

Abriendo espacios para la educación en valores y ética en tiempos de tecno-utopías: Una propuesta formativa

Carlos Enrique Zerpa
Departamento de Ciencia y Tecnología del Comportamiento
Universidad Simón Bolívar
czarpa@usb.ve

Introducción

Es posible que exista consenso en que la vieja aspiración de la humanidad de llegar al futuro parece ya haber ocurrido; en efecto, no hay duda alguna, en la era post-pandemia del COVID-19, que el mundo imaginado hace ya seis o más décadas, de grandes innovaciones técnico-mecánicas y electrónicas, es algo de lo real de la experiencia cotidiana. El internet de las cosas (IoT, por sus siglas en Inglés) es un ejemplo de ello. Y es que los avances de la tecnología parecen ir a una velocidad tan alta que inclusive es posible perder la pista a lo que se está haciendo. Las máquinas, sistemas y dispositivos electrónicos se hacen cada vez más técnicamente avanzados, poderosos e “*inteligentes*”, y también son cada vez más deseados y demandados, lo cual influye en la rápida evolución de innovaciones de aún mayor sofisticación, de mayor poder de cómputo y mayor inteligencia artificial (IA). En tal sentido, la evolución de las interfaces entre la persona y tales creaciones ha permitido avanzar hacia el diseño de entornos cada vez más amigables, fáciles de utilizar y adaptables a las necesidades y preferencias del usuario. Pero también ha implicado el surgimiento de nuevos desafíos frente a la deriva hacia fines y usos no deseados.

Esta realidad contiene dos grandes aristas referidas al ser humano: la primera, tiene que ver con quien idea, diseña y programa el dispositivo o la máquina; la segunda, tiene que ver con quien la usa con algún fin. En ambos casos el sujeto interactúa con el dispositivo, el sistema o la máquina y en esa interacción, similarmente a como ocurre en la interacción entre personas, se presentan riesgos éticos que se han constituido como problemas de considerable gravedad. González & Martínez (2020) señalan que estos riesgos pueden categorizarse en cuatro (4) campos: 1) *invasión de los entornos* (aprovechamiento de debilidades para hacer que un sistema inteligente cometa errores, o tome el control de otro sistema a partir de tales, entre otros; 2) *intimidad* (reemplazo de identidad, publicación de falso contenido adaptado a las características del usuario, control a partir del conocimiento de datos de comportamiento o de las preferencias

de un usuario, entre otras); 3) *seguridad* (digital, física y política); y 4) *identidad* (identificación de personas para fines no lícitos o cuestionables). Un ejemplo serían los “bots” de respuestas falsas de Twitter; o los usos potencialmente ilícitos que en contextos de aprendizaje y enseñanza se le puede dar a innovaciones como el “ChatGPT”. Al respecto, Brundage et al (2018, p. 4) refieren que:

“... (los) ingenieros en inteligencia artificial deberían tomar seriamente la naturaleza de doble uso de su trabajo, lo que permite consideraciones para influir en las prioridades de investigación y normas, y llegar de manera proactiva a los actores relevantes cuando las aplicaciones dañinas son previsibles”.

En otras palabras, afirman que las aplicaciones de la inteligencia artificial pueden tener riesgos éticos, pero que en los actuales momentos tales riesgos pueden anticiparse y por tanto prevenirse o intervenir; además señalan que una forma de actuar sobre ellos consiste en la involucración de todas las partes interesadas de forma tal que se elaboren “estándares éticos” para el desarrollo y despliegue de sistemas de IA, que signasen empresas, investigadores e implementadores de estos sistemas. Todo ello como componente de una *tecno-ética* que “plantea cuestiones específicas en relación con las implicaciones éticas de la tecnología, los artefactos tecnológicos y otros subproductos” (Muñoz, 2022, p. 1). O bien como un marco regulatorio para esa creciente fusión entre lo físico, digital y biológico que puede transformar la esencia misma de la humanidad (Porcelli, 2020); por ejemplo, documentos como las comunicaciones que la comisión europea ha dirigido al parlamento europeo referidas al desarrollo de una inteligencia artificial centrada en el humano y que amerita un plan coordinado sobre la misma (Comisión Europea, 2018a, 2018b).

No obstante, “estándares éticos”, “códigos de ética”, “normas” resultan en contenidos de conducta regulatoria cuyo acatamiento en definitiva dependerá de la *voluntad* del agente, es decir, del diseñador de la inteligencia artificial en cuestión y de sus usuarios, y que pueden resultar “vacíos”, con solo un alcance lógico, e inclusive derivar en interpretaciones diferentes o contradictorias y relativistas, por lo cual como documentos de buena intencionalidad por sí mismos no cumplirían su función. Se requiere ir más allá y considerar las raíces del problema. Al respecto, desde una perspectiva cognitiva-evolutiva (p. e.: Phillips, 1977) el que una persona sea capaz de comprender y asumir tales modelos de comportamiento apropiado dependerá fundamentalmente tanto del momento de desarrollo del ciclo vital en el que se encuentre el

sujeto, como de la oportunidad que éste haya tenido de someterse a experiencias formativas con significado psicológico. La intersección de ambas condiciones ocurre en un contexto que llamamos “educación” y es por tanto el espacio o entorno crítico que provee la oportunidad de desplegar tareas formativas entre las que destacan dos clases: 1) tareas científicas, y 2) tareas éticas (o morales filosóficas) (Uhl, 2001). Las tareas éticas implican una transmisión de los saberes orientativos, valorativos y normativos de generaciones anteriores que se encuentran objetivados en las instituciones y que sólo puede llegar a los individuos a través de un proceso de “*re-subjetivización*” (Uhl, 1997). Desde esta perspectiva, Uhl afirma que para que se integre como parte de la vida cotidiana, presente en todos sus actos, la ética debería convertirse en un hábito, una “... propiedad vivencial de la personalidad del individuo o una disposición psíquica relativamente duradera para un determinado tipo de comportamiento” (p. 224). Por tanto, son estas últimas, las tareas éticas de la educación, las que interesan aquí dado que la solución al desafío de enfrentar el uso y abuso de la inteligencia artificial en el futuro está en la educación en valores y ética que se pueda dar a cada persona en el presente, para que se constituya en un sujeto con una “*personalidad moral*” en cuyo mapa psicológico o *anatomía* debe constituirse el carácter, los valores, el razonamiento y la emoción, y que por tanto pueda supeditar su voluntad a un criterio moral propio, y ejercer así una acción moral éticamente apropiada (Berkowitz, 1995). En otras palabras, estas tareas buscan formar un sujeto con *autonomía moral* ejercida a través del diálogo y la acción y desarrollada a partir de la interrelación humana (Sepúlveda, 2003). Autonomía entendida como esa cualidad del sujeto que “cumple con los requisitos de la experiencia del deber y no la reduce a nada distinto de ella misma” (Gracia, 2012, p. 1). En el sentido kantiano, se refiere al *autogobierno* de uno mismo por el deber, que es el único criterio estrictamente moral. (Schneewind, 2005).

La pregunta obligada que surge a partir de lo anterior tiene que ver con el asunto de lo procedimental: ¿cómo lograr el ideal de abrir un espacio de formación en valores que sirva para enfrentar los desafíos éticos que surgirán en la interacción persona-IA en un escenario utópico futuro? De allí se desprende el objetivo de este trabajo: explicar los aspectos conceptuales y estructurales de una propuesta de educación y formación en valores y ética que sirva para ello. Al respecto, la propuesta que se presenta resulta en el planteamiento de una praxis (cómo lo es la educación) en la que se toman en cuenta los aportes de al menos cuatro pensadores en temas de educación en valores, ética y moral: 1) los principios de la educación moral del futuro planteados

por Fritz Osers (1992): (a) estimular el nivel inmediatamente superior de desarrollo moral; b) sensibilizar a los niños y adolescentes respecto a los temas morales, por ejemplo, temas de cuidado, de justicia, de veracidad, de tolerancia; (c) apoyar y hacer posible la conducta prosocial; y (d) establecer un clima moral positivo, todo lo cual el autor cristaliza en lo que ha llamado los *mandamientos sobre educación moral*; 2) la perspectiva cognitiva-evolutiva del desarrollo del juicio moral de Lawrence Kohlberg (1992); 3) los planteamientos acerca de la eficacia de los medios de educación moral de Sigfried Uhl (1997); y 4) los aportes acerca del desarrollo de la conducta pro-social de Robert Roche (2010). En tal sentido, a continuación se presenta un modelo de base a una propuesta educativa integrada durante el desarrollo del ciclo vital, para abrir espacios a la educación y formación en valores y ética, y cuyo propósito es educar al sujeto que en ambientes futuros, como el laboral, tendrá interacción frecuente con máquinas y sistemas de inteligencia artificial y no (solo) con personas, y que pretende moldear su personalidad para enfrentar los desafíos que estas interacciones implican, previniendo sus consecuencias indeseadas.

El modelo de la propuesta formativa: un esquema de implantación.

El modelo representa un esquema simplificado del diseño de una intervención educativa que inicia en el espacio privado de la familia y se extiende hasta el contexto de la organización, es decir, a través de diferentes sistemas en los que ocurren procesos de educación en valores, ética y moral, de naturaleza formal o no formal; también se trata de una propuesta evolutiva, que abarca diferentes periodos sensibles del desarrollo humano, desde la infancia hasta la edad adulta y que coinciden con los momentos en que las personas son actores sociales de los contextos referidos; además de complejidad progresiva, en cuanto a las estrategias y tácticas instruccionales de posible implementación en los contextos que se consideran, desde las menos complejas (la adquisición e identificación de valores), hasta las más complejas (práctica de la responsabilidad social). En otras palabras, se trata de la generación de un espacio formativo multicontextual y transversal en el ciclo de vida, que integra estrategias y tácticas instruccionales de efectividad comprobada en la psicología y la educación, busca explicar el cómo es posible enseñar (o aprender) contenidos y prácticas de orden ético-valorativo, y cuyo propósito último es la conformación de sociedades más justas, plurales y democráticas y de personas con autonomía

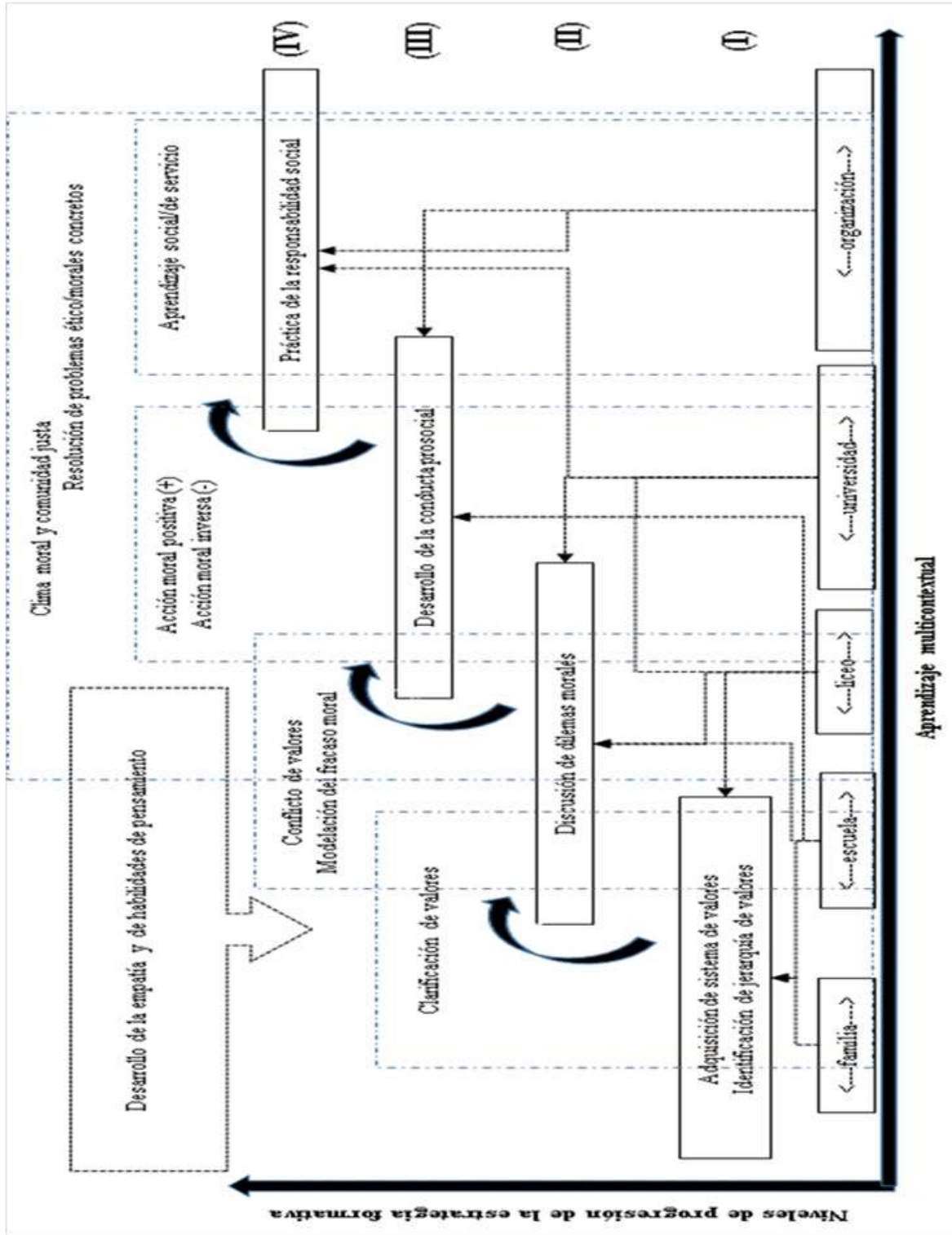
moral. El modelo formativo puede observarse en la figura 1. Para comprender el modelo o esquema de implantación de la propuesta formativa, transversal, es necesario precisar en sus componentes estructurales:

En primer lugar, en el eje de las abscisas, se encuentran los diferentes *contextos de aplicación*, considerados en un horizonte temporal indefinido, y que abarcan el entorno no formal de educación de la familia, los entornos formales educativos (escuela, liceo, universidad) y el entorno híbrido, formal y no formal de la organización. En estos espacios de interacción humana presencial o virtual, ocurrirán los eventos de aprendizaje que impliquen las tareas éticas de la educación.

En segundo lugar, en el eje de las ordenadas, una progresión de la *complejidad de la estrategia formativa*. Cada etapa instruccional se considera requisito para la siguiente puesto que se articulan a condicionantes del desarrollo evolutivo de la persona que es prudente incorporar: evolución del desarrollo del pensamiento formal complejo y de las habilidades sociales, desde períodos del desarrollo previos en una secuencia que se considera universal e invariable.

En tercer lugar la estrategia formativa, compuesta por *cuatro grandes etapas*, que pueden solaparse: a) *Adquisición de un sistema de valores e identificación de la jerarquía en la que se organizan*, experiencias formativas que ocurren de forma privilegiada tanto en la familia como en la escuela, pero que esencialmente se extienden hasta el liceo y la universidad. Desde la perspectiva de Schwartz (1992), los valores a adquirir e identificar pueden considerarse como metas y motivaciones que sirven de principios-guía en la vida de las personas y se organizan en dos categorías según si sirven a intereses individuales o colectivos: 1) poder, logro, hedonismo, estimulación y autodirección, son tipos de valores que sirven a intereses individuales; 2) benevolencia, tradición y conformidad, sirven a intereses colectivos; 3) universalismo y seguridad son tipos de valores que sirven a ambos intereses y están situados en el límite entre estos dos. b) *Discusión de dilemas morales*, en los contextos educativos de la escuela, el liceo y la universidad; una estrategia en la que se debate una situación extrema de conflicto moral (o situación en la que un agente se ve confrontado con dos obligaciones morales que le instan a actuar) en la que el agente no puede seguir un curso de acción que esté conforme con sus dos obligaciones en conflicto (Realpe, 2001). Para que un conflicto moral tenga el carácter de ser un auténtico dilema moral y no simplemente un aparente dilema moral, ninguna de las obligaciones en conflicto es en efecto más fuerte o logra invalidar a la otra obligación.

Figura 1.
 Modelo de formación en valores y ética para un escenario tecno-utópico futuro.



c) El desarrollo de la conducta pro-social, una experiencia que si bien inicia en la familia, cristaliza operacionalmente en los entornos del liceo, la universidad y la organización. En el modelo de Roche (2010), la conducta pro-social hace referencia a comportamientos que, sin la búsqueda de recompensas externas, extrínsecas o materiales, favorecen a otras personas o grupos, según los criterios de estos, o metas sociales objetivamente positivas y que aumentan la probabilidad de generar una reciprocidad positiva de calidad y solidaria en las relaciones interpersonales o sociales consecuentes. La ayuda, cooperación y solidaridad son prácticas que conducen a su desarrollo. d) Finalmente, la práctica de la responsabilidad social, como estrategia de mayor complejidad y que se ejerce tanto en la universidad como en la organización y que enfatiza las dimensiones de buena gobernabilidad, gestión de los impactos medioambientales y sociales, diálogo con los grupos de interés y la rendición de cuentas, y alianzas para participar en el desarrollo sostenible (Vallaeys et al, 2009).

En cuarto lugar, las *tácticas específicas* para el trabajo en diferentes contextos formativos: a) clarificación de valores en la familia y la escuela (establecer cuáles son y qué importancia tienen); b) conflicto de valores y modelación de la experiencia-vivencia del fracaso moral de otros, tanto en la escuela, como en el liceo y la universidad (comprender y juzgar también los errores morales de otros con base a la perspectiva *desde* el otro); c) la acción moral positiva (+) e inversa (-) en los contextos del liceo y la universidad (aprender lo que hay que hacer y aprender también *lo que no* hay que hacer); d) la creación de la comunidad justa y la resolución de problemas ético/morales concretos, en la escuela, el liceo, la universidad y la organización (promover comunidades educativas plurales y democráticas); y e) el aprendizaje social/de servicio tanto en la universidad como en la organización.

En quinto lugar, las estrategias auxiliares para el desarrollo de la capacidad de comprensión del otro (empatía) y habilidades de pensamiento (análisis, comparación, inferencia) que fundamentalmente se ejercitan en la familia, la escuela y el liceo y que constituyen un basamento cognitivo para el funcionamiento óptimo de todo lo demás.

Finalmente, las conexiones entre los componentes que hacen referencia a las relaciones entre los contextos de aprendizaje y las macroestrategias formativas que pueden solaparse en las etapas de progresión estratégica y en los contextos. Implícitamente, y conforme a las tendencias mismas de la educación del futuro, se encuentra tácita la posibilidad de que los diferentes

contextos formativos incluyan a las tecnologías de información y comunicación aplicadas a la educación, como medios adicionales/alternativos a la presencialidad, por ejemplo, centrada en la realidad aumentada, que servirán para distribuir contenidos instruccionales, participar de contextos de aprendizaje valorativos-éticos-morales simulados virtualmente, reforzar conocimientos a través de campañas en redes sociales, fomentar el trabajo colaborativo, la co-construcción de saberes éticos y enfrentar experiencias transformadoras.

El resultado de tal proceso educativo transversal sería articulable a lo que refiere Sepúlveda (2003, p. 28) al hacer referencia a los enfoques cognitivo-evolutivos del desarrollo moral:

“Se construye un sentido ético que define al sí mismo en términos de una jerarquización de valores y una integración social efectiva, de forma autónoma, (...), en el cual se lograría una mayor diferenciación e integración (...), en la medida que la persona tendría la capacidad de autolegislarse, al imponerse a través de la razón un principio de justicia universal...”

Es decir, se trata de la clase de autonomía necesaria para enfrentar los desafíos que las tecno-utopías presentarán a las personas en el futuro próximo y que permitiría decir “no” a las transgresiones éticas-valorativas en la interacción entre sistemas con IA y las personas.

Prospectiva y conclusiones

El propósito de este documento era explicar la formulación conceptual-estructural de un modelo para la apertura de espacios en la formación en valores y la educación ética, que permitiera la constitución de la personalidad moral necesaria para el escenario tecno-utópico del futuro. La propuesta de modelo resulta ser de carácter transversal pues está enfocada en diferentes periodos sensibles del desarrollo evolutivo de las personas, que se implementaría a través de estrategias instruccionales y tácticas de complejidad creciente, en función de la etapa evolutiva y del contexto particular en el que la persona se eduque, incluyendo la familia y la organización, pues son espacios de aprendizaje permanente.

La propuesta formativa transversal supera la necesaria pero no suficiente elaboración de estándares o normas de comportamiento característica de decálogos o códigos de ética e integra diferentes medios de educación moral, a la manera de un plan coordinado para enfrentar los

desafíos éticos que implicará la adopción creciente de la inteligencia artificial en diferentes contextos futuros.

En un escenario tecno-utópico en el que la interacción humana también incluirá a las máquinas con inteligencia artificial en la vida cotidiana, en los contextos íntimos familiares y en entornos sociales y laborales como la escuela, el liceo, la universidad y la organización, se hace impostergable en los momentos actuales buscar alternativas efectivas, como la que se propone, para la formación de las personas en esas tareas educativas que implican cuestiones ético-valorativas, y que permitan el desarrollo de la autonomía moral necesaria tanto para anticiparse a los problemas derivados del diseño y uso de tales tecnologías, como para generar la acción remedial y la intervención cuando las circunstancias éticas así lo exijan.

Referencias

- Berkowitz, M. (1995). Educar la persona moral en su totalidad. *Revista Iberoamericana de Educación: Educación y democracia* 1 (8), 73-101.
<https://rieoei.org/historico/oeivirt/rie08a03.pdf>
- Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., (...) & Amodei, D. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*. (1^{era} Ed.). Future of Humanity Institute, University of Oxford, Arizona State University. <https://arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf>
- Comisión Europea (2018a) Comunicación de la Comisión al Parlamento Europeo, al Comité Económico y Social Europeo y al Comité de las Regiones: Inteligencia Artificial para Europa. *COM(2018) 237 final*. Bruselas. <https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=COM%3A2018%3A237%3AFIN>
- Comisión Europea (2018b) Comunicación de la Comisión al Parlamento Europeo, al Comité Económico y Social Europeo y al Comité de las Regiones: Plan Coordinado sobre la Inteligencia Artificial. *COM(2018) 795 final*. 7 de diciembre de 2018. Bruselas. <https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=COM:2018:795:FIN>
- Gracia, D. (2012). La construcción de la autonomía moral. *Revista del Hospital Italiano de Buenos Aires*, 32 (1).
https://www.hospitalitaliano.org.ar/multimedia/archivos/noticias_attachs/47/documentos/11750_10-bioetica2%2032-1-12%20OK.pdf

- Kohlberg, L. (1992). *Psicología del desarrollo moral*. Bilbao: Desclée de Brouwer.
- Muñoz, a. (2022). *Futuro, ética y digitalización: una propuesta de introducción de la tecnética en la asignatura Tecnología e Ingeniería*. Trabajo de grado no publicado, Departament de Ciències de la Computació, Universitat Politècnica de Catalunya. Barcelona, España.
- Osers, F. (1995). Futuras perspectivas de la educación moral. *Revista Iberoamericana de Educación. Educación y democracia* 8 (2), 1-26.
<https://rieoei.org/historico/oeivirt/rie08a01.htm>
- Phillips, J. (1997). *Los orígenes del intelecto según Piaget*. Barcelona: Fontanella.
- Porcelli, A. M. (2020). La inteligencia artificial y la robótica: sus dilemas sociales, éticos y jurídicos. *Derecho global. Estudios sobre derecho y justicia*, 6 (16), 49-105.
<https://doi.org/10.32870/dgedj.v6i16.286>
- Realpe, S. (2001). Dilemas morales. *Estudios Gerenciales*, 17 (80), 83-113.
http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0123-59232001000300004&lng=en&tlng=es
- Roche, R. (2010). *Prosocialidad: nuevos desafíos. Métodos y pautas para la optimización creativa del entorno*. Buenos Aires: Ciudad Nueva.
- Schneewind, J. B. (2005). *The invention of autonomy*. Cambridge: University Press.
- Schwartz, S. (1992) Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. En: M.P Zanna (Ed.) *Advances in experimental social psychology*, vol 25, pp. 1-65. San Diego: Academic Press.
- Sepúlveda, M. G. (2003). Autonomía moral: Una posibilidad para el desarrollo humano desde la ética de la responsabilidad solidaria. *Revista de Psicología de la Universidad de Chile*, 12 (1), 27-35. <https://www.redalyc.org/pdf/264/26400102.pdf>
- Uhl, S. (1997). *Los medios de educación moral y su eficacia*. Barcelona: Herder.
- Uhl, S. (2001). Die aufgaben der allgemeinen pädagogik. Eine klassifikation der gängigen auffassungen. *Zeitschrift für Erziehungswissenschaft*, 4, 61–82
<https://doi.org/10.1007/s11618-001-0006-2>
- Vallaes, F., de la Cruz, C., & Sasia, P. M. (2009). *Responsabilidad social universitaria: manual de primeros pasos*. México: McGraw-Hill.