

**92**

Serie de documentos  
de trabajo del IIEP

ISSN 2451-5728

JULIO | 2024

# Hacia un nuevo esquema de incentivos en el contexto de la inteligencia artificial

Sandra Maceri  
Agustín Coll



Autores

**Sandra Maceri**  
smaceri@conicet.gov.ar

Universidad de Buenos Aires. Facultad de Ciencias Económicas. Centro de Investigaciones en Epistemología de las Ciencias Económicas. Instituto Interdisciplinario de Economía Política. Buenos Aires, Argentina.  
CONICET. Instituto Interdisciplinario de Economía Política. Buenos Aires, Argentina.

**Agustín Coll**  
agustinjcoll@gmail.com

Universidad de Buenos Aires. Facultad de Ciencias Económicas. Centro de Investigaciones en Epistemología de las Ciencias Económicas. Instituto Interdisciplinario de Economía Política. Buenos Aires, Argentina.  
CONICET. Instituto Interdisciplinario de Economía Política. Buenos Aires, Argentina.

---

**Como citar**

---

Maceri, S. y Coll, A., (2024). Hacia un nuevo esquema de incentivos en el contexto de la inteligencia artificial. *Serie Documentos de Trabajo del IIEP*, 92, 1-36. <https://ojs.econ.uba.ar/index.php/DT-IIEP/issue/view/501>

---

Los Documentos de Trabajo del IIEP reflejan avances de investigaciones realizadas en el Instituto y se publican con acuerdo de la Comisión de Publicaciones. Los autores son responsables de las opiniones expresadas en los documentos.

---

Coordinación editorial

Ed. Hebe Dato

Corrección de estilo

Ariana Lay y Ed. Hebe Dato

Diseño

DG. Vanesa Sangoi

---

El Instituto Interdisciplinario de Economía Política IIEP UBA CONICET, reconoce a los autores de los artículos de la Serie de Documentos de Trabajo del IIEP la propiedad de sus derechos patrimoniales para disponer de su obra, publicarla, traducirla, adaptarla y reproducirla en cualquier forma. (Según el art. 2, Ley 11.723).



Esta es una obra bajo Licencia Creative Commons  
Se distribuye bajo una Licencia Creative Commons Atribución-NonComercial-Compartir Igual 4.0 Internacional.

## Hacia un nuevo esquema de incentivos en el contexto de la inteligencia artificial

Inteligencia artificial  
Incentivos  
Motivación

Este escrito examina la intersección entre la inteligencia artificial (IA) y las teorías de motivación humana, con un enfoque en las implicaciones para el futuro del trabajo y los sistemas de incentivos económicos. Analizamos cómo la rápida evolución de la IA desafía las teorías económicas tradicionales sobre la motivación, las cuales se basan principalmente en incentivos extrínsecos. A través de una revisión de la literatura y del análisis de estudios empíricos, mostramos la importancia creciente de la motivación intrínseca en un entorno laboral cada vez más automatizado. Al respecto, exploramos el efecto crowding out, donde los incentivos extrínsecos pueden disminuir la motivación intrínseca, y sus implicaciones para el diseño de sistemas de incentivos. Particularmente, presentamos dos escenarios potenciales: una simbiosis entre el trabajo humano y la IA, y una transformación radical que podría hacer que el trabajo humano sea opcional en muchos sectores. Concluimos que es necesario reevaluar y adaptar los modelos económicos y sistemas de incentivos para reflejar una comprensión más matizada de la motivación humana en el contexto de la IA. Este estudio subraya la importancia de un enfoque interdisciplinario para la transición hacia un futuro donde la relación entre los humanos, el trabajo y la tecnología deberá ser redefinida.

## Towards a new incentive scheme in the context of artificial intelligence

Artificial intelligence  
Incentives  
Motivation

This paper examines the intersection of the artificial intelligence and theories of human motivation, focusing on implications for the future of work and economic incentive systems. We analyze how the rapid evolution of AI challenges traditional economic theories of motivation, which primarily rely on extrinsic incentives. Through a literature review and analysis of empirical studies, we demonstrate the growing importance of intrinsic motivation in an increasingly automated work environment. In this regard, we explore the crowding out effect, where extrinsic incentives can diminish intrinsic motivation, and its implications for designing incentive systems. In particular, we present two potential scenarios: a symbiosis between human work and AI, and a radical transformation that could make human work optional in many sectors. We conclude that it is necessary to re-evaluate and adapt economic models and incentive systems to reflect a more nuanced understanding of human motivation in the context of AI. This study highlights the importance of an interdisciplinary approach for the transition to a future where the relationship between humans, work and technology will have to be redefined.

JEL CODE O31

# Índice

05	Introducción
06	Estado de la cuestión
09	Incorporando los avances psicológicos en la teoría económica moderna
15	Efecto crowding out
19	La transformación radical de la inteligencia artificial en perspectiva histórica
22	Inteligencia artificial como complemento humano
27	Inteligencia artificial como reemplazo humano
30	Conclusión
32	Referencias bibliográficas

## Introducción

La inteligencia artificial (IA) es de suma importancia en la historia de la innovación tecnológica. Este estudio se propone analizar el entendimiento de la disciplina económica sobre la motivación humana en el contexto de la transformación laboral que puede estar próxima a suceder, producto de la inteligencia artificial.

Como profundizaremos a lo largo del trabajo, existen distintas teorías sobre la motivación y el correcto uso de incentivos, tanto desde economía como desde psicología. Dado el cambio que puede significar la llegada de capital no humano capacitado, resulta necesario analizar estas teorías sobre incentivos y evaluar la evidencia de cada una de ellas, para que los modelos puedan adaptarse a la transformación laboral resultado de las nuevas tecnologías.

La economía ha basado tradicionalmente su entendimiento del ser humano en una visión que privilegia la motivación extrínseca, particularmente los incentivos monetarios. Sin embargo, la evidencia proveniente de la psicología y la economía del comportamiento sugiere que esta visión es incompleta y puede ser contraproducente en ciertos contextos (Deci *et al.*, 1999; Pink, 2011). La teoría de la autodeterminación, desarrollada por Ryan y Deci (Ryan & Deci, 2000), propone que existen tres necesidades psicológicas universales que son esenciales para el fomento de la motivación intrínseca y el bienestar psicológico: autonomía, competencia y relaciones sociales.

Un fenómeno a considerar es el efecto *crowding out*, en el cual la introducción de incentivos extrínsecos puede disminuir la motivación intrínseca (Frey & Jegen, 2001). Este efecto tiene implicaciones significativas para el diseño de sistemas de incentivos en el contexto laboral, especialmente en un futuro donde el rol del empleado humano puede cambiar de manera radical.

La integración de la IA en el entorno laboral tiene el potencial tanto de complementar como de reemplazar ciertas funciones humanas. Por un lado, la IA puede aumentar las capacidades humanas, permitiendo a los trabajadores centrarse en tareas que requieren creatividad, pensamiento estratégico y habilidades interpersonales (Jarrahi, 2018). Por otro lado, existe la posibilidad de que la IA reemplace completamente ciertos tipos de trabajos, lo que podría llevar a transformaciones sociales y económicas profundas (Zhang *et al.*, 2023).

Se abordarán ambos escenarios y se abogará por una necesidad de adaptar la teoría predominante a la hora de evaluar los incentivos humanos independientemente de cual de dichos escenarios se cumpla. Este estudio se propone explorar estas cuestiones en profundidad, analizando las teorías económicas y psicológicas sobre la motivación humana, evaluando la evidencia empírica disponible, y considerando las implicaciones de la revolución de la IA para el futuro del trabajo y los sistemas de incentivos tradicionales.

La exposición de los temas se organizará de la siguiente manera. En primer lugar, examinaremos las teorías de motivación tanto en economía como en psicología, analizando

las implicaciones de las divergencias entre los supuestos económicos tradicionales y los hallazgos psicológicos recientes en relación con la motivación humana. Posteriormente, contextualizaremos la revolución de la IA desde una perspectiva histórica, comparándola con revoluciones previas y evaluando el impacto de transformaciones sociales y productivas. Argumentaremos que los cambios potenciales derivados de la IA podrían ser incomparables a los de revoluciones anteriores, y presentaremos dos escenarios prospectivos sobre la evolución del trabajo.

El primer escenario explorará la posibilidad de una simbiosis entre el trabajo humano y no humano, analizando cómo esto podría modificar el papel del trabajador actual y las implicaciones para los sistemas de incentivos. El segundo escenario examinará la viabilidad de un ingreso básico universal facilitado por los avances en IA y sus consecuencias socioeconómicas. El presente artículo buscará exponer que el cumplimiento, aunque sea parcial de alguno de estos escenarios, implica la necesidad de un cambio en las teorías económicas que sea coincidente con las transformaciones tecnológicas y sociales en curso.

### **Estado de la cuestión**

No se registran antecedentes, hasta nuestro conocimiento, de un análisis del entendimiento de la disciplina económica sobre la motivación humana en el contexto de la transformación laboral que puede estar próximo a producirse producto de la inteligencia artificial. Como ahondaremos a lo largo del trabajo, y a modo de resumen en esta sección, existen distintas teorías sobre la motivación y el correcto uso de incentivos, tanto desde economía como desde psicología. Dado el cambio que puede significar la llegada de capital no humano capacitado, resulta necesario analizar estas teorías sobre incentivos y evaluar la evidencia de cada una de ellas, para que los modelos puedan adaptarse a la transformación laboral producto de las nuevas tecnologías.

Proveniente de la corriente utilitarista de Jeremy Bentham, la economía se ha manejado con una visión de un ser humano que persigue el placer y huye del dolor. Esto es lo que se entiende en economía por utilidad, noción central, similar a felicidad según la tradición utilitarista. Bentham remarco la importancia de la utilidad dado que servía como métrica para realizar acciones, haciendo énfasis en el entendimiento que debían tener los gobernantes de las fuerzas del placer y del dolor para poder realizar cambios en la política (Charles-Leija *et al.*, 2018).

Jevons sostuvo, ante la imposibilidad de medir el placer y dolor pertenecientes a cada ser humano, que es el consumo lo que otorga placer, convirtiéndolo en un proxy de utilidad (Charles-Leija *et al.*, 2018), en función de las necesidades de cada individuo. Esta concepción, junto con el entendimiento del *homo economicus* como un humano que solo tiene un deseo, el de riqueza, provoca que la economía entienda al ser humano como uno que está motivado principalmente por razones extrínsecas, particularmente monetarias.

Así, la utilidad representa una noción cuya dilucidación fue largamente concebida como un menester ajeno al ámbito económico, recayendo principalmente en el terreno de la psicología. A modo ilustrativo, Stigler & Becker (1977) esgrimieron que no es potestad de la economía escrutar las preferencias, sino meramente aceptarlas como un dato apriorístico. En consonancia, Friedman (1953) percibía las utilidades como factores exógenos, circunscribiendo el rol del economista al análisis de las consecuencias emanadas de tales preferencias, mientras que la elucidación de su génesis quedaba bajo la órbita de los psicólogos.

La economía entiende la motivación humana principalmente como algo que puede ser manipulado a través de incentivos externos que buscan modificar un determinado comportamiento. En psicología esto es lo que se entiende como motivación extrínseca, y es la motivación que se determina como variable en la mayoría de los modelos económicos.

Frey & Jegen (2001) analizan el efecto de las intervenciones externas según las teorías económicas preponderantes. Los autores identifican tres situaciones distintas en un contexto de principal-agente que ilustran la complejidad de este fenómeno. La teoría económica convencional, representada por autores como Alchian y Demsetz o Fama y Jensen, postula que las intervenciones externas, en especial las recompensas monetarias, deberían incrementar el rendimiento al elevar el costo marginal de la holgazanería o aumentar el beneficio marginal del desempeño. Este efecto, denominado "efecto precio", se basa en la premisa de que la motivación intrínseca es constante o inexistente.

Sin embargo, la investigación empírica ha revelado una dinámica más compleja. En algunos casos, las intervenciones externas pueden tener el efecto opuesto, socavando la motivación intrínseca y, por ende, reduciendo el rendimiento. Este fenómeno se conoce como efecto de *-crowding out-* o desplazamiento. La interacción entre estos dos efectos —el efecto de precio relativo y el efecto de desplazamiento— determina el resultado final de una intervención externa en el rendimiento (Frey & Jegen, 2001). Pero a pesar de la amplia utilización del sistema de incentivos, la evidencia acerca de este tipo de motivación no es alentadora. Por ejemplo, incentivos monetarios, o premios en búsqueda de aumentar la motivación extrínseca en estudiantes, ha dado como resultado un peor desempeño de los alumnos (Kohn, 1993, List *et al.*, 2018).

Visaria *et al.* (2016) presentan evidencia empírica sobre los efectos de los incentivos en la asistencia escolar y el rendimiento académico en escuelas no formales de barrios marginales en la India. Durante el período de incentivo, se observó un aumento significativo en la asistencia de los estudiantes, particularmente aquellos con una asistencia inicial alta. Pese a ello, una vez finalizado el período de incentivos, los estudiantes con asistencia inicial baja mostraron una disminución en su asistencia y rendimiento académico en comparación con el grupo de control.

Además, los investigadores encontraron que los incentivos no solo afectaron la asistencia, sino también el interés de los estudiantes por el material escolar y su optimismo y confianza

en sus propias habilidades. Entre los estudiantes con baja asistencia inicial, el incentivo se asoció con una menor motivación intrínseca, menor interés en los temas escolares y una visión más pesimista de sus propias capacidades. Estos efectos negativos persistieron después de que se retiraron los incentivos, sugiriendo que los incentivos pueden tener consecuencias no deseadas a largo plazo, especialmente para aquellos estudiantes que se esperaba beneficiaran más de los incentivos. Del mismo modo, a la hora de incentivar un determinado comportamiento ya sea en niños (Lepper *et al.*, 1973) o en adultos (Deci, 1971, Ariely *et al.*, 2009b), pagar para realizar una acción (ya sea de manera monetaria o vía premios), contrario a lo que podría predecirse según la teoría tradicional, provoca que el individuo se encuentre menos interesado en realizar dicha actividad, lo que disminuye su rendimiento, productividad y creatividad. Esto provoca que, por ejemplo, un bono salarial termine resultando en una disminución en la productividad (como muestra Pink, 2011), así como en qué tanto disfruta la actividad que está realizando (Kamenica, 2012). Lo mismo se ha observado en la productividad del sector público, donde incentivos monetarios no mejoran el rendimiento de los empleados (Harris-McLeod, 2013).

A pesar, pues, de que las empresas incurren en costos tales como un bono por desempeño o un aumento de sueldo esperando que se transforme en una inversión, este acto termina siendo contraproducente (Coll y Maceri, 2023). Esto es en parte indiferente de la cuantía del incentivo monetario propuesto.

Gneezy y Rustichini (2000b) presentan estudios de laboratorio y de campo en el cual quienes no reciben ningún premio mostraron mayor esfuerzo que quienes recibían un pequeño pago. Con pagos más grandes los efectos se replican. La Reserva Federal de Estados Unidos, en un estudio llevado a cabo por Ariely *et al.* (2009b) en India (donde, dado el menor ingreso, la intervención pudo ser llevada a cabo con montos relevantes para los sujetos de estudios) observaron que la performance de los individuos estaba negativamente correlacionada con un incentivo monetario alto.

Esto ha sido replicado en múltiples países y contextos. Por ejemplo, en la investigación del aprendizaje y la motivación, Murayama *et al.* (2010) concluyeron el modo en que las recompensas externas pueden socavar la motivación intrínseca al desplazar el enfoque de la tarea misma a la obtención de la recompensa, disminuyendo el disfrute y el compromiso personal con la tarea.

La teoría económica tradicional del principal – agente sugiere que una intervención externa (como una recompensa monetaria) incrementará el rendimiento laboral al aumentar el costo marginal de no trabajar o el beneficio por llevar a cabo el trabajo (Frey & Jegen, 2001). En este contexto la motivación intrínseca se considera implícitamente constante o inexistente, lo que arroja una conclusión obvia: la intervención externa vía incentivos monetarios mejora el rendimiento laboral. Esto, sin embargo, no tiene el apoyo total de la evidencia empírica obtenida por las ciencias del comportamiento.



En su libro *Drive: The Surprising Truth About What Motivates Us* (2009), Pink recolecta evidencia para desafiar las nociones tradicionales sobre motivación, particularmente en el contexto laboral. Pink explica que el sistema tradicional de recompensas y castigos, o lo que él denomina motivación "si-entonces" (if-then), es efectivo únicamente para un conjunto limitado de tareas, principalmente aquellas que son simples, rutinarias y no requieren pensamiento creativo.

Según los estudios detallados por Pink, y apoyados por las afirmaciones de psicólogos de renombre en las ciencias de la motivación (como Deci y Ryan, autores que profundizaremos a lo largo del presente trabajo), las recompensas extrínsecas, como los incentivos monetarios, funcionan bien para tareas que pueden realizarse de manera automática o que requieren poco esfuerzo cognitivo. En estos casos, el efecto de precio relativo tradicional se mantiene: un mayor incentivo monetario generalmente resulta en un mayor esfuerzo o productividad.

Para tareas que involucran cualquier nivel de complejidad cognitiva, creatividad o pensamiento conceptual, Pink argumenta que los incentivos extrínsecos son ineficaces a la vez que podrían ser contraproducentes. En estos casos, la introducción de recompensas monetarias puede estrechar el enfoque y reducir la capacidad de resolución de problemas, limitando el pensamiento creativo necesario para encontrar soluciones innovadoras.

He aquí una anomalía: la introducción de un incentivo monetario, contrario a lo que predice la teoría del principal – agente, puede resultar tanto en individuos menos motivados como también en sujetos con peor rendimiento. Las ciencias del comportamiento son una alternativa viable para incorporar los avances psicológicos a los modelos tradicionales. Este enfoque alternativo puede ser la base de nuevos esquemas de incentivos para que la disciplina se adecue a los nuevos cambios en las relaciones laborales producto de la inteligencia artificial.

### **Incorporando los avances psicológicos en la teoría económica moderna**

La distinción entre tipos de tareas y su relación con diferentes formas de motivación tiene implicaciones significativas para el diseño de sistemas de incentivos. Investigar sobre la motivación se centra en entender qué impulsa a las personas a tomar ciertas decisiones y cómo estas decisiones afectan a las dinámicas de mercado, de las organizaciones o de otros sectores de la economía. En este caso, el foco de análisis es el motivo que lleva a la acción humana (el significado etimológico de la palabra proviene de “motus” o “causa del movimiento”).

En la actualidad son los economistas del comportamiento quienes basan sus explicaciones en hipótesis empíricas sobre cómo piensan y actúan los seres humanos más allá de deducciones utilizadas bajo supuestos de racionalidad. Asimismo son ellos quienes, con este objeto de estudio, retoman la tradición psicológica como parte de la disciplina

económica, la cual, si bien casi abandonada, había sido parte de la economía desde sus comienzos (Bruni & Sugden, 2007).

En “The road not taken: How psychology was removed from economics, and how it might be brought back” (2007), Bruni y Sugden desarrollan como, antes de Pareto, la economía neoclásica se basaba en la psicología teórica y experimental, similar a la economía del comportamiento actual. Los autores exploran la manera en que el "giro paretiano" eliminó los conceptos psicológicos de la economía, basando la teoría económica en principios de elección racional. En el mismo trabajo exponen el modo en que la economía del comportamiento intenta revertir este cambio, reintegrando enfoques psicológicos para explicar los fenómenos económicos. Los autores adjudican a Pareto el hecho de que, al enfatizar sobre la elección racional y la utilidad marginal, la economía se haya vuelto más matemática y abstracta, alejándose del “realismo” que supone desarrollar las teorías entrelazándose con los avances psicológicos recientes.

En el mismo sentido, Easterlin (2021) también sostiene que es Pareto el responsable de un cambio profundo en la disciplina económica. En efecto, la economía dejó de preocuparse por el bienestar, para enfocarse solamente en la toma de decisiones a través de la matematización del comportamiento humano.

Sin embargo, Pareto afirma:

La fundación de la economía política, y, en general, de toda ciencia social, es evidentemente psicológica. Puede llegar el día en que seamos capaces de decidir las leyes de las ciencias sociales a partir de los principios de la psicología. (Cervellati, 2018, pág. 94, traducción propia).

Siguiendo la misma línea, sostiene John Maurice Clark:

El economista puede intentar ignorar la psicología, pero es imposible hacerlo al estudiar la naturaleza humana, ya que la economía es una ciencia del comportamiento humano. Cualquier concepción de la naturaleza humana que el economista adopte es una cuestión psicológica, y cualquier concepción del comportamiento humano que adopte implica supuestos psicológicos, ya sean explícitos o no. Si el economista toma su concepción del hombre del psicólogo, su trabajo constructivo puede tener alguna posibilidad de permanecer puramente económico. Pero si no lo hace, no evitará la psicología; más bien, se verá obligado a crear la suya propia, y será una mala psicología.” (Clark, 1918, pág. 4, traducción propia).

La economía basó su comprensión del ser humano en el consenso psicológico de entonces -el comportamiento humano era entendido como aquel que se llevaba a cabo por causas biológicas- y en la recompensa o castigo que brindaba la acción en sí. El mismo Darwin adjudicó la explicación de las acciones humanas en reacciones instintivas evolutivas (Ryan

*et al.*, 2019). De esta manera, acciones tales como comer, dormir o reproducirse son llevadas a cabo por instintos, más allá del esquema de incentivos o las circunstancias que giren alrededor del individuo llevando a cabo la acción.

Otras teorías, tales como el conductismo, contemplaron que, en otras acciones, el rol de los factores externos en los que se encuentra cada individuo modifica considerablemente su curso de acción. En base al premio (recompensa) o pena (castigo) que brinde un determinado curso de acción, el individuo actuará en una u otra dirección. Esto es lo que hoy se entiende como motivación extrínseca, es decir, acciones que se llevan a cabo no por la acción en sí sino por las consecuencias obtenidas por ella. Se trata, pues, de un comportamiento impulsado por la búsqueda de un objetivo, no inherente al comportamiento y que depende en gran medida del entorno externo (Wu & Lu, 2013).

Una acción motivada extrínsecamente es, entonces, aquella que es realizada debido a las consecuencias que se obtendrán producto de la acción. Las acciones que son un medio para un fin (un trabajo por el dinero que se obtiene al finalizar el mes, por ejemplo) son acciones motivadas por causas externas, y los procesos neurológicos subyacentes a este tipo de comportamiento son distintos a aquellos que son llevados a cabo por razones intrínsecas a la actividad (Pink, 2011).

En sintonía con la tradición utilitarista, si la consecuencia de una acción conlleva la obtención de placer por parte de quien la ejecuta, el sujeto se verá motivado (extrínsecamente) a realizarla. Si, por el contrario, la consecuencia implica alguna forma de dolor, ésta será interpretada en tanto castigo, configurándose como un desincentivo para la conducta prevista. El *homo economicus* obtiene deleite al satisfacer su anhelo de riqueza, por lo que las acciones percibidas como vías para la consecución de dicha riqueza serán las predilectas por este agente según la predicción económica tradicional. En aras de inducir que un individuo se conduzca de una manera determinada, ofrecer una compensación pecuniaria a cambio de una acción será concebida como la fuerza motivante más poderosa, únicamente superada por las razones biológicas, para modificar el comportamiento humano.

El *homo economicus* es un sujeto motivado principalmente debido a razones externas. Esta concepción del ser humano conlleva, pues, una noción lógica: el individuo, egoísta y racional, actuará siguiendo su propio interés. Esto es, siguiendo su propio placer, que obtiene primordialmente vía el consumo y la acumulación de riqueza. ¿Cómo modificar el comportamiento del *homo economicus*? A través de incentivos externos. El dinero emerge como el principal candidato, dado que tenerlo implica saciar su principal deseo. Con él podrá consumir, actividad a través de la cual obtendrá placer, que bajo el prisma utilitarista es felicidad.

No es difícil ver las consecuencias de esta línea de pensamiento. Para hacerlo no es necesario ir a la teoría o pasar mucho tiempo buscando realizar un análisis deductivo de las implicancias del ser humano y la motivación entendidas de esta manera, sino que

resulta suficiente con observar el sistema actual de incentivos. El incentivo monetario es visto como el incentivo de mayor ponderación a la hora de motivar un determinado comportamiento. Ofrecer dinero (o algún otro tipo de recompensa extrínseca) a cambio de un comportamiento específico o para lograr un mejor desempeño en ese mismo comportamiento es usual no solo en las prácticas empresariales sino en la manera de intentar incentivar a una persona en general, ya sea en el sector público, privado o educativo.

Es cierto que existen pocas cosas más medibles y de más fácil acceso para el encargado de diseñar el sistema de incentivos que el uso de dinero o algún otro tipo de premio para motivar un comportamiento particular. Este mecanismo simplifica la realidad de manera correcta, permitiendo gran facilidad a la hora de llevar a cabo políticas o entender el funcionamiento del ser humano en general. Sin embargo, como hemos explicado en la sección anterior, esto no siempre se cumple. Las razones detrás de este suceso, y de que la teoría económica en su mayor parte no se haya puesto al día con la teoría psicológica, se explican debido a que, al desviarse de ella, la economía perdió de vista los últimos avances que desarrolló la psicología.

Si bien los economistas suelen asumir que incentivos monetarios mejoran el rendimiento, los psicólogos afirman que puede suceder el efecto opuesto (Gneezy & Rustichini, 2000a). Del mismo modo, por el lado de los castigos, una intervención económica que busca servir como pena para desincentivar un comportamiento específico, puede no producir resultados, o producir resultados contrarios a los deseados. Uno de los experimentos más populares en economía del comportamiento, en cual se buscó investigar el impacto de las sanciones económicas en el comportamiento humano, así lo muestra.

En el trabajo de Gneezy & Rustichini (2000a) se introdujo una multa a los padres que llegaban tarde a recoger a sus hijos del jardín. Contrario a la predicción esperada, donde las multas deberían reducir la incidencia de llegadas tardías, los investigadores observaron un aumento en el número de padres que llegaban tarde una vez impuesta la multa. La interpretación de los investigadores sobre la evidencia obtenida fue que la imposición de una sanción económica cambió la percepción de los padres sobre llegar tarde, pasando de ser considerado un mal comportamiento social a simplemente un servicio por el que se podía pagar. Los padres, entonces, sabiendo el precio por su mal comportamiento, lo aceptaron, y comenzaron a llegar aún más tarde y con mayor frecuencia que previo a la intervención por parte de los economistas.

Evidencia de este tipo muestra la necesidad de entender con mayor precisión la motivación detrás del comportamiento humano. En efecto, lo que ocurre en economía, y la razón detrás de que intervenciones económicas puedan fallar en lograr su objetivo deseado es que, al contemplar la motivación extrínseca como único factor más allá de motivos biológicos o de supervivencia a la hora de modificar el comportamiento, la disciplina está dejando de lado un gran avance de la psicología. La motivación extrínseca es solo el principio de la

motivación humana, y en realidad existe una fuerza mucho mayor (Pink, 2011). Como desarrollaremos a lo largo de este trabajo, consideramos que hay razones para creer que la transformación laboral y social producto del avance tecnológico implicará la necesidad de actualizar el sistema tradicional de incentivos.

A pesar de que la economía se haya dejado de preocupar, en gran medida, por los aspectos psicológicos, la psicología, obviamente, siguió su camino. Harlow (1950) fue uno de los primeros que, realizando un experimento de laboratorio en primates, observó cómo, sin necesidad de instaurar recompensas, los objetos de estudio comenzaron a intentar resolver un rompecabezas de manera voluntaria. Llevar a cabo una actividad difícil, que requería esfuerzo sin recibir nada a cambio y sin satisfacer ninguna necesidad biológica, no tenía una explicación teórica. Reconocido por el propio Harlow, sus hallazgos iban en contra de la literatura de la época, que contemplaba en las razones biológicas o de recompensas y castigos, los únicos motivos de acción del comportamiento.

Los estudios de Harlow dieron origen al concepto de motivación intrínseca. El investigador sospechó, y experimentos posteriores lo confirmaron, que los primates encontraban motivación para resolver rompecabezas en el disfrute mismo de la acción, lo cual era un incentivo suficiente para llevar a cabo dicho comportamiento. Las argumentaciones de Harlow sobre la incapacidad de la teoría tradicional para explicar fehacientemente temas como la curiosidad o los comportamientos exploratorios desafiaron la teoría ortodoxa sobre la motivación humana. Este desafío conllevó un debate dentro del campo de la motivación (Ryan *et al.*, 2019): entre quienes buscaban defender la teoría tradicional de recompensas y castigos, y quienes buscaban explicaciones alternativas ante los avances empíricos de la disciplina.

Investigaciones posteriores, principalmente Deci (1971), marcaron un punto de inflexión al mostrar que la aplicación de recompensas a actividades inherentemente interesantes implicaba un efecto perjudicial en el comportamiento de "libre elección" posterior de los participantes, dependiendo de cómo se otorgaran dichas recompensas. Deci encontró que si las recompensas, como el elogio, aumentaban la percepción de competencia en la actividad realizada, la motivación intrínseca se mantenía. Sin embargo, si las recompensas indicaban una percepción de un control externo, la motivación intrínseca se veía socavada. Lo que encontró el psicólogo fue que las recompensas afectan la motivación intrínseca en función de su relación con las necesidades psicológicas de autonomía (locus de causalidad percibido) y competencia o eficacia. Algo crucial de estos experimentos fue la demostración de que la motivación podía manipularse no con las recompensas y castigos impuestos sino, en cambio, modificando las condiciones que apoyan o frustran necesidades psicológicas como la autonomía o competencia (Ryan *et al.*, 2019).

Los descubrimientos de Deci tenían su utilidad, en primer lugar, por explicar lo que hoy entendemos es un tercer impulso, más frágil, pero también más intenso que los biológicos o los externos. Además, los resultados de los experimentos no solamente mostraban que

había una tercera manera de motivar a los seres humanos, sino que quienes estaban motivados de manera intrínseca tenían un mejor rendimiento que aquellos motivados por la obtención de una recompensa o la evitación de un castigo. Deci (Pink, 2011) aconsejó entonces que aquellos interesados en desarrollar y potenciar la motivación intrínseca en niños, empleados, estudiantes, etc., no deberían enfocarse en sistemas de control externo, como las recompensas monetarias.

No se trata únicamente de que la motivación intrínseca existe como algo separado de motivos biológicos o extrínsecos; se trata, además, de que quienes realizan su comportamiento por razones de este tipo son más creativos, persistentes y rinden comparativamente mejor que quienes encuentran su motivo en razones externas (Deci & Ryan, 2000). Estos descubrimientos, junto con estudios posteriores, desembocaron en la teoría de la autodeterminación (Deci & Ryan, 1985). Esta teoría es un marco teórico basado en la evidencia empírica producto de las investigaciones que sugiere que existen tres necesidades psicológicas universales que son esenciales para el fomento de la motivación intrínseca y el bienestar psicológico: autonomía, competencia y relaciones sociales.

La teoría de la autodeterminación ha ganado un lugar clave en el campo de la motivación, convirtiéndose en una concepción aceptada por las ciencias del comportamiento acerca de las necesidades que debe satisfacer el ser humano. Propone que la satisfacción de las necesidades básicas de autonomía, competencia y relaciones sociales es fundamental para fomentar la motivación intrínseca, lo que a su vez resulta en individuos más productivos y capaces. Este enfoque expande las suposiciones tradicionales de la economía que considera al ser humano como un agente racional que maximiza su utilidad basándose únicamente en incentivos extrínsecos.

Esta teoría complementa las debilidades ya mencionadas que sufren los incentivos puramente extrínsecos o monetarios. El deseo de riqueza propio del *homo economicus* no se encuentra simplemente ausente. En este escenario puede incluso intentar directamente en las probabilidades de obtención de los otros anhelos propios del *homo sapiens*, al menos según la teoría psicológica actual.

Comprender a cabalidad la motivación intrínseca tiene implicaciones trascendentales, por ejemplo, al diseñar políticas públicas y prácticas organizacionales. Los programas de incentivos ya sean estatales o empresariales pueden estructurarse para fomentar la autonomía y la competencia, en lugar de menoscabar la motivación intrínseca. Esto sugiere un giro respecto a los enfoques tradicionales basados en recompensas y castigos, hacia estrategias que cultiven la satisfacción de necesidades psicológicas fundamentales para el ser humano.

Del mismo modo, Ryan (Pink, 2011), co-creador junto con Deci de la teoría de la autodeterminación, sostiene que, al intentar mejorar la productividad laboral, las compañías erróneamente recurren a incentivos basados en recompensas y castigos. Si bien las recompensas son necesarias, "cuanto más discretas sean, mejor. Cuando la gente utiliza

recompensas para motivar, es cuando se desmotivan más" (pp. 38-39, traducción propia). En cambio, Deci y Ryan afirman que deberíamos enfocar nuestros esfuerzos en crear entornos que permitan florecer nuestras necesidades psicológicas innatas.

En el ámbito empresarial quienes han diseñado un sistema de incentivos en busca de aumentar la motivación intrínseca han notado cómo sus tácticas resultaban exitosas, en vistas del aumento de productividad experimentado (para ejemplos ver Pink, 2011). La aplicación de estos avances también se ha dado en la teoría, con el intento de construcción de modelos que incluyan la motivación intrínseca como una variable más. Sin embargo, la teoría económica tradicional no diferencia las distintas fuentes de motivación dado que son interpretadas como preferencias, siendo la motivación extrínseca la única que forma parte de los modelos tradicionales, dejando a la motivación intrínseca como exógena, dada la dificultad que conlleva su medición (Frey & Jegen, 2001).

En definitiva, al referirnos a la motivación intrínseca, aludimos al deseo de satisfacer las tres necesidades psicológicas propuestas por Ryan y Deci en un entorno que coincida con las preferencias individuales, las cuales la economía suele considerar como factores exógenos (Charles-Leija *et al.*, 2018). De este modo, una actividad será llevada a cabo por un motivo intrínseco si la propia acción, más allá de la recompensa o las consecuencias, genera disfrute o placer, sensaciones que provienen mayormente de ver cumplidas sus necesidades psicológicas o de realizar algo acorde a las preferencias personales.

### **Efecto *crowding out***

La diferenciación entre motivación intrínseca y extrínseca no es algo innecesario, como quizás podría parecer. La motivación intrínseca es inobservable mientras que la extrínseca es fácilmente maleable y cuantificable. Suponer constante o exógena al primer tipo de motivación para concentrarse en la motivación extrínseca como la variable a controlar de aquellos encargados de diseñar el sistema de incentivos puede parecer una decisión razonable.

El problema, sin embargo, ocurre vía dos procesos. En primer lugar, como ya fue mencionado, individuos motivados de manera extrínseca tienen una performance diferente a los motivados de manera intrínseca, siendo estos últimos más eficientes, creativos y con mejor capacidad de resolución de problemas, entre otros (para ejemplos ver Achor, 2010). Esto implica que un individuo motivado de manera extrínseca no es un sustituto perfecto de uno motivado de manera intrínseca.

Si se opta por no controlar la variable intrínseca de la motivación de un empleado ésta no puede compensarse con una suba en un salario u otro intento de recompensa que funcione como motivación extrínseca, porque la productividad deseada en una gran cantidad de tareas será menor. En trabajos que impliquen creatividad o aunque sea un uso cognitivo rudimentario (Pink, 2011), empleados motivados de manera extrínseca no serán igual de capaces que aquellos motivados de manera intrínseca. Análogamente, un mismo empleado

que ha perdido su motivación intrínseca no volverá a sus niveles de productividad anteriores si se compensa la pérdida de este tipo de motivación a través de un bono o aumento salarial.

Se infiere, pues, una necesidad de no tratar a la motivación intrínseca como dada. El potencial y las repercusiones de la motivación intrínseca son mucho más poderosos para modificar o mejorar la calidad del comportamiento de los individuos que el simple incremento del dinero que recibe el empleado. Además, los costos de aumentar la motivación extrínseca son directos y muchas veces más costosos que el que implica aumentar el segundo tipo de motivación, la intrínseca.

A pesar de lo ya mencionado, resulta en cierta parte lógico el alto énfasis a la motivación extrínseca por parte de la teoría y la práctica empresarial. A diferencia de la intrínseca, la motivación extrínseca puede modificarse fácilmente. Un aumento salarial, un bono por desempeño o cualquier otro tipo de premio servirán para aumentar este tipo de motivación. Ahora bien, si se quiere lograr lo mismo en el caso de la motivación intrínseca esto implicará aumentar algo que no puede medirse con facilidad, por lo que la magnitud de análisis y del éxito de la intervención resultarán difícilmente cuantificables, razón por la cual la economía ha optado en cierta parte por catalogarla meramente como exógena (Frey & Jegen, 2001), siguiendo a Friedman (Charles-Leija *et al.*, 2018) y a la economía posterior a Pareto (Easterlin, 2021).

No obstante, los incentivos monetarios se han erigido como un eje de controversia en la economía experimental y otras disciplinas afines como la psicología (Bardsley *et al.*, 2010). Se trata de una polémica con sólidos fundamentos empíricos. Pareciera que al relegar la motivación intrínseca a un plano inferior y menguar las posibilidades de que forme parte de una intervención, no solo se desaprovechan poderosos impulsores, sino que el empleo de la motivación extrínseca termina por forjar precisamente el efecto opuesto al perseguido. Este fenómeno es conocido como efecto *crowding out* en el ámbito de la motivación, el cual será desarrollado a continuación.

Cuando ya existe motivación intrínseca, en trabajos que requieren creatividad o incluso una mínima cantidad de esfuerzo cognitivo, elevar los niveles de recompensa del sujeto que está llevando a cabo la acción tiene un efecto perjudicial sobre el rendimiento (Ariely *et al.*, 2009b). Realizar una acción por el premio, provoca una reducción en el enfoque, limita la capacidad para la resolución de problemas e implica la pérdida de la motivación intrínseca que era la responsable de que se llevara a cabo la acción previa a la intervención (Pink, 2011). La motivación extrínseca “desplaza” a la motivación intrínseca, produciendo el llamado efecto *crowding out*, de sobre justificación (Lepper *et al.*, 1973) o el costo oculto de la recompensa (Lepper & Greene, 1978). Esto, en realidad, se sospecha desde el principio. Los estudios de Harlow ya mencionados no tienen su novedad simplemente en el acto de los primates sin el ofrecimiento de una recompensa. Siguiendo la línea de experimentos, Harlow ofreció recompensas a aquellos primates que ya estaban resolviendo rompecabezas simplemente



porque lo estaban disfrutando. Contrario a la teoría tradicional los primates realizaron la acción con menos frecuencia y tuvieron un peor desempeño en la tarea (Pink, 2011).

La idea de que incentivos monetarios eliminan la motivación para llevar a cabo una acción es considerada una gran anomalía en el campo económico dado que la teoría tradicional predice exactamente el comportamiento opuesto, presuponiendo un comportamiento acorde al efecto del precio relativo (Frey & Jegen, 2001). Pero la evidencia proveniente de psicología y de la economía del comportamiento es alta, y desde los estudios de Harlow muchas replicas han sido realizadas, en su mayor parte con éxito.

Deci *et al.* (1999) encontraron, en su meta-estudio, que cuando a los individuos se les ofrecían recompensas tangibles por realizar una actividad que ya encontraban interesante, su motivación intrínseca para realizar esa actividad disminuía. Este fenómeno se observó especialmente cuando las recompensas se presentaban como contingentes al desempeño, lo que llevaba a los participantes a sentir que su comportamiento estaba siendo controlado externamente, reduciendo así su interés intrínseco en la tarea.

Del mismo modo, Gneezy y Rustichini (2000a) investigaron el efecto de las recompensas monetarias en la asistencia de niños a clases. Los investigadores encontraron que, aunque inicialmente las recompensas monetarias se tradujeron en un incremento de la asistencia, este efecto se desvaneció con el tiempo y, una vez eliminadas las recompensas, la asistencia cayó a niveles incluso más bajos que los observados antes de la introducción de las recompensas. Ariely *et al.* (2009a) observaron lo mismo (un incremento en la motivación extrínseca disminuye la intrínseca y con ello la motivación total de querer realizar el comportamiento) en los comportamientos del tipo prosocial dado el cambio en la imagen de la acción realizada. Estos hallazgos son acordes a los mencionados por Titmuss (1970), quien había comenzado las mismas discusiones, inicialmente de manera teórica, afirmando que pagar para incentivar la donación de sangre iba a disminuirla, dado que la acción iba a ser percibida no como una altruista sino motivada por incentivos monetarias.

Las intuiciones de Titmuss no estaban relacionada con los avances que se estaban llevando a cabo en psicología (Frey & Jegen, 2001), pero llegaban a una conclusión similar. Incentivos extrínsecos (siendo el más común el monetario) puede decrecer el deseo de llevar a cabo la acción, especialmente si la acción ya era deseada antes de la intervención. Además, el rendimiento disminuirá y será peor que quienes no hayan sido, en este caso, víctimas de dicha intervención. Es decir, un incentivo monetario (aumento en la motivación extrínseca) disminuirá la motivación total (dado que disminuye la motivación intrínseca) y el rendimiento (debido a las diferencias entre una acción efectuada por motivos intrínsecos en comparación con una llevada a cabo por razones extrínsecas).

En el ámbito de la economía aplicada, la efectividad de los mecanismos de incentivos dentro de las estructuras organizacionales también ha sido objeto de debate. Central a este discurso es la prevalencia del modelo de *pay for performance* (pago por desempeño), que ha emergido como una herramienta predominante para la búsqueda de un rendimiento

superior por parte de los empleados. Esta metodología se arraiga profundamente en la aplicación práctica del efecto precio. Así, la creencia es que los incentivos entre el principal y el agente estarán alineados de manera óptima cuando el esquema de remuneración se basa en el desempeño o al menos se premia la productividad mediante aumentos salariales o bonos por rendimiento laboral (en otros ámbitos sucede lo mismo con premios por rendimiento escolar, por ejemplo). Esta serie de pensamientos ha influenciado gran parte del esquema de incentivos moderno.

En "Punished by Rewards," Kohn (1993) presenta varios casos que evidencian cómo los incentivos extrínsecos pueden afectar negativamente la motivación intrínseca en el ámbito laboral. Por ejemplo, el realizado por Deci en una empresa de alta tecnología. En este estudio, Deci encontró que cuando los empleados recibían recompensas monetarias por completar tareas, su motivación intrínseca disminuía significativamente. Inicialmente, los empleados mostraban un alto nivel de interés y satisfacción en su trabajo, pero con la introducción de recompensas monetarias, comenzaron a ver sus tareas como obligaciones en lugar de actividades interesantes y satisfactorias.

Otro estudio citado por Kohn muestra el impacto de los bonos financieros en la motivación de los empleados. Los resultados mostraron que los empleados que recibieron incentivos financieros por su desempeño experimentaron una disminución en su motivación intrínseca para realizar su trabajo. A pesar de que los bonos estaban destinados a aumentar el rendimiento, en realidad llevaron a una reducción en la satisfacción laboral y el interés por las tareas que los empleados solían disfrutar.

Kohn también menciona las investigaciones de Teresa Amabile sobre la creatividad en el lugar de trabajo. Amabile concluyó que los incentivos financieros pueden tener un impacto negativo en la creatividad de los empleados. Amabile observó que cuando se ofrecían recompensas monetarias por trabajos creativos, los empleados producían trabajos de menor calidad y demostraban menos creatividad. La presencia de recompensas externas llevó a una disminución de la motivación intrínseca, lo que a su vez inhibió el pensamiento creativo. Este hallazgo es particularmente relevante en entornos laborales donde la innovación y la creatividad son esenciales para el éxito a largo plazo.

Al incurrir en un bono o aumento salarial las empresas incurren en un costo, esperando que se transforme en una inversión para obtener empleados más motivados y productivos. Sin embargo, el resultado final es que el aumento en la motivación extrínseca provoca una disminución en la motivación intrínseca, lo que disminuye la capacidad total del empleado. Para hacer factible esta posibilidad los empleados deben percibir que su trabajo tiene un propósito, más allá del de proporcionar una vida digna a través del intercambio dinero-mano de obra que se propicia a la hora de firmar un contrato laboral. En trabajos que requieren creatividad, innovación y esfuerzo cognitivo, el empleado no necesita solamente sentirse seguro económicamente sino también ver cumplidas las necesidades psicológicas

básicas de competencia, relación y autonomía para experimentar niveles mayores de vitalidad, automotivación y bienestar (Ryan, 2009).

## **La transformación radical de la inteligencia artificial en perspectiva histórica**

La inteligencia artificial (IA) representa un hito en la historia de la innovación tecnológica. Para comprender plenamente el alcance y las implicaciones de este fenómeno, es crucial situarlo en el contexto más amplio de la evolución tecnológica y económica.

Históricamente, la humanidad ha experimentado varias revoluciones industriales, cada una de las cuales ha transformado profundamente la estructura de la economía, las relaciones laborales y la organización social. La Primera Revolución Industrial, que comenzó a finales del siglo XVIII, se caracterizó por la introducción de la máquina de vapor y la mecanización de la producción. Este período vio el surgimiento de las fábricas, la urbanización masiva y el nacimiento de la clase obrera industrial. La Segunda Revolución Industrial, a finales del siglo XIX, trajo consigo la producción en masa, la electricidad y nuevas formas de comunicación como el telégrafo. Estas innovaciones no solo aumentaron drásticamente la productividad, sino que también alteraron fundamentalmente la naturaleza del trabajo y la estructura de las sociedades.

Las revoluciones marcaron el inicio de las tecnologías que ahorran mano de obra (LST, por sus siglas en inglés), que reemplazaron el trabajo humano con máquinas. Este cambio condujo a aumentos significativos en la productividad y la eficiencia (Zarifhonarvar, 2023). La Tercera Revolución Industrial, también conocida como la revolución digital, comenzó a mediados del siglo XX con la introducción de la computación y la automatización a gran escala. Esta era vio el nacimiento de internet, la globalización de la producción y el surgimiento de la economía del conocimiento. Cada una de estas revoluciones ha traído consigo desafíos y oportunidades, desplazando ciertas formas de trabajo mientras creaba otras nuevas.

A lo largo del siglo XX, las LST continuaron evolucionando de manera notable, impulsadas por la expansión de la informática y la tecnología de la información. Esta evolución tecnológica llevó a la adopción generalizada de robots y sistemas de automatización en diversos sectores industriales y de servicios, transformando profundamente la naturaleza del trabajo y la producción (Zarifhonarvar, 2023).

Es en este contexto histórico donde debemos situar el actual cambio radical de la IA. La historia de la IA como campo de estudio se remonta a mediados del siglo XX, con el término "inteligencia artificial" acuñado en 1956 durante la histórica conferencia de Dartmouth. Sin embargo, el desarrollo de la IA ha sido un proceso no lineal, caracterizado por ciclos de entusiasmo y períodos de estancamiento, conocidos en la industria como "inviernos de la IA".

Los primeros años de la IA, desde la década de 1950 hasta la de 1970, estuvieron marcados por altos niveles de optimismo. Se hicieron avances significativos en áreas como el procesamiento simbólico y los sistemas expertos, lo que llevó a predicciones audaces sobre

el futuro de la inteligencia de las máquinas. Sin embargo, las limitaciones de la tecnología de la época pronto fueron evidentes, lo que condujo a un período de desencanto y reducción de la financiación en las décadas de 1970 y 1980.

A pesar de estos reveses, la investigación en la IA continuó experimentó un resurgimiento en la década de 1990 y principios de 2000. Este renacimiento fue impulsado en gran medida por el aumento exponencial de la potencia computacional y el desarrollo de nuevas técnicas de aprendizaje automático. Los avances en áreas como el reconocimiento de voz, la visión por computadora y el procesamiento del lenguaje natural comenzaron a demostrar el potencial práctico de la IA en una amplia gama de aplicaciones.

Es en la última década cuando hemos presenciado un salto cualitativo en las capacidades de la IA, particularmente en el campo de la IA generativa. Con ChatGPT a la cabeza, un sinnúmero de modelos de lenguaje que han demostrado capacidades sorprendentes en la generación de texto, traducción y comprensión del lenguaje natural han llegado a las manos del público. Estos sistemas pueden analizar y clasificar información existente, además de que pueden generar contenido nuevo y original, planteando preguntas muy interesantes, y a veces inquietantes, sobre la naturaleza de la creatividad y la inteligencia.

Paralelamente a los avances en la IA generativa, el campo de la robótica ha experimentado progresos significativos. La integración de IA avanzada en sistemas robóticos ha dado lugar a máquinas cada vez más autónomas y adaptables. Si bien el rol de la IA en el “mundo real” es aún menos preponderante que en el “mundo digital”, esto parece ser simplemente una cuestión de tiempo (la robótica se encuentra más atrasada que la IA generativa), y no de una imposibilidad de este tipo de tecnologías de irrumpir también en las actividades físicas que hoy llevan a cabo los seres humanos.

En su artículo "Learning from Ricardo and Thompson: Machinery and Labor in the Early Industrial Revolution, and in the Age of AI", Acemoglu & Johnson (2024) analizan el efecto de la automatización producto de la Primera Revolución Industrial, en los trabajadores y los salarios. La industria del algodón en el Reino Unido fue una de las primeras en experimentar una amplia adopción de maquinaria. Las innovaciones en el hilado y tejido, como la introducción del telar mecánico, aumentaron significativamente la productividad. Desde ya, estos avances tecnológicos también tuvieron consecuencias adversas para los trabajadores.

Ricardo sostuvo que la maquinaria beneficiaría a los trabajadores al aumentar la demanda de mano de obra, aunque observaciones posteriores llevaron a Ricardo a revisar su opinión. En la tercera edición de "Principles of Political Economy and Taxation" en 1821 (Acemoglu & Johnson, 2024), Ricardo reconoció que la maquinaria podría desplazar a los trabajadores y reducir la demanda de trabajo, especialmente cuando la automatización no iba acompañada de la creación de nuevas tareas que incrementaran la productividad marginal del trabajo.

Este caso de los tejedores manuales de algodón puede servir para ilustrar el impacto negativo de la automatización. Antes de la introducción generalizada del telar mecánico, los tejedores manuales disfrutaban de salarios relativamente altos y un considerable grado de autonomía en su trabajo. Pese a ello, la llegada de los telares mecánicos resultó en un desplome de los salarios reales y un deterioro de las condiciones laborales. Aun con la creciente productividad en la industria textil, los tejedores manuales vieron cómo sus ingresos disminuían a medida que eran reemplazados por maquinaria más eficiente.

Thompson, en su obra "The Making of the English Working Class," (Acemoglu & Johnson, 2024), argumentó que la automatización durante la Revolución Industrial desplazó a los trabajadores, como también empeoró las condiciones laborales. Los trabajadores fueron forzados a trabajar en fábricas insalubres y bajo una estricta supervisión, perdiendo así su autonomía y capacidad de negociación. La falta de poder político y la incapacidad para organizarse colectivamente también contribuyeron a que los trabajadores no pudieran exigir una parte justa de las ganancias de productividad (Acemoglu & Johnson, 2024).

El avance de la IA promete ser algo distinto a las anteriores revoluciones. Lo que distingue a la IA de las revoluciones industriales anteriores es su velocidad de adopción, la amplitud de su impacto y la naturaleza del trabajo que está transformando. Mientras que las revoluciones anteriores se desarrollaron a lo largo de décadas, los avances en IA están ocurriendo a un ritmo veloz, con nuevas capacidades que surgen casi mensualmente.

En cuanto a la amplitud de su impacto, la IA tiene el potencial de transformar prácticamente todas las industrias y aspectos de la vida humana. Empero, contrario a otras revoluciones, en esta ocasión las nuevas tecnologías (junto con las nuevas oportunidades que estas tecnologías proporcionan) se encuentran al alcance de todo usuario con un dispositivo electrónico.

Con los avances actuales, los primeros estudios sobre los efectos de la IA afirman que podrían conducir a una mejora del 20 al 80% en la productividad, lo cual se contrapone con la energía de vapor, la más fundamental de las tecnologías de uso general en la Primera Revolución Industrial, que al ser instalada en una fábrica tan solo mejoró la productividad entre un 18 y un 22% (Mollick, 2024).

El impacto esencial de la IA es su llegada masiva a manos de prácticamente todos los usuarios. A diferencia de otras revoluciones tecnológicas, en esta ocasión las empresas a cargo de la innovación han dado acceso a las nuevas tecnologías a todo usuario con acceso a internet. Esto permite una adopción más rápida y una integración veloz con el resto de las tareas cotidianas que se llevaban a cabo antes de los avances en los Grandes Modelos de Lenguaje como ChatGPT.

Además, el potencial de crecimiento exponencial de la IA añade un elemento de imprevisibilidad a esta "nueva revolución". Debido a su capacidad de aprendizaje y mejora continua, las tecnologías de IA tienen el potencial de evolucionar a un ritmo que supera

nuestra capacidad de predicción, lo que podría conllevar a avances disruptivos que transformen radicalmente nuestras sociedades en formas que aún no podemos imaginar.

Con esto en mente, el conjunto de escenarios posibles como resultado del avance de la IA puede dividirse, de manera simplificada, en dos. Por un lado, quienes argumentan que las nuevas tecnologías elevarán la productividad a niveles tan altos que existirá la posibilidad y, fundamentalmente, la necesidad, de crear algo similar a un ingreso básico universal.

Pero también existe un escenario en el cual la IA no reemplazará al trabajador humano. En efecto, existen razones para creer que la IA puede permitir un protagonismo humano incluso más activo que el actual, aunque con un cambio en su rol y una transformación radical de la cotidianeidad laboral.

Ambos escenarios requieren una readaptación de los incentivos monetarios. Comenzaremos por analizar esta última posibilidad.

### **Inteligencia artificial como complemento humano**

Sería un error interpretar el poder de la IA únicamente en términos de sustitución del trabajo humano. La IA también tiene un enorme potencial para aumentar y complementar las capacidades humanas, creando nuevas formas de colaboración entre humanos y máquinas. En última instancia, la revolución de la IA no es simplemente una cuestión de tecnología, sino un fenómeno que está redefiniendo nuestra comprensión de la inteligencia, la creatividad y el trabajo.

Quizás lo más significativo es la naturaleza del trabajo que la IA está comenzando a realizar. Las LST, que han sido una característica de las revoluciones industriales anteriores, adquieren una nueva dimensión. A diferencia de las revoluciones anteriores, las cuales principalmente automatizaron tareas físicas y repetitivas, la IA está demostrando capacidad para realizar tareas cognitivas complejas que antes se consideraban exclusivamente del dominio humano. Esto plantea preguntas importantes sobre el futuro del trabajo, la educación y el papel del ser humano en una economía. Contrario a lo que podría especularse, los primeros trabajos en ser reemplazados por la IA están siendo los creativos, y no los automáticos (Mollick, 2024).

Si bien tradicionalmente se consideraba que la IA era puramente lógica y carecía de aspectos de imaginación, los avances recientes han demostrado que las tecnologías de la IA pueden generar información, participar en discusiones aparentemente emocionales y ser intensamente creativas. Por ejemplo, investigadores de la Universidad de Wharton, organizaron un concurso de generación de ideas en el que se enfrentó ChatGPT-4 a estudiantes de una clase de innovación. Los resultados mostraron que ChatGPT-4 generó ideas más numerosas, económicas y de mejor calidad que los estudiantes, y que la intención de compra de los jueces externos fue mayor para las ideas generadas por la IA (Girotra *et al.*, 2023).

La investigación de Boussioux *et al.* se centró en un concurso de *crowdsourcing* en el que se pidió a las personas que generaran ideas de negocio basadas en la reutilización, el reciclaje o el intercambio de productos como parte de la economía circular. Al comparar estas ideas con las generadas por ChatGPT-4, se puso en evidencia que el nivel de calidad general era similar, pero la IA fue considerada mejor en cuanto a viabilidad e impacto (Boussioux *et al.*, 2024).

Lo mismo ocurre en el estudio de Bellemare-Pepin *et al.* (2024), quienes comparan a la creatividad divergente de LLM de vanguardia como ChatGPT-4 y 100.000 humanos. Utilizando métricas computacionales los autores obtienen evidencia a favor de la IA. En tareas creativas como escritura creativa y asociación divergente los Grandes Modelos de Lenguaje vuelven a superar las capacidades humanas, esta vez no acorde a jueces sino a *benchmarks* de creatividad.

Con todo, pensar que la IA viene simplemente a reemplazar los trabajos creativos sería un error. Al contrario, la IA democratiza el acceso a la creatividad, la innovación y la tecnología de punta. Las diferencias genéticas de coeficiente intelectual (IQ) y personalidad pueden a veces jugar un rol determinístico en el éxito laboral de los individuos. Quienes tienen la suerte de ser creativos o inteligentes gozan de una ventaja comparativa a comparación de aquellos que no encuentran en ese tipo de tareas sus principales fortalezas.

Del mismo modo, quienes han sido afortunados de tener una buena educación o de proceder de una familia de alto poder adquisitivo cuentan con facilidades que permiten un mejor desempeño a lo largo del periodo laboral. La IA, por el contrario, viene a subir el piso y permitir que los empleados que no disponen de esas ventajas puedan utilizar las habilidades de la IA para poder competir con quienes sí lo hacen.

El hecho de que la IA esté disponible para todos los empleados permite que haya una igualdad de oportunidades. Cualquier empleado, indistintamente de su situación económica o geográfica, cuenta en su computadora con las últimas tecnologías de inteligencia artificial. Con un correcto uso, cualquier empleado, indistintamente de si es o no creativo, de si tiene o no conocimiento general o buenos niveles de educación, cuenta en su computadora con una tecnología que podrá suplir sus fallas en este apartado, aumentando de esta manera su potencial laboral respecto a productividad, creatividad y velocidad a la hora de realizar trabajo.

En lugar de un mero reemplazo de mano de obra, la integración de la IA en el día a día laboral puede permitir un aumento de las posibilidades humanas, aunque por supuesto, también un cambio de rol del empleado humano. Nos referimos a la entrada de otro tipo de trabajador (no humano) al entorno laboral. No nos referimos solamente a máquinas que llevan a cabo tareas repetitivas, hablamos también de una IA que puede jugar un rol preponderante en la resolución de problemas, la innovación, la creatividad, la productividad y todo tipo de trabajos que hoy en día hace el humano.

Este cambio modifica por completo el rol humano y llama a una nueva estructura organizativa del entorno laboral. Algo muy similar ocurrió en el enfrentamiento de 1997 entre Deep Blue y Garry Kasparov, que marcó un hito en la historia de la IA y la comparación con las habilidades humanas. La victoria de Deep Blue sobre Kasparov se consideró inicialmente como el triunfo definitivo de la inteligencia artificial sobre la humana en el ajedrez. Este evento, sin embargo, catalizó una nueva era de colaboración entre humanos y máquinas, gracias a la complementariedad entre fortalezas y debilidades que presenta la relación humano-máquina.

La paradoja de Moravec, propuesta por el científico Hans Moravec en la década de 1980, plantea una observación contraintuitiva sobre la relación entre la inteligencia artificial y las capacidades humanas. Esta paradoja postula que las tareas que son relativamente fáciles para los humanos a menudo resultan extremadamente difíciles de replicar en sistemas de IA, mientras que las tareas que son desafiantes para los humanos pueden ser realizadas con relativa facilidad por las máquinas. La paradoja se manifiesta claramente en el contraste entre las habilidades cognitivas de alto nivel y las habilidades sensoriomotoras básicas. Por ejemplo, mientras las computadoras han demostrado una capacidad sobresaliente para jugar al ajedrez, realizar cálculos complejos y procesar vastas cantidades de datos, tareas que requieren un considerable esfuerzo cognitivo para los humanos, acciones aparentemente simples como reconocer objetos en diversos contextos, manipular objetos con destreza o navegar en entornos no estructurados, que los humanos realizan sin esfuerzo consciente, han resultado ser sorprendentemente difíciles de implementar en sistemas de IA.

Agrawal (2010) desarrolla una explicación para esta paradoja, la cual radica en la evolución de la cognición humana. Las habilidades que consideramos "simples" son el resultado de millones de años de evolución y están profundamente arraigadas en nuestras estructuras neuronales más primitivas. En contraste, las habilidades cognitivas de alto nivel, como el razonamiento abstracto o el cálculo matemático, son evolutivamente recientes y se basan en estructuras cerebrales más nuevas. Paradójicamente, estas habilidades más recientes son más fáciles de formalizar y, por lo tanto, de replicar en sistemas artificiales.

En el ajedrez las computadoras sobresalen en cálculos tácticos, mientras que los humanos destacan en pensamiento estratégico y planificación a largo plazo, por lo que Kasparov (Epstein, 2021) introdujo el "ajedrez avanzado", que aprovechó esta complementariedad. En este formato, cada jugador humano se emparejaba con una computadora, permitiendo a los humanos centrarse en la estrategia general mientras delegaban los cálculos tácticos a la máquina. Esta colaboración resultó en un nivel de juego superior al que podían alcanzar humanos o máquinas por separado.

Kasparov (Epstein, 2021), quien apenas un mes antes había derrotado contundentemente a un oponente con un marcador de 4-0 en un encuentro tradicional, se vio forzado a conformarse con un empate 3-3 en este nuevo formato: reconoció que su ventaja en el



cálculo táctico, desarrollada a lo largo de años de entrenamiento especializado, había sido efectivamente neutralizada por la máquina. En consecuencia, en un contexto donde los participantes humanos podían concentrarse primordialmente en aspectos estratégicos, Kasparov se encontró compitiendo en igualdad de condiciones con jugadores que previamente no se consideraban sus pares.

Un ejercicio por el estilo no implicaba delegar la totalidad del trabajo en manos de la computadora. El rol del humano en materia estratégica seguía siendo crucial. En palabras de Kasparov (Epstein, 2021, pág. 29, traducción propia), “la creatividad humana era aún más primordial en estas condiciones, no menos”.

El aspecto táctico del ajedrez, que tradicionalmente requiere años de experiencia y práctica intensiva, frecuentemente iniciada a una edad temprana para alcanzar niveles de élite, se vio neutralizado por la capacidad computacional de la máquina. No obstante, esta delegación de las responsabilidades tácticas a la IA, un área en la que las computadoras demuestran una clara superioridad sobre los humanos, permitió a los jugadores redirigir su atención hacia tareas de igual o mayor importancia estratégica, donde las capacidades cognitivas humanas mantienen una ventaja comparativa significativa. La redistribución de enfoque resalta la complementariedad potencial entre las habilidades humanas y la inteligencia artificial en contextos de toma de decisiones complejas.

La evolución de esta idea llevó al "ajedrez libre" o "*freestyle*": equipos compuestos por múltiples humanos y computadoras competían entre sí. Sorprendentemente, equipos de jugadores aficionados con computadoras estándar lograron derrotar a supercomputadoras especializadas y a equipos de grandes maestros con asistencia computacional.

El éxito de estos equipos "centauros" (combinaciones de humanos y computadoras) demostró que la clave no residía en la habilidad individual para jugar al ajedrez, sino en la capacidad de los humanos para dirigir eficazmente a las computadoras y sintetizar la información para formar estrategias generales. Este enfoque permitió alcanzar niveles de juego sin precedentes.

El caso de Anson Williams en el torneo de ajedrez libre de 2014 es particularmente ilustrativo. Williams, un jugador de ajedrez aficionado, fue el capitán del equipo ganador del torneo. Si bien no poseía altas habilidades de ajedrez, tenía experiencia en computación y habilidad para integrar información en tiempo real para tomar decisiones estratégicas (Epstein, 2021). Así, un jugador sin altas habilidades, especialmente en el ajedrez, pudo ganar el torneo gracias a su capacidad de toma de decisiones en tiempo real y el trabajo en equipo junto con otros humanos y máquinas especializadas.

Según Jarrahi (2018), los procesos de toma de decisiones en las organizaciones se caracterizan típicamente por tres desafíos principales: incertidumbre, complejidad y ambigüedad. En este contexto, la IA, con su mayor capacidad de procesamiento de información y enfoque analítico, puede extender la cognición humana al abordar la

complejidad. Por su parte, los humanos siguen ofreciendo un enfoque más holístico e intuitivo para lidiar con la incertidumbre y la ambigüedad en la toma de decisiones organizacionales.

La complementariedad y relación simbiótica con la IA implica la posibilidad de que el humano vea aumentada su inteligencia (*Aumented Intelligence*) en lugar de ser simplemente sobrepasados por las capacidades de las nuevas tecnologías. En este sentido, el papel del trabajador humano en el entorno laboral futuro no es simplemente el de ser reemplazado por la IA, sino el de evolucionar hacia un rol de colaboración más sofisticado con los sistemas de IA.

Jarrahi (2018) argumenta que mientras la IA puede ser superior en el manejo de tareas analíticas y en el procesamiento de grandes cantidades de datos, los humanos retienen una ventaja comparativa en situaciones que requieren pensamiento holístico y visionario. Esto es particularmente relevante en los niveles más altos de las organizaciones, donde la planificación estratégica a menudo implica mayores niveles de ambigüedad e incertidumbre.

En este nuevo escenario, el trabajador humano no se vuelve obsoleto, sino que se reposiciona como un "entrenador" o "director" de los sistemas de IA. Así, las habilidades más valiosas en el mercado laboral futuro no serán necesariamente las relacionadas con tareas específicas que la IA puede realizar: serán las meta-habilidades que permiten a los humanos trabajar eficazmente con y junto a los sistemas de IA. Estas habilidades incluyen la capacidad de pensar críticamente sobre los *outputs* de la IA, la habilidad para integrar información de múltiples fuentes (tanto humanas como artificiales), y la capacidad de tomar decisiones estratégicas basadas en una comprensión holística del contexto organizacional y social.

En la actualidad la integración de la IA en el entorno laboral ha demostrado tener un impacto significativo en la productividad de los empleados en diversos sectores. Fitri *et al.* (2023) analizan los resultados del uso de IA en productividad y capacidad de aprendizaje, con resultados notables. En el ámbito del servicio al cliente, por ejemplo, los agentes que utilizaron herramientas de IA experimentaron un aumento del 13.8% en su capacidad para manejar consultas por hora. También se observó un incremento del 59% en la producción de documentos por hora observado entre los profesionales de negocios que utilizaron IA para tareas de redacción rutinarias. Los desarrolladores, por su parte, al estar respaldados por IA lograron completar un 126% más de proyectos por semana.

Independientemente de las mejoras inmediatas en la productividad, la IA también ha demostrado su capacidad para acelerar significativamente los procesos de aprendizaje y desarrollo de competencias. El estudio reveló que los agentes de servicio al cliente que utilizaron herramientas de IA alcanzaron en dos meses niveles de competencia que normalmente requerían ocho meses de formación. Esta aceleración en la curva de aprendizaje beneficia a los empleados al permitirles adquirir nuevas habilidades más

rápidamente, y representa una ventaja considerable para las organizaciones, que pueden contar con personal altamente capacitado en un tiempo mucho menor.

De esta manera, la IA permite mejorar las capacidades del individuo y libera al trabajador del trabajo automático y repetitivo gracias a la automatización. El empleado puede ocuparse de tareas en las que se encuentra más capacitado y que es especialmente útil. Para que esto suceda de manera correcta y para que el *output* humano sea tan valioso como lo será el *output* de la IA, en las tareas en las que cada uno es especialmente bueno, puede no bastar simplemente con empleados humanos motivados de manera extrínseca.

Como se ha expuesto, en trabajos que requieren creatividad, innovación o resolución de problemas, los individuos motivados de manera intrínseca rinden mejor que aquellos motivados por incentivos externos. Debido al efecto *crowding out* un aumento de la motivación extrínseca puede perjudicar a la motivación intrínseca, y con ello el rendimiento del empleado (Coll y Maceri, 2023). Los modelos económicos tradicionales suelen no contemplar este tipo de situaciones. Los modelos tampoco suelen contemplar la incursión de una tecnología del carácter transformador que tiene la IA para el trabajo en general, y para el rol del humano en particular.

Reiteramos que la falta evidencia sólida respecto de la eficacia de los incentivos monetarios y no monetarios en un correcto desempeño en un contexto en el cual la IA es utilizada en conjunto con las actividades humanas. Si este escenario se cumple, al menos parcialmente, examinar la motivación humana en este tipo de contextos será necesario para obtener modelos más eficaces.

Loureiro *et al.* (2023) avalan que, cuando se implementa adecuadamente, la IA puede mejorar la eficiencia operativa, potencialmente aumentar el bienestar de los empleados al permitirles participar en trabajo más gratificante y aumentar el significado que perciben gracias a su trabajo. Sin embargo, Gull *et al.* (2023) observan que la IA puede contribuir a inseguridad laboral, debido al riesgo de que los avances tecnológicos reemplacen al empleado humano en el mercado laboral. Estos miedos, si bien fundados, no son el único escenario posible, como hemos explicado anteriormente. En contraposición, la IA sí puede representar un riesgo para el mercado laboral humano, tanto debido a un reemplazo de la máquina por el humano, como debido a la nula necesidad de que el humano participe de la producción de bienes y servicios.

### **Inteligencia artificial como reemplazo humano**

El incentivo monetario debe existir, y debe ser percibido como justo por quien se está buscando incentivar, pero, más allá de la percepción y certeza de justicia, y de contar con las necesidades primarias satisfechas, en trabajos no automáticos, que implican algo de creatividad o de aunque sea ligero uso cognitivo, el incremento de la motivación extrínseca se traduce en una disminución de la motivación intrínseca que, a su vez, afecta negativamente el desempeño (Pink, 2011).

Ahora bien, si la tarea requiere, aunque sea un uso cognitivo rudimentario, incrementar el incentivo monetario solo tiene efecto hasta un punto en el que se satisfacen las necesidades básicas y sentimientos de justicia del agente (es decir, el agente debe percibir que lo que está recibiendo por la tarea es equitativo). Una vez superado ese umbral, persistir en el dinero como principal herramienta resulta ineficiente, y hasta se corre el riesgo de forjar efectos contrarios a los deseados. En una realidad donde los trabajos automáticos son (y serán aún más) delegados a la IA y a las nuevas tecnologías, el modelo preponderante deberá ser otro. Cuando esto suceda la motivación intrínseca cobrará aún más relevancia de la que ostenta en la actualidad. Ver una película, jugar un deporte o la gran mayoría de hobbies suelen ser acciones que se realizan no por la consecuencia que conllevan sino por el fin en sí mismo en el que son percibidos. ¿Qué sucede con el trabajo? Algunos trabajos, especialmente los que implican tareas repetitivas o que tienen poca autonomía (Pink, 2011) parecieran ser inevitablemente medios para el fin de la supervivencia. En este tipo de trabajos el incentivo monetario ha demostrado ser exitoso. Estos trabajos, sin embargo, van camino a desaparecer.

La economía concibe al trabajo principalmente bajo esta óptica: algo poco placentero que se realiza como un medio para alcanzar el tan ansiado consumo, es decir, como algo que genera una pérdida en la utilidad (Kohn, 1993). No debería sorprender que, por ejemplo, en Estados Unidos, 9 de cada 10 empleados estarían dispuestos a percibir una menor remuneración a cambio de desempeñar un trabajo más significativo (Achor *et al.*, 2018). El rol del dinero en el contexto laboral pareciera ser menos preponderante del que ha asumido la economía tradicional (Kohn, 1993; Pink, 2011).

Lo relevante es la posibilidad de que el rol del dinero para motivar un determinado comportamiento se vea relegado. La participación cada vez más activa de empleados no humanos en el mercado de trabajo puede provocar, tal como anticipamos, la implementación de algo similar a un salario básico universal.

Estos temas aparentemente utópicos deben ser tenidos en cuenta por quienes se encuentran en la vanguardia de la innovación tecnológica. Elon Musk (Isaacson, 2023) sugirió que, dados los avances tecnológicos y las posibilidades como el reemplazo de la mano de obra por la inteligencia artificial, sumado a la eventual implementación de un ingreso básico universal propiciado por el impulso económico de las nuevas tecnologías, existe la factibilidad real de que “el trabajo se convierta en algo opcional” (pág. 427, traducción propia).

Esta posibilidad tiene la posibilidad de no encontrarse en un futuro lejano, dada el desarrollo exponencial de las nuevas tecnologías (Kurzweil, 2005). Aquí entra en juego un efecto atribuido a Moore, con la ley que lleva su nombre. La ley de Moore (1998) es un fenómeno que ha moldeado profundamente el panorama tecnológico durante más de medio siglo. Su premisa fundamental, el incesante aumento exponencial en la densidad de

transistores en los microprocesadores, ha sido la fuerza impulsora detrás de la aceleración sin precedentes del progreso computacional.

A primera vista, esta progresión exponencial puede parecer abstracta, pero sus implicaciones son profundas y tangibles. Esencialmente, la ley de Moore dicta que la capacidad de procesamiento se duplica aproximadamente cada dos años, lo que significa que la tecnología avanza a un ritmo exponencial en lugar de lineal.

En el cambio radical de la actual IA los avances parecieran ir incluso a un ritmo más acelerado. En los Grandes Modelos de Lenguaje, Ho *et al.* (2024) observan que la capacidad computacional requerida para alcanzar un determinado nivel de habilidad de los modelos se reduce a la mitad en un lapso de tiempo de entre 5 a 14 meses (en contraste con los 24 meses que predice la ley de Moore). Avances que parecían inalcanzables, se vuelven factibles en cuestión de años o incluso meses (Kurzweil, 2005; Altman, 2021). La capacidad de procesar enormes cantidades de datos, entrenar modelos complejos y realizar cálculos sofisticados a velocidades vertiginosas está acelerando el ritmo de la innovación en la IA. La progresión exponencial desafía las nociones intuitivas acerca de la velocidad de los cambios que se aproximan en la civilización.

El avance del progreso tecnológico implica la urgente necesidad de hacerle frente a estas omisiones del sistema actual de incentivos. Al fin y al cabo, el futuro al que hacemos referencia no es lejano.

En la misma dirección Sam Altman (2021), CEO de OpenAI, la empresa detrás de ChatGPT, anticipó, más de un año antes del lanzamiento público de su exitoso producto, la llegada de un cambio socioeconómico de magnitudes difíciles de cuantificar gracias a las nuevas tecnologías en las cuales se está jugando un papel importante a desarrollar. Asimismo, remarcó que “si las políticas públicas no se adaptan en consecuencia, la mayoría de las personas terminarán en peor situación que hoy” (Altman, 2021, p. 1, traducción propia). Altman auguró que la revolución tecnológica es imposible de detener, advirtiendo que, si bien este progreso tecnológico traerá mayores niveles de riqueza, esto no provoca necesariamente una mejora en el bienestar general. Para Altman, el futuro cambiará de manera abrupta, producto del progreso tecnológico exponencial:

El mundo cambiará tan rápida y drásticamente que será necesario un cambio igualmente drástico de política para distribuir esta riqueza y permitir que más personas puedan llevar la vida que desean [...] Debido a que estamos al comienzo de este cambio tectónico, tenemos una rara oportunidad de girar hacia el futuro. Ese giro no puede simplemente abordar los problemas sociales y políticos actuales; debe diseñarse para la sociedad radicalmente diferente del futuro próximo. Los planes políticos que no tengan en cuenta esta transformación inminente fracasarán por la misma razón por la que los principios organizativos de las

sociedades preagrarias o feudales fracasarían hoy (Altman, 2021, pág. 1, traducción propia).

Este cambio también debe producirse en economía. En un escenario donde el trabajo automático y fácilmente replicable dejará de estar en manos del humano, el rol del trabajador tendrá que transformarse. La pérdida de este tipo de trabajos implica un cambio fundamental en el escenario laboral actual, teniendo en cuenta la desaparición del tipo de actividades para el cual está diseñado el sistema de incentivos del sistema económico tal como lo conocemos.

No se debe tomar a la ligera la posibilidad de un escenario mundial de estas características. Al menos, empresas como Google o la misma OpenAI, quienes se encuentran en la frontera de este progreso tecnológico, no lo hacen (Cerón Uribe, 2023; Zhang *et al.*, 2023). La transformación será en el entorno laboral y educativo, pues el avance de la IA tiene el potencial de transformar por completo cómo se llevan a cabo las tareas de aprendizaje y enseñanza (Latif *et al.*, 2023).

Las tareas repetitivas no serán las únicas eliminadas del mercado laboral debido al progreso de la IA. Existen pronósticos que demuestran que estas nuevas tecnologías tienen el potencial de reemplazar a los empleos con salarios más altos (Zhang *et al.*, 2023). Así, las actividades no automáticas, pero que eran llevadas a cabo gracias al dinero que se percibía por ellas, verán este componente motivacional gravemente afectado.

Si un escenario en el cual la motivación extrínseca tiene un rol menos preponderante debido a que las necesidades básicas han sido cubiertas, la evidencia mencionada en las primeras secciones de este trabajo puede ser de gran utilidad. Si los trabajos automatizables, en efecto, se automatizan, la eficacia del esquema de incentivos actual se verá disminuida.

Es plausible suponer que esta transformación, fruto de la innovación tecnológica, expondrá con mayor nitidez las omisiones del sistema de incentivos. Aspirar a un futuro de tal naturaleza implica también reflexionar sobre las consecuencias que éste acarreará. Si el ingreso básico universal u otras políticas análogas llegan a convertirse en realidad, la necesidad de abordar seriamente los temas expuestos en el presente trabajo será una conveniencia extraordinaria. Es más, constituirá una necesidad ineludible para proseguir con el progreso de la civilización.

## **Conclusión**

El presente escrito ha explorado la intersección entre la IA y las teorías de motivación humana, con un enfoque particular en las implicaciones para el futuro del trabajo y los sistemas de incentivos económicos. A lo largo del análisis realizado, hemos arribado a varias conclusiones:

-La IA representa un cambio radical en la historia de la innovación tecnológica, distinguiéndose de las revoluciones industriales anteriores por su velocidad de adopción, amplitud de impacto y capacidad para realizar tareas cognitivas complejas.

-Las teorías económicas tradicionales sobre la motivación humana, basadas principalmente en incentivos extrínsecos, han demostrado ser insuficientes y, en ocasiones, contraproducentes, especialmente en tareas que requieren creatividad y pensamiento complejo. La evidencia empírica respalda la importancia de la motivación intrínseca, tal como se describe en la teoría de la autodeterminación, la cual enfatiza la satisfacción de las necesidades psicológicas de autonomía, competencia y relaciones sociales.

-Las consecuencias de no incluir a la psicología a la hora de realizar supuestos sobre la motivación humana se amplifican, producto del efecto *crowding out*. Los incentivos extrínsecos pueden disminuir la motivación intrínseca, planteando desafíos significativos para el diseño de sistemas de incentivos en un futuro laboral cada vez más automatizado.

-La IA tiene el potencial tanto de complementar como de reemplazar ciertas funciones humanas en el lugar de trabajo, lo que podría llevar a dos escenarios principales: una simbiosis entre el trabajo humano y no humano, o una transformación radical que podría hacer que el trabajo humano sea opcional en muchos sectores. Ambos escenarios requerirían un nuevo sistema de incentivos.

-Es necesario, entonces, reevaluar y adaptar nuestros modelos económicos y sistemas de incentivos para reflejar una comprensión más matizada de la motivación humana en relación con la IA.

Para finalizar, cabe destacar que la investigación aquí expuesta es un trabajo en construcción, y a futuro buscamos contribuir a la falta de evidencia señalada, analizando la variación de los incentivos en un contexto de plena adopción de la IA. De todas maneras, la evidencia actual sugiere que es crucial que los responsables políticos, los líderes empresariales y los investigadores trabajen en conjunto para diseñar políticas y prácticas que fomenten la motivación intrínseca y el bienestar humano en un mundo cada vez más automatizado. Esto podría implicar la exploración de conceptos como el ingreso básico universal, la reestructuración de los sistemas educativos para enfatizar las habilidades que complementan la IA, y el rediseño de los lugares de trabajo para maximizar la colaboración humano-IA.

A medida que la humanidad avanza hacia este futuro incierto pero potencialmente transformador, será crucial adaptar nuestras teorías económicas y prácticas organizacionales para reflejar una comprensión más completa de la motivación humana, aprovechar plenamente el potencial de la IA, mejorar la productividad, la satisfacción laboral y la calidad de vida

## Referencias bibliográficas

Achor, S. (2010). *The happiness advantage: How a positive brain fuels success in work and life*. Crown Currency.

Achor, S., Reece, A., Kellerman, G. R., & Robichaux, A. (2018). 9 out of 10 people are willing to earn less money to do more-meaningful work. *Harvard Business Review*, 96(6), 82-89. <https://hbr.org/2018/11/9-out-of-10-people-are-willing-to-earn-less-money-to-do-more-meaningful-work>

Acemoglu, D., & Johnson, S. (2024). *Learning from Ricardo and Thompson: Machinery and labor in the early industrial revolution, and in the age of AI*. National Bureau of Economic Research. <https://www.nber.org/papers/w32416>

Agrawal, K. (2010). *To study the phenomenon of the Moravec's paradox*. <https://doi.org/10.48550/arXiv.1012.3148>

Altman, S. (2021). *Moore's law for everything*.

<http://tcw.org/left/Short%20Stories/Moore's%20Law%20for%20Everything.pdf>

Ariely, D., Bracha, A., & Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American economic review*, 99(1), 544-555. <https://doi.org/10.1257/aer.99.1.544>

Ariely, D., Gneezy, U., Loewenstein, G., & Mazar, N. (2009). Large stakes and big mistakes. *The Review of Economic Studies*, 76(2), 451-469.

Bardsley, N., Cubitt, R., Loomes, G., Moffatt, P., Starmer, C., & Sugden, R. (2010). Incentives in Experiments. In *Experimental Economics: Rethinking the Rules* (pp. 244–285). Princeton University Press. <http://www.jstor.org/stable/j.ctt7sgpt.9>

Bellemare-Pepin, A., Lespinasse, F., Thölke, P., Harel, Y., Mathewson, K., Olson, J. A., Bengio, Y., & Jerbi, K. (2024). Divergent Creativity in Humans and Large Language Models. <https://doi.org/10.48550/arXiv.2405.13012>

Boussioux, L., Lane, J. N., Zhang, M., Jacimovic, V., & Lakhani, K. R. (2024). *Generative AI and Creative Problem Solving*. Harvard Business School Technology & Operations Mgt. Unit Working Paper No. 24-005. <http://dx.doi.org/10.2139/ssrn.4533642>

Bruni, L., & Sugden, R. (2007). The road not taken: How psychology was removed from economics, and how it might be brought back. *The Economic Journal*, 117(516), 146-173. <https://doi.org/10.1111/j.1468-0297.2007.02005.x>

Cerón Uribe, J. F. (2023). Entrevista a Juan Felipe Cerón Uribe, desarrollador de ChatGPT en Open AI, por Fabrizio López de Pomar. *Futuro Hoy*, 4(1), 11-16. <https://futurohoy.ssh.org.pe/>



Charles-Leija, H., Aboites, G., y Llamas, I. (2018). Una revisión de aportaciones que contribuyeron al estudio de la utilidad y la felicidad en la economía. *Análisis económico*, 33(84), 57-76. [https://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S2448-66552018000300057](https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2448-66552018000300057)

Clark, J. M. (1918). Economics and modern psychology: I. *Journal of Political Economy*, 26(1), 1-30. <https://www.journals.uchicago.edu/doi/10.1086/253060>

Cervellati, E. M. (2018). Behavioral Re-Evolution: How Behavioral Economics has Evolved and is Evolving. In *The Behavioural Finance Revolution: A New Approach to Financial Policies and Regulations*. Edward Elgar Publishing Limited, Cheltenham (UK).

Coll, A. J. y Maceri, S. B. (2023). Las consecuencias indeseadas del dinero desde el punto de vista de la psicología y la Economía de la Felicidad. *Cultura Económica*, 41(105), 67-84. <https://doi.org/10.46553/cecon.41.105.2023.p67-84>Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and Social Psychology*, 18(1), 105-115 <http://dx.doi.org/10.1037/h0030644>

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of research in personality*, 19(2), 109-134. [https://doi.org/10.1016/0092-6566\(85\)90023-6](https://doi.org/10.1016/0092-6566(85)90023-6)

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627-668. <https://doi.org/10.1037/0033-2909.125.6.627>

Easterlin, R. A. (2021). *An Economist's Lessons on Happiness: Farewell Dismal Science!* Springer Nature.

Epstein, D. (2021). *Range: Why generalists triumph in a specialized world*. Penguin.

Fitri, D., Ratnasari, S. L., Suyanto, & Sultan, Z. (2023). *Enhancing employee productivity through technology system AI-based approaches*. The 6th International Seminar on Business, Economics, Social Science, and Technology (ISBEST) 2023, 3, 77-82.

Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of economic surveys*, 15(5), 589-611. <https://doi.org/10.1111/1467-6419.00150>

Friedman, M. (1953). *The methodology of positive economics*. University of Chicago Press.

Girotra, K., Meincke, L., Terwiesch, C., & Ulrich, K. T. (2023). *Ideas are dimes a dozen: Large language models for idea generation in innovation*. <https://dx.doi.org/10.2139/ssrn.4526071>

Gneezy, U., & Rustichini, A. (2000a). Pay Enough or Don't Pay at All. *Quarterly Journal of Economics*, 115, 791-810. <https://doi.org/10.1162/003355300554917>

----. (2000b). A fine is a price. *The journal of legal studies*, 29(1), 1-17. <https://www.journals.uchicago.edu/doi/10.1086/468061>

Gull, A., Ashfaq, J., & Aslam, M. (2023). AI in the Workplace: Uncovering Its Impact on Employee Well-being and the Role of Cognitive Job Insecurity. *International Journal of Business & Economic Affairs (IJBEA)*, 8(4). <https://doi.org/10.24088/IJBEA-2023-84007>

Harlow, H. F. (1950). Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys. *Journal of comparative and physiological psychology*, 43(4), 289. <https://psycnet.apa.org/doi/10.1037/h0058114>

Harris-McLeod, E. (2013). Incentives for public service workers and the implications of crowding out theory. *Public Policy and Governance Review*, 4(2), 5-21. <https://ppgreview.ca/wp-content/uploads/2013/04/ppgr-volume-4-issue-2-2-emily-harris-mcleod.pdf>

Ho, A., Besiroglu, T., Erdil, E., Owen, D., Rahman, R., Guo, Z. C., Atkinson, D., Thompson, N., & Sevilla, J. (2024). Algorithmic progress in language models. <https://doi.org/10.48550/arXiv.2403.05812>

Isaacson, W. (2023). *Elon Musk*. Penguin.

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577-586. <https://doi.org/10.1016/j.bushor.2018.03.007>

Kamenica, E. (2012). Behavioral economics and psychology of incentives. *ANNUAL REVIEW OF ECONOMICS*, 4(1), 427-452. <https://doi.org/10.1146/annurev-economics-080511-110909>

Kohn, A. (1993). Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes. Houghton Mifflin Company.

Kurzweil, R. (2005). *The singularity is near*. In *Ethics and emerging technologies* (pp. 393-406). Palgrave Macmillan UK.

Latif, E., Mai, G., Nyaaba, M., Wu, X., Liu, N., Lu, G., ... & Zhai, X. (2023). Artificial general intelligence (AGI) for education. arXiv preprint arXiv:2304.12479.

Lepper, M. R., & Greene, D. (1978). The hidden costs of reward: New perspectives on the psychology of human motivation. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and Social Psychology*, 28(1), 129. <https://doi.org/10.1037/h0035519>

List, J. A., Livingston, J. A., & Neckermann, S. (2018). Do financial incentives crowd out intrinsic motivation to perform on standardized tests? *Economics of Education Review*, 66, 125-136. <https://doi.org/10.1016/j.econedurev.2018.08.002>

Loureiro, S. M. C., Bilro, R. G., & Neto, D. (2023). Working with AI: can stress bring happiness? *Service Business*, 17(1), 233-255. <https://doi.org/10.1007/s11628-022-00514-8>

- Mollick, E. (2024). *Co-Intelligence: Living and Working with AI*. Penguin Publishing Group.
- Moore, G. E. (1998). Cramming more components onto integrated circuits. *Proceedings of the IEEE*, 86(1), 82-85. <https://doi.org/10.1109/JPROC.1998.658762>
- Pink, D. H. (2011). *Drive: The surprising truth about what motivates us*. Penguin.
- Ryan R. (2009). Self-Determination Theory and Wellbeing. *Social Psychology*, 84, 822-848.
- Ryan, R. M., Bradshaw, E., Deci, E. L., Sternberg, R., & Pickren, W. (2019). *A history of human motivation theories*. The Cambridge handbook of the intellectual history of psychology, 391-411.
- Ryan, R. M., Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well being. *American Psychologist*, 55, 68-78.
- Stigler, G. J., & Becker, G. S. (1977). De gustibus non est disputandum. *The American economic review*, 67(2), 76-90.
- Titmuss, R. M. (1970). *The gift relationship* (vol. 220). Allen & Unwin.
- Visaria, S., Dehejia, R., Chao, M. M., & Mukhopadhyay, A. (2016). Unintended consequences of rewards for student attendance: Results from a field experiment in Indian classrooms. *Economics of Education Review*, 54, 173-184. <https://doi.org/10.1016/j.econedurev.2016.08.001>
- Wu, J., & Lu, X. (2013). Effects of Extrinsic and Intrinsic Motivators on Using Utilitarian, Hedonic, and Dual-Purposed Information Systems: A Meta-Analysis. *Journal of the Association for Information Systems*, 14(3). <https://aisel.aisnet.org/jais/vol14/iss3/1>
- Zarifhonarvar, A. (2023). *Economics of ChatGPT: A Labor Market View on the Occupational Impact of Artificial Intelligence*. ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg. <https://hdl.handle.net/10419/268826>
- Zhang, C., Zhang, C., Li, C., Qiao, Y., Zheng, S., Dam, S. K., & Hong, C. S. (2023). *One small step for generative ai, one giant leap for agi: A complete survey on chatgpt in aigc era*. arXiv preprint arXiv:2304.06488. <https://doi.org/10.48550/arXiv.2304.06488>

## Sobre los Documentos de Trabajo

La serie de Documentos de Trabajo del IIEP refleja los avances de las investigaciones realizadas en el instituto. Los documentos pasan por un proceso de evaluación interna y son corregidos, editados y diseñados por personal profesional del IIEP. Además de presentarse y difundirse a través de la página web del instituto, los documentos también se encuentran disponibles en la biblioteca digital de la Facultad de Ciencias Económicas de la Universidad de Buenos Aires, el repositorio digital institucional de la Universidad de Buenos Aires, el repositorio digital del CONICET y en la base IDEAS RePEc.



## INSTITUTO INTERDISCIPLINARIO DE ECONOMÍA POLÍTICA

I I E P

Universidad de Buenos Aires | Facultad de Ciencias Económicas

Av. Córdoba 2122 1º y 2º piso (C1120 AAQ)  
Ciudad Autónoma de Buenos Aires, Argentina  
+54 11 5285-6578 | [www.iiep.economicas.uba.ar](http://www.iiep.economicas.uba.ar)

  @IIEP\_OFICIAL