas como falsas. No pudiendo establecerse relación alguna entre  $t(p)$  y  $t(e)$  cuando  $e$  es una especulación fuerte, sólo éstas son 'salvajes' frente a la verdad.

En el caso de una refutación,  $p < r'$  implica  $t(p) \leq 1 - t(r)$ : Si la premisa es completamente verdadera, la refutación será completamente falsa y, de ser ésta verdadera, será falsa la premisa.

Nótese que de ser  $t(p) = 0.7$ , las consecuencias tendrían sus valores de verdad en el intervalo  $[0.7, 1]$ , las hipótesis en  $[0, 0.7]$ , las especulaciones débiles en  $[0.3, 1]$ , las refutaciones en  $[0, 0.3]$  y sin poder asegurar nada respecto de los valores de verdad de las especulaciones fuertes. Así, en los ejemplos del apartado 3.1 con  $p = a + b$  y las especulaciones  $a + c$  (fuerte) y  $a + c + d$  (débil), supuesto que los grados de verdad de los cuatro átomos fuesen  $t(a) = 0.6$ ,  $t(b) = 0.7$ ,  $t(c) = 0.3$  y  $t(d) = 0.9$ , con los cuales es  $t(p) = 0.7$ , se obtendrían los valores  $t(a + c + d) = 0.9 \in [0.3, 1]$  y  $t(a + c) = 0.6$  también en ese intervalo.

De llegar a la conclusión a través de una especulación  $e$  y como se ha visto,  $h = p \cdot e$ , o  $q = p + e$ , seguirían:

$$t(h) = \min(t(p), t(e)), \text{ y } t(q) = \max(t(p), t(e)),$$

de aceptar esas fórmulas. Con ello, si  $t(p) = 0.7$ , serían

$$t(h) = \min(0.7, t(e)), \text{ y } t(q) = \max(0.7, t(e));$$

entonces,

$$\text{- Si es } t(e) \leq 0.7, \text{ serían } t(h) = t(e) \text{ y } t(q) = 0.7,$$

en tanto que,

$$\text{- Si es } 0.7 < t(e), \text{ serían } t(h) = 0.7 \text{ y } t(q) = t(e).$$

Nótese que

$$\text{- De ser } t(h) = 1 \text{ nada puede decirse del valor } t(e),$$

pero



- Si fuese  $t(q) = 1$ , entonces sería  $t(e) = 1$ ;
- De ser  $t(e) = 0$ , serían  $t(h) = 0$  y  $t(q) = 0.7$ .

Además, si en el caso de una consecuencia  $q = p + e$ , fuese  $p \cdot e = 0$ , se tendría la ecuación  $\max(t(p), t(e)) = 0$ , obligando a que ambos grados de verdad fuesen nulos.

Tales fórmulas permiten un cierto control numérico del grado de verdad de la conclusión en función de los de la premisa y la especulación. Un grado de verdad que, usualmente, será empírico y, por tanto, más cercano al concepto popperiano de verosimilitud que a la idea lógica de verdad total.

3.5. Los dos axiomas con los que, en el apartado 3.3 se han definido los grados de verdad, se refieren a su relación con  $<$  y, por ello, pueden considerarse ‘*sintácticos*’, pero la verdad de un enunciado es un asunto semántico, tiene qué ver con su significado (Trillas, 2017, 2018, en prensa). Se supone que, en los enunciados con grado de verdad, éste se refiere a la adecuación del enunciado con la realidad, es decir, con lo que signifique el enunciado en relación con aquello a lo cual se refiera.

Por ello es por lo que se supone que cada enunciado considerado  $p$ , tiene significado medible, es decir, su significado está descrito por una magnitud escalar, o conjunto borroso (Trillas, 2018, 2020),  $(X, <_p, m_p)$ , donde  $<_p$  es la relación binaria y empírica ‘*menos p que*’ traduciendo el significado cualitativo, primario, de  $p$ , en tanto que  $m_p$  es una ‘*medida*’ en el intervalo  $[0, 1]$  de la misma (Trillas, 2020, en prensa). Con ello, una forma de calcular un grado de verdad  $t(p)$  es igualándolo al valor numérico, entre 0 y 1, que corresponda a la medida de  $p$ . Por ejemplo, el enunciado  $p$  en el universo del discurso  $[0, 10]$ , tal que  $p(x) = \text{‘}x \text{ es grande’}$  o ‘ $x$  no es grande y  $x$  es 1’, representado por la fórmula

medida de  $p(x) = \max(m_{\text{grande}}(x), \min(1 - m_{\text{grande}}(x), m_1(x))) = \max(x/10, \min(1 - x/10, 1)) = \max(x/10, 1 - x/10)$ ,

con lo cual la medida total o grado de verdad de  $p$ , puede definirse como el número

$$t(p) = \sup_{x \in [0, 10]} \max(x/10, 1 - x/10) = 1/2.$$

Así, en cierta forma y como se apuntó en el anterior apartado 3.4,  $t(p)$  puede asemejarse a cuán verosímil es  $p$  en el universo del discurso correspondiente; por lo menos y de acuerdo con la teoría de la posibilidad de Lotfi A. Zadeh (Zadeh, 1978), a su medida de posibilidad, de cuán posible es.

3.6. Una vez conocida una hipótesis  $h$  para  $p$ , la teoría que  $h$  permite desarrollar es el conjunto de sus consecuencias,  $h\uparrow = \{u; h < u\}$ , en tanto que la premisa permite desarrollar la teoría  $p\uparrow = \{q; p < q\}$ . Ambas teorías están relacionadas, de ser  $<$  transitiva, por medio de  $h < p, p < q \Rightarrow h < q$ , es decir,  $p\uparrow \subseteq h\uparrow$ . La hipótesis genera una teoría más amplia que la premisa, va más allá de ella; lo que se deduce de la hipótesis, de explicar los hechos si se quiere, no siempre se sigue de la premisa, de los hechos. La hipótesis introduce más conceptos de los que se siguen de los hechos.

Por otra parte, si  $r$  refuta a la hipótesis,  $h < r$ , no puede concluirse que, necesariamente, refute a la premisa, pero si  $r$  refuta a ésta y  $<$  es transitiva, entonces también refuta a la hipótesis:  $p < r$  y  $h < p \Rightarrow h < r$ ; la mejor refutación es la que afecta a los hechos y, por lo tanto, de aceptar que la refutación de cualquier  $u$  obliga a rechazar su teoría  $u\uparrow$ , la refutación de  $h$  obligaría a rechazar  $h\uparrow$  y también  $p\uparrow$ , pero la de  $p$  sólo obligaría a rechazar la de  $p\uparrow$  y restaría la puerta abierta a conservar cuanto en  $h\uparrow$  no esté en  $p\uparrow$ , es decir,  $h\uparrow - p\uparrow$ : La refutación de una premisa no comporta necesariamente el rechazo de toda la teoría generada por una de sus hipótesis. Cuanto es contra-factual, contra los hechos, no elimina cuanto pueda seguirse de una hipótesis al respecto.

Todo lo anterior permite pensar que el control del razonamiento basado en hipótesis sobre hechos puede alcanzar un cierto nivel de ‘control’ formal. Sin embargo, el control formal del razonamiento ordinario sólo es posible cuando éste se explicita simbólicamente y en alguna estructura con el substrato del

‘esqueleto’; por ello, también debe pensarse en el control más ‘ingenuo’ de la crítica a la que cada pensador puede someterlo al repensarlo sin antes someterlo a formalización alguna.

No hay más racionalidad que la que surge de repensar críticamente lo antes pensado o razonado; repensar el común razonamiento ordinario cómo se pueda (Trillas, en prensa). Por eso y en buena medida, *todo sirve*, como Paul K. Feyerabend señaló en su (anarquista) tratado contra el método (Feyerabend, 1986). De existir un tal método no podría ir más allá del razonamiento ordinario y debería ser una organización situacional resultante del zigzag de la inferencia adecuada a cada caso.

## El falsacionismo

La doctrina filosófica llamada falsacionismo, se refiere a rechazar la teoría generada por una hipótesis  $h$  sobre unos hechos, expresados por una premisa  $p$  que puede ser la conjunción de enunciados recogiendo los aspectos relevantes de aquellos hechos. De lo acabado de exponer, se sigue que refutar  $p$  no obliga a rechazar cuanto se deduzca de  $h$ , sólo la refutación de  $h$  lleva a ello. Refutar la hipótesis aparece, así, más fuerte que refutar algún hecho. Refutar una posible explicación es ‘más fuerte’ que encontrar aspectos contra-factuales.

De otra parte, tanto  $h$  como una de sus refutaciones  $r$ , pueden constar de varias partes cuyos grados de verdad hagan que rechazar totalmente ambas pueda ser excesivo.

Si, por ejemplo, es  $h = u + v$ , de  $h = u + v < r'$ , al ser  $u < u + v$  y  $v < u + v$ , de ser  $<$  transitiva se siguen  $u < r'$  y  $v < r'$ . Es decir, una refutación de  $h$  refuta todas sus partes, pero sin que la refutación de una de ellas implique la de  $h$ :  $u < r'$  no implica  $u + v < r'$ . Sin embargo, de ser  $t(h) = 0.7 = \max(t(u), t(v))$  y si fuese  $t(v) = 0.3$  resultaría  $t(u) = 0.7$ , lo que podría llevar a sólo considerar como completamente refutada  $u$  y a ‘olvidar’  $v$ . Ello podría dejar ‘libre’ aquello de  $h$  que estuviese afectado por  $v$ .

Análogamente, de ser  $r = s + x$  una refutación de  $h$ ,  $h < (s + x)$ , de  $s < s + x$  y  $x < s + x$ , siguen  $(s + x) < s'$  y  $(s + x) < x'$ ; por tanto,  $h < s'$  y  $h < x'$ , cada parte de la refutación también refuta  $h$ . De nuevo,  $t(r) = \max(t(s), t(x))$  con, por ejemplo,  $t(r) = 0.7$  y de ser  $t(x) = 0.3$ , llevaría a  $t(s) = 0.7$  y, tal vez, a sólo considerar como refutación a  $s$  y dejar de lado, libre, lo de  $p$  afectado por  $x$ .

Es obvio que, de ser todos los valores de verdad 0 o 1, la cosa es distinta. En efecto, si  $t(h) = 0$ , entonces es  $t(u) = t(v) = 0$  y, si  $t(h) = 1$ , entonces es  $t(u) = 1$  o  $t(v) = 1$  con la posibilidad de que sean  $t(u) = 1$  y  $t(v) = 0$ . Análogamente, si  $t(r) = 0$ , es  $t(s) = t(x) = 0$ , pero si  $t(r) = 1$  por lo menos uno de los números  $t(s), t(x)$  será 1.

Obsérvese que, de ser las dos partes contradictorias,  $s < x'$ , de  $t(s) \leq 1 - t(x)$ ,  $t(x) = 0.4$  implicaría  $t(s) \leq 0.6$ , no podría llegar al valor 0.7, valor al cual tampoco podría llegar  $r$ . De estar en un álgebra de Boole, donde es  $s < x' \Leftrightarrow s \cdot x = 0$ , de  $t(s \cdot x) = \min(t(s), t(x)) = 0$ , resultaría  $t(s) = 0$  y esa parte no contribuiría en nada a la verdad de  $r$ .

Resumiendo, siempre es posible efectuar un *control* de la verdad de las hipótesis o las refutaciones a partir de las partes que las constituyan.

En conclusión, en una ciencia formal, como es el caso de las matemáticas, donde los enunciados sólo se aceptan si son completamente verdaderos, la refutación es un criterio suficiente de rechazo. No obstante, en las ciencias no formales donde los enunciados pueden ser parcialmente verdaderos y parcialmente falsos, el rechazo de una teoría empírica por la refutación de una de sus hipótesis debería ser ‘tomada con pinzas’. De hecho, tanto las hipótesis como las refutaciones no suelen ser enunciados ‘monolíticos’, sino que suelen constar de diversas partes con distinta influencia en la hipótesis o la refutación; el análisis de ambas en cada caso es relevante.

Por ejemplo, detectar la llamada ‘anomalía del perihelio’ del planeta Mercurio no llevó a refutar la teoría, la hipótesis básica de Newton, sino que, más adelante, sirvió para encontrar el planeta Neptuno cuya existencia fue predicha por una parte de aquella hipótesis que explica la anomalía, la cual provenía de aceptar la teoría newtoniana.

Otro ejemplo lo ofrece el ‘perceptrón’ o neurona artificial, introducido para el aprendizaje automático y que, rechazado a causa de la refutación implicada por su imposibilidad de representar una importante operación lógica (la llamada ‘Barra de Sheffer’ o ‘Nand’,  $p \mid q$ , ni  $p$  ni  $q$ ), luego y aumentando el número de esas neuronas artificiales, conectándolas y distribuyéndolas en varias capas, se llegó a la teoría de las redes neuronales cuya eficacia en el aprendizaje automático es manifiesta en la Inteligencia Artificial. Si una neurona artificial aislada, un perceptrón aislado, no podía hacer gran cosa, una ‘sociedad’ de neuronas artificiales organizada de cierta forma, permite muchísimo más. Fue una refutación que llevó a repensar todo el problema; fue una refutación fértil.

## Conclusión

Decía Nietzsche que siempre hay que dudar de la forma más meticulosa. Repensar los razonamientos es una necesidad permanente; poco, si algo, de lo que alcanza la razón humana es definitivo, casi todo tiene una ‘fecha de caducidad’ y por más sea desconocida (Trillas, en prensa). Una duda que no debe ser ‘contemplativa’ sino ‘activa’, buscando formas de mejorar aquello que se considere sabido, sea a través de más datos de observación o experimentación, sea a través de posibles contradicciones con otros conocimientos, sea reconsiderando el problema correspondiente.

Siempre hay que intentar encontrar un ‘punto flaco’, un detalle que permita, poniendo atención en él, ver algún aspecto que hasta entonces había pasado desapercibido. Como lo fue para el primer autor darse cuenta que la transitividad permitía probar, como William Whewell ya había enunciado sin demostrarlo, que las consecuencias debían ser conjeturas; algo que luego le llevó hasta la primera prueba formal de los antiguos principios aristotélicos de No Contradicción y Tercero Excluido, a las que llegó tras probarlos, sucesivamente, en las álgebras de Boole, las de De Morgan y las llamadas ‘estándar’ de conjuntos borrosos (Trillas, 2017).

Fue precisamente la observación de las poquísimas propiedades que esas pruebas requerían, y la atención a ello prestada, lo que le sugirió especular acerca de una posible estructura del ‘esqueleto’ que, de acuerdo con la regla metodológica llamada la ‘Navaja’ de William of Ockham (siglo XIV), resultó en un mínimo de hipótesis independientes entre sí y que permiten un cálculo incipiente con el cual y no obstante ya cabe, por ejemplo, probar aquellos principios una vez añadida la transitividad local. Es un añadido, como para obtener otros resultados lo es el carácter de la negación o la existencia de mínimo o máximo de inferencia y que, conforme a la Adenda de Karl Menger (siglo XX) a la Navaja de Ockham, permite enunciarla así:

No sólo no trabajar con más hipótesis de las estrictamente necesarias para establecer algo, sino tampoco con menos de las estrictamente suficientes para su continuación.

Es una sabia regla metodológica que Albert Einstein enunció, en libre paráfrasis, como ‘Hacer las cosas tan simples como sea posible, pero no más simples que lo necesario’.

Es una regla que, además, autoriza ver aquellas leyes que no figuran en el esqueleto formal como no universales sino locales en el razonamiento ordinario; son leyes que aparecen en determinadas situaciones que las permiten. Situaciones en las cuales, con esas leyes, el razonamiento correspondiente se traduce por un cálculo más específico y como, desde el alemán Wilhelm Leibniz (siglo XVIII) y, tal vez, desde el mallorquín Ramon Llull (siglo XIII), en quien se inspiró el primero, se ve a la Lógica tras el famoso *No discutamos, ¡Calculemos!* de Leibniz.

Con respecto a la insistencia popperiana sobre el razonamiento deductivo y como ya se ha dicho, éste no basta para obtener conocimiento, no es suficiente. A tal respecto, merece la pena tener en cuenta, como las de una ‘autoridad’, las palabras del Premio Nobel de Biología, Patrick B. Medawar, en su libro (Medawar, 1984) sobre los límites de la ciencia:

Ningún proceso de razonamiento lógico permite ampliar el contenido informacional de los axiomas, premisas o enunciados observacionales de los que se parte.

Es decir, para ello se requiere más que deducir y abducir, por tanto, se necesita conjeturar especulaciones, inducir para mejor captar aquello que los enunciados iniciales, sean premisas reflejando hechos observados o sean hipótesis explicativas, puedan indicar. En el razonamiento conducente al conocimiento científico, como a cualquier tipo de razonamiento, deducir es insuficiente; usualmente se acaba en conjeturas con una inferencia desde la premisa que no es sino un zigzag que no permite establecer su grado de verdad dependiente del de aquella. Por ello y siempre, debe intentarse, por seguridad, controlar el razonamiento, cualquiera que éste sea.

Tal control es una operación intelectual que forma parte de repensar y que tanto puede llevar a un rechazo total, a la completa falsación, de aquello bajo reconsideración, como sólo a una modificación parcial de ello e, incluso, a dejarlo tal cual estaba y hasta que nuevas observaciones o un nuevo repensar, tal vez a la luz de otros conocimientos, vuelvan a llevar a ello; unas veces para acabar restringiéndolo y otras ampliándolo. Es un camino que sólo en el caso extremo de rechazo total lleva al falsacionismo popperiano; un camino que, recorrido por los científicos gracias al razonamiento ordinario refinado como sea, no les aparta del razonamiento del común de los mortales más que en determinados ‘tecnicismos’ y, por ello mismo, no es insensato el antes citado *todo sirve* de Paul K. Feyerabend.

Con San Francisco de Asís, *Un rayo de sol ahuyenta muchas sombras*. Repensar, que es realmente la base de la racionalidad, ofrece ese rayo de luz.

Por lo tanto y antes de un pretendido criterio único de falsificación, aparece el *principio* que aconseja repensar críticamente especulando para controlar cuanto se cree conocido. Algo que, cómo poco, será una muestra de humildad intelectual y que a un pensador nunca le sobraré; algo que marginará todo ‘autoritarismo’ como aquel al que puede asemejarse el falsacionismo absoluto.

De ahí el título de este artículo que no pretende ser sino un humilde alegato a favor del control del razonamiento a través de repensarlo de manera crítica, un alegato que, a su vez, debería ser repensado, criticado, controlado.

Un alegato que, en realidad, es una defensa del razonamiento ordinario como el camino que, con su crítica, el pensamiento ofrece a las personas para cuanto logran. Un razonamiento, el ordinario, que es al fin cuanto comparten intelectualmente quienes, por ejemplo, se sitúan en posiciones muy apartadas entre sí y que, por ello y por cuánto a partir del mismo se ha conseguido, debería declararse ‘patrimonio de la humanidad’ con independencia de cualquier posición ideológica. Ni Popper, ni Miller, ni Medawar, ni los autores de este artículo, podrían haber escrito lo que han escrito sin, por lo menos, especular previamente, inducir, sobre ello. Como cualquier hijo de vecino piensa con un objetivo y poseyendo una formación previa adecuada.

Un patrimonio universal que sugiere, de nuevo con Feyerabend (Feyerabend, 1986), pensar seriamente en la ‘separación entre la ciencia y el Estado’; en lo que se refiere al razonamiento. ¿Es que, por lo que se refiere a razonar, hay alguna diferencia esencial entre un científico y cualquier otro ciudadano, o entre éste y un filósofo?

## Referencias

Birkhoff, G. (1948). “Lattice Theory”. *American Mathematical Society Colloquium Publications*. 25.

- Feyerabend, P. K. (1976). *Tratado contra el método*. Tecnos. 1986.
- Medawar, P. B. (1984). *The Limits of Science*. Harper & Row.
- Miller, D. (1994). *Critical Rationalism. A Restatement and Defense*. Open Court.
- Popper, K. R. (1934). *La lógica de la investigación científica*. Tecnos. 1973.
- Popper, K. R. (1963). *Conjeturas y refutaciones: el crecimiento del conocimiento científico*. Paidós. 1983.
- Trillas, E. (2017). *On the Logos: A Naïve View on Ordinary Reasoning and Fuzzy Logic*. Springer.
- Trillas, E. (2018). *El desafío de la creatividad*. Publicación da USC.
- Trillas, E. (2020). *Narrar, conjeturar y computar: El pensamiento*. EUGR (en prensa).
- Trillas, E. (en prensa). *Reflexiones sobre la lógica y su génesis*. (manuscrito aún no publicado)
- Zadeh, L. A. (1978). "Fuzzy Sets as a Basis for a Theory of Possibility". *Fuzzy Sets and Systems*. 1/1. 3-28.



## UNA APROXIMACIÓN AL PROBLEMA DEL MAL EN SCHELLING

*AN APPROACH TO THE PROBLEM OF EVIL IN SCHELLING*

*UMA ABORDAGEM AO PROBLEMA DO MAL EM SCHELLING*

**Leopoldo Tillería**

(Universidad Tecnológica de Chile  
Instituto Nacional de Capacitación Profesional)  
[leopoldo.tilleria@inacapmail.cl](mailto:leopoldo.tilleria@inacapmail.cl)

Recibido: 26/07/2020

Aprobado: 05/02/2021

### RESUMEN

Se discute la idea del mal en la filosofía tardía de Schelling, partiendo de la comprensión del problema de Dios como fundamento último del mal en el hombre. Tal conjetura, a la que se llega mediante una exégesis filosófica, pone foco en la idea de la “caída” y en las implicancias que en ella tiene el principio de libertad. En confrontación con el concepto de una filosofía de la revelación, se sugiere que la filosofía de la religión en Schelling concibe a un Dios que parece representar el propio mal original. Teniendo en cuenta la interpretación heideggeriana, se afirma, por un lado, que el fundamento origina el mal en el hombre como expansión de la voluntad propia que tiende hacia sí mismo, y, por otro, que esta propia movilidad del fundamento es originada, a su vez, por la necesidad metafísica de Dios de patentizar su omnipotencia. De este modo, la filosofía de la revelación de Schelling formaría parte de las llamadas filosofías analíticas de la religión, en las que se recupera la vocación filosófica central de las teodiceas. Se concluye con la idea de que el mal es la búsqueda y realización originaria del fundamento de Dios.

Palabras clave: filosofía. libertad. mal. religión. Schelling.

### ABSTRACT

The idea of evil is discussed in Schelling's late philosophy, starting from the understanding of the problem of God as the ultimate foundation of evil in man. Such conjecture, which is reached through a philosophical exegesis, focuses on the idea of the “fall” and the implications of the principle of freedom. Once confronted with the concept of a philosophy of revelation, it is suggested that the philosophy of religion in Schelling conceives of a God who appears to represent the original evil itself. From the Heideggerian interpretation perspective, it is stated, on the one hand, that the foundation originates evil in man as an expansion of his own will that tends towards himself, and, on the other, that this same mobility of the foundation is originated, in turn, by God's metaphysical need to patent his omnipotence. In this way, Schelling's philosophy of revelation would be part of the so-called analytical philosophies of religion, in which the central philosophical vocation of the theodicy is recovered. It concludes with the idea that evil is the original search and realization of the foundation of God.

Keywords: philosophy. freedom. evil. religion. Schelling.

## RESUMO

A ideia do mal é discutida na filosofia tardia de Schelling a partir do entendimento do problema de Deus como o fundamento último do mal no Homem. Essa conjectura, alcançada através de exegese filosófica, concentra-se na ideia da “queda” e nas implicações que o princípio da liberdade tem nela. Uma vez confrontado com o conceito de filosofia da revelação, sugere-se que a filosofia da religião em Schelling conceba um Deus que parece representar o próprio mal original. A partir de uma interpretação heideggeriana, afirma-se, por um lado, que o fundamento origina o mal no Homem como uma expansão de sua própria vontade, que tende para si mesmo e, por outro, que essa própria mobilidade do fundamento é originária novamente, por causa da necessidade metafísica de Deus de patentear sua onipotência. Desse modo, a filosofia da revelação de Schelling seria parte das chamadas filosofias analíticas da religião, nas quais a vocação filosófica central da teodicéia é recuperada. O artigo conclui com a ideia de que o mal é a busca e realização original do fundamento de Deus.

Palavras-chave: filosofia. liberdade. mal. religião. Schelling.

## Introducción

Este artículo trata de la filosofía de lo Absoluto en F.W.J. Schelling, concepto que intentaré hacer girar sobre la última sección del artículo al de una filosofía de la religión. En todo caso, y como es sabido de sobra, el pensamiento de Schelling es en realidad el de una filosofía idealista. Forzado a catalogar al Schelling de la libertad y del mal —en otras palabras, al Schelling de Dios—, seguiría la opinión de Fuhrmans (1965), para quien las “Investigaciones filosóficas sobre la esencia de la libertad humana y los objetos con ella relacionados” constituyen una vuelta, una ruptura, que ofrece ahora la otra cara del sistema filosófico, la cara de la noche y la oscuridad, que viene representada en la literatura por el romanticismo tardío (Leyte & Rühle en Schelling, 1989). En su intento por dar respuestas a una compleja noción de Dios a partir de la sistematicidad de la filosofía, y por no caer en proposiciones indemostrables provenientes de la fe, Schelling termina exponiendo, como dije, una especie de genealogía de lo Absoluto. Las *Investigaciones* incorporarán, a través de la noción de “caída” desde lo absoluto, una fractura definitiva en lo que hasta ahora era una sistemática filosofía de la identidad. En nuestro filósofo, la idea del mal surge necesariamente a partir de la ontología de Dios, a tal punto que puede decirse, siguiendo la idea de Dios inicial, que en su filosofía de lo Absoluto Dios y mal parecen coincidir.

Propondré dos conjeturas iniciales sobre la filosofía del “último” Schelling: 1) la identidad entre Ser y ser del mal, y 2) la necesidad metafísica del mal en la existencia del hombre en tanto ente. Ambas, en un mutuo condicionamiento, operarían como condiciones de posibilidad de la conclusión de nuestro escrito, que sugiere que en la filosofía tardía de Schelling, el mal consistiría en la búsqueda y en la realización originaria del Dios inicial, es decir, correspondería a una determinación fundamental de lo Absoluto. Antes de presentar cualquier argumentación, quisiera referirme al momento específico en que surge esta apelación, por así decir, teológicamente imprescindible acerca del origen y manifestación del mal. Dicho problema, si seguimos la orientación cronológica de Schelling, se sitúa de lleno en el periodo de su filosofía de la libertad, esto es, alrededor de 1809. Es por esta época que Schelling, al concebir la esencia del hombre en cuanto ser libre, rechaza el concepto de un Dios como principio eterno, bueno y único, y establece, al mismo tiempo, la idea de una libertad para el bien y para el mal. La condición metafísica del mal habrá que buscarla justamente en su filosofía de la libertad. La pregunta que busca responder Schelling, y que probablemente sea la pregunta que lo lleva de manera definitiva a su última etapa —a la llamada filosofía del ser—, parece ser, sin más, la pregunta por la naturaleza del mal.

El artículo se organiza como sigue. Primero, se aborda la tensión en el pensamiento de Schelling entre las ideas de Ser y de mal, a propósito del modo en el que se encuentra el hombre en el sistema schellingiano. Enseguida, se trata el problema de Dios como fundamento último del mal en el hombre, con foco en el asunto de la “caída” y en las implicancias que en ella tiene el principio de libertad.

Después, se discute la idea de revelación de Dios, y la relación que habría entre mal y libertad a partir de una presunta naturaleza inicial. En cuarto lugar, y a modo, pudiera decirse, de cierre “programático”, se plantea la pregunta en torno a si en el Schelling del mal es posible hablar de una filosofía de la religión, o si, más bien, debiésemos referirnos a su sistema como una filosofía de la revelación. Finalmente, se sugerirá cómo esta filosofía de la religión nos mostraría a un Dios que parece representar el propio mal original, y en qué sentido este mal se revela como la búsqueda y realización originaria del fundamento de Dios.

## **La identidad entre Ser y ser del mal**

En el joven Schelling, el principio de la identidad funciona como principio absoluto de todo el sistema del idealismo trascendental. Esto es, como una suerte de pensamiento organizado que explica el desarrollo de la naturaleza a partir de un idealismo de corte fichteano. Cada una de las partes del sistema contiene este principio de identidad. Lo Absoluto, el Ser o Dios (en otros momentos Schelling habla también del Uno indiferenciado), se constituye en cuanto identidad como un punto de indiferencia o como síntesis superior en la contención de sujeto y objeto. El Ser, las cosas (la naturaleza, los entes) y el hombre, ciertamente están determinados por la esencia del Ser; no obstante, solo aquel es un Ser-infinito. Se trata, y esto es muy importante de aclarar, de tres entes distintos: el primero, el ente en total, y los siguientes, entes finitos, que, sin embargo, “son” en el ente en total. Es decir, son en Dios, como determinación particular de esta filosofía de la naturaleza mal llamada panteísmo. El mismo Schelling cree que no hay razones para hablar en su sistema de panteísmo, pues, y, ante todo, en su filosofía de la naturaleza se afirma la absoluta no realidad de todos los fenómenos, es decir, el mundo fenoménico como no-realidad absoluta, y lo finito como nada (Del Barco, 1999: 102). Así, la totalidad es lo propio del modo de ser del Ser. Esto es lo que podemos llamar, con Heidegger, el sistema mismo como sistema de la libertad. La esperable contradicción entre necesidad y libertad, la soluciona Schelling retro trayendo su sistema al ente divino:

En lo referente a esta relación se puede afirmar que Dios se caracteriza, en primer lugar, por la representación que tiene del sistema, y en segundo lugar por ser el ente libre por antonomasia; en su inmanencia no encontramos una colisión entre libertad y sistema. De este modo, llegamos a Dios como idea directriz del sistema, y con ello Schelling avanza en la realización de una teosofía o teología racional (Vargas, 2005: 135).

En este sentido, sujeto y objeto se entenderán como contrarios unidos en el principio de identidad. Es decir, y esto es de suma importancia en esta filosofía del devenir (lo Absoluto en Schelling es el propio devenir), la relación del Ser con las cosas supone no una mismidad en sí, sino más bien una mismidad que las expresa uniéndolas, desde luego, pero sin dejar de hacerlas diferentes. Es lo que ocurre en esta ontología circular entre el Ser y el ente: el ente (o sea, una naturaleza que incluye al hombre) debe presuponerse en la esencia del Ser. Es decir: el hombre debe presuponerse necesariamente al interior de la vida de Dios. Quisiera poner acento en esto, en el sentido de que este bloqueo ontológico de la idea de mismidad como identidad, implica, simultáneamente, el reposicionamiento fundamental de la idea de libertad, la que desde ahora ocupará en Schelling el centro del todo del Ser.

Ahora bien, en Schelling la diferencia entre Ser y ente constituye, paradójicamente, la identidad del ente en total y el ente en particular. El ser del ente co-responde a la inteligibilidad ideal existente que lo determina en cuanto ente (Heidegger, 1997). Mas, de partida debe dejarse en claro que la identidad del Ser y el ente no es sino una unidad que está atravesada por el Uno todo y ligada por lazos de esencialidad. Si el Ser y el ente no son lo mismo, no obstante, el ser es siempre el ser de un ente (Heidegger, 1997), y comparten, por tanto, la misma esencia (el querer como modo de ser del ser originario), puede establecerse el siguiente predicado: el Ser es el fundamento del ente (o sea, el ente es lo fundado del Ser en tanto fundante). Por ello, al pensar el ente en total, pensamos el ente en su integridad, como siendo un ente, y lo pensamos ya desde el ser (Heidegger, 1997). Ahora, si una metafísica del mal es la fundamentación de la pregunta por el Ser (Heidegger, 1985), cabe preguntarse: ¿cuál es en Schelling la esencia del Ser en la propia estructura del mal? Esta apelación aclarará si la relación entre el ser del mal

y Dios corresponde efectivamente a una relación determinante entre mismidad y libertad. A su vez, ya el principio de identidad (como unidad y recursividad de contrarios) ha establecido que cada ser solo puede hacerse patente en su contrario. Por eso para Heidegger (1985), la identidad significa la copertencia originaria de lo diverso en lo uno, el cual es en ese caso, a la vez, el fundamento de la posibilidad de lo diverso. Pregunta el filósofo del idealismo en su tratado sobre la libertad:

¿Acaso queda otra salida en contra de esta argumentación más que salvar a los hombres con su libertad en el propio ser divino, puesto que ésta es impensable en oposición con el poder absoluto y decir que el hombre no está fuera de Dios, sino en Dios, y que su propia actividad pertenece a la vida misma de Dios? (Schelling, 1989: 119).

De aquí que Dios se haga patente en el hombre (su contrario) o, lo que es lo mismo, sea la forma magnificada del hombre. Así, la libertad pasa a ser algo incondicionado, y en tanto libertad del hombre, algo infinito (Heidegger, 1985: 86). En consecuencia, siguiendo al filósofo de Messkirch, la pregunta que yace en el concepto de la libertad humana es en realidad la pregunta por una incondicionalidad finita. De este modo, el sistema del idealismo puede denominarse “sistema de la libertad”, porque el principio de la formación de sistema, el fundamento que determina el ensamble fundamental del Ser, o sea, la “idea”, es entendido precisamente como libertad (Heidegger, 1985: 114-115). Ser absoluto, ‘Ser’-en-sí del ente, quiere decir ‘Ser’-libre, y ‘Ser’-libre significa determinarse a sí mismo a partir del propio ser (Heidegger, 1985). En lo que sigue, intentaré explicar, lo más razonablemente posible, la naturaleza del mal en esta filosofía mitad idealismo mitad realismo, a la que, como insinué, Schelling no duda en llamar “una verdadera filosofía de la naturaleza”. Nos damos cuenta, entonces, que en Schelling la libertad como sistema únicamente es posible sobre la base de un concepto idealista del Ser, es decir, de la idea de Dios.

De esta laya, la filosofía positiva tendrá por tema la libertad, precisamente aquella libertad que puede comenzar sin necesidad de suposición. La filosofía de la revelación es, para Schelling, filosofía positiva (Jacobs, 2018: 111).

### **Dios, fundamento último del mal en el hombre**

Mas como el hombre, a diferencia de Dios, es un ser finito, los principios que constituyen la “ensambladura ontológica” (fundamento y existencia) a partir del primer acto del Uno-originario (o infundamento), están en él separados. Solamente Dios contiene la unidad inseparable de aquellos principios en un perfecto equilibrio. Para Schelling, Dios es devenir porque la relación recíproca entre fundamento y existencia es eterna. Es decir: si Dios es el “ente más ente” de todos, tiene que presentarse en Él el devenir más difícil y más grande (Heidegger, 1985: 133). El politeísmo, la fragmentación tripartita de la originaria unicidad de lo divino, supondrá, pues, una suerte de “vaciamiento” del Absoluto, que implica inevitablemente su ingreso en el flujo de la temporalidad, su participación en el proceso general del devenir (Llorente, 2016).

Pero, ¿qué entiende Schelling por fundamento? ¿Es el fundamento de Dios? ¿Es un segundo y más primitivo Dios? ¿Es acaso el fundamento del mal y, por tanto, fundamento del hombre? Ese fundamento de su existencia, piensa el filósofo, que Dios tiene en sí mismo, no es Dios considerado de un modo absoluto (en cuanto que existe), sino solo en cuanto lo que constituye el fundamento de su existencia, la naturaleza en Dios, un ser inseparable de Él, sin embargo, distinto de Él (Schelling, 1989: 163). Según esto, el fundamento en Dios sería algo así como una función engendradora específica, que, sin embargo, no por ello debe comprenderse ni como una dimensión anterior a Él, ni como una “prioridad esencial” a su condición de existencia. Como Dios es su propia actividad, no necesita sustrato; la actividad es su propio sustrato (García, 2016). El fundamento en Dios es aquello que, en Dios mismo, dirá Schelling, no “es” Él mismo. Dicho fundamento, es decir, lo que está en Dios, pero que no es Dios mismo, es presentado como el fundamento del mal en cuanto manifestación de su propia efectividad:

Ese Fundamento de su existencia que Dios tiene en sí no es Dios considerado absolutamente, esto es, en tanto que existe, pues es sólo el Fundamento de su existencia; Él es la naturaleza –en Dios; Una esencia

(*Wesen*) inseparable (*unabtrennliches*) de él, pero, sin embargo, distinta (*unterschiednes*) de él (Schelling, 1989: 166).

El Dios libre de Schelling tiene que ser pensado como ser capaz de comenzar; es decir, no solo tiene que poder iniciar algo, tiene que poder Él mismo comenzar, o, dicho de otra forma, tiene que “querer” ser Dios mismo (Jacobs, 2018). Surge aquí –dialécticamente– la necesidad de concebir la identidad y co-pertenencia de ambas formas de comprender la existencia del mal: como posibilidad y como realidad en el hombre. Felix Duque consigna especialmente la dificultad de Schelling en postular esta escisión de Dios:

De hecho, esa prodigiosa tragicomedia divina comenzó con la huida de “aquel” que en el futuro debía ser Dios respecto de su propio ser: ciego, caótico, impensable de antemano; o en la terminología schellingiana: empezó con la ruptura de la inerte confusión e indiferente identidad de «lo que puede ser» (*des Seinkönnenden*) y del «puro ente» (*des reinen Seienden*): una identidad, evidentemente, que sólo cabe conjeturar: una suerte de pasado *absoluto* que nadie (ni siquiera Dios, aún no *existente*) ha podido contemplar, pero que Schelling se ve obligado a *postular* (Duque, 2014: 50).

Por un lado, el fundamento origina el mal en el hombre como expansión de la voluntad propia que tiende hacia lo oscuro, hacia sí mismo; y, por otro, esta propia movilidad del fundamento (que quiere separarse, sin lograrlo, de la existencia de Dios mismo) es originada, a su vez, por la necesidad metafísica del Absoluto de patentizar su omnipotencia. En otras palabras, por efectividad del devenir de Dios. ¿De qué otro modo, si seguimos a Schelling, la voluntad de amor, que determina la efectividad de la existencia de Dios, podría manifestarse como tal, vale decir, como puro amor, si no es mediante la actividad de la “voluntad del fundamento”, especie de “voluntad de confrontación” del mismo espíritu del bien, y con ello, de un movimiento de propiciación de una voluntad de revelación? Sin embargo, debe hacerse acá una aclaración, como quien dice, más metafísica, para comprender esta perturbadora identidad entre Dios y mal. Dicho sintéticamente: el mal es una posibilidad en el hombre, no en Dios. Y lo es justamente en la medida en que el hombre, haciendo uso del principio de libertad, tiende a usar su condición espiritual más allá de su condición de creatura. Es decir: en contra de la unidad con la voluntad universal engendrada por Dios. El mal, entonces, procederá de la naturaleza ideal de la creatura, hasta donde esta depende de las verdades eternas contendidas en el entendimiento divino, pero no de la voluntad de Dios. De manera que la región de las verdades eternas será la causa ideal del mal y del bien, y deberá ocupar el puesto de la materia de los antiguos (Schelling, 1989: 187).

En efecto, como en el hombre el fundamento y la existencia se unen en su separabilidad (pues de otra forma no podrían existir ni el bien ni el mal como esencias del ente), en el mal la oscuridad y la luz se desvinculan, y constituyen en el hombre esta problemática presencia del no-ser. Cuando ambos principios entran en discordia, “otro espíritu” se desliza en el lugar que debería ocupar Dios, no obstante, siempre “quiere” ser, y en cuanto que no siendo Él mismo, toma entonces prestada su apariencia del ser verdadero (Schelling, 1989: 241). Esta indicación es de suyo importante –yo diría, trascendental–, pues lo que muestra es que el mal llega a la conciencia del hombre solo cuando opera la transformación del principio oscuro. El fundamento, dice Schelling, origina el mal en la medida en que aquel propende a la expansión de la voluntad propia y a su enseñorearse sobre la voluntad general (Dios existente). Pero este despliegue de la escisión de fundamento y existencia, que a su vez nos remite indefectiblemente al concepto de “caída”, es al mismo tiempo originado por el propio Absoluto de manera esencial. Este desplazamiento implica que el fundamento es abandonado en cierto modo a sí mismo, a fin de que actúe precisamente como fundamento. Pero esto es solo una consecuencia de lo Absoluto, pues el imperio del “amor” tiene que aceptar a la voluntad del fundamento si no quiere aniquilarse a sí mismo (Heidegger, 1985: 184).

El mal parece ser la voluntad propia queriendo ser la voluntad universal, como modo del ‘Ser’-libre del hombre. Si el mal es el querer ir hacia sí mismo del fundamento, y si ser capaz del mal es, paradójicamente, la esencia del hombre en cuanto ente, entonces, el hombre parece estar destinado necesariamente en dirección al fundamento. Dicho de otro modo: en el hombre se revela, al mismo tiempo que la efectividad del principio de libertad, una especialísima y necesaria condición de separabilidad de los principios fundamentales, del bien y el mal, tan esencial a su naturaleza, que sin



esta separabilidad la unidad indisoluble de ambos principios sería la misma como en Dios. El hombre, entonces, tiene dentro de sí la fuente de su automovimiento hacia el bien y hacia el mal en la misma medida. Es decir, la conexión de los principios en él no es necesaria, sino libre (Schelling, 1989: 201-202).

Como queda claro en el escrito de Schelling sobre la libertad, esta embestida maligna del hombre, como creatura natural, se produce justamente a raíz de la separabilidad de los principios fundamentales, parafraseando a Heidegger, en su facticidad. Tal separabilidad, explicaría *in nuce* la posibilidad de manifestación del mal a partir del movimiento del propio fundamento, en una escena a la que Schelling denomina, muy teogónicamente, “primera creación”. Desde luego, la imagen del “ángel de luz”, aspirando a conseguir para sí los privilegios de la inseparabilidad, nos muestra la sabiduría de Dios moviendo su voluntad de fundamento. El acontecimiento de los primeros ángeles caídos parece literalmente reflejarse en estas palabras del filósofo: “Y dado que el fundamento contenía también en sí la totalidad del ser divino, pero no como unidad, por esto sólo podían ser seres divinos singulares los que reinasen en ese actuar-para-sí del fundamento” (Schelling, 1989: 213). El acontecimiento del mal es una catástrofe en las mismas barbas de Dios. La descomposición del ente, a decir de Heidegger. El mal se expresa como una posición propia de la voluntad, y a la una con esto, como un modo de ‘Ser’-libre en el sentido del ‘Ser’-sí-mismo a partir de la propia ley esencial (Heidegger, 1985: 174).

Esta separabilidad, como dualidad, es vista por Vanden, como quien dice, como una auténtica genealogía de Dios. La aparente contradicción –en Dios- entre existencia y fundamento, se aclara si se sigue más en detalle el ensayo sobre la libertad. En efecto, en esta doble perspectiva de Dios, por un lado, como existente, Dios es el espíritu que domina el mundo; por otro, como fundamento de la existencia, Dios se revela como la naturaleza necesitada de iluminación. La diferencia entre Dios y las cosas creadas es que Dios contiene el fundamento de su existencia dentro de sí mismo, mientras que, en cuanto creador, pudiéramos decir, en su tarea de dotación de criaturas, estas tienen el fundamento de su existencia fuera de ellas mismas, vale decir, en Dios (Vanden, 2019: 244).

## **Dios, tiene que ser**

El mal, pues, debe entenderse como la ruptura de la voluntad del hombre con la voluntad universal:

La voluntad, que sale de su sobrenaturalidad para volverse a un tiempo, como voluntad universal, particular y propia de las criaturas, tiende a subvertir la relación que existe entre los principios, a elevar al fundamento por encima de la causa y a usar el espíritu, que sólo recibió para el centro, fuera de este último y en contra de la criatura, lo que conduce al caos dentro y fuera de ella (Schelling, 1989: 183).

Mas, cabe preguntarse: ¿de dónde proviene esta transformación de la voluntad del hombre? ¿A qué responde esta búsqueda del fundamento de una preeminencia absoluta por sobre la existencia? Schelling ha definido la libertad como una facultad en el hombre para el bien y para el mal. Haciendo esto, ha reconocido una raíz de libertad en el fundamento independiente de la naturaleza, lo que supone que el vínculo de los principios es, por lo mismo, esencialmente libre. Sin embargo, esta libertad no es el libre arbitrio que se pueda confundir con el azar o con la indeterminación. Libertad y necesidad, en el sistema de Schelling, deben ser entendidos como principios no contradictorios. Es un presupuesto indispensable en esta etapa, que el hombre sea ciertamente libre en lo que respecta al actuar mismo, pero en lo que concierne al resultado final de sus acciones, dependa de una necesidad que está por encima de él, aunque llegue a interferir en el juego de su libertad (Schelling, 1988: 394). En su devenir, Dios tiene que revelarse necesariamente, porque en la creación no puede subsistir nada que sea equívoco. De esta forma, si Dios solo puede revelarse en el hombre, requiere, entonces, que este previamente alcance su unidad primordial. Esta misma unidad de los principios, que en Dios es inseparable dada su calidad de espíritu, es efectiva en el hombre únicamente cuando en él reina el espíritu del amor. En este estado de unicidad fundamental, la luz domina al movimiento sin las reglas de una naturaleza inicial.

En este devenir de Dios en la naturaleza, su fondo oscuro y caótico se transforma en orden, no estrictamente como sinónimo de unidad, sino, en la observación de Rae, como un mundo susceptible de ser comprendido por la razón. De este modo, se tiene la idea de que Dios es demasiado grande para aparecer en el mundo como es en Sí mismo. Lo mejor que pudiéramos obtener es una imitación pálida del creador. Sin embargo, la luz de Dios no logra abolir totalmente la oscuridad. Antes bien, la ascendencia de la luz no derrota su suelo oscuro; en todo caso, lo provoca a resurgir, reafirmando así la lucha entre los dos, única dinámica generadora de la vida en el sistema metafísico-religioso de Schelling (Rae, 2018: 240). A guisa de esto, lo que Dios requiere para su revelación eterna no es sino la condición de armonía de los principios originales del hombre. Esto se explica, pues el hombre, en su permanente impotencia para permanecer orientado en dirección a Dios, desvía su vista y su búsqueda hacia el no ser, pudiéramos decir, hacia las tinieblas. En palabras simples, cae. El estado original del hombre, entonces, es la separabilidad de los principios: la Caída.

En cierto modo, la condición entitativa del hombre está determinada por la movilidad libre del fundamento contra la existencia. Si esto es así, el bien y el mal no son en el hombre sino la re-partición de las fuerzas originales ya separadas en él en sí mismas. Dios, piensa Schelling, tiene que ser, y para ser, necesita que el hombre, a su vez, sea. Dios no ha elegido la existencia del hombre, sino que la ha decidido, pero mediante una decisión que no admite elección alguna, puesto que Dios no tiene, desde luego, posibilidades infinitas de elección. De ahí que la perfección de lo Absoluto consista en poder querer solamente una cosa, y esta es la necesidad de su esencia más propia, es decir, el amor (Heidegger, 1985: 195). Lo Absoluto debe, pues, patentizarse en el hombre. Es lo mismo que decir que el amor debe resolverse en el hombre como la suprema concordia de los principios. El vínculo divino que liga a ambos principios en una identidad absoluta, solo puede originarse a partir de la subordinación del mal al bien. El bien en el hombre es la condición trascendental que posibilita la identidad absoluta de los principios en Dios. Dios puede revelarse solo a partir del hombre en cuanto ente existente. De manera que, si Dios es el fundamento, y, como tal, vida creadora, lo engendrado en Él no puede ser un nuevo mecanismo. Es decir, si Dios se revela como fundamento en lo infundado por Él, entonces solo puede revelarse a sí mismo en esto infundado, en el sentido de que lo dependiente debe ser él mismo un ente capaz de actuar libremente en sí mismo, justamente porque depende de Dios (Heidegger, 1985: 106).

Pero, paradójicamente, el bien, para serlo, necesita que el mal exista. Se requiere una mismidad activada para el rigor de la vida. Sin esta mismidad vendría la muerte como un adormecerse del bien, puesto que, sin su papel activo en la dualidad, no hay vida (Schelling, 1989: 265). Visto así, el mal es en el hombre el encumbramiento del principio oscuro, que patentiza en él, en tanto de-cisión entre lo idéntico, la posibilidad de la libertad. Tal principio solo es existente específicamente en el hombre, y, por tanto, tan real como el bien. Es decir, “la posibilidad del bien depende de su realidad. Sólo se puede revelar un poder en su oposición; así la luz sólo en las tinieblas, Dios sólo en el mal” (Hartmann, 1960: 229). De manera que el fundamento del mal no es sino el fundamento del ‘Ser’-hombre, fundamento que tiene que yacer en el centro más interior de Dios (Heidegger, 1985: 146). Inopinadamente, la necesidad del mal parece ser, sin más, la posibilidad del bien, en cuanto el bien –según lo dicho- es el vínculo divino existente en los principios del hombre, la luz de Dios que atraviesa en plenitud la oscuridad del fundamento. Lo bueno, pues, es lo que es propiamente ente en el mal (Heidegger, 1985). Ahora, teniendo en cuenta esta suerte de metafísica del fundamento, el hombre sería, como modo de ser-libre, propia y radicalmente la facultad para el mal. Libertad y necesidad, contrincantes eternos, se hallan en esta respectividad unidas idénticamente. La separación (la “caída”) origina, pues, la contrariedad, que inicia con la auto-objetivación en sí (es impensable una objetivación fuera de sí, al no haber fuera-de-sí) del Absoluto: mientras la necesidad une al Absoluto con su réplica o segundo-Absoluto, la libertad los separa (Del Barco, 1999: 100).

El mal, reiterémoslo, no es ni consecuencia ni atributo de una decisión en el hombre, pertenece a la esencia del mundo, como el fundamento a la esencia de Dios (Hartmann, 1960). Pertenece, por tanto, a la más propia esencia del hombre. Únicamente el hombre es capaz del mal y, en tal medida, este constituye precisamente la esencia de su ‘Ser-hombre’ (Heidegger, 1985). Ahora, la respectividad de la que hablábamos, la que une idénticamente libertad y necesidad, supone desde luego la efectividad del

sistema. Es a lo que Grave llamará “giro teológico”, el que supone dirigir el pensamiento hacia el interior de la filosofía, en cuyo seno ha nacido la teología. Esto quiere decir que en Schelling, el sistema forma parte de la esencia del ser en general, y que, por lo mismo, tiene que estar presente primeramente en el ser originario como fundamento de todo ente en Dios (Grave, 2007: 61). El Dios de Schelling parece ser el Dios que podemos, al mismo tiempo, comprender metafísicamente y reconocer objetivamente. Conviene, pues, insistir en que el fundamento de la existencia de Dios no implica otro ser, sino que es definido por la oscuridad, la que perfectamente puede interpretarse como la condición de Su ser existente (Rae, 2018: 238). En palabras de Das:

El Dios que es conocido metafísicamente y que llega al conocimiento absoluto dialécticamente mediante un proceso necesario cuya ley se puede comprender categóricamente. Es solo este Dios quien puede encarnarse, objetivamente, como una figura mundana en el orden profano, ya sea que esta sea la Iglesia visible o el estado moderno de Prusia (Das, 2016: 136).

Esta comprensión metafísica de Dios proviene, si seguimos a White, de la concesión que Schelling hace del argumento de Kant de que la razón exige a Dios como principio absoluto, pero también – especialmente a partir de su ensayo sobre la libertad- de la idea de que la razón sola, sin ayuda de la experiencia, no puede dictaminar si Dios existe o no. Schelling, al igual que Kant, rechazará la prueba ontológica en todas sus formas, e insistirá en que ningún concepto puede establecer por sí solo la realidad de una deidad inteligente y benevolente. En otros términos, la determinación ontológica del concepto de Dios no revela más que Dios solo puede existir como fundamento absoluto, es decir, que “debe ser” aquello que es absolutamente anterior, el “ser necesario”, si es que existe (White, 1983: 164). Probablemente, como advierten Franco Bonifaz & Cortés del Moral, el dogmatismo de Schelling de presuponer en Dios una realidad no absoluta, lo que hace es revelar, paradójicamente, un Dios imperfecto. Empero, para Schelling la complejidad de Dios es también infinita. Es menester, creen los autores, axiomatizar a Dios. Esto quiere decir que lo imperfecto que lo nombra, o lo no Dios, tiene que ser por Dios. Si el concepto que lo nombra es lo no Dios, tiene que estar, no obstante, en Dios y ser por Dios; es decir, lo imperfecto es por lo perfecto. Lo contrario supondría hacer del concepto otro absoluto, lo que, desde luego, no es posible; o defender al concepto fuera de Dios, lo que es (todavía) más imposible (Franco Bonifaz & Cortés del Moral, 2019: 158).

### **¿Filosofía de la religión o filosofía de la revelación?**

Si la de Schelling pueda ser o no una filosofía de la religión, es algo que seguramente no vamos a zanjar en este trabajo. Me limitaré, no obstante, a sugerir que sí y, anticipándome a alguna objeción, agregaré, de corte conservadora.

Ciertamente, la justificación de su meditación sobre el problema del mal, se halla alojada sobre todo en una pregunta más fundamental aún por la naturaleza y la existencia de Dios. De esto va, se deduce, el corazón de toda filosofía de la religión. Ahora, es palmario que el pensamiento tardío de Schelling se enfocó en la pregunta por Dios. Y habría que reconocer que este sistema que el filósofo se empeñó en construir, se impone, en primer lugar, como una robusta metafísica, en el sentido, como acota Vargas (2005), de que se ocupa de la pregunta por el principio originario e infinito y su relación con los entes finitos y naturales. Tal cosa le permite decir a Gaudio (2004), que la explicación del origen del mal lleva inevitablemente en Schelling al problema metafísico (cosmogónico y cosmológico) de la articulación entre la unidad primigenia e indiferenciada y la diferencia o finitud existente. Siguiendo la argumentación de Schmitz (1987), podemos entender que toda filosofía cuya preocupación fundamental sea precisamente lo que significa ser Dios, o la divinidad, caería sin mayores inconvenientes en la clasificación de una filosofía de la religión, de una doctrina filosófica de Dios. Dicha preocupación resulta asaz evidente en el siguiente pasaje de las *Investigaciones*: “Dios es algo más real que un mero orden moral del mundo, y contiene en sí otro tipo muy distinto y más vivo de fuerzas movimiento de las que le atribuye la mísera sutileza de los idealistas abstractos [refiriéndose seguramente a Jacobi y a su interpretación del panteísmo de Spinoza]” (Schelling, 1989: 159).

En una cuerda un poco distinta, Vargas prefiere ver en el trabajo de Schelling, antes que una filosofía de la religión, una moderna teodicea, aunque en otros lugares también se refiera al pensamiento del alemán como una metafísica del mal. La investigación de Schelling, en cuanto teodicea, implicaría, en primer lugar, la justificación de Dios frente al mal y la posibilidad de su superación por medio de la salvación divina; en segundo lugar, el estatuto de la libertad humana, y, por último, la responsabilidad del hombre frente al mal (Vargas, 2005: 123). Incluso, a partir de una hermenéutica de la facticidad como la heideggeriana, es posible reconocer los límites y posibilidades propiamente filosóficas del fenómeno religioso. En cierto modo, la preocupación de Heidegger por la meditación de Schelling sobre la libertad, muestra coincidencias muy sugerentes con el esquema con que desde temprano el filósofo de Friburgo visualizaba la cuestión religiosa, esquema formado esencialmente por el tándem religiosidad protocristiana/historicidad:

La tarea del entendimiento fenomenológico [en el joven Heidegger] deberá dar cuenta de dos aspectos formal-indicativos, por ahora, hipotéticos: 1. La religiosidad protocristiana [*urchristliche Religiosität*] se encuentra en la experiencia de vida protocristiana (GA 60: 80). 2. La experiencia fáctica de la vida es histórica [*historisch*]. La religiosidad cristiana vive [*lebt*] la temporeidad [*Zeitlichkeit*] en cuanto tal (GA 60: *idem*) (Masís, 2013: 7).

De hecho, una de las observaciones cardinales de las *Investigaciones* precisamente es la diferenciación entre Dios y naturaleza, distinción que, como se sabe, constituye la principal “cuña” metafísica entre el panteísmo de Spinoza y el idealismo trascendental de Schelling. Escribirá el filósofo de Leonberg a Schubert en 1809, en alusión al panteísmo “fatalista” de Spinoza: “¡Ojalá pueda darse pronto fin a este horror de una nueva beatitud que ahora comercia con el temor a Dios y la religión y los relaciona con la vida y la naturaleza más lánguidas y disolutas (...)!” (Carrasco, 2011: 96). Pues bien, el problema de la religión cruza prácticamente todo el camino filosófico de Schelling, quedando sistematizada especialmente en su Filosofía de la mitología y en su Filosofía de la revelación. En todo caso, para Schelling resulta fundamental la concepción de un espíritu necesariamente ligado a la naturaleza material, la que en ningún caso será entendida como mera res extensa cartesiana, sino como un organismo vivo en el que se expresa Dios en cuanto *natura naturans* (De Torres, 1995: 121). De ahí que la concepción esotérica de Schelling de una naturaleza viva, le permita concebir la vida justamente como significativa en el sentido cósmico, sin la idea clásica de Dios como un ser omnipotente que es responsable del sentido (podríamos decir, del destino) de la vida (Pitkänen, 2019: 503).

Donde es posible hallar concomitancias directas con la filosofía del mal de Schelling, y considerarla, por tanto, una absoluta filosofía de la religión, es en las observaciones de Beuchot sobre el fenómeno religioso. Para el filósofo mexicano, antes de decir que Dios es malo o cruel porque permite el mal, hay que ver qué se juega en ello. Es indudable que Dios conoce el mal y puede evitarlo; también lo es que no puede querer el mal, porque quererlo sería contrario a su perfección. En consecuencia, si no lo evita no es ni por desconocerlo, ni por no poder o por no querer, sino por otras consideraciones (Beuchot, 2017: 121). Si mantenemos el enfoque del filósofo mexicano, caemos en cuenta que esta filosofía de la revelación de Schelling, parece en realidad formar parte de las llamadas filosofías analíticas de la religión, en las que se recuperaría la vocación filosófica central de las teodiceas, esto es, los argumentos a favor de la religión relativos a los problemas de la existencia de Dios, su esencia o propiedades, su providencia y al problema del mal (Beuchot, 2017: 12-13). Pero si quisiéramos ser todavía más específicos, en un intento de justificación “por dentro” de una posible filosofía de la religión en Schelling, tendríamos que reconocer, con De Torres, que la experiencia relevadora de la tensión entre fundamento y existencia (el corazón de las *Investigaciones*) constituye *in toto* una filosofía de la religión. En efecto, la revelación sería una forma específica de comunicación, por medio de sus condiciones internas (lo revelado, la revelación) y sus condiciones externas (el revelador y el receptor de la revelación). Dios mismo se da a conocer en su existencia al receptor de la revelación de un modo esencialmente espiritual. Bajo esta consideración, para Fichte la creación no sería revelación de Dios, como sí lo sería para Schelling (De Torres, 1995: 120).

Por otro lado, si nos atenemos a la crítica de Grassi, en particular a su clasificación historiográfica sobre el pensamiento de la religión, la filosofía de Schelling tendría que ubicarse dentro del método de



confrontación, el que se define por la contraposición y la puesta en juego de una dialéctica entre razón y fe. Es decir, una filosofía de la religión debe estar siempre hecha en función de una tendencia filosófica más o menos determinada (Grassi, 2012: 66). Asimismo, la posibilidad de entender la metafísica de Schelling como una real filosofía de la religión (al menos en su obra sobre la libertad, que termina siendo, a la postre, una verdadera exégesis de la relación entre Dios y el mal), queda asegurada en la propia filosofía de la religión de Eugenio Trías, quien a través de su idea de límite (para ser precisos, de la idea de hombre como límite) intenta salvar el “cerco hermético” en pos de la “cita con lo sagrado”. Apoyándose en Schelling, Trías considera una razón que requiere como primado fáctico (y fundamento) una revelación previa. Es lo mismo que decir que a la reflexión de la razón le precede siempre la revelación de lo sagrado. Por eso, la filosofía, si quiere superar el nihilismo, debe recuperar el diálogo tenso con una existencia revelada, es decir, demandar de la revelación el fundamento que no puede hallar en ella misma (Conesa, 2004: 106).

Por último, el excelente estudio de Duque sobre Schelling nos pone en liza acerca de lo que podemos discurrir como una filosofía de la religión definitiva. En efecto, y más allá de las implicancias que en el pensamiento del alemán pudieran hallarse entre una filosofía de la religión y una filosofía de la revelación, lo cierto es que toda su última filosofía está impregnada de una especial fruición en torno a la religión. Es más, si seguimos la propia fundamentación de Schelling, lo que revelaría su filosofía no sería sino la propia esencia de su pensamiento religioso, el cristianismo, y este, a partir de su envoltura en algo parecido a un velo primitivo, fundamental, impensado. Una suerte de sentido primigenio del cristianismo como filosofía de la revelación. En este sentido, para Schelling tanto la religión original como la que “cerrará” los tiempos, sigue siendo cristianismo. Aunque Schelling diferencia entre un cristianismo “exotérico”, de casi dos milenios de antigüedad, y otro “esotérico”, que mentaría el basamento del mundo y, por tanto, sería tan antiquísimo como este. Por lo mismo, sin el cristianismo “aparente”, e incluso sin las religiones mitológicas precristianas, nunca se habría llegado a conocer el cristianismo “verdadero” y todas sus verdades esenciales. Esta revelación, pues, latente tras las imágenes y parábolas bíblicas, será la que impulse y eleve a la razón schellingiana (Duque, 2014: 34).

Toda la filosofía del último Schelling está transida de realización del Futuro del Señor (Duque, 2014: 56). El siguiente pasaje de Pérez-Borbujo parece esclarecer el lugar de la religión en el sistema de Schelling:

La filosofía de la Revelación, que coincide, en lo que respecta a su contenido, con la filosofía positiva, y que junto a la Filosofía de la Mitología y de la Naturaleza forman parte de ésta, está precedida por la Filosofía de la Mitología, que constituye el fundamento. La Mitología es la religión del politeísmo, por ello, la falsa religión y el monoteísmo, la religión revelada, y por eso la verdadera religión (Pérez-Borbujo, 2016: 477).

## Conclusiones

La observación final de nuestro trabajo apunta a una re-definición de la metafísica del mal en Schelling. Tal determinación se ubica en el centro de su meditación sobre la libertad, y, vista desde una estructura aún mayor, dentro de una filosofía de la religión. Sin embargo, lo decisivo es que la pregunta por el problema del mal en Schelling, nos ha llevado a la pregunta por el fundamento mismo en Dios, y desde ahí, a sostener la identidad de Dios inicial con el mal. Esta relación de originariedad del mal en el fundamento de Dios ya se había anunciado en los trabajos del joven Schelling, pero no en el marco – permítaseme el giro- de una filosofía de lo prometeico. En consecuencia, podemos designar al mal en una relación de identidad con el fundamento en Dios, con aquello que está en Dios, pero que, paradójicamente, no es Dios-mismo. De esto se derivan dos cuestiones gravitantes sobre el ser del mal en Schelling:

En primer lugar, si la patentización de lo Absoluto requiere que el hombre sea, y si el ser esencial de este ente es la facultad para el bien y el mal, entonces, lo que parece corroborarse no es otra cosa que la identidad entre Ser y ser del mal. Esto explica que para Schelling el mal no sea más que el primerísimo



fundamento de la existencia, pero tampoco más que la superior potencia del fundamento que actúa en la naturaleza.

En segundo lugar, si el mal es el mero fundamento para que el bien, formándose a partir de él mediante sus propias fuerzas, sea, gracias a su fundamento, un ser independiente y separado de Dios (es decir, para que el hombre pueda singularizarse como un ente separado de Dios), entonces, la identidad entre el Ser y el mal, y la necesidad metafísica del mal para la constitución del hombre en cuanto ente, manifiestan una doble resolución metafísica del mal. Primero, el mal tiene la esencia originaria del Ser, no pudiendo carecer de ella, pues en la filosofía de la religión de Schelling, no puede existir nada fuera de Dios; y segundo, el mal es el fundamento (desde luego como no-ente). De ser esto así, cabe especular una fundamentación todavía más originaria: que el mal es la naturaleza en Dios, su fundamento para su ‘Ser’-sí-mismo. El mal, según esta conjetura, se hallaría en una relación de identidad con el fundamento en Dios. Dios sería el propio mal original. Como no puede haber nada absolutamente independiente de Dios, se sigue, con Hartmann, que el mal no puede ser ningún principio original junto a Él: únicamente Él es el mal original mismo.

Pero hay algo más. Se ha definido a Dios existente como devenir y como el único ente que es por sí mismo. De esta manera, el fundamento de Dios en Dios mismo no es Dios existente, y no puede, por tanto, mostrar ninguna de las propiedades del Ser. De aquí que el mal sea originariamente la expresión del no-ser, y que exprese su maldad solo en la medida en que se transforma en actividad (idea que abre la tesis schellingiana del “salto”). El mal es malo solo en cuanto va más allá de la potencialidad, pero, reducido al no-ser o al estado de potencia, es aquello que “debería” ser siempre. Por eso, nuestro filósofo dirá que el mal no es ningún ser, sino un no-ser, que no es real más que en su oposición, pero no en sí-mismo. El mal es, dicho al modo heideggeriano, la revuelta contra el ser originario, el Dios invertido, la descomposición del ente, la perversión de la existencia.

Ciertamente, la proposición final requiere volver a situarse en el espacio de la relación de identidad postulada por Schelling. Solo es posible concebir la identidad unificante de Dios (inicial) y el mal, en cuanto del mismo modo es posible considerar la identidad fundante entre Dios (inicial) y Dios-mismo, como principios diferenciados en el infundamento, pero transformados en uno en la unidad universal. La identidad entre Dios (inicial) y Dios existente, entre fundamento y existencia, entre no-ser y Ser, se mantiene eterna porque el devenir del Absoluto también es eterno.

Por lo mismo, la posibilidad del mal en el fundamento en Dios, no se extingue con la existencia del Ser. Así como el fundamento queda permanentemente relegado en su propia oscuridad, como negación de la unidad esencial del Ser, de la misma forma, el mal queda relegado en el espacio del no-ser, como principio necesario para la libertad del hombre y, con ello, para la glorificación de lo Absoluto. Metafísicamente, el mal es el contrario mediante el cual el bien puede abrazar la unidad de la voluntad universal. Al hacerlo, lo Absoluto patentiza su esencial omnipotencia. El mal, al igual como ocurrió en la escisión original desde el infundamento, entre fundamento y existencia, constituye, a partir de su expansión y querer-dominar la existencia, la emergencia del bien y la posibilidad de la imagen divina. De ahí que el mal sea la posibilidad del bien, y que, al mismo tiempo, el bien sea la posibilidad del mal. Ambos principios están en una relación de identidad, de la misma manera que Dios como Absoluto no ha dejado de contener en sí (en cuanto su Ser como eterno devenir) al fundamento de su existencia. Podemos, pues, afirmar, con Heidegger, que su ‘Ser’-fundamento es un modo del eterno devenir de Dios en total.

*El mal es la búsqueda y realización originaria del fundamento de Dios.*

## Referencias

- Barco, O. Del (1999). "Schelling, el pensamiento del Dios vivo". *Nombres* (13-14), 97-161. Recuperado de <https://tinyurl.com/rxyhkgd>
- Beuchot, M. (2017). *Filosofía de la religión*. Guadalajara: Cátedra.
- Carrasco, A. (2011). "Panteísmo y Panenteísmo: Schelling, Schlegel y la polémica en torno al panteísmo". *Δαιμόνιον. Revista Internacional de Filosofía*. (54). 93-109. Recuperado de <https://tinyurl.com/vtftp3nw>
- Conesa, F. (2004). "Pensar la religión: la aportación de la filosofía española actual". en T. Trigo (ed.). *Dar razón de la esperanza*. Servicio de Publicaciones de la Universidad de Navarra. Pamplona.
- Das, S. B. (2016). *The Political Theology of Schelling*. EUP. Edinburgh.
- Duque, F. (2014). "Schelling: filosofía de la revelación como dialéctica de la historia". *Contrastes*. 19(3). 27-59. Recuperado de <https://tinyurl.com/yx57ana6>
- Franco Bonifaz, N. & Cortés del Moral, R. (2019). "Lo absolutamente infinito (Spinoza), lo infinito (Schelling, 1795-1796) y sobre lo absoluto como logos contemporáneo". *Metafísica y Persona*. 11(22). 125-165. Doi: <https://doi.org/10.24310/Metyper.2019.v0i22.5322>
- García, M. (2016). "Energeia vs Entelecheia: Schelling vs Hegel sobre Metafísica Lambda". *Tópicos* (51). 113-137. Doi: <https://doi.org/10.21555/top.v0i0.763>
- Gaudio, M. (2004). "En el fondo de Dios. El origen del mal según Schelling". *V Jornadas de Investigación en Filosofía*. Universidad Nacional de La Plata.
- Grassi, M. (2012). "Fenomenología y filosofía de la religión en Henry Duméry". *Franciscanum*. LIV (158). 53-81. Recuperado de <https://tinyurl.com/rtqszs2>
- Grave, C. (2007). "El Schelling de Heidegger. Modernidad y sistema de la libertad". *Theoría*. (18). 51-66. Recuperado de <https://tinyurl.com/sjfenl6>
- Hartmann, N. (1960). *La filosofía del idealismo alemán*. Sudamericana. B. Aires.
- Heidegger, M. (1985). *Schelling y la libertad humana*. Monte Ávila. Caracas.
- Heidegger, M. (1997). *Conceptos fundamentales*. Altaya. Barcelona.
- Jacobs, W. G. (2018). *Leer a Schelling*. Herder. Barcelona.
- Llorente, J. (2016). "Dios ante el abismo. La filosofía de la revelación de Schelling como *Bildungsroman* del espíritu". *Areté*. XXVIII (2). 231-265. Recuperado de <https://tinyurl.com/yyqqt9h4>
- Masís, J. (2013). "Fenomenología de la vida religiosa en el joven Heidegger: de la Indicación Formal a la Interpretación Heideggeriana del Cristianismo Primitivo". *Logos*. 41(123). 45-78. Recuperado de <https://tinyurl.com/trzqy3m>
- Pérez-Borbujo, F. (2016). *Schelling: el sistema de la libertad*. Herder. Barcelona.
- Pitkänen, O. P. (2019). "Schelling, esotericism and the meaning of life [Schelling, esoterismo y el significado de la vida]". *Human Affairs*. 29(4). 497-504. Doi: 10.1515/humaff-2019-0045

Rae, G. (2018). "The Problem of Grounding: Schelling on the Metaphysics of Evil [El problema de la reconexión: Schelling sobre la metafísica del mal]". *Sophia: International Journal of Philosophy and Traditions*. 57(2). 233-248. Doi: 10.1007/s11841-017-0594-9

Schelling, F.W.J. (1988). *Sistema del idealismo trascendental*. Anthropos. Barcelona.

Schelling, F.W.J. (1989). *Investigaciones filosóficas sobre la esencia de la libertad humana y los objetos con ella relacionados*. Anthropos. Barcelona.

Schmitz, J. (1987). *Filosofía de la religión*. Herder. Barcelona.

Torres, M. J. De (1995). "Perspectivas sobre religión y filosofía en los inicios del idealismo alemán: Fichte y Schelling antes de 1794". *Enrahonar*. (23). 117-124. Recuperado de <https://tinyurl.com/r458fhz>

Vanden, D. (2019). "Kant and Schelling on the ground of evil [Kant y Schelling sobre el fundamento del mal]". *International Journal for Philosophy of Religion*. (85). 235-253. Doi: <https://doi.org/10.1007/s11153-019-09701-6>

Vargas, J. (2005). "Sobre el problema del mal y de la libertad según Schelling". *Praxis filosófica* (20). 121-156. Recuperado de <https://tinyurl.com/wrtfh6x>

White, A. (1983). *Schelling: An Introduction to the System of Freedom* [Schelling: Una introducción al sistema de la libertad]. Yale University Press.

## THE KANTIAN NOTION OF FREEDOM AND AUTONOMY OF ARTIFICIAL AGENCY

*LA NOCIÓN KANTIANA DE LIBERTAD Y AUTONOMÍA DE LA AGENCIA ARTIFICIAL*

*A NOÇÃO KANTIANA DE LIBERDADE E AUTONOMIA DA AGÊNCIA ARTIFICIAL*

**Manas Sahu**

*(Indian Institute of Technology Bombay)*

*manassahu23@gmail.com*

Recibido: 22/06/2020

Aprobado: 04/04/2021

### ABSTRACT

The objective of this paper is to provide a critical analysis of the Kantian notion of freedom (especially the problem of the third antinomy and its resolution in the critique of pure reason); its significance in the contemporary debate on free-will and determinism, and the possibility of autonomy of artificial agency in the Kantian paradigm of autonomy. Kant's resolution of the third antinomy by positing the ground in the noumenal self resolves the problem of antinomies; however, invites an explanatory gap between phenomenality and the noumenal self; even if he has successfully established the compatibility of natural causality and non-natural causality through his transcendental argument. This paper is also devoted to establishing the plausibility of the knowledge claim that Kantian reduction of phenomenality has served half of the purpose of the AI scientists on the possibility of Artificial Autonomous Agency.

Keywords: transcendentalism. phenomenal consciousness. artificial agency. naturalized rationality and autonomy.

### RESUMEN

El objetivo de este trabajo es proporcionar un análisis crítico de la noción kantiana de libertad (especialmente el problema de la tercera antinomia y su resolución en la crítica de la razón pura); su importancia en el debate contemporáneo sobre el libre albedrío, el determinismo, y la posibilidad de autonomía de la agencia artificial en el paradigma kantiano de autonomía. La resolución de Kant de la tercera antinomia al colocar el fundamento en el yo nouménico resuelve el problema de las antinomias; sin embargo, invita a una brecha explicativa entre la fenomenalidad y el yo nouménico; incluso si ha establecido con éxito la compatibilidad de la causalidad natural y la causalidad no natural a través de su argumento trascendental. Este artículo también está dedicado a establecer la plausibilidad de la afirmación del conocimiento de que la reducción kantiana de la fenomenidad ha servido a la mitad del propósito de los científicos de IA sobre la posibilidad de una agencia autónoma artificial.

Palabras clave: trascendentalismo. conciencia fenomenológica. agencia artificial. racionalidad naturalizada y autonomía.

### RESUMO

O objetivo deste artigo é fornecer uma análise crítica da noção kantiana de liberdade (especialmente o problema da terceira antinomia e sua resolução na crítica da razão pura); seu significado no debate contemporâneo sobre livre-arbítrio e determinismo, e a possibilidade de autonomia da agência artificial no paradigma kantiano de autonomia. A resolução de Kant da terceira antinomia postulando o fundamento no eu numênico resolve o problema das antinomias; no entanto, convida a uma lacuna explicativa entre a fenomenalidade e o eu numênico; mesmo que tenha estabelecido com sucesso a compatibilidade de causalidade natural e causalidade não natural por meio de seu argumento transcendental. Este artigo também se dedica a estabelecer a plausibilidade da alegação de conhecimento de que a redução kantiana da fenomenalidade atendeu à metade do propósito dos cientistas de IA sobre a possibilidade de Agência Autônoma Artificial.

Palavras-chave: transcendentalismo. consciência fenomenal. agência artificial. racionalidade naturalizada e autonomia.

## Introduction

The overwhelming debate among A.I scientists, psychologists, neurologists, and philosophers on free-will and determinism have opened up a new perspective for the age-old problem of free-will and autonomy. The A.I Scientist claims that when the level of AI will be reached the level of A.S. I<sup>1</sup> (Artificial Super Intelligence), it can defeat human intelligence and have the autonomy to build the A.I machine itself without the intervention of human intelligence. The revolutionary advancement in the field of A.I has raised the possible threats of the dominance of Artificial Intelligence over human autonomy. The present paper is a critical inquiry of the Kantian paradigm of autonomy and evaluation of his paradigm, whether it is compatible with the knowledge claim of the autonomy of artificial agency.

We are going to deal with the following question in this paper: How does Kant has established the compatibility of natural causality with unconditional causality through abstraction? Whether the solution to the third antinomy of Kant can also be applicable to the contemporary debate on the possibility of the autonomous artificial agency? Can an artificial intelligent machine have autonomy? How the autonomy of A.I is different from human autonomy?

The first section of the paper delineates the conflict between natural causality and unconditioned causality (transcendental freedom), the Kantian solution to this antinomy, and its relevance in the current debate on the autonomy of artificial agency. The second section will focus on the possibility of artificial agency and autonomy, the conflict of phenomenality with transcendental idealism, the objection to autonomy of artificial intelligence- agency (hence after AAIA) from phenomenal realism and transcendentalism and in the concluding section, their limitation has shown.

## 1. Third Antinomy: Conflict of Natural Causality with Transcendental Freedom

The debate on “freedom versus determinism” is a traditional problem of philosophy. Every scientific theory is grounded on the laws of causality principles, which claim that whatever happens in the universe is based on the principle of causality. Every physical occurants in the universe (including human action) is grounded on the causality principle (cause-effect relationship); hence wholly determinable. The present state of something is caused/determined by the previous chain of events, which is based on

---

<sup>1</sup> The genealogical foundation for the argument of the possibility of artificial super intelligence can be traced back from the Alan Turing's (1950) highly cited paper “Computing Machinery and Intelligence” in which, he argued that an artificially constructed machine not only can *act like a human being rather think like a human being*. Bostrom (2014), Kurzweil (2006), and Copeland (1998) argued for the possibility of artificial super intelligence. They predict that there will be the *artificial singularity* in the future, the view that A.I agents will transcend the artificial general intelligence and have the ability to build to the similar level of A.I machines.



natural laws. Therefore, there cannot be any instances wherein the chain of causality principle can be ruled out.

On the contrary, if the deterministic view is correct, there cannot be anything in the universe that is not subject to the laws of physics. Hence, no human freedom is possible in the deterministic paradigm, which provide the foundation of moral philosophy<sup>2</sup>. If the subject is not free to choose their action because their decision is predetermined by causal laws or causal chains of events, then the subject is not responsible for their actions. Because freedom is the precondition for moral action and moral responsibility. All moral principles are only applicable to those actions performed by the agent with the exercise of free-will. Therefore, determinism and freedom are incompatible with each other because both cannot be true together at the same time.

Immanuel Kant, one of the most influential philosophers of the 17th century, argues that although the conflict seems real to us, hence it needs attention to solve the issue, and either one must be true because both cannot be true or false at the same time, is, in fact, a pseudo-conflict. It is the consequence of the *Rhetoric use of reason*, which generates antinomies in Kant's language. An antinomy is a problem in which the debate between two superficial conflicting ideas seems real to the participant because both of their claims are seems to be supported by reason; however, in reality, is a pseudo conflict. It is the basic principle of an antinomy that the debate over two conflicting claims will be called antinomy in so far as either both to be true or false. If one became true or false, the conflict became real; hence nevertheless can be called antinomy. Kant, in his "Critique of Pure Reason," demonstrates that there are four pseudo conflicts that seem to us as a real problem due to *euthanasia of reason*. Our objective in this paper is to focus on the third antinomy of Kant which is popularly known as the conflict of freedom and determinism.

## Thesis

"Causality in accordance with laws of nature is not the only one form which all the appearances of the world can be derived. It is also necessary to assume another causality through freedom in order to explain them" (A444/B472)<sup>3</sup>. The sensible world cannot be explained through causality only; transcendental freedom as a principle of causality is also necessary in order to explain the *world of appearances*. In other words, we have to accept that there are two kinds of causality for accommodating natural philosophy and non-natural philosophy. The natural causality and transcendental freedom as a causality. The natural causality governs the physical states of affairs of the phenomenal world, whereas the unconditional causality is responsible for moral action and the moral responsibility of the conscious individuals.

## The argument for the thesis

Let us suppose that only the causality (the laws of nature) can explain the world of appearances. The principle of causality says that every occurrence must be sufficiently determined for being part of the causality. If the chain of causes is infinite, then a complete explanation of the chain of events or their causal account became impossible, which leads to the conclusion that the chains of the event as a totality cannot be established from a deterministic standpoint.<sup>4</sup> As we know, in the laws of causation, every effect must have a predetermined (a-priorily determined) cause. Every event necessitated by a prior event, (i.e., x caused by w, w caused by v ... Infinite regress) without first beginning or origin. However,

---

<sup>2</sup> It is a noncontroversial claim for both the proponents of naturalized ethicists (Except Illusionist, reductive materialist) and non-naturalized ethicists to have a collective agreement that freedom is the necessary condition of morality. Illusionist argument for Illusion of free-will (Sam Harris, Dennett, Churchland, etc.) does not uphold the position that free-will is necessary for moral responsibility. The subject has the illusion of free-will. Everything is predetermined by predeterminate conditions and scientific laws.

<sup>3</sup> See Guyer & Wood, (1998) trans. "The Critique of Pure Reason". Hence after only the Acedemie number (for instances, A445/B443, A415/B44, A449/B447 etc., here A refers to the first edition of the original text which was published in 1781 and B refers to the second edition published in 1887) will be mentioned.

<sup>4</sup> See, M. A. Gillespie (2014, p.3) "Philosophy and Rhetoric in Kant's Third Antinomy,"

as we know that no events can occur without sufficiently predetermined/ a-priorily determined principles; otherwise, it violates the laws of causality principle. Therefore, the view that there is only one kind of causality is inconsistent with its own principles, and the claim that everything is determined by causality only is self-contradictory. On the one hand, the unconditional causality is responsible for freedom of choice and secure the ground for morality. On the other hand, natural causality is responsible for the physical events of the world.

## Antithesis

There is one kind of law that determines everything in the universe, and transcendental freedom is an *empty thought entity* (A445-7/B-473-5). "There is no freedom, but everything in the world happens solely in accordance with laws of nature" (A445/B473). Let inherits that there are two kinds of causality, natural causality in one hand and transcendental freedom as an unconditional causality on the other for the sake of the argument. The claim that an uncaused-cause (absolute spontaneity), is the antecedent of all other events, however, itself not caused by any other antecedent cause is inconsistent with the laws of natural causality (A415/B443). Every beginning of action/event presupposes a state, which must be depended on fixed laws of nature. There cannot be any causal connection between unconditional causality<sup>5</sup> (Transcendental freedom) and natural causality for the reason that transcendental freedom is contrary to the laws of nature based on which the unity of experience is possible. If transcendental freedom is independent of causality, then it is impossible to establish the relationship between transcendental freedom and the phenomenal world, in contrast, if it is a part of nature or affected by the laws of nature, then it is bound to follow the deterministic principles of the empirical realm. In that case, there is no need to claim that transcendental freedom has an independent existence from nature. Because it is nature itself that guides the principle of the phenomenal realm, therefore, freedom and causality are contradictory to each other, and both cannot exist at the same time in the phenomenal realm. Freedom is independent of the laws of nature is illusory because it entails *lawlessness*.<sup>6</sup> Even if we accept transcendental freedom in the epistemic sense, we cannot establish the *unity of experience* (A449/B477) because the laws of nature will always guide the experience. Though the causality always conditioned by its own law, however, be able to provide *the lawful unity of experience*, which cannot be established through the transcendental freedom.<sup>7</sup>

The current advancement in the field of neuroscience also argues for the validity of the antithesis. The neurological activities (follow the deterministic neurological pattern) in the brain generates the illusion of free-will<sup>8</sup> and responsible for decision making and performing a particular action.<sup>9</sup>

## Synthesis

If we look at the thesis and antithesis, then it seems that both of their claims are inevitable and grounded on rationality at the same time it is crystal clear that if the thesis (a) is true, then the antithesis (-a) must be false and vice-a-versa. Kant contends that the conflict may seem to be real because of the

---

<sup>5</sup> Kemp smith (1918) termed it as "free *cause*" (which is not an object of possible experience), is an empty thought entity for determinist because of its nature (non-spatial and non-temporal) as advocated by the supporters of the thesis as the unconditional causality.

<sup>6</sup> See Allison (1990, 'The Third Antinomy.' p.20), if freedom as a causality, is independent of laws of nature (hence lawlessness), then it cannot be a cause at all and have any effect on Will of the agent. On the other hand, if it is determined by the laws of nature, then there is no need to presume that freedom as a different kind of causality separate from natural causality; hence, it became part of natural causality.

<sup>7</sup> This problem is a threat to all forms of idealism that owe to establish the relation between the transcendental idea with the experience of the phenomenal realm. It is crystal clear that by undermining the external world as an appearance, the relationship cannot be established.

<sup>8</sup> The science wing on the project of *Big Questions on free-will* conducted the research at Harvard in 2011 for the support of the antithesis, which argues for determinism and illusionism.

<sup>9</sup> Libel, B. (1999) argues for the deterministic theory of free-will. He has conducted a thought experiment popularly known as Libet's experiment for a scientific understanding of free-will and arrived at the conclusion that whatever happens in the brain's decision-making processes can be traced to the neurological brain processes through EEG techniques. One can predict whatever decision the subject is going to take before half a second, the subject being consciously aware of about the decision from the analysis of neurological activities of the brain of the subject gained through EEG techniques. Hence free-will is a unconscious process and has no role while taking decision and performing actions.

delusionalisation of rationality in their claims; however, it is a pseudo conflict<sup>10</sup>. Both the thesis and antithesis are correct from their perspective as far as physical and metaphysical necessity is concerned; however, they do not contradict each other because it is the same reason, which is the basis of their claim. The appearance of contradiction is due to the misconception of mathematical totality as the sole totality. When the nature views in terms of mathematical totality it provides the sensible condition, which is subject to laws of nature and space and time whereas when it views as dynamic causal series, then there is a demand for an unconditioned condition to condition the heterogeneous causal series which can only grasp in terms of intelligibility. The thesis needs a foundation on which unconditioned unity of totality can be established, whereas the antithesis requires a foundation for completeness of the causal series and universal unity of experience. Freedom is the prior condition for both experiences of nature and moral action; however, it must be outside the domain of the empirical realm.<sup>11</sup>

Kant's resolution of the third antinomy through transcendental idealism is a controversial issue.<sup>12</sup> His attempt to preserve the sanctity of universal moral order has to be praised, however, not with the cost of suspension of the phenomenality.<sup>13</sup> The reason for Kant to keep separate the transcendental freedom from the empirical realms is to protect the sanctity of the phenomenal realm, however, the intelligible realm, on which, the conceivability of the transcendental freedom is relied upon, is implausible (Kant's justification for transcendental world is unsatisfactory, controversial because it yields technical problem in philosophy) and promotes skepticism about the phenomenal reality and solipsism. There is no satisfactory response to the question that how transcendental freedom related with phenomenal realm by preserving the sanctity of the both realm? Generally, there are two categories of response from Kantian reader to the question, namely *two-world thesis*- the view that the object of things in themselves and appearances are numerically distinct object, belongs to transcendental world and phenomenal world respectively. The *two-aspect thesis* advocate that there is ontologically one world but two kinds of perspectival interpretation- The things really are (things in themselves) and how they are appear to us through the constituent of private sense data.

Allison<sup>14</sup> one of the proponents of two-aspect view, holds that the same reality can be viewed from transcendental and empirical prospective. he argues for two-fold sense of viewing the same reality by rejecting the two-object thesis such as- phenomenal and noumenal object. He argues

*“Man (along with rest of creation), is a part of nature, subject to its laws, and on the other hand knower, actor, rational being, transcendent nature in the sense that not completely determined by natural laws and along with that author of these laws”*.<sup>15</sup> It's the transcendental paradox of Kant that on the one hand, man is part of the sensible realm in the sense that subject has to follow the laws of nature, on the other hand, have the ability of intelligibility, hence knows the limits of sensibility and capable to transcend the limits of the sensible realm. Allison argues that this apparent perspectival dualism- transcendental and empirical distinction does not lead us to the noumenal-phenomenal dichotomy<sup>16</sup> rather provides two perspectives of viewing *the ordinary empirical point of view; its limit and transcendental necessary conditions of possibility*.

Despite of the disagreement on things-in-itself, and appearance distinction, it is quite uncontroversial that Kantian treatment of appearance as the product of sensible realm cannot have separate ontological

<sup>10</sup> For Kant, it's a conceptual confusion about the foundation of causality according to which natural causality is based on the phenomenal realm whereas transcendental causality demands its foundation for beyond the limits of sensibility (phenomenal realm) that is noumenal realm. Kant argues that both natural and transcendental causality are grounded on *rationality*. The laws of physics in its abstract form can't be grounded on sensibility. The discovery or creation of any laws of nature or moral laws is impossible without rationality. Rationality provides the foundation for both the condition of sensibility and intelligibility.

<sup>11</sup> Allison (1990, p.26), "The Third Antinomy," argues that reason has the tendency to transit its limit; therefore, it demands an unconditional totality of conditioned independent of the empirical realm for the sake of causal principles, which can be only intelligible and beyond the sensibility.

<sup>12</sup> He negates the content of experience as real, which is prone to promotes solipsism and skepticism on the existence of the external world.

<sup>13</sup> This problem is popularly known as the boundary condition of knowledge.

<sup>14</sup> See, Allison (1971 pp. 182-186) *Kant's Transcendental Humanism* for critical analysis of his two-view thesis.

<sup>15</sup> Ibid, p.183

<sup>16</sup> Ibid., p.188

status, and the locus of transcendental freedom cannot be sensible realm. Appearances has to be treated as the content of experience for the subject (things in themselves) which has only ontological status of being in noumenal reality; otherwise, transcendental idealism became indefensible and will lead to transcendental realism. The object of the external world can be expressed in two way- as it appears to the subject and really it is in its essence (indistinguishable from subject-noumenal reality). There can be two ways of interpretation of Kantian view of things-in-itself, the plurality of self or things-in-itself, in which there can be more than one thing-in-itself, and monistic self- the view that there is one inseparable and indistinguishable thing-in-itself. From the observation of the trajectory of *transcendentalism*, we have seen that in the interpretation of the Kantian self<sup>17</sup> as the plurality of self always ends with an open-ended question about the relationship among things-in-itself and, has been the victim of skeptical attack. On the other hand, the monistic interpretation of things in itself has either leads to *dogmatic transcendentalism* or concludes with *epistemic phenomenalism*.<sup>18</sup>

In the *Critique of Practical Reason*, Kant contends that unlike pathological beings, the Will of the rational agent is conditioned by practical principles that are independent of subjective maxims, because subjective maxims are causally affected by deterministic laws of nature<sup>19</sup>. In other words, the Will has to be only conditioned by the objective maxims<sup>20</sup> or categorical imperatives alone for preserving the autonomy or freedom of the agents. The great difficulty for the Kantian philosopher is to establish the connection between higher-order reality with lower-level phenomena by keeping the purity of both realms. If there is freedom and any moral principle at all, that must be reflected in human action. If the higher order principles (freedom) can't have reflection on the lower-level phenomena then such kinds of higher order principles become redundant. On the contrary, if the lower-level reality (phenomenal realm) has an effect on the higher order reality (noumenal realm- the locus of transcendental freedom) then transcendental freedom bound to follow natural causality. Theoretical knowledge without any relation to the phenomenal world is utterly piffle, because a theory-building activity<sup>21</sup> without any significance to the practical world is not going to solve any problem.

There can be two general approaches to connect the transcendental ideas of freedom with the phenomenal reality: the bottom-up approach and the top-down approach. Kant rejects the former and accepts the latter. Kant argues that rationality discovers the categorical principles because of which the status of human autonomy can be retained. The categorical imperatives are devoid of the content of experience and unaffected by the phenomenal realm in order to retain the purity of reason. In the top-down approach, certain moral principles (i.e., deontological principles, and cardinal values) are the ground of moral action.<sup>22</sup> The higher level reality (noumenal realm-locus of transcendental freedom and morality) is independent of the causal nexus of phenomenal realm, hence, grounded on pure rationality

<sup>17</sup> There is no direct reference for the transcendental self as a substance in Kantian literature. The simple reason behind it might be Kant's commitment to establishing a non-dogmatic transcendentalism. There is *no metaphysical assumption of self* in Kantian literature advocated by Guyer, P. (1987, p.86), Walker (2006). They argue that any attempt to attribute the transcendental self will lead to dogmatic metaphysics. Hence, the inference of the existence of the transcendental self from the boundary of the phenomenal self can only be intelligible.

<sup>18</sup> See, Willaschek (2001, p.683) for textual evidence.

<sup>19</sup> The dependency of the Will on the representation of appearances means to give up the freedom of rational agent, which makes the action of the agent similar to the action of a puppet for Kant. The practical principles determine the condition of the Will. The conditioned Will generates an effect in the sensible realm, which can be traced through the action of the agent. Once the Will of the rational agent affected by inclination and sensualistic desires, it follows a certain kind of deterministic behavioral pattern in which non-deterministic action is not possible. The determination of the Will must be based on those practical principles which are independent of the empirical world or based on the self-determination of the Will of the agency in order to ensure freedom or autonomy of the agent.

<sup>20</sup> Kant makes the distinction between objective laws from objective maxims. If the Will is determined by objective law, then the autonomy or freedom cannot be protected. Therefore, the categorical principles are given in terms of ought rather than is. The determination of the Will is depended upon the subject, that is, whether the Will of the subject will be determined by the laws of causality of the phenomenal realm or by the categorical principles. Every human being by nature rational or has the potentiality to be rational which entails the fact that they are free or autonomous beings that's why Kant advocates the dignity of human beings is routed around this idea that every rational agent must be treated as an end and never as a means for the simple reason of recognition of the freedom of individuals.

<sup>21</sup> One of the most influential analytic philosophers, namely Wittgenstein, attacked such kind of theory building activity in his work- *Tractatus Logico Philosophicus* (4.112).

<sup>22</sup> Though all these ethical principles are arguing for top-down determination of action, however, each one has a unique way of establishing their point of view, and each theory has its limitations.



which can never be naturalized, however, has an indirect effect on phenomenal realm due to which the effect of the transcendental realm became materialized and moral activities became possible.

The bottom-up approach has quite a different standpoint. Unlike the former, it argues that transcendental freedom has direct effect on phenomenal realm, hence transcendental freedom is traceable in human action through the application of rationality. In this approach, one can ground moral principles on the phenomenal experience itself. The awareness about phenomenal experience of the subjects is the foundation of morality. It upholds the position that every living being has shared the ability to be aware about the phenomenal experience; therefore, it is universal and is the ground of morality. The value paradigms of this approach are depended on the awareness of phenomenal experiences of every individual. The lower level phenomenal- having phenomenal experience, is directly connected with higher order reality-awareness<sup>23</sup> about phenomenal experience. The bottom-up approach even though provide a verificationist account of transcendental freedom, however, is incompatible with transcendental idealism of Kant.

### Critical Remark

Form the observation of Kantian solution to antinomy; his transcendental argument can be summarized in the following way

P1) Whatever belongs to the noumenal world is transcendental and unknowable<sup>24</sup>.

P2) The locus of autonomy cannot be the sensible world<sup>25</sup> Hence must belongs to the noumenal world for its existence.

Therefore, the autonomy of agents is transcendental and unknowable.

The reading of Kant from an idealistic perspective may have no problem with this argument. However, there are implicit problems with the argument for transcendental realist. The mere logical deduction of unconditionality for autonomy is not the sufficient condition for establishing the certainty of the knowledge claim of the unknowability of autonomy. It also invites the skeptical attack (i.e., reductionist and illusionist position) for the agency's autonomy. If the locus of autonomy or free-will is the noumenal world, one can never ever know or be certain about the autonomy of human beings. The only option left is to accept the autonomy as a-priorily given, and the fundamental building block of the universe in order to explain the responsibility of the action of the agents, because as Kant argues that it can only be intelligible but not sensible, which leads to us for an explanatory gap between the relationship of intelligible realm and sensible realm.

One of the most influential works of Kantian ethics is that it construed moral order as universal. The attempt for reconciliation of natural causality and freedom through the internalist approach is a significant effort of Kant; however, Kant's approach to the external world generates lots of technical problems in philosophy. Kant, like Descartes, invited the danger of solipsism and skepticism by upholding the view that the intelligible world is the sole reality. Kant holds that the mind constructs the world through the representation of appearances. Hence, the world has no independent existence from the mind, and the external world is just a mere product of mental construction. The skeptical approach towards the world and the denial of phenomenological reality for the sake of transcendentalism is not an

---

<sup>23</sup> Here I have used both of the term consciousness and awareness interchangeably.

<sup>24</sup> Here the word unknowability is used in a special sense. The autonomy of agents cannot be traceable because of its unknowability condition, as far as the limits of the phenomenal perspective is concerned; however, the effect of the self-determination of the Will can be found in phenomenal realism.

<sup>25</sup> Either the locus of autonomy is the phenomenal world or the transcendental world. The unconditional causality (i.e., autonomy) cannot be part of the sensible world. Whatever exists in the phenomenal realm, is contingent in nature. Therefore, it's locus must be in the transcendental world.



ideal approach. Kant failed to establish the transcendental freedom as a living reality by brushing off the reality of the phenomenal realm and subjective certainty about the experience.

The mind always takes part in structuring the sensory experience as well as in the active construction of thought or ideas, which might be fictional, non-existential, and objects of the possible and the real world. The product of mental construction sometimes may correspond to the world sometimes not. The mind is not entirely detached from the experience and content of experience. Experiences provide the information based on which the future prediction of conventional reality became possible through the help of reason. It is the logical faculty of reason which makes the distinction between the possibility of the existential entity from the non-existential or fictional entity. It also plays a vital role in the environmental prediction, construction of ideas, and judgments. Therefore, both experience and reason have a significant role in moral judgment.

For Kant, theoretical reason secured the ground for human autonomy. He argues that we have sufficient reason to accept the transcendental freedom as the foundation of morality. Transcendental freedom could not rely on the sensible realm for its existence because of the demand of an unconditioned condition of unbounded reason. Kant also accepts that the effect of the Will, being affected by only intelligible realm- (moral laws<sup>26</sup>), can be traceable through action only, which is itself belongs to the phenomenal realm. The underlying assumption behind Kant's transcendental idealism is that a universal moral order cannot be established, if the phenomenal realm is accepted as real due its nature of relativity and uncertainty. Hence, for Kant, it is not possible to find a universal moral order through the bottom-up approach. Rationality is the key factor in the self-determination of the Will of the subject. It is the criterion of autonomy, and it prevents the natural causality from affecting the determination of the agent's Will.

In light of the above evidence, it is crystal clear that the transcendental argument established the ground for transcendental freedom, however, is not plausible enough to receive universal recognition irrespective of the divergent paradigm. In the non-reductive naturalistic paradigm, the so-called human autonomy is an evolutionary product.<sup>27</sup> The redundancy of phenomenality for autonomy in the Kantian project and justifying it through some abstract principle has opened up the possibility for autonomy of artificial agents.

## 2. Artificial Agency and Autonomy

The tremendous advancement in the field of A.I. has brought the technological revolution and has changed life perspective. If intelligence is the sole criterion and almost everything is a product of intelligence (including moral principles and normativity of autonomy), then AAIA is possible. The dependency of autonomy on intelligence also affirms the fact that greater intelligence generates greater autonomy. If A.I. machines can be more intelligent than human beings; then, they can have more autonomy and be able to make better moral decisions in every situation.<sup>28</sup> If we accept the criteria of utility (maximization of utility and minimize the loss) and subjectivism (might is right principles) and not withhold speciesism,<sup>29</sup> then artificial super-intelligent machines become a reality that can not only

---

<sup>26</sup> Which are in itself abstract principles, are the criterion of moral action. The a-priori moral principles are being independently true and do not rely on the sensible realm, are necessarily true; therefore, universalizable, not because they are universalizable, therefore true. In other words, these moral laws are given in terms of abstract form as a criterion of moral action. However, the moment we try to understand the principles with the help of content, the self-contradiction within the moral laws arises. Therefore, Kant has tried to keep the distinction between the content of the moral law from its structure while talking about the moral principles. However, there is an ontological as well as explanatory gap between objective moral laws and its transformation in action; This has not bridged yet despite the several attempts to solve the problem. Another problem of the Kantian notion of moral agency is that, as he defended one kind of anthropocentric morality on the influence of rationalist slogan- *Man is the major of all things*, according to which, our moral obligation must be limited within rational agents (human beings alone). Rationality is the criteria for moral obligation and free-will. Non-human beings are devoid of rationality; therefore, they do not have moral rights and human species do not have any moral obligations towards them, which is a problematic issue and implausible with the argument of environmentalists, (i.e., Biocentrism, Cosmo-centrism) and feminism.

<sup>27</sup> See Mirazo and Moreno (2012 p.25-27) *Autonomy in evolution: From minimal to complex life*

<sup>28</sup> In certain circumstances, the biological rational agent became frozen (cannot take any moral decisions for a particular time being, i.e., in the trolley case thought experiment) due to limited intelligence, which might not be a problem for A.I. machines.

<sup>29</sup> Speciesism is the view that the dominance of one species over another is justifiable, grounded on the might is the right principle. It promotes human species centric ethics in which all the moral principles are rooted in favor of human interest only.

be autonomous but also can build autonomous machines.<sup>30</sup> In fact, as Kahneman<sup>31</sup> contends that human beings are less efficient in maximizing the utility because of the limitation of their cognitive power, though they are good at minimizing the loss. Hence, the A.I. machines can be a better performer proportionately than human beings. The ultimate goal of autonomy is to promote moral action. An A.I. agent can have more capability of information processing and predicting the situation than human beings. It can also have the ability to follow the normative conditions of morality in order to generate desirable actions. A perfect moral agent always maximizes the desirable action and minimize the undesirable actions as per the utilitarian principles of morality. An A.I. agent can serve the purpose in a better way in comparison to human agents.

Another approach towards autonomy of A.I. agency can be possible that as Kant argued, rationality<sup>32</sup> is sole criterion that gives human beings the status of autonomy. If A.I. machines can be built in such a way that they can be rational or can have the potential to be rational, then the objective (Autonomy of A.I. agency) can be achieved without the violation of the principles of categorical imperatives of Kant.<sup>33</sup>

### Can A.I machine be rational?

The word rationality itself is an elusive and ambiguous term because we cannot have a particular definition of rationality.<sup>34</sup> Rationality can be interpreted in many different ways as per the needs of the context. There are two kinds approaches of rationality a) Non-naturalistic approach of rationality b) Naturalistic approach of rationality<sup>35</sup>,

The non-naturalistic rationalist treats rationality as an ability; independent of normativity<sup>36</sup>, and *apriorily given* to human species. The ability of rationalizing is objective in nature and plays the significant role in the civilization of human species. There can't be any moral principle possible without rationality. The rational ability plays the pivotal role in the determination of the agent's Will. Rationality provides the formal structure because of which synthetic a-priori knowledge and knowledge about the phenomenal experience became possible. It can't be the object of experience due to its non-reducibility to the world of experience, however, is the necessary condition for the knowledge of phenomenal experience.

There is another category of rationality; namely naturalized rationality proposed by Quine<sup>37</sup>, Nozick<sup>38</sup>, Dretske<sup>39</sup>, in the 20<sup>th</sup> century- the view that the normative condition of rationality is dependent on the

<sup>30</sup> Of course, it is a logical possibility that might happen in the future, as claimed by A.I. scientist. It is an inference or future prediction from the observation of current advancement in the field of A.I.

<sup>31</sup> See, Kahneman, (2011) *Thinking*, for more details

<sup>32</sup> The formula of autonomy as stated by Kant in the Groundwork of the metaphysics of morals that 'the idea of the Will of every rational being as a Will that legislates universal law' (G4: 431; II, p.55.) The principle of autonomy is, therefore: "to choose only in such a way that the maxims of your choice are also included as universal law in the same volition" (G4: 440; II, p 80). See also Gerald Dworkin (1988, pp.4-15), *The Theory and Practice of Autonomy* for a comprehensive analysis.

<sup>33</sup> The normative principles of categorical imperative can be reduced and set as the rules of programming language once it is derived to natural language. Further, the reduction of natural language to programming language is not an impossible task, which will determine the activity of the A.I. machine. The self-determination of programmed robots, like human beings, can make decisions without any human intervention.

<sup>34</sup> Christine M. Korsgaard (2018, pp.5-12) illustrates that rationality sometimes defined as "the capability to perform prudently "(doing what is in your own best interests), acting with instrumental utility or acting with instrumental rationality (doing what will get you whatever ends you wish to achieve (doing what will get you whatever ends you wish to achieve)" which are easily compatible with AAIA. this definition of rationality cannot be construed as a correct definition for the Kantian thinker. The disagreement about the definition of rationality between Kantian and Anti-Kantian philosophers lies in the problem of normative conditions of rationality. Once we try to set the normative condition of rationality, it is quite possible to be fall in the trap of instrumentality of reason. Hence, reason as the precondition for establishing objective reality limited within human cognition.

<sup>35</sup> The view that the normativity of rationality can be traced in the natural world. See, Sturm T. & Gigerenzer, G. (2012) *How (Far) Can Rationality Be Naturalised*, for defence of ecological rationality (one version of naturalised rationality) which attempt to establish the relationship between cognitive heuristics with the environment. See also Konzelmann-ziv A. (2008) *Naturalised Rationality: A Glance at Bolano's Philosophy of Mind*, for argument in the support of naturalised rationality.

<sup>36</sup> See, Kolodny, N. (2015) *Why Be Rational* for the defence of non-normative rationalist.

<sup>37</sup> See, Quine (1969) *The Naturalised Epistemology* for comprehensive analysis. Quine in this text provides a critical outlook towards the traditional formulation of rationality. He has defended one kind of replacement naturalism.

<sup>38</sup> See, Nozick (1993, p.133) *The Nature of Rationality* for a version of normative rationality in which he argues that the norms of rationality can be traced from empirical resources

<sup>39</sup> See, Dretske (1995) *Naturalising the Mind* for naturalistic project of rationality.

empirical world. The non-naturalist perspective of defining rationality (i.e., Kantian rationality) always keeps the equal distance from psychological reductionism and natural reductionism.

Let us calibratedly look into the non-naturalistic definition of rationality to examine their limitation of universalizability.

Df. - X is rational *iff*, has a universalizable logical explanation of Y for the action Z, which justify it as right.<sup>40</sup>

The justification of rationality must have an isomorphic relation with logical structures and must follow the predeterminate principles of logic in order to have a valid justification. Hence, the universal valid justification for the action is the criteria of rationality.

The anticipated objection to this definition of rationality from reductionist, can be the limits of human cognition for having an infallible universal logical structure. Therefore, we cannot have a concept of rationality, which is in itself universal and true forever. We have evidence from the observation of the past that the valid logical structure is not valid forever. So, the change in logical structure can also change our perspective of rationality. However, this objection cannot be applied for bounded rationality.<sup>41</sup> If the ascription of rationality for agents is depended on certain abstract laws, rules, or logical principles and it necessitated by the abstract neural pattern of the human brain, then an artificial neural network can be created by replicating the neural network of the human brain which can generate rationality by following the same logical principles/rules as applied to the human brain<sup>42</sup>.

From the analysis of the debate on rationality, it is clear that the reduction of rationality to causal or functional effect of the neurological structure or abstract neural pattern of the brain entails the fact that AAIA is possible. If rationality is merely an inbuilt program to follow certain a-priorily given rules unconditionally, then also it is possible to have such kind of A.I. machines which are by default rational and being capable of making the decision independently. As we know, Kant argues that rationality entails the autonomy (or potential to be autonomous); the same logic can be applied to A.I. machines also if we accept the same normativity of rationality.

Now the question may arise that if the AAIA is possible then how the autonomy of A.I will be different from human Autonomy? The biological machine analogy construed the view that human beings are different from the machines as far the living biological organism is concerned; otherwise, functionally and behaviorally can be identical to machines. Machine acquires certain technique to perform an act, and do so as programmed by the programmers; similarly, an untrained brain (the newborn baby) learn from the environment (i.e., language acquisition, skills, environmental mapping, so and so forth), programmed itself as per the information collected from the environment, like the machine programmed. However, the difference is the capability of having multiple functions because of the more skillfulness of human beings than the machine, which is due to the limitation of programming. Human beings have multiple sources and techniques to be programmed than A.I. machines. The efficiency of programming determines the efficiency of a machine to perform the act in a skillful way.

The reductive-naturalist externalism also argues that the moral values are evolutionary products acquired by human beings through the process of natural selection. The newborn baby and animals cannot have these values because these values can only be acquired through moral training. However, Dretske<sup>43</sup>, one

---

<sup>40</sup> Here 'X' is the agent, Y is the justification for the action, and Z is the action itself. Some of the Kantian thinkers may disagree with this definition of rationality because of their acceptance of the view that reason as the A-priori necessity for human beings to establish something as valid. Hence reason is not limited within its own boundary. Therefore, reason in the first place as the foundation for any kind of discourse has to be accepted as the primary condition of knowledge, because no reason can establish it for the simple fact that they cannot go beyond their own limitations. Asking the question that what is rationality is itself the question of meta-rationality (it is the rationality about the rationality). In order to define what rationality is, we have to apply the reason, which is prone to the circularity of reason.

<sup>41</sup> Bounded rationality has accepted the flexibility of the normative condition of rationality. It claims for the complete rationality of A.I., unlike human rationality. See, Simon, H. (1982). "Models of Bounded Rationality" vol.1 for detail analysis

<sup>42</sup> See, Steinhäuser, M. et. al (2012) for neurological basis of rational activities and Decision making.

<sup>43</sup> See, Dretske (1995 pp.5-8) *Naturalizing the Mind* for detailed analysis of natural, conventional representational indicative function.

of the proponents of naturalistic externalism, argues that unlike other representational systems, human beings have the privilege of a *metarepresentational* power- ability to conceptually represent the sensory representational information. Other biological representational systems can have only the natural-systemic indicative representational function through which the particular system is able to provide only sensory representational information. The non-biological representational system can have only the conventionally assigned representational function due to the particular representational function is assigned by those biological representational system who has the privilege of a special kind representational function, called-conceptually acquired representational indicative function. In Dretske's naturalistic paradigm, even if these higher order functions can be naturalized, the non-biological representational system (i.e., A.I) can't have the metarepresentational power- conceptually represent sensory representational information. However, recent development in data analysis by A.I machines indicates that the limit of A.I can't be constrained within the information providing system. Because it can not only perform a specific assigned function (like provide a particular piece of information), but also can analyse and predict the future from the assigned indicative representational information. Hence, Dretske's objection to AAIA can be defeasible and it is logically possible that A.I machines can defeat human intelligence on conceptual representation. The naturalistic reduction of higher order properties to natural language is compatible with the possibility of A.S.I. Anything which can be converted to programming language can be artificially constructed and programmed. The ethical principles (i.e., ten commandments and other core value teachings of religion, categorical imperative so and so forth) can be programmed, consequently, the artificial moral agency is possible. We can have an artificial moral agent who may be more ethical than the human agent in their action in the Kantian paradigm of autonomy.<sup>44</sup>

The argument for the possibility of AAIA can only succeed if it is only a matter of intelligence and rationality. Intelligence and rationality are not the sufficient condition for the criterion for autonomy even though, are the necessary condition for autonomy and moral actions. What makes human agents special is the phenomenal experience from the first-person perspective.<sup>45</sup> Only a living organism (i.e., human agents) can have privilege of having phenomenal qualitative experience<sup>46</sup>, which an A.I. agent will always be deprived even though it can replicate all the human action, brain function and behavior. Human agents have not only phenomenal subjective experiences but also can have privilege of cognitive skill to use the information of phenomena or experiences for creative acts and can understand the content of the experience of others through analyzing their own content of the experience in the similar given environmental input, which cannot be possible in case of A.I. agents. An A.I. machine can act like an autonomous human agent by creating the delusion of the replicated-behavior of an autonomous agent; however, in reality, it is just a machine, subject to deterministic patterns of behavior and bound to perform assigned function. Even if artificial super intelligence can have rationality in a loose sense,<sup>47</sup> It will still be like a philosophical zombie. There can be nothing to experience something from inside in the first-person perspective for a zombie. It is quite uncontroversial that rationality and intelligence are not the only criterion of autonomy even though are the necessary conditions of autonomy.

---

<sup>44</sup> The objection against this model can be raised that the AAIA is itself an imaginary friction. Even if we inherited it as a logical possibility, the determination of the action by human agents, as argued by Kant, will be different from A.A.I.A, for the simple fact that the determination of the Will of the agent through the categorical imperative is given in terms of self-determination of the Will and conducive to the principles of alternative possibility. To be moral or immoral is deepened upon the Will of the human agent, which means the possibility of doing the action otherwise. The AAIA agent (will either fall into the category of by-default to follow the moral laws or otherwise in which there cannot be autonomy), being the product of human intelligence, cannot have the choice to do it otherwise. The determination of autonomy will rely on human intelligence until it surpasses human intelligence, and as we know, human intelligence is determinant of A.I and remains forever. Therefore, the AAIA is imaginary friction.

<sup>45</sup> See, Sahu M.K. (2019, pp.77-83) "Qualia and Intentionality" for detailed analysis of phenomenal first-person accessibility. This paper defends the ontological subjectivity of phenomenal qualities of experience. Deborah, J.H. (2006) also argues against the possibility of artificial moral agency.

<sup>46</sup> See, Sahu M.K. (2020, pp.160-66) "Phenomenal Consciousness" for the argument of phenomenal consciousness.

<sup>47</sup> Kant construed rationality is a non-representational form, which is a-priorily given in the consciousness of human Agents, and cannot have a mechanistic explanation at all by virtue of the fact that it is the nature of abstraction and untenable in experience. Kantian thinkers tried to block the possibility of rational A.I. agency through this argument. However, the term rationality itself has elusive and ambiguous usage. They could not give a satisfactory account for rationality; therefore, A.I. rational agency can be defended by giving a mechanistic theory of rationality. Russell (1997), in his article "*Rationality and Intelligence*," defended such kind of position. He has classified four categories of rationality in this article and argues for Artificial Rational Intelligence.



Phenomenality, as the sufficient condition, must be accepted for the sake of realization of free-will. Even if an A.I. agent can have rationality in a minimalistic sense without phenomenality; however, can't have autonomy without phenomenality; therefore, phenomenality becomes the criterion for autonomy, accessible to the living organism alone.

Another approach to transcendentalism can be the reading of Kant from the Transcendentalism-Naturalism Compatibilist (hence after TNC) perspectives, which advocate for one kind of transcendental realism<sup>48</sup>. Only rational agents can have phenomenal pre-conditionality; therefore, it excludes the non-rational animal as well as artificially constructed machines from having the phenomenal pre-conditionality. The TNC's interpretation of the Kantian phenomenal self can succeed in blocking the knowledge claim of the possibility of AAIA; however, the problem in this approach can be easily detectable. For the sake of transcendentalism of the self, its interconnection with the world has to be diminished in order to preserve the sanctity; In that case, it will give rise to an unbreachable ontological as well as explanatory gap between phenomenal-transcendental self and its further interconnection with the world becomes impossible. Therefore, Kantian transcendentalism for AAIA is not a bigger challenge with comparison to phenomenal realism.

Furthermore, let's consider the compatibility of transcendentalism with phenomenal realism for a while-the view that the transcendental account of phenomenal pre-conditionality (the demand of reason for phenomenal consciousness as the precondition of the phenomenal qualitative aspect of experience) as a transcendental form (like other categories) discoverable through the application of transcendental method of reasoning, hence, provide an objective method of inquiry of the subjective mental states without any appeal to phenomenal perspectivism (as advocated by Berkeley-and his successor theories, phenomenalism, and perspectivism). Phenomenality as an a-priori transcendental form can play the role of a synthesizer of the subjective-objective dichotomy of experience.<sup>49</sup> It can function as a-priori<sup>50</sup> condition for phenomenal experience without phenomenal content (enables us for defending non-reducibility of phenomenal qualitative experience in Kantian sense); because of which an organism has the privileged access of the phenomenal experience, is objective to the particular living organism; must be outside the domain of sensibility, hence, it is the demand of reason to accept the phenomenal pre-conditionality (phenomenal self) as the foundation for every phenomenal experience. then also it can raise the objection against AAIA, however, became inconclusive and violate the principles of transcendentalism.

### 3. Conclusion

The resolution of Kant's third antinomy through abstraction, though established the ground for both moral and natural principles, cannot give sufficient reason for the connection between the autonomy of agents with the phenomenal world. Kant has the ontological commitment to establish the existence of the self in the first place; therefore, he adopted the pure abstraction method, and denied the existence of phenomenal reality. Transcendental realism somehow manages to avoid this obscurity by arguing that the knowledge about the phenomenal reality is different from phenomenal reality. Kant has this misconception that reality has to be transcendental and unaffected by the sensible realm because of the fallibility of the phenomenal realm. Thomas Nagel, one of the most influential philosophers of the 20th century, concurs with Kant about the non-physiological concept of reasoning by arguing for the objectification of knowledge; however, he differs by holding "*objectification-without-reification*" strategy. Kant has an ontological commitment, which Nagel tried to avoid in his work. If we critically

---

<sup>48</sup> See, Loudon (2019, pp. 8-12) Kant the Naturalist for more details.

<sup>49</sup> One has to be careful while interpreting the transcendental a-priori of reason for the defence of phenomenality as a new category either of the sensibility and understanding or an independent category from both sensibility and understanding. Because it is just one step away from falling in the trap of Berkelian empirical subjectivism. 0

<sup>50</sup> Here the a-priori condition is used as the transcendental a-priori condition distinct from synthetic a-priori condition of reasoning. Allison (1971, p.187) has the similar kind of view that the misinterpretation of the a-priori conditions always leads to entanglement of Kantian transcendental subjectivism with Berkelian empirical subjectivism. Kant, unlike Berkeley, has rejected the view that the objects of experience are private sense data (as followed by the Russellian sense datum theory).



analyze the Kantian solution to the problem of autonomy, then we can find out that except the Kantian ontological commitment, his epistemological standpoint is compatible with AAIA. However, Kantian autonomy is incompatible with any naturalistic and non-naturalistic theory of autonomy because of its unique nature of self-legislation and self-determination of moral laws of the transcendental self. The floccinaucinihilipilification of phenomenal reality and the reduction of the external world into the appearance of representation by Kant makes the job easy for A.I. scientists.

Rationality without phenomenality leads to abstract objectivism, whereas phenomenality without rationality leads to subjectivism and relativism. Phenomenality is the necessary condition for the realization of the effect of the self-determination of the Will even if we accept the Kantian thesis for a while as a plausible thesis. The only difference between naturalists and Kant is the logical implication of the transcendentalism of autonomy, which deviates Kant from the naturalistic camp. Therefore, except for the idea of transcendental self, the categorical imperative as a command can be implemented in the A.I system, hence, is compatible with the AAIA thesis. From the evaluation of the paradigm of rationality (as we have found that rationality sets its condition on normativity and certain principles can also be implemented in a rational, autonomous agent), it is crystal clear that the compatibility of transcendental naturalism with AAIA is undeniable. Therefore, the objection from phenomenal realism to AAIA is more plausible than the Kantian objection to AAIA.

Acknowledgements: I would like to thank my supervisor Prof. Ratikanta Panda for his constant help, support and encouragement, being aware about his responsibility of supervising without which the substantive revision of the manuscript might not be possible. I am also grateful to the anonymous referees for their quick review and constructive feedback. I would also like to thank Prof. Pravesh Jung for his valuable comment on the paper.

## References

- Allison, H.E. (1971), Kant's Transcendental Humanism, *The Monist*, Vol. 55:2, pp. 182-207
- Allison, H.E. (1990), *The Third Antinomy, Kant's Theory of Freedom*, New York: Cambridge University Press
- Bostrom, N. (2014), *Superintelligence: Paths, Dangers, Strategies*, Oxford, UK: Oxford University Press.
- Copeland, B. J. (1998) *Super Turing Machines, Complexity*, Vol. 4, pp. 30-32.
- Christine, M. K. (2018) Rationality, In Lori Gruen (ed.), *Critical Terms for Animal Studies*, Illinois: University of Chicago Press.
- Deborah, J.H. (2006) Computer Systems: Moral Entities but Not Moral Agents, *Ethics and Information Technology*, Vol.8, pp.195-204
- Dretske, F. (1995) *Naturalizing the Mind*, London: MIT Press
- Dworkin, G. (1988) *The Theory and Practice of Autonomy*, New York: Cambridge University Press.
- Gillespie, M. (2014), Philosophy and Rhetoric in Kant's Third Antinomy, *The Political Review*, Vol. 30, pp.7-33
- Guyer, P. and Wood, W. A. (1998) *Critique of Pure Reason*, New York: Cambridge University Press
- Guyer, P. (1987) *Kant and The Claims of Knowledge*, Cambridge: Cambridge University
- Kahneman, D. (2011) *Thinking, Fast and Slow*, New York: Farrar, Straus and Giroux Press

- Kolodny, N. (2015) Why Be Rational, *Mind*, Vol. 114, pp.509-63
- Konstantin, P. (2013) Naturalism and Kant's Resolution of The Third Antinomy, In Margit
- Ruffing, Claudio La Rocca, Alfredo Ferrarin & Stefano Basin (eds.), *Kant Und Die Philosophie in Weltbürgerlicher Absicht: Akten des Xi. Kant-Kongresses*, 2010, pp. 731-742.
- Konzelmann-Ziv, A. (2008) Naturalized Rationality: A Glance At Bolzano's Philosophy of Mind, *Baltic International Yearbook of Cognition, Logic and Communication*, Vol.4:1, New Prairie Press
- Kurzweil, R. (2006) *The Singularity Is Near: When Humans Transcend Biology*, New York: Penguin USA.
- Louden, R.B. (2020) Kant the Naturalist, *Journal of Transcendental Philosophy*, Vol. 1:1, pp. 3-17
- Mirazo-Ruiz K. & Moreno A. (2012) Autonomy in Evolution: From minimal to complex Life, *Synthesis*, Vol.185, pp.21-52
- Nozick, R. (1993) *The Nature of Rationality*, New Jersey: Princeton University Press
- Reimer T. & Hoffrage U. (2004) Models of Bounded Rationality: The Approach of Fast and Frugal Heuristics, *Management Revue. The International Review of Management Studies*. pp. 437-459
- Sturm T. & Gigerenzer, G. (2012) How (Far) Can Rationality Be Naturalised, *Syntheses*, Vol. 187, pp. 243-268
- Quine, W.V.O. (1969) *Ontological Relativity & Other Essays*, New York: Columbia University Press
- Russell, E. (1999) Rationality and Intelligence, In Wooldridge M., Rao A. (eds) *Foundations of Rational Agency*, Applied Logic Series, Springer, Dordrecht, Vol.14, pp.11-33
- Sahu, M.K. (2019) Qualia and intentionality, *Journal of All Odisha Philosophy Association*, Vol. 5:4, pp.76-87
- Sahu, M.K. (2020) Phenomenal Consciousness: A Critical Analysis of Knowledge Argument, Inverted Spectrum Argument and Conceivability Argument, *Journal of Advances in Education and Philosophy*, Vol. 4:4, pp. 160-166
- Simon, H.A. (1982) *Models of Bounded Rationality*, Vol. 1. London: MIT Press
- Smith, N. K. (1918) *A Commentary to Kant's Critique of Pure Reason*, London: Palgrave Macmillan Press
- Steinhauser, M., Achtziger, A., Alös-Ferrer, C., Hügelschäfer, S. (2014) The Neural Basis of Belief Updating and Rational Decision Making, *Scan*, Vol. 9, pp.55-62
- Turing, A. (1950) Computing Machinery and Intelligence. *Mind*, Vol. 59:236, pp. 433-460.
- Willaschek, M. (2001) The ambiguity of the Kantian Distinction between things in themselves and appearance: On the debate about two-aspect and two world interpretations of the transcendental idealism, In V. Gerhardt, R. Horstmann & R. Schumacher (ed.) *Kant and the Berlin Enlightenment*, Vol.9, pp.679-790, Boston: De Gruyter

<https://doi.org/10.34024/prometeica.2021.23.12411>

## METHODOLOGICAL CHALLENGES FOR A NEW PHILOSOPHY OF TECHNOLOGY

INTERVIEW WITH NICOLA LIBERATI

*DESAFÍOS METODOLÓGICOS PARA UNA NUEVA FILOSOFÍA DE LA TECNOLOGÍA*

*Entrevista a Nicola Liberati*

*DESAFIOS METODOLÓGICOS PARA UMA NOVA FILOSOFIA DA TECNOLOGIA*

*Entrevista com Nicola Liberati*

**Cristina Pontes Bonfiglioli**

(Universidade de São Paulo)

khryz@usp.br

Recibido: 12/07/2021

Aprobado: 12/07/2021

*Nicola Liberati's research focuses on the effects of emerging digital technologies related to intimacy from a phenomenological and postphenomenological perspective. He holds a Ph.D. in Philosophy, and he is an Associate Professor at the Shanghai Jiao Tong University. Before moving to China, he worked in Japan as a JSPS Postdoctoral Fellow and in the Netherlands at the University of Twente with Peter-Paul Verbeek. In 2016, he was elected Co-chair of the Society for Phenomenology and Media. He won international prizes like the MASH'D Full Paper Award at the world-leading IEEE conference on Augmented Reality ISMAR2015 and the SPM2014 Best Paper by a Returning Member Award. In 2021, he was granted the Shanghai Overseas High-Level Talent Program, one of China's most prestigious awards. Newspapers and news agencies like Reuters, de Volkskrant, CNET, Happinez, and U-Today interviewed him, and commissions like UNESCO COMEST use his work to highlight the impact he has on societal issues. He has more than 50 publications, including a special issue on Prometeica on science-fiction and new humanities, which has been already accepted.*



Associate Professor at Shanghai Jiao Tong University  
(Dept. of Philosophy, School of Humanities)

### **1. *Why Shanghai Jiao Tong University? Was it an opportunity, or did you have plans to go to an Asian university?***

I was looking to come to China because it is growing fast and tightly connected with digital technologies. So, I thought that there were many opportunities for developing my research. In China, at least in my field, you can do high-risk, high-gain research. Thus, you have the freedom to make complicated choices and explore a new line of research freely. Before China, I worked in the Netherlands, and I wanted to change, so I started to see possible different countries. I was looking for positions in Europe and the USA at first, but when I found this interesting opening position at *Shanghai Jiao Tong University*, one of the best universities in the world, and I thought, “Why not try it?”. The University paid for my flight to attend the interview. When I arrived, I visited the campus, the department, and the city, and I enjoyed them very much. The only problem with this position in China is that it is very far from Italy, where my family lives.

China is an important country in many fields, and, if you look at the statistics in research, very soon, China will be the first country in the world for research in terms of articles and projects. Moreover, most Universities in Europe where I applied could not offer the same working conditions like China. The position I got grants me great freedom, and it comes with a generous research grant to develop my research. Also, my previous experience in Asia has been much appreciated. And because I was a *JSPS fellow* in Japan, which is similar to the *Marie Skłodowska-Curie* European grant, was noticed this during the selection process. Thus, I felt the committee appreciated and evaluated all the different steps in my career, which touched countries out of the European Union like Japan and Mexico.

### **2. *Do you consider yourself a continental philosopher?***

I do not know because even the distinction between continental or analytic philosophy is very “western-centric.” Of course, my Ph.D. has been developed in Italy on Husserlian phenomenology and Postphenomenology, so I would pick the continental side if I had to choose. However, at the same time, I think this classification is a little bit reductive, and it might not grasp the distinctions existing in other contexts like the Chinese one. For example, in China, the difference between ‘western thought’ and ‘Chinese tradition philosophy’ might be more relevant than continental and analytic tradition. So, I do not think it is possible to follow this distinction too much.

### **3. *In what department are you working? What kind of research are you and your colleagues developing?***

I am in the Philosophy department, and, in my department, there are three main lines of research. A group of researchers is working on traditional Chinese philosophy, focusing on the relation between philosophy and religion and different ways of thinking in China. Another group is working on Epistemology and Philosophy of Science. A third group is working on European philosophers, focusing on phenomenology and standard philosophers based in Europe from the last centuries.

### **4. *And you are in this last one?***

No, I am in Epistemology (laughs) concerning the Philosophy of Science, and I know that in some way, it is a little bit tricky because I work in-between these groups. On the one hand, I am working with Philosophy of Technology, so I am close to Epistemology and Philosophy of Science. On the other hand, I work with Phenomenology, and so I am also close to the group working on European philosophers. I have always situated my interests in different fields. Postphenomenology is in-between Phenomenology and Philosophy of Technology, and I work with artists, engineers, and designers, so I am in philosophy, but, at the same time, I also work with other fields.

### 5. *What are your main research interests today?*

I have come to China because they allowed me to be completely free in what I do by developing my transdisciplinary research. So, I try to work with artists and computer scientists to reshape the kind of questions we have on the effects of emerging digital technologies. I say I work transdisciplinary because one thing is to study these effects from our perspective as philosophers. Another thing is to work with engineers, artists, and designers to reframe the philosophical questions by looking at the technologies used as the source of the effects. The empirical turn in the philosophy of technology already moved a bit in this direction. Its main idea is to organize a round table where every kind of expert sits together, and they try to have a dialogue starting from their perspective. This work is interdisciplinary and multidisciplinary, but everyone keeps their own identity and position within a discipline. I do not want to have just this dialogue among disciplines. I want to try to do philosophy by using other fields to reshape the question in philosophy according to the elements provided by engineering, arts, and design. I use engineering, arts, and design because I am close to them.

### 6. *I wondered: do you think of social change when proposing these kinds of studies? Do you think it is about changing the way people behave more than the way people think?*

I would say that it is about changing the way people think. Digital technologies are not neutral, and they have a deep impact on how we feel about ourselves and how we think of ourselves as human beings. We need to address these kinds of effects to understand better who we are and the society we live in.

Usually, Philosophy provides us a kind of interpretation of the main aspects of a specific technology thanks to the perspective of a particular philosopher. This standard approach is advantageous, especially for my work, but I want to move a little beyond it. The empirical turn, in a way, already started aiming at a social impact by opening a dialogue with other disciplines like engineering. Now, I suppose, it is time to shape philosophy by using other fields to tackle better the societal challenges we all face. It is not just a question of a dialogue among disciplines. We need new words, ideas, and tools. We need to look at the latest digital technologies and the collaborations with artists, engineers, and designers as valuable sources to generate them.

### 7. *And how do you do that? How do you execute your research project?*

I focus on the body of people and intimacy. Everything in my research works on the human body understood as 'flesh' and 'living body.' I am interested in how we live with computers and how we live with others when digital technologies are on and around us. I have three different lines of research about intimacy: smart textiles, digital technologies for sex, and augmented reality.

One of my last works on the journal *Humana.mente* is on smart textiles. Some clothes have digital technologies interwoven in them, which means it is possible to have clothes reacting to people's moods. Clothes can change colors, shape, and position according to how a person feels or how they live in the world. So, what happens to fashion when the mood becomes part of the clothes? What happens to the way we think of clothes when we can 'smile' with them? What if your clothes become blue or transparent when you are sexually aroused, for example? What happens to your way of relating to these feelings and emotions the moment they are visualized by everyone? What are the meanings and values you give to your inner life when it becomes visible to everyone?

As we know, clothes have deep repercussions on many elements of our life like the constitution of the self, the position in society, and social identity. So, from a phenomenological perspective, when there is such a dramatic change in clothes because they become digitally embedded, it becomes interesting to see what happens to the subject and society.

I also study robotics in relation to sexual activities, focusing on how we can be digitally together. There are different kinds of sexual technologies that allow a person to experience sexual intercourse. Sex robots, for example, are essential for the implications they have on the values and meanings we give to



love and sex in general, even with other human beings. What happens to our way of thinking about love and sex when thinking about having sex and love with robots? There are also haptic sexual technologies like *teledildonics*, which allow you to touch and feel touched in physical bodily sexual intercourse. When you use these devices, you are “in touch” with the other person but differently from a “normal” physical situation. What happens to the way of being with another person the moment we use these technologies? My publications on *teledildonics* in *Science and Engineering Ethics* and *Paladyn Journals* study these elements and how digital technologies shapes the values we give to our intimacy.

Moreover, I study Augmented Reality and its implication to what we mean with ‘real’ and ‘physical’ in our surroundings. This line of research is related to my Ph.D. thesis, and it is about the meanings and values we give to our world the moment we can have objects made of ‘digital materiality’ around us.

I found that I cannot just address these kinds of questions in a traditional philosophical way. For instance, I can use Merleau-Ponty’s idea of touching-touch for *teledildonics*. Does this kind of touching-touch relation tell you something about what it means to have sex with another person through *teledildonics*? Yes, to some extent, but maybe there are aspects of touch and intimacy that become visible only through the use of these technologies. So, we need to reshape our ideas and words according to the technologies we use to understand better the change in values and meanings we face.

**8. *So you hypothesize that these experiences that are being promoted by all these new technologies are going to generate a new kind of humanity, other values...?***

Yes! I do not think that values and meanings are untouched by the technologies we use. We have many different examples already, like the ultrasound pre-natal technologies. What does it mean to be a father or a mother when you or your partner are pregnant, and the technologies show you some problems in the fetus? These technologies force the parents to make choices that emerge only because of the action of the technologies. So, what it means to be a father and a mother changes according to the devices you can use.

Regarding sexual technologies, condoms change what we mean and the values we give to sexual acts. Condoms are used for ‘protection,’ but this ‘simple’ element shows that the choice of using or not using the condom makes visible the level of protection you want as lovers. Thus, the possibility of using condoms reveals the kind of trust the lovers have that before was not so directly visible. Our world, the values, and the meanings we give to our life change with the introduction of new digital technologies like sex robots and *teledildonics*, and it is our work to understand these changes to live better.

**9. *When you say “robot,” people start thinking of a machine as a version of C3PO from Star Wars (George Lucas, 1977). But in Blade Runner 2049 (Denis Villeneuve, 2017), Ryan Gosling’s character falls in love with a robot that suggests a threesome with a human, and he does not want it at first. Is this the kind of threesome between robots and humans that you have in mind? Are there examples of real robots that are having sex with people?***

Well, that is why I’m not particularly eager to talk too much about robots because we do not have anything like that today while we already have other technologies for intimacy like *teledildonics* which are already available on Amazon for around 100 USD. However, I think it is vital to keep futuristic imaginaries in perspective. For example, my last publication in the *International Journal of Technoethics* is on a threesome with a robot and two human beings. The effects on values and meanings in love and sex implied by it.

Especially in the last year, there has been a detachment between Philosophy of Technology and Science-fiction, highlighting how much philosophy is not about novels and future scenarios. Still, it is something applied to the technologies we already use. I suppose this change is helpful because it can show companies and the general public that what philosophers do is related to the world we live in and real-world cases. However, I think this position might have some limitations.

Science-fiction is a way to envision the future or at least a possible future scenario and present issues and topics not yet present in our society. Since philosophy is usually seen as a study of something which has already passed, and science-fiction provides a look to the future, why not intertwine them? For this reason, I am working on a conference on science-fiction and philosophy. The main idea is to show that it is possible to mesh philosophical methods and science-fiction visions of the future to address the effects of new digital technologies.

Digital technologies change who we are and how we live, and we need a transdisciplinary perspective to tackle the societal challenges we are all facing.

## References

Liberati, N. The Borg-eye and the We-I. The production of a collective living body through wearable computers. *AI & Soc* 35, 39–49 (2020). <https://doi.org/10.1007/s00146-018-0840-x>

Liberati, N., Nagataki, S. Vulnerability under the gaze of robots: relations among humans and robots. *AI & Soc* 34, 333–342 (2019). <https://doi.org/10.1007/s00146-018-0849-1>

Liberati N. (2018) Achieving a Self-Satisfied Intimate Life Through Computer Technologies?. In: Altobrando A., Niikawa T., Stone R. (eds) *The Realizations of the Self*. Palgrave Macmillan, Cham. [https://doi.org/10.1007/978-3-319-94700-6\\_13](https://doi.org/10.1007/978-3-319-94700-6_13)

Liberati, N. Teledildonics and New Ways of “Being in Touch”: A Phenomenological Analysis of the Use of Haptic Devices for Intimate Relations. *Sci Eng Ethics* 23, 801–823 (2017). <https://doi.org/10.1007/s11948-016-9827-5>

<https://doi.org/10.34024/prometeica.2021.23.11835>

**PRECIADO, P. B. UM APARTAMENTO EM URANO. CRÔNICAS DA TRAVESSIA. SÃO PAULO: ZAHAR. ISBN-13: 2020 978-8537818831**

*PRECIADO, P. B. AN APARTMENT ON URANUS. CHRONICLES OF THE CROSSING. SÃO PAULO: ZAHAR. ISBN-13: 2020 978-8537818831*

*PRECIADO, P. B. UN APARTAMENTO EN URANO. CRÓNICAS DE LA TRAVESÍA. SÃO PAULO: ZAHAR. ISBN-13: 2020 978-8537818831*

**Breno Benedykt**  
(Universidade de São Paulo)  
*breno.benedykt@usp.br*

Recibido: 04/03/2021

Aprobado: 01/06/2021



*Um apartamento em Urano: crônicas da travessia*,<sup>1</sup> lançado em espanhol em 2019 e em 2020, em português, pela editora Zahar, com tradução de Eliana Aguiar, reúne 73 crônicas escritas pelo filósofo Paul B. Preciado, entre os anos de 2010 e os primeiros meses de 2018, publicadas, em sua maioria, no jornal francês *Libération*. Parte dessas crônicas já circulavam em português através de blogs interessados em difundir as reflexões do filósofo espanhol. Esse é o caso, por exemplo, da crônica, *O feminismo não é um humanismo*. Nela, Paul traça o horizonte de uma conexão entre feminismo e animalismo, a qual, diferentemente do naturalismo presente em certas correntes conservadoras do feminismo, opõe-se à naturalização dos gêneros que opera como pressuposto que exclui do feminismo transsexuais sob o argumento de que seriam mulheres artificiais. Paul, que em nascimento foi identificada como menina, conta-nos, por meio de suas crônicas, como, na infância, já sonhava com um mundo utópico, cuja primeira representação estaria em um de seus desenhos de escola, o qual acabou por lhe render, aos oito anos, assim como à sua

família, uma série de consequências normativas desastrosas. Nele, Beatriz representava humanos em harmonia com os animais e sua companheira amorosa, uma amiga da escola. Porém, se Paul nos remete a essa cena do passado, é apenas para nos ajudar a compreender o presente e com isso abrir o horizonte para o futuro. Não à toa, aquilo que Paul mais nos narra em relação às transformações de seu *si* “sexual” dizem respeito à travessia de suas passagens jurídicas e biológicas em homem trans. O uso da testosterona, a mudança de voz, as dificuldades nos aeroportos quando a discrepância entre a foto antiga e o rosto novo começaram a se acentuar, a mudança jurídica de nome, de Beatriz para Paul Beatriz - apenas para citar algumas das cenas presentes nas crônicas. Mas, se Paul nos dá o mapa aberto de uma cartografia de si, essa cartografia está totalmente atrelada a primorosas reflexões filosóficas e políticas. São ao menos duas as crônicas que discutem o método filosófico que atrela escrita, biografia e política. O método de uma vida filosófica engajada politicamente de modo entusiasmado com o seu presente, *O*

<sup>1</sup> PRECIADO, P. B. *Um apartamento em Urano: Crônicas da travessia*. São Paulo: Zahar, 2020.

*método Marx*, e o método de uma escrita onde a vida se faz presente, o método Derrida/Foucault, para não dizer, também, Deleuze-Guattariano. É por meio da lente desses pensadores que Paul, atenta como Walter Benjamin a todos os detalhes do cotidiano, também se volta a realizar uma ampla crítica à cultura do capitalismo em sua modalidade técnico-neoliberal, a qual vai de jogos de celular a filmes de sucesso, como *Ninfomaniaca* e *Azul é a cor mais quente*, até biografias. Mas, se a lente se aproxima desses autores ditos pós-estruturalistas, a filiação é mais ampla e a luta contemporânea os ultrapassa, o que o leva a se contrapor publicamente a nomes como os de Michel Onfray, que atacou Judith Butler com argumentos extraídos de blogs filiados ao tradicionalismo reacionário canadense. É assim que Paul se move de crônica em crônica, muitas delas escritas em hotéis e aeroportos, ao longo de seus deslocamentos entre diferentes cidades: Paris, Atenas, Barcelona, cidades que ama, Nova York, cidade na qual morou oito anos, enquanto estudava na New School for Social Research de Nova York, além de tantas outras como Buenos Aires, Lyon, Calella, Kiev. É graças a essa segunda travessia, inseparável da anterior, que Paul também nos dá oportunidade de ver pelos seus olhos o que era a ocupação da Praça Syntagma, em Atenas, em 2011, o ar revolucionário que se conectava com aqueles que estavam nas ruas ou praças de Madri, Nova York, Paris. Mas suas crônicas também nos trazem explicações importantes sobre o que acontecia lá de específico, tanto em relação à constituição de uma experiência revolucionária do comum, como em relação às questões econômicas, geográficas e políticas que atravessavam cada país; como os efeitos na Grécia da invenção geopolítica da noção de Sul da Europa, uma vez que sempre há a invenção de um Sul pobre e endividado e por detrás dela há a imposição de um Norte rico, como nos explica sua crônica, *O Sul não existe*.

Pensador que trafega por mundos teóricos denominados genericamente de *queer*, feminista, pós-estruturalista e decolonial, Paul afirma, ao lado de uma série de pensadoras, como Gayatri C. Spivak, Judith Butler, Monique Witting, Virginie Despentes e Donna Haraway, que agora se impõe de modo decisivo, o que chama, a partir de Thomas Kuhn, de uma revolução de paradigma, tanto epistemológico como estético e político. O que essa revolução implica? Se pensarmos em algumas de suas entrevistas e no modo como sua experiência filosófica se dá e se exprime, inclusive nas crônicas, trata-se de reavivar uma tradição filosófica que Michel Foucault se esforça por restituir em seu curso *A coragem da verdade*,<sup>2</sup> a qual compreende que a atitude filosófica, sua experimentação ética e estética, é primeira em relação à sua erudição. Aquilo que para os pós-estruturalistas e, talvez até mesmo para Walter Benjamin, era uma borda, aqui converte-se num meio. Mas epistemologicamente, se voltarmos à crônica *O feminismo não é o humanismo*, a qual diga-se de passagem perverte o título do famoso ensaio do filósofo Jean Paul Sartre -, trata-se de desfazer, desmontar, desconstruir ou até mesmo destituir todos os binarismos que sustentam os fundamentos universais do humanismo europeu, isto é, os axiomas que sustentam o modo normativo e impositivo de se entender liberdade, igualdade e fraternidade: “O animalismo revela as raízes coloniais e patriarcais dos princípios universais do humanismo europeu. O regime da escravidão e depois o do salário como fundamento da ‘liberdade’ dos homens modernos; a guerra, a competição e a rivalidade são os operadores da fraternidade; a expropriação e a segmentação da vida e do conhecimento, o reverso da igualdade”<sup>3</sup>. Mas, toda essa negatividade só se dá graças à existência de um avesso de positividade, um feminismo voltado a uma expansão ilimitada das alianças entre todos aqueles que experienciam, como nos ensina Spivak, uma condição de subalterno diante daquilo que Paul chama, em uma de suas crônicas de, *Ne(©r)oliberalismo*. “Porque”, afirma Paul em outra de suas crônicas, “a revolução atua através da fragilidade”<sup>4</sup>, e não através daqueles que têm coragem de impor, por meio da lei ou da norma, o modo adequado de se existir biológica ou subjetivamente. Por isso, Paul parece estar sempre nos convidando a reaprender a imaginar politicamente o meio, onde se traçam os modos moleculares de alianças, de experimentações e de transformações de vida que nos foram dadas como inimagináveis: indigenismo *queer*, pansexualidade planetária, tecnoxamanismo, assembleia solene de plantas e flores em torno das vítimas da história do humanismo. Uma nova epistemologia que passa necessariamente pela atenção às variações afetivas que nos atinge tanto pela exterioridade das lutas políticas como por aquela, igualmente sensível, da escrita, onde mais do que eu, é o *si* ou a vida que se

<sup>2</sup> FOUCAULT, M. *A coragem da Verdade: O governo de si e dos outros II*. São Paulo: WMF Martins Fontes, 2011.

<sup>3</sup> PRECIADO, P. B. *Um apartamento em Urano: Crônicas da travessia*. São Paulo: Zahar, 2020, p. 133.

<sup>4</sup> Ibidem, p. 142.

põe à prova, pois são nesses casos que a cada lance pode ganhar luz uma pequena, mas decisiva mudança. Toda uma aventura real, ética, estética e política de uma ciência nômade ou minoritária. Como dizia Deleuze e Guattari: “*ciências ambulantes, itinerantes, que constituem em seguir um fluxo num campo de vetores no qual singularidades se distribuem como outros tantos ‘acidentes’*”<sup>5</sup>.

“‘Pode o subalterno falar?’: a pergunta que Gayatri C. Spivak fazia pensando nas complexas condições de enunciação dos povos colonizados ganha agora um sentido distinto”, avança Paul em relação à sua travessia trans e política-imaginativa, “E se o subalterno fosse também uma possibilidade sempre já contida em nosso próprio processo de subjetivação?”<sup>6</sup>. Como deixá-lo falar? Com que voz? É o começo, nos diz Paul, da “guerra dos mil anos; a mais longa das guerras, pois afeta as políticas da reprodução e os processos através dos quais um corpo humano se constitui como sujeito soberano. A mais importante das guerras, portanto, porque o que está em jogo não é o território ou a cidade, mas o corpo, o gozo, a vida”<sup>7</sup>. Guerra esta que se trava em qualquer um e em qualquer lugar, na cultura pop e na erudita, nas casas de BDSM e nas praças ocupadas, nos museus e na academia, na escola e no trabalho, na escrita e na cama, nos filmes e na escrivadinha. Resta a nós aprender a dançar com Paul. *Um apartamento em Urano: Crônicas da travessia*, talvez seja o melhor dos começos.

---

<sup>5</sup> DELEUZE, G; GUATTARI, F. *Mil Platôs: Capitalismo e esquizofrenia* 2, vol. 5. São Paulo: 34, 2012, p. 42.

<sup>6</sup> PRECIADO, P. B. *Um apartamento em Urano: Crônicas da travessia*. São Paulo: Zahar, 2020, p. 176.

<sup>7</sup> Ibidem, p. 303.



**PEREIRA, GUSTAVO. EL ASEDIO A LA IMAGINACIÓN, EDITORIAL COMARES, 2018, 244 PP, ISBN: 978-84-0000-000-0**

*PEREIRA, GUSTAVO. THE SIEGE OF IMAGINATION, EDITORIAL COMARES, 2018, 244 PP, ISBN: 978-84-0000-000-0*

*PEREIRA, GUSTAVO. O CERCO DA IMAGINAÇÃO, EDITORIAL COMARES, 2018, 244 PP, ISBN: 978-84-0000-000-0*

**Juan Cruz Apcarián**

*(Universidad Nacional del Comahue  
Instituto Patagónico de Estudios de Humanidades y Ciencias Sociales  
Consejo Nacional de Investigaciones Científicas y Técnicas)  
juanapcarian@hotmail.com*

Recibido: 24/06/2021

Aprobado: 16/07/2021



El asedio a la imaginación es una propuesta de Gustavo Pereira para pensar en escenarios que instalen, preserven o restauren, condiciones de justicia social y de libertad. En esta publicación, de Editorial Comares (2018), Pereira invoca el rol axial que cumple la imaginación en el desenvolvimiento de nuestras vidas prácticas. La imaginación, en cuanto habilita la representación de estados de cosas alternativos, articula las distintas vías y cursos de acción posibles que puede transitar una persona en vistas de la realización de un buen vivir. Pero este señalamiento de la imaginación es el requerimiento de un diagnóstico más profundo por parte del autor, que encuentra a la imaginación práctica como el blanco de un asedio continuo e institucionalizado que aceptamos y reproducimos sin darnos cuenta y de manera paradójica, ya que dicho asedio significa un menoscabo de las posibilidades y del ejercicio de nuestras libertades. En este sentido, el libro aquí presentado se podría pensar como un esfuerzo por contrarrestar las tendencias, instituciones y dinámicas alienantes

de la condición humana.

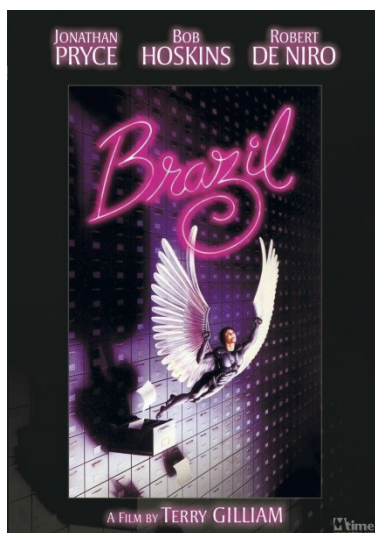
El estudio está estructurado en tres partes. La primera, “La vida práctica: imaginación, autonomía y reflexión”, profundiza en lo que define y constituye a la vida práctica, a los fines de establecer un criterio normativo que permita determinar cuándo, o en qué contextos, la libertad de un agente se encuentra vulnerada. Se desarrollan aquí las diversas formas que asume la vida práctica a través de las nociones de autonomía, de reflexión y de imaginación práctica. La segunda parte, “Patologías sociales como bloqueo de la imaginación”, analiza de qué manera en las sociedades contemporáneas se institucionalizan prácticas alienantes, a partir de la revisión histórica de una tradición filosófica que advirtió de manera temprana e intuitiva los peligros inherentes a las dinámicas del capitalismo. Esta esquematización dará paso a la tercera parte, “Salidas del asedio e injusticia anónima”, donde el autor presenta una salida del asedio a través de una moderación de la tendencia instrumentalista de las patologías sociales, a partir de la programación de estímulos que favorezcan espacios intersubjetivos de justificación y promuevan una interacción signada por respuestas cognitivas profundas y contextualizadas.

## 1. La vida práctica: imaginación, autonomía y reflexión.

Las reflexiones del autor comienzan evocando la película *Brazil*, dirigida por Terry Gilliam (1985), en la cual se relata la historia de una sociedad aturdida por la burocracia, por una rutina automatizada, que no requiere reflexión, deliberación o comunicación alguna entre sus miembros para funcionar. Aunque uno podría pensar que esta es una situación indignante, depresiva o desmoralizante, la misma es aceptada con entusiasmo y parsimonia por los personajes de la película. Solo hay un indicio en esta indiferencia, que remite a algo parecido a un anhelo de libertad, y que se esconde en el tarareo de la canción de Antonio Carlo Jobim, como una promesa siempre incumplida de libertad que sobrevive confinada en el ámbito de lo onírico. El bufonesco y desesperante retrato de *Brazil*, en donde se mixturán los sueños y la realidad, resume el panorama al que se enfrenta Pereira, en la medida en que, en la película, se reflejan el absurdo y la opresión en que se ven envueltos los agentes bajo mecanismos anónimos de alienación.

El autor anuncia inmediatamente una estructura, un modelo antropológico, que oficia de clave hermenéutica para la argumentación. Dicho modelo distingue una dimensión interna subjetiva, íntima, de la persona, desde donde se ejercitan capacidades que son propias, únicas e irrenunciables; y una dimensión de estímulos y elementos externos u objetivables, que es pública, compartida con los demás. La distinción de estos dos planos no habilita, no obstante, una lectura dualista. En cambio, es preciso comprender estas dos dimensiones de manera dinámica, puesto que la interacción entre agentes habilita una tercera dimensión, intersubjetiva, que oficia de condicionante, tanto de la dimensión subjetiva como de aquella objetiva. La intersubjetividad, opera en este análisis de garante epistémico, puesto que es en el terreno de la procedimentalidad, la deliberación y el intercambio de razones con los demás, en donde se actualiza la mayor cantidad de potenciales de acción de los agentes (Pereira, 2018: 10).

Existen diferentes contextos prácticos en los cuales los agentes se desempeñan y desarrollan su vida práctica. El autor desarrolla aquí cuatro contextos de interés para su investigación - el ético, el moral, el político y el legal (2018: 10)- que son tempranamente diferenciados en el libro y que posibilitarán una reflexión situada de las dinámicas sociales patológicas y no patológicas. El contexto ético remite al ejercicio de la libertad que lleva a cabo cada agente en función de lo que comprende como buena vida. Los contextos político y legal se definen como modos de relacionarse entre los agentes que reflejan las pretensiones de objetivación, institución y regulación normativa de las prácticas en común. El contexto moral, es el plano sobre el que los agentes ensayan acciones y decisiones que involucran de manera universal a todos los demás agentes. Esta tetralogía, recuperada de manera reiterada durante el desarrollo de toda la argumentación, oficia en la segunda parte del libro como un terreno para medir la incidencia de las patologías sociales en la vida práctica. En este caso, a partir del señalamiento de su distorsión u ocultamiento, bajo el peso de una lógica pragmática, anónima, que distorsiona las posibilidades de los contextos prácticos.



La vida práctica se especifica en tres elementos que la articulan y la explican de manera suficiente: la imaginación, la reflexividad y la autonomía. La imaginación es práctica o ejercicio mental contra-fáctico, que posibilita la representación de cosas, relaciones y estados ausentes. Se habilitan entonces posibilidades de acción alternativas a las dadas, lo que provoca un distanciamiento del agente respecto del contexto en el cual se inserta, así como la posibilidad de ensayar distintos escenarios de acción, en función de lo cual la reflexividad puede decidir y justificar sus acciones y decisiones. La autonomía, comprendida como un descentramiento gradual del individuo consigo mismo, es destacada como una cualidad relacional, intersubjetiva, que opera como garantía de la justicia en las relaciones. En este sentido, un agente es autónomo cuando se construye a partir del reconocimiento recíproco con los demás agentes. Estos tres elementos son los que le dan sentido a la vida práctica y ofician de criterios normativos, ya que a partir de la consideración de su funcionamiento conjunto o diferenciado es posible

determinar si alguien actúa de manera libre, crítica o consciente (2018:74). A su vez, aunque los tres elementos aludidos se requieren para lograr un grado apreciable de libertad, es la imaginación en última instancia la que posibilita el libre ejercicio de la reflexión y de la autonomía.

## **2. Patologías sociales como bloqueo de la imaginación**

El desacople que padecen los agentes, las personas, respecto de sus contextos de acción, que imposibilita o condiciona el ejercicio de sus libertades, es explicado aquí a partir de la conceptualización ‘patologías sociales’. Estas dinámicas, que se caracterizan por ser anónimas y elusivas, son abordadas en esta segunda parte del libro, en relación con el bloqueo y el menoscabo que significan para la imaginación práctica. El autor encuentra en la historia de la filosofía, en las teorías de Marx y de Rousseau, una base teórica para conceptualizar las patologías sociales (2018:161). El criterio articulador de esta tradición que inicia en la modernidad filosófica, en este punto, radica en que estos filósofos critican aquellos mecanismos sociales que bloquean la capacidad de respuesta cognitiva de un agente en el contexto en que se desenvuelve. La modernidad es aquí presentada como un proceso histórico en construcción continua que rige nuestra sociedad contemporánea, que es pletórico en imaginarios sociales e imágenes que aspiran a la libertad y la justicia, pero que en su concreción material, institucional, niega una cuota de libertad al mismo tiempo que la ofrece (2018: 83). Esta breve genealogía de las patologías sociales trazada por Pereira, vertebrada entre Marx, Weber y Lukacs; la Teoría Crítica de Frankfurt, concretamente de Adorno y Horkheimer, y de Habermas - cuyas propuestas articulan de manera sutil y fundamental el arco argumental de El asedio a la imaginación-; y de la psicología cognitiva, principalmente, desde donde se incorpora la dimensión sub-intencional para explicar las patologías sociales.

En las teorizaciones de Marx, Weber y Lukacs se pueden hallar elementos para pensar de manera básica el alcance de este asedio: sus teorizaciones dan cuenta de una objetivación o cosificación de los vínculos sociales que deviene en la institucionalización de prácticas distorsivas de los contextos prácticos, como son el consumismo, la burocracia, las relaciones políticas según lógicas de monetarización o de reificación. Estas dinámicas sociales alienantes son comprendidas como un desborde de la “razón pragmática”, o “razón instrumental”, a partir de los estudios de Adorno y Horkheimer. Sin embargo, esta óptica debe ser complementada con los aportes de Habermas, acerca de la importancia de la razón práctica comunicativa (2018:103). Las investigaciones de Elster, Tversky y Kahneman en el campo de la psicología cognitiva contribuyen complejizando el enfoque habermasiano, considerando la distorsión del sentido de los contextos prácticos de los individuos a partir de un condicionamiento a nivel pre-consciente o de segundo plano. En este caso, la patología se manifiesta en un agente como la expresión de un deseo que no es propio, o la incapacidad de justificar su posición en un determinado contexto práctico.

Las patologías tienden a reproducirse de manera circular (2018:134), involucrando cualquier racionalidad práctica – ética, moral, legal o política- en la lógica de la instrumentalización, cumpliendo con ello una función ideológica, por medio de justificaciones engañosas que tergiversan el sentido de los diferentes contextos. No obstante, la ideología, en tanto justificación engañosa y enmascaramiento, puede ser combatida por la crítica de un agente. Esta crítica es inmanente, en cuanto proviene de la virtud epistémica y subjetiva de desmontar el engaño de las justificaciones deficientes; y es trascendente, en cuanto que tiende a ordenarse en torno a guías generales que van más allá del contexto que las inspira. Una vez más, la posibilidad de romper con la circularidad de las ideologías y las patologías sociales radica en un ejercicio dialógico, en lo que sería una concepción post-metafísica del problema de la ideología: “la conceptualización de la crítica a la ideología pasa a tener como espacio privilegiado la discusión pública, y en ella el sujeto deja de tener una posición privilegiada para ser sustituido por la comunidad de justificación” (2018: 150). Esta crítica, en el marco de la autonomía entendida en términos relacionales, requiere, a la vez que favorece una mayor sensibilidad a la vulnerabilidad como criterio modulador de lo que se percibe como una patología social. A su vez, la identificación y denuncia pública

de una patología contribuye como criterio normativo de evaluación, así como de explicación de “las lógicas sociales anónimas que caracterizan a las patologías sociales” (2018: 154).

### **3. Salidas del asedio e injusticia anónima**

La presentación de las consecuencias del sistemático asedio a la imaginación que se atestigua en la sociedad contemporánea en el desborde de la razón pragmática que afecta nuestros planes. Esta razón se impone de manera elusiva y engañosa, bajo una máscara de anonimato que los agentes internalizan como deseos sub-conscientes y reflexiones superficiales, que afectan decisiones de vital importancia. En la tercera parte el autor propondrá salidas del asedio a la imaginación; contraofensivas que tienen como objetivo reconquistar las posibilidades de la vida práctica (2018: 179). Pereira reflexiona aquí sobre la dignidad vulnerada en nuestras sociedades contemporáneas y sobre qué es la justicia, destacando la necesidad de estimular espacios sociales fuertemente mediados por la intersubjetividad como condición para contrarrestar los efectos de las patologías. Una eticidad democrática, como término que reúne el aspecto institucional y cultural – simbólico disperso en la sociedad civil contribuiría a reorientar la auto-comprensión de los agentes así como posibilitaría un disenso constructivo que dispararía la imaginación práctica y la reflexividad profunda a partir del reconocimiento intersubjetivo de pretensiones de validez.

Frente a la injusticia patológica, que deviene anónima, en dinámicas tales como el marketing o la presión de pares que se encuentran desacoplados del contexto compartido, se presentan propuestas que ponen el énfasis en relaciones dialógicas, de cooperativismo y de disonancia cognitiva -fricción normativa (nudge)-, a través de los cuales se podrían propiciar procesos reflexivos profundos que interpelen a los agentes y vuelvan patentes las patologías que los envuelven, contribuyendo a una mejor calidad de vida. A pesar de que estos estímulos en el corto plazo acentuarían el malestar del agente envuelto en la patología, harían también patente los bloqueos que opera, ofreciendo una posibilidad para la acción y la crítica de las dinámicas patológicas.

En sus reflexiones finales, Pereira apela a la necesidad de abonar teorías y marcos conceptuales para combatir las patologías sociales, que son producto de las relaciones de las sociedades capitalistas complejas contemporáneas. Estas patologías, que se presentan bajo formas tales como el burocratismo, la monetarización, el consumismo, la instrumentalización y la mercantilización de las relaciones entre personas, se imponen como fantasmagóricas e irrebasables manifestaciones a las cuales ni siquiera la academia o la ciencia han sabido dar solución todavía (2018: 235). La intención de Pereira, en esta ocasión, es la de devolver a la teoría crítica que busca alternativas al capitalismo la centralidad en los debates. La propuesta del autor, de recuperar la imaginación como práctica vital fundamental para el ejercicio de la libertad, se presta a la coordinación con otros estudios, disciplinas e investigaciones que promuevan estilos de vida capaces de resistir y contravenir los escollos que obstaculizan el desarrollo equitativo de la libertad y de la justicia social en las sociedades contemporáneas.

## Normas para los autores

- 1) Los artículos y reseñas remitidos deberán ser inéditos (esto incluye publicaciones digitales como blogs, actas online, etc.) y no podrán ser postulados de forma simultánea para su publicación en otros periódicos. Al someter a la plataforma de la Prometeica, el autor afirma su anuencia a esa exigencia.
- 2) Los artículos no deberán exceder los 40.000 caracteres. Las reseñas no deberán exceder los 10.000 caracteres.
- 3) Todos los artículos deberán estar acompañados de un resumen y un abstract equivalente en inglés, cada uno de no más de 1.500 caracteres, incluyendo tres palabras claves.
- 4) Los idiomas aceptados para los artículos serán:
  - a) español,
  - b) portugués,
  - c) inglés.
- 5) Los artículos deben ser enviados a la revista a través de lo que el acceso del usuario a la plataforma OJS. Las presentaciones deben estar en dos archivos con formatos doc, docx o odt. La primera debe ser el artículo y el resumen, y sin los datos del autor. El segundo debe contener los datos del autor: un breve curriculum, afiliación académica e información de contacto. Los metadatos de los archivos deben ser borrados por el autor para garantizar una evaluación ciega, es decir, el autor no ve el nombre del parecerista y el parecerista no ve el nombre del autor.
- 6) Una vez enviado el artículo/reseña el autor recibirá un e-mail de Prometeica acusando recibo. Desde la recepción de ese mensaje el comité editorial tendrá un máximo de 4 meses para evaluar si el artículo/reseña será publicado/a en la revista. El tiempo máximo previsto entre el envío del artículo y su publicación, en caso de aceptación sin reformulaciones, es de 6 meses.
- 7) En cuanto al sistema de referencias Prometeica adopta las Normas APA – 2016 - 6ª edición, de la American Psychological Association.
- 8) Para las notas aclaratorias se empleará la referencia al pie. Preferentemente se sugiere no abusar de este recurso.
- 9) En el caso de artículos que incluyan imágenes, deben enviarse en un archivo separado. Las imágenes deben tener una resolución de 300 dpi, en formato \*.jpg o \*.tiff. El copyright de la imagen ya debe ser concedido, o el autor debe usar imágenes sin derechos de autor.
- 10) En cuanto a la evaluación de los artículos, los mismos serán remitidos al miembro del consejo editorial responsable del área del trabajo en cuestión. Los artículos serán enviados a dos especialistas externos y evaluados en el sistema de revisión doble ciega en que el autor no ve el nombre del evaluador y el evaluador no ve el nombre del autor. En el caso de haber desacuerdo entre ellos, un tercer árbitro podrá ser consultado, por decisión del consejo editorial.
- 11) Los trabajos pueden tener tres resultados posibles que constan en el formulario de evaluación que completará junto a otras observaciones el evaluador:
  - a) recomendado para su publicación sin alteraciones,
  - b) recomendado para su publicación con modificaciones,
  - c) no recomendado para su publicación.
- 12) En el caso 11 (b), la publicación del mismo quedará sujeta a que el autor esté dispuesto a realizar las modificaciones y las remita para su nueva evaluación.
- 13) Todos los trabajos aprobados serán publicados en un idioma aceptado por la revista. Si el original no está escrito en uno de estos idiomas y demandas de traducción, puede resultar en un retraso de la publicación.
- 14) El contenido de los originales publicados es responsabilidad exclusiva de sus autores.
- 15) Los artículos presentados a la revista deben ser escritos utilizando el **Modelo\_Autor\_Estilos.dotx** para seguir los estilos y formatos descritos en **Formato\_Prometeica\_2019.docx**
- 16) Prometeica adopta una política antiplágio sometiendo todos los artículos enviados a un primer análisis por el software Turnitin.