

Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), enero-febrero 2026,
Volumen 10, Número 1.

https://doi.org/10.37811/cl_rcm.v10i1

BUSCANDO LA CAUSALIDAD EN EL TRASTORNO DE ANSIEDAD

SEARCHING FOR CAUSALITY IN ANXIETY DISORDERS

Silday Natividad Veleces Soriano
Tecnológico Nacional de México, México

José Antonio Montero Valverde
Tecnológico Nacional de México, México

Miriam Martínez Arroyo
Tecnológico Nacional de México, México

Juan Miguel Hernández Bravo
Tecnológico Nacional de México, México

Eduardo de la Cruz Gamez
Tecnológico Nacional de México, México

Buscando la Causalidad en el Trastorno de Ansiedad

Silday Natividad Veleces Soriano¹

mm24320018@acapulco.tenm.mx

<https://orcid.org/0009-0008-2113-8881>

Tecnológico Nacional de México Campus
Acapulco
México

José Antonio Montero Valverde

jose.mv@acapulco.tecnm.mx

<https://orcid.org/0009-0000-5357-3257>

Tecnológico Nacional de México Campus
Acapulco
México

Miriam Martínez Arroyo

miriam.ma@acapulco.tecnm.mx

<https://orcid.org/0000-0002-5685-1731>

Tecnológico Nacional de México Campus
Acapulco
México

Juan Miguel Hernández Bravo

juan.hb@acapulco.tecnm.mx

<https://orcid.org/0000-0003-4892-3368>

Tecnológico Nacional de México Campus
Acapulco
México

Eduardo de la Cruz Gamez

jose.mv@acapulco.tecnm.mx

<https://orcid.org/0000-0003-3318-788X>

Tecnológico Nacional de México Campus
Acapulco
México

RESUMEN

De acuerdo con datos de la Organización Mundial de la Salud (2025), el trastorno de ansiedad es una enfermedad mental que afecta a más de mil millones de personas a nivel mundial. Por su parte, la Secretaría de Salud de México (SSA, 2024), reportó que este trastorno mental representó el 52.8% de los casos de salud mental atendidos en el año 2024. En este trabajo, se busca determinar una relación causal entre las variables que causan la manifestación de este trastorno. Para esta finalidad, se utilizaron algoritmos básicos de descubrimiento causal, específicamente: i) PC (Peter Spirtes y Clark Glymour), ii) GES (Algoritmo de Búsqueda Voraz de Equivalencias), y iii) LiNGAM (Modelo Acíclico No Gaussiano Lineal). Asimismo, se utilizó el conjunto de datos manejado por Fathi & Ahmadi (2020), el cual usa información de variables demográficas, emocionales y físicas, considerando un total de 28 variables. Los hallazgos encontrados al utilizar los algoritmos mencionados, determinan coincidencias de causalidad en algunas variables. La similitud entre las coincidencias, así como la orientación y estimación de la relación causal entre las variables en los grafos, fue valorada por diferentes métricas. Las coincidencias encontradas determinan las variables con mayor probabilidad de ser factores causales en el trastorno de ansiedad. De ser comprobables, a través de un experto humano, estas relaciones, son de gran importancia para intervenciones específicas a diferentes trastornos.

Palabras clave: descubrimiento causal, PC, GES, LiNGAM, trastorno de ansiedad

¹ Autor principal

Correspondencia: mm24320018@acapulco.tenm.mx

Searching for Causality in Anxiety Disorders

ABSTRACT

According to data from the World Health Organization (2025), anxiety disorder is a mental illness that affects more than one billion people worldwide. For its part, the Mexican Ministry of Health (SSA, 2024) reported that this mental disorder represented 52.8% of the mental health cases treated in the year 2024. In this work, it is sought to determine a causal relationship among the variables that cause the manifestation of this disorder. For this purpose, basic causal discovery algorithms were used, specifically: i) PC (Peter Spirtes and Clark Glymour), ii) GES (Greedy Equivalence Search Algorithm), and iii) LiNGAM (Linear Non-Gaussian Acyclic Model). Likewise, the dataset handled by Fathi & Ahmadi (2020) was used, which uses information from demographic, emotional, and physical variables, considering a total of 28 variables. The findings obtained by using the aforementioned algorithms determine causal coincidences in some variables. The similarity among the coincidences, as well as the orientation and estimation of the causal relationship among the variables in the graphs, was evaluated by different metrics. The coincidences found determine the variables with the highest probability of being causal factors in anxiety disorder. If they can be verified through a human expert, these relationships are of great importance for specific interventions for different disorders.

Keywords: causal discovery, PC, GES, LiNGAM, anxiety disorder

*Artículo recibido 12 diciembre 2025
Aceptado para publicación: 16 enero 2026*



INTRODUCCIÓN

El trastorno de ansiedad se presenta cuando una persona experimenta miedo o preocupación de manera intensa. Estos sentimientos van acompañados de síntomas físicos, conductuales y cognitivos. Este trastorno es el resultado de una compleja interacción entre factores sociales, psicológicos y biológicos (OMS, 2025).

El impacto de la ansiedad en la salud física se ha estudiado ampliamente y se ha demostrado que puede influir negativamente en varios sistemas del cuerpo. Desde el aumento de la frecuencia cardíaca hasta el deterioro de la función inmunológica. Los efectos de la ansiedad pueden contribuir a enfermedades graves a largo plazo.

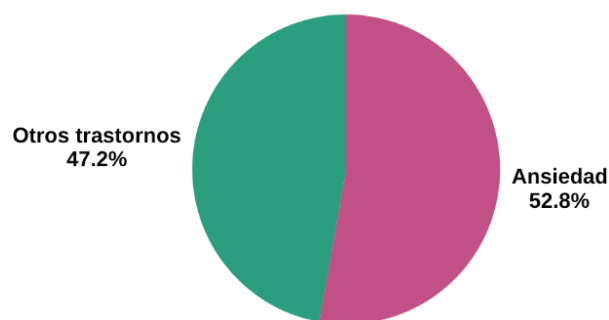
Por otro lado, los efectos resultantes de la ansiedad también se presentan en conductas de evitación que se pueden agravar al punto de llegar al aislamiento social impactando en la vida diaria, académica, laboral o personal de manera negativa.

Según la Organización Mundial de la Salud (2025), los trastornos de salud mental prevalecen en más de mil millones de personas a nivel mundial y condiciones como la ansiedad y la depresión no solo son sus principales formas de manifestación, sino que están causando un gran gasto de recursos humanos y económicos. Estos trastornos afectan a personas de todas edades, desde niños hasta adultos mayores y representan la segunda causa principal de discapacidad a largo plazo. Aún más grave es el hecho de que el suicidio permanece como un resultado grave por la falta de atención oportuna a estos trastornos.

Si bien afecta a todos, las personas que se encuentran en el rango de edades correspondientes a la adolescencia hasta la adultez, son un grupo muy vulnerable debido a la presión académica o laboral.

En México, se estima un alto porcentaje de casos atendidos, siendo el trastorno de ansiedad el más frecuente, por encima del 52% según datos de la Secretaría de Salud de México (2024) (Figura 1).

Figura 1. Prevalencia del trastorno de ansiedad en México (Fuente: Secretaría de Salud de México (2024)).



Estudios realizados en varios países demostraron que el trastorno de ansiedad como consecuencia de la pandemia del COVID-19 ha aumentado notablemente. Del Moral Chica et al. (2024) señala que el aislamiento social, la incertidumbre económica y el temor al virus han llevado a un aumento significativo de los casos de ansiedad y depresión.

Este trabajo busca determinar la estructura causal de las variables asociadas al trastorno de ansiedad mediante la aplicación de algoritmos de descubrimiento causal.

Relevancia clínica y social. Se pueden diseñar intervenciones efectivas para prevenir el trastorno de ansiedad, identificar factores de riesgo modificables con mayor precisión y optimizar tratamientos personalizados basados en perfiles de riesgo individual. Estas intervenciones preventivas pueden reducir los costos relacionados a los tratamientos. Además, el desarrollo e implementación de algoritmos de IA causal en el campo médico respecto al trastorno de ansiedad podría abrir nuevas áreas de investigación en ciencia de datos y salud.

En este trabajo utilizamos el trastorno de ansiedad social (TAS) como caso representativo dentro del grupo de los trastornos de ansiedad. El TAS comparte con otros trastornos de ansiedad síntomas generales importantes de miedo, evitación y malestar, así como factores de riesgo comunes, como la historia familiar de trastornos emocionales, entre otras. Por esta razón, en el marco teórico hacemos referencia al trastorno de ansiedad de forma general, pero las evidencias empíricas que presentamos en este trabajo provienen de un conjunto de datos específico de personas con TAS. Así pues, cuando nos referimos a trastorno de ansiedad en este texto, hablamos principalmente del TAS evaluado en este estudio.

METODOLOGÍA

Este trabajo divide la metodología en las siguientes etapas. Primero se menciona todo lo relacionado a los datos de entrada, después se aborda el tema de determinación de causas y, finalmente, la evaluación del modelo.

Datos de entrada. El conjunto de datos utilizado en este trabajo está integrado por diferentes tipos de variables, entre las cuales destacan las siguientes: 1) variables sociodemográficas 2) variables conductuales, 3) variables físicas y 4) miedos específicos relacionados con la ansiedad social, siguiendo la estructura propuesta por Fathi & Ahmadi (2020).

La población que se estudia en este trabajo es la que pertenece al rango de edades entre los 14 a los 65 años. Como se sabe, los cambios físicos y psicológicos asociados a la adolescencia y posteriormente a la adultez, sumados con las responsabilidades escolares, sociales y laborales, hace que este grupo sea de especial importancia estudiarlo.

Aunque conceptualmente este trabajo se encuentra en el marco amplio de los trastornos de ansiedad, el conjunto de datos utilizado corresponde a personas evaluadas por trastorno de ansiedad social. Por tanto, las relaciones causales identificadas describen principalmente al subtipo específico y cualquier extensión a otros trastornos de ansiedad debe hacerse considerando esta información.

El conjunto de datos original fue utilizado en una tarea de aprendizaje automático; cuenta con 214 instancias y 30 columnas de datos, sin embargo, algunas de ellas no son necesarias para la presente investigación. Por lo mismo, se eliminó la columna 'id' por ser irrelevante y 'hasSAD' por ser la variable de clase (objetivo), innecesaria para el descubrimiento causal.

Dentro del conjunto de datos, se encuentran dos variables importantes basadas en test de ansiedad. Variable “LSAS” la cual representa el valor obtenido en la Escala de Ansiedad Social de Liebowitz. Esta escala es una herramienta muy utilizada que consta de 24 preguntas. Cada una de ellas consiste en una situación a considerar y que debe ser puntuada tanto con relación a la ansiedad o miedo que supone y la frecuencia con que se evita (lo que proporciona una puntuación entre 0 y 3 para cada cuestión). La puntuación máxima posible es por tanto de 144 y la mínima de 0 (Liebowitz, 1987).

La variable “SPIN” representa el resultado de la aplicación del cuestionario Inventario de Fobia Social diseñado para evaluar la ansiedad social en población clínica y general. Fue desarrollado por Connor et

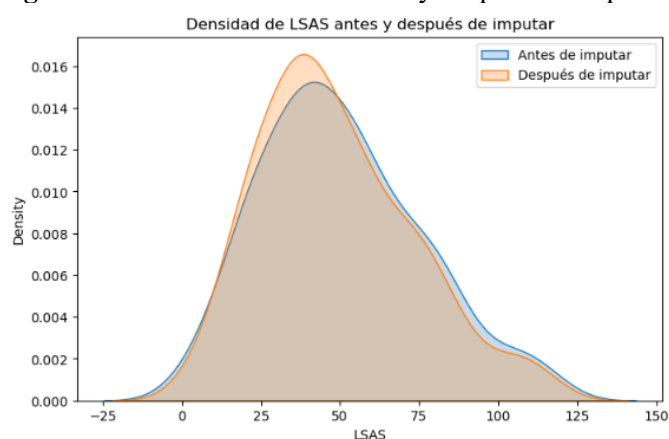
al. (2000) y es una de las escalas que más se utilizan en investigación psicológica debido a que es breve, confiable y tiene muy buena capacidad de detección. Consta de 17 ítems y cada uno consta de una escala de 0 a 4. Mide 3 dimensiones principales: miedo, evitación y manifestaciones fisiológicas (Connor et al., 2000).

Inicialmente, la verificación y limpieza de datos fue una etapa crítica para garantizar la correcta aplicación de los algoritmos.

Preprocesamiento: Aunque la eliminación de instancias con valores faltantes suele ser una estrategia común, se optó por implementar el algoritmo MICE (Multiple Imputation by Chained Equations). Esta decisión se fundamenta en las limitaciones de métodos tradicionales (como la imputación por media o mediana), los cuales reducen la varianza al reemplazar los valores ausentes por una constante fija. Dicha reducción artificial distorsiona las correlaciones entre variables y puede alterar las dependencias estadísticas esenciales para el funcionamiento de los algoritmos causales.

Posteriormente, se validó la imputación para asegurar que la distribución original de los datos no sufriera alteraciones significativas. Para ello, se analizó el histograma de la variable LSAS (la cual presentaba los datos faltantes). Como se aprecia en la Figura 2, no hubo cambios considerables que comprometieran la integridad del estudio.

Figura 1. Densidad de LSAS antes y después de imputar.



Fuente: Elaboración propia.

Las técnicas utilizadas en este trabajo son algoritmos que han sido probados en diferentes estudios causales, y sus resultados generalmente han sido satisfactorios. Como en el trabajo de Singh, V. K., &

Sharma, M. (2025) donde se aplican técnicas de inferencia causal como PC, GES y LiNGAM al análisis de datos de colocación de estudiantes para identificar relaciones de causa y efecto entre factores académicos y no académicos, y los resultados de empleabilidad.

Mientras que en el estudio de Cotta, G. A., Runge, J., & Drouard, M. (2023) se aborda la identificación de los impulsores causales más robustos para modelos de machine learning en ciencias ambientales donde emplea los algoritmos de descubrimiento causal PC o PCMCI (una extensión de PC para series de tiempo) para inferir partes del grafo causal demostrando cómo el descubrimiento causal mejora la generalización y la robustez de los modelos predictivos en el dominio climático.

Determinación de causas. El descubrimiento causal, también conocido como aprendizaje de estructuras causales, detona un amplio conjunto de métodos destinados a recuperar información sobre mecanismos causales a partir de datos. Existen 4 familias principales de métodos de descubrimiento causal: Basados en restricciones, basados puntuaciones, funcionales, basados en gradientes y otros que incluye métodos que no encajan en las cuatro categorías principales.

A continuación, se ofrece una breve descripción de las técnicas utilizadas:

El algoritmo PC. Un ejemplo clásico de algoritmo basado en restricciones es el algoritmo PC (Sprites & Glymour, 1991). Su nombre proviene de los nombres de sus creadores: Peter Sprites y Clark Glymour. El algoritmo PC es una variante del algoritmo IC propuesto anteriormente por Verma & Pearl (1990). Se utiliza el test gsq (G^2 Likelihood Ratio Chi-squared) ya que no requiere que los datos sean normales y se comprobó que los datos no siguen una distribución normal aplicando el test Shapiro-Wilk.

El algoritmo GES. Los métodos basados en puntuaciones funcionan generando iterativamente grafos candidatos, evaluando lo bien que cada uno explica los datos y seleccionando el mejor. Un ejemplo bien conocido de método basado en puntuaciones es la Búsqueda de Equivalencia Codiciosa (GES, por sus siglas en inglés) introducida por David Maxwell Chickering.

El algoritmo es un procedimiento de dos etapas. Primero, genera los bordes y, a continuación, poda el gráfico. La primera etapa del GES comienza con un gráfico desconectado. A continuación, el algoritmo añade bordes de forma iterativa, calculando la puntuación en cada paso. Esto continúa hasta que la puntuación ya no puede aumentarse. En la segunda etapa, el algoritmo comienza a podar los bordes existentes para ver si la puntuación puede mejorarse aún más. Todos estos cálculos se realizan de forma

codiciosa (de ahí su nombre). Al igual que el algoritmo PC, el GES es sensible a la confusión oculta. También es posible que no pueda orientar todos los bordes, lo que le proporcionaría una clase de equivalencia de Markov de posibles grafos (de ahí su nombre, una vez más) (Chickering, 2003).

El algoritmo LiNGAM. Fue introducido por primera vez por Shohei Shimizu y sus colegas en 2006. El método original utiliza el análisis de componentes independientes (ICA) para recuperar información sobre el proceso de generación de datos. Su variante posterior, DirectLiNGAM (Shimizu et al., 2011) aprovecha un modelo lineal y una medida de independencia basada en el núcleo. Las dos hipótesis principales en las que se basa LiNGAM son:

- No hay factores de confusión ocultos.
- Todos los términos de error (o todos menos uno) son no gaussianos.

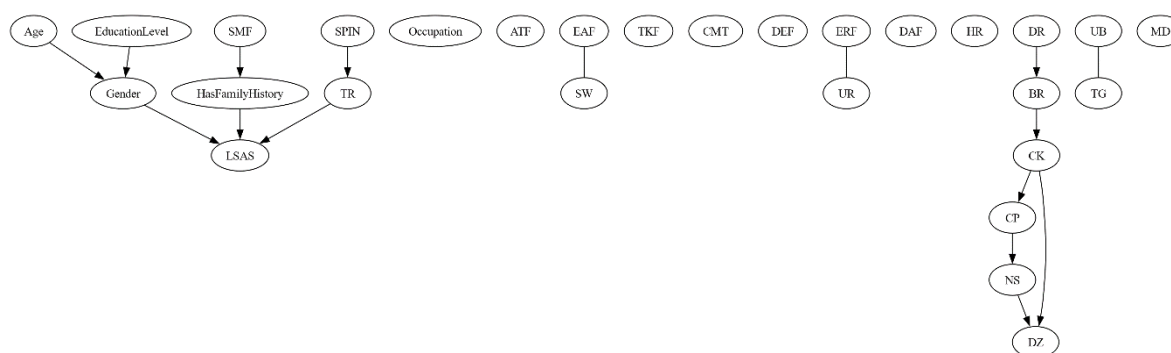
Dicho esto, se han propuesto varias extensiones de LiNGAM que permiten aplicar el modelo a escenarios con factores de confusión ocultos (Hoyer et al., 2008) o ciclos (Lacerda et al., 2008).

Evaluación de modelos. De acuerdo con Tsamardinos et al. (s. f.) la distancia estructural de Hamming hace una comparación directamente entre la estructura de las redes aprendidas y las originales, y su uso está más orientado al descubrimiento que a la inferencia. Se define entre dos PDAG como el número de los siguientes operadores necesarios para hacer que los PDAG coincidan: añadir o eliminar una arista no dirigida, y añadir, eliminar o invertir la orientación de una arista. Así pues, un algoritmo se puntúa añadiendo un punto por aprender un PDAG con una arista no orientada adicional y un punto por no orientar una arista que debería haber sido orientada.

RESULTADOS Y DISCUSIÓN

Resultado del algoritmo PC. Como resultado de aplicar el algoritmo PC en el conjunto de datos, se obtuvo la estimación de variables mostrada en el siguiente grafo acíclico dirigido (Figura 4).

Figura 2. Grafo causal generado por PC.



Fuente: Elaboración propia.

La estructura aprendida por el algoritmo, representada en la figura 4, revela múltiples relaciones de dependencia condicional e independencia entre las variables del estudio.

Las variables Age (Edad) y EducationLevel (Nivel educativo) actúan como variables exógenas que influyen sobre Gender (Género). A su vez, el género se muestra como un padre directo de LSAS, lo que implica una influencia directa de las características demográficas sobre el puntaje de ansiedad social.

La variable SMF muestra una relación causal directa hacia HasFamilyHistory (Historial Familiar), la cual a su vez incide directamente en LSAS. Esto sugiere una ruta hereditaria o de antecedentes familiares clara.

Se observa una cadena causal donde SPIN influye sobre TR, y esta última actúa como mediadora directa hacia LSAS.

Se encontró una estructura causal compleja. La cadena sigue la ruta $DR \rightarrow BR \rightarrow CK$. A partir de la variable CK, se observa una bifurcación que converge finalmente en la variable DZ.

Una influencia directa: $CK \rightarrow DZ$.

Una influencia mediada: $CK \rightarrow CP \rightarrow NS \rightarrow DZ$.

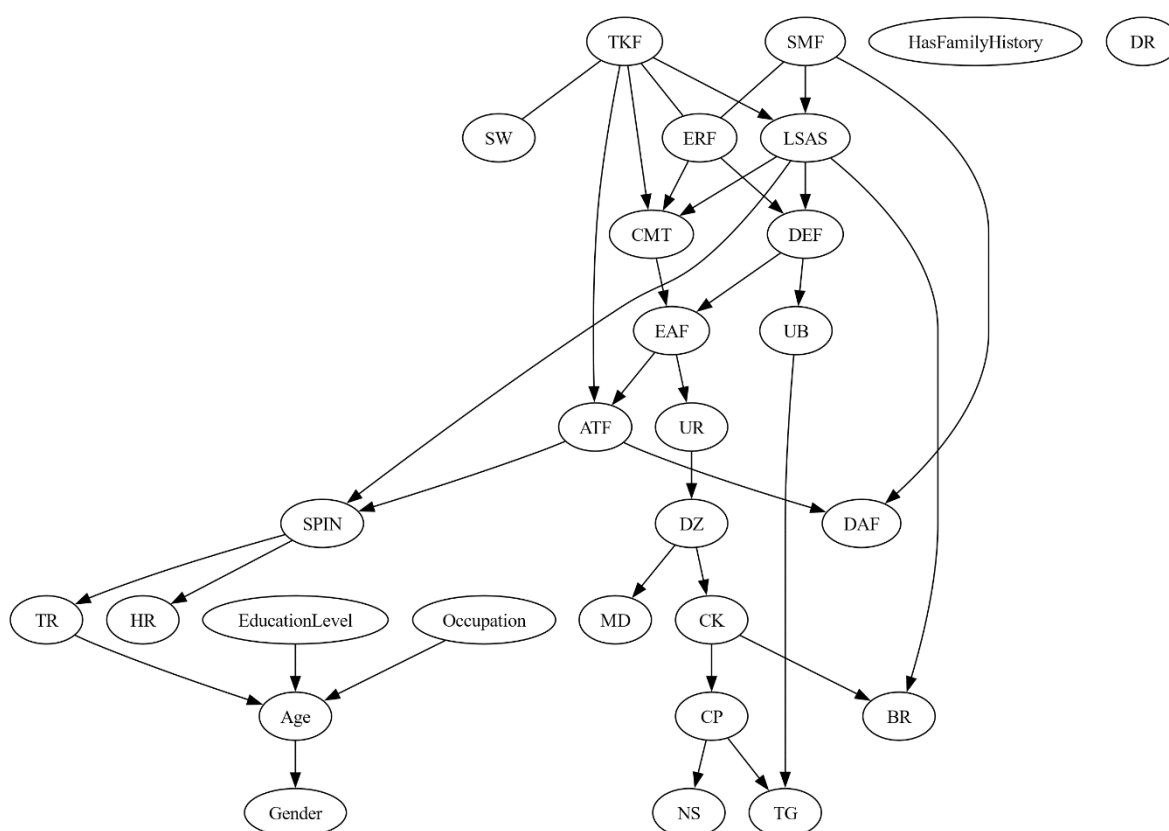
Esta estructura sugiere un mecanismo complejo de regulación entre estas variables, donde CK juega un rol central como distribuidor de efectos.

El algoritmo detectó relaciones simples de causa-efecto aisladas del resto del sistema, tales como $EAF \rightarrow SW$, $ERF \rightarrow UR$ y $UB \rightarrow TG$.

Por otro lado, variables como Occupation (Ocupación), ATF, TKF, CMT, DEF, DAF y MD aparecen como nodos desconectados. Esto indica que, bajo los criterios de independencia condicional del algoritmo aplicado, estas variables no mostraron una influencia causal significativa sobre el resto de las variables en este conjunto de datos específico.

Resultado del algoritmo GES. Respecto a los resultados del algoritmo GES (Greedy Equivalence Search). En este caso, los datos se transformaron para que sean adecuados para aplicar este método causal.

Figura 3. Grafo causal generado por GES.



Fuente: Elaboración propia.

En este modelo, la variable LSAS deja de ser únicamente un resultado final y se convierte en un nodo transmisor. Recibe influencia directa de SMF. El nodo LSAS influye activamente sobre otras variables como CMT, DEF y DAF. Esto implica que la ansiedad social no es solo un síntoma pasivo, sino un factor que desencadena o modifica otras características psicológicas o de comportamiento en el individuo.

Se observa una ruta causal extensa y secuencial que conecta la ansiedad con otros factores: LSAS → DEF → EAF → ATF → SPIN. Esta secuencia es un hallazgo crítico: sugiere que el nivel de ansiedad

social (LSAS) es el precursor que altera una serie de factores intermedios (DEF, EAF, ATF) que finalmente determinan el puntaje en el inventario de fobia social (SPIN). A su vez, el SPIN impacta en TR y HR.

Las variables TKF y SMF aparecen en la parte superior del grafo como causas raíz.

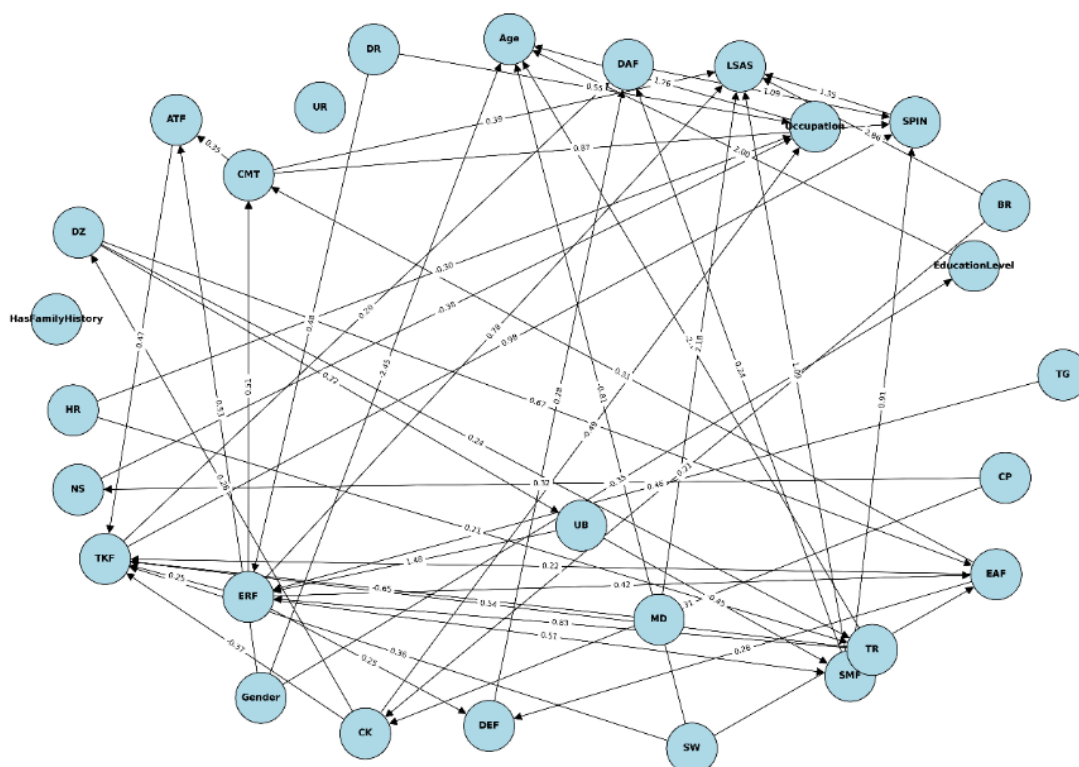
- TKF tiene un impacto múltiple, influyendo en ramas separadas hacia SW, ERF y CMT.
- SMF actúa como detonante directo del LSAS y también afecta a la rama fisiológica (BR).

A diferencia del grafo anterior, aquí la variable Occupation (Ocupación) sí está conectada. Se muestra como una causa de Age (Edad), junto con EducationLevel. Aunque la dirección causal Ocupación → Edad puede parecer contraintuitiva (y podría indicar una correlación inversa o un artefacto de los datos donde ciertos rangos de edad se agrupan en ciertas ocupaciones), es relevante notar que el algoritmo ya no considera la ocupación como irrelevante.

En la parte inferior derecha, persiste un clúster fisiológico, pero con diferente estructura. La variable DZ influye sobre MD y CK. A su vez, CK desencadena una cadena hacia CP, NS y TG.

Resultados del algoritmo LiNGAM. El algoritmo LiNGAM requiere de un tratamiento de datos previos debido a la naturaleza del conjunto de datos que se utiliza. Por lo tanto, se ajustaron variables del conjunto de datos para poder aplicar este algoritmo.

Figura 4. Grafo causal generado por LiNGAM.



Fuente: Elaboración propia.

A diferencia de los modelos PC y GES, LiNGAM identifica direcciones causales explícitas basadas en la no-Gaussianidad de los datos, lo que permite observar rutas más definidas entre grupos de variables. Variables como Gender, Age, EducationalLevel y Occupation aparecen en posiciones periféricas, pero actúan como nodos de entrada, influyendo en distintos componentes del sistema emocional y clínico. Esto sugiere que el contexto social y económico predispone indirectamente al desarrollo de síntomas de ansiedad social a través de factores intermedios.

Variables como ERF, TKF, CK y HR se localizan en posiciones altamente conectadas, recibiendo y enviando múltiples influencias. Esto indica que los síntomas fisiológicos funcionan como mediadores entre las características personales y los síntomas clínicos de ansiedad.

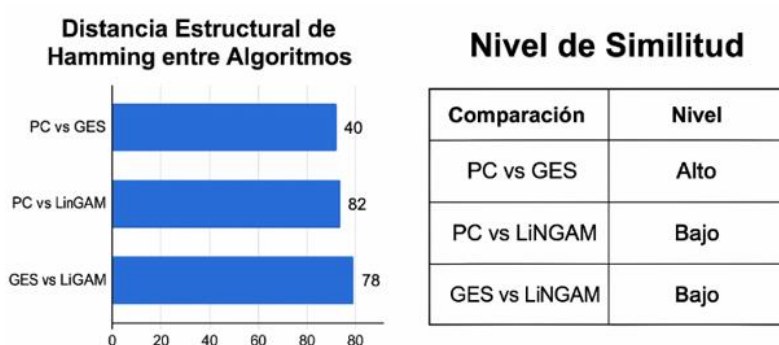
Las escalas de ansiedad social (SPIN, LSAS) aparecen como nodos receptores, lo cual es consistente con su función: sintetizan el estado emocional general después de que interactúan los factores fisiológicos, sociodemográficos y conductuales.

Variables como HasFamilyHistory y DZ muestran conexiones dirigidas hacia nodos emocionales y fisiológicos. La gran cantidad de rutas cruzadas e interacciones indica que la ansiedad emerge de múltiples vías causales simultáneas.

Distancia Estructural de Hamming. Una vez teniendo resultados de los diferentes algoritmos, se realizó una evaluación de los grafos obtenidos. Esto con la finalidad de observar las similitudes entre los resultados de los algoritmos. El método de evaluación de grafos es la Distancia Estructural de Hamming (DEH).

Los resultados obtenidos en esta evaluación son los siguientes:

Figura 5. Gráfica DEH entre los algoritmos.



Fuente: Elaboración propia.

Para este método, el indicador de similitud de grafos es un número que entre más bajo sea, más similares son los grafos entre ellos; por el contrario, entre más alto sea el número, menos similitudes hay. Como se puede observar, entre el grafo obtenido por el algoritmo PC y el grafo obtenido por GES hay más similitud. Mientras que, por ejemplo, el grafo obtenido por PC y el resultado obtenido por LiNGAM son muy diferentes.

Finalmente, se pudo determinar a través de los algoritmos que existen variables exógenas como edad, educación, ocupación y género. Son variables que afectan a síntomas de manera directa o indirecta.

Respecto a GES y LiNGAM, el nodo LSAS es fuertemente influido por los nodos SPIN, CMT y DEF. Mientras que PC al ser más conservador detecta menos relaciones, pero refuerza algunas claves como SPIN → TR y varias conexiones básicas.

El análisis de coincidencias nos permitió trazar una cadena clínica específica. A diferencia de las conexiones dispersas, este camino muestra una progresión lógica de síntomas físicos que se siguen unos a otros (tabla 1).

Tabla 1. Comparativa de una cadena específica en cada algoritmo.

Algoritmo PC		Algoritmo GES		Algoritmo LiNGAM	
Regulador	Regulado	Regulador	Regulado	Regulador	Regulado
DR	BR	DZ	MD	BR	CK
BR	CK	DZ	CK	CK	DZ
CK	CP	CK	BR	DZ	TR
CK	DZ	CK	CP	CP	CK
CP	NS	CP	NS	CP	NS
NS	DZ	CP	TG	CK	TKF

Sensación de asfixia (CK) → Dolor en el pecho (CP) → Malestar gastrointestinal y náuseas (NS)

Resulta notable que el algoritmo haya captado la lógica física de un ataque de pánico. La secuencia descubierta muestra cómo la sensación de asfixia (CK) escala a un dolor en el pecho (CP), para desembocar finalmente en malestar gastrointestinal y náuseas (NS). Podemos observar, también, que la dificultad respiratoria (BR) y los mareos (DZ) están relacionados como causas o efectos en todos los casos resultantes y con sentido lógico clínico.

Las relaciones causales varían según el algoritmo, sin embargo, la convergencia entre grafos permite afirmar que el modelo causal es estable, coherente y teóricamente justificable, proporcionando un marco sólido para comprender los factores que configuran la ansiedad.

Una limitación importante de este estudio es que el modelo causal se construyó a partir de un solo subtipo de trastorno de ansiedad, el trastorno de ansiedad social. En consecuencia, los patrones causales encontrados no pueden generalizarse de manera automática a otros trastornos de ansiedad, como el trastorno de pánico o el trastorno de ansiedad generalizada. No obstante, dado que comparten síntomas y factores de riesgo comunes, los resultados ofrecen hipótesis útiles sobre mecanismos que podrían estar

presentes en el espectro de los trastornos de ansiedad y que deberían ser contrastadas en muestras clínicas más diversas.

CONCLUSIONES

Este estudio tuvo como propósito construir un modelo computacional que ayudara a entender cómo se relacionan, de manera causal, las variables implicadas en el trastorno de ansiedad social. Al combinar los algoritmos PC, GES y LiNGAM pudimos generar y contrastar diferentes grafos causales, lo que permitió organizar de forma más clara la interacción entre factores sociodemográficos, antecedentes familiares, miedos situacionales y manifestaciones físicas de la ansiedad.

Desde el punto de vista clínico y de la salud pública, los resultados apuntan a que intervenir sobre la fobia social medida por SPIN y sobre ciertos miedos situacionales específicos podría producir cambios indirectos en el conjunto de síntomas físicos, permitiendo un tratamiento mucho más completo. El perfil demográfico juega un papel fundamental. Por ello, es importante crear planes de prevención y tratamiento personalizados, que respeten las diferencias de edad, género y formación académica.

La experiencia confirma que estos algoritmos son una vía eficaz para descifrar la complejidad de la ansiedad social. Gracias a ellos, podemos diseñar intervenciones mucho más específicas y con un respaldo científico sólido.

Anexo 1 diccionario de variables y conjunto de datos

A continuación, se describen las variables utilizadas en el análisis causal, correspondientes al conjunto de datos base.

Variables Sociodemográficas y de perfil:

- Edad: Variable numérica que indica la edad del paciente en años.
- Nivel educativo: Variable ordinal (1-6) que clasifica el grado de estudios: 1) Secundaria, 2) Diploma, 3) Pregrado, 4) Licenciatura, 5) Maestría, 6) Posgrado.
- Género: Variable dicotómica (1 = Masculino, 0 = Femenino).
- Ocupación: Variable categórica: 1) Estudiante, 2) Miembro de facultad, 3) Empleado, 4) Independiente, 5) Desempleado.
- Antecedentes familiares (hasfamilyhistory): Variable binaria que identifica la presencia de historial familiar de ansiedad o depresión (1 = Sí; 0 = No).

Variables de miedo situacional (Escala 0-10): Valores más altos indican mayor nivel de temor.

- ATF: Miedo a estar en el centro de atención.
- EAF: Miedo a comer frente a otras personas.
- TKF: Miedo a hablar en público.
- CMT: Miedo a asistir a fiestas.
- DEF: Miedo a comer o beber en lugares públicos.
- SMF: Miedo a interactuar o tener contacto con extraños.
- ERF: Miedo a entrar en una habitación donde hay otras personas sentadas.
- DAF: Miedo a tener desacuerdos con extraños.

Variables de sintomatología fisiológica (Binarias): (1 = Presencia del síntoma; 0 = Ausencia)

- HR: Palpitaciones.
- SW: Sudoración.
- TR: Temblor.
- DR: Boca seca.
- BR: Dificultad para respirar.
- CK: Sensación de asfixia.
- CP: Dolor en el pecho.
- NS: Malestar gastrointestinal o náuseas.
- DZ: Mareo o debilidad.
- UR: Sensación de irrealidad.
- UB: Miedo a perder el equilibrio.
- MD: Miedo a volverse loco.
- TG: Entumecimiento o sensación de hormigueo.

Escala global:

- SPIN (Social Phobia Inventory): Cuestionario (0-68); a mayor puntaje, mayor fobia social.
- LSAS (Liebowitz Social Anxiety Scale): Escala (0-144) que mide ansiedad y evitación social.



REFERENCIAS BIBLIOGRÁFICAS

- Causal-learn. (s.f.). PC — causal-learn 0.1.3.6 documentation. Read the Docs. https://causal-learn.readthedocs.io/en/latest/search_methods_index/Constraint-based%20causal%20discovery%20methods/PC.html
- Delgar, B. G., López, E. F., García, L. L., Rivas, M. À. F., & Miguel, S. L. (s.f.). Ansiedad: causas, síntomas, tratamiento. PortalCLÍNICA. Clínic Barcelona. <https://www.clinicbarcelona.org/asistencia/enfermedades/trastornos-de-ansiedad>
- Fathi, S., & Ahmadi, M. (2020). Data for: Development and use of a clinical decision support system for the diagnosis of social anxiety disorder [Conjunto de datos]. Mendeley Data. <https://doi.org/10.17632/4jycfwhb4g.2>
- Organización Mundial de la Salud. (8 de septiembre de 2025). Trastornos de ansiedad. <https://www.who.int/es/news-room/fact-sheets/detail/anxiety-disorders>
- Py-Why. (s.f.). Causal-learn: Causal Discovery in Python [Software]. GitHub. <https://github.com/py-why/causal-learn/>
- Qin, C., & Gai, W. (2025). A causal inference study of circulating metabolites mediating the effect of Obesity-Related indicators on the incidence of anxiety disorders. *Brain and Behavior*, 15(7). <https://doi.org/10.1002/brb3.70653>
- Renero, J., Ochoa, I., & Maestre, R. (22 de enero de 2025). REX: Causal Discovery based on Machine Learning and Explainability techniques. arXiv. <https://arxiv.org/abs/2501.12706>
- Salesforce. (s.f.). GES for Tabular Causal Discovery. Salesforce CausalAI Library 1.0 documentation. https://opensource.salesforce.com/causalai/latest/tutorials/GES_Algorithm_Tabular.html
- Saxe, G. N., Bickman, L., Ma, S., & Aliferis, C. (2022). Mental health progress requires causal diagnostic nosology and scalable causal discovery. *Frontiers in Psychiatry*, 13. <https://doi.org/10.3389/fpsyt.2022.898789>
- Tsamardinos, I., Brown, L. E., & Aliferis, C. F. (2006). The Max-Min Hill-climbing Bayesian Network Structure Learning Algorithm. *Machine Learning*, 65(1), 31–78. <https://doi.org/10.1007/s10994-006-6889-7>



- Del Moral Chica, J., Sánchez Camín, M., Sora, V., Morales Sánchez, J. C., Alquezar Domene, D., & Sánchez Fernández, F. (2024). Incremento de los problemas de salud mental tras la pandemia del COVID-19. *Ocronos*, 7(9), 855. <https://revistamedica.com/incremento-problemas-salud-mental-pandemia-covid/>
- Liebowitz, M. R. (1987). Social phobia. In D. F. Klein & B. G. Rabkin (Eds.), *Anxiety: Modern Problems of Pharmacopsychiatry* (Vol. 22, pp. 141–173). Basel: Karger.
- Connor, K. M., Davidson, J. R. T., Churchill, L. E., Sherwood, A., Foa, E., & Weisler, R. H. (2000). A psychometric assessment of the Social Phobia Inventory (SPIN). *Primary Care Companion to the Journal of Clinical Psychiatry*, 2(6), 231–235. <https://doi.org/10.4088/pcc.v02n0604>
- Chickering, D. M. (2003). Optimal structure identification with greedy search. *J. Mach. Learn. Res.* 3, 507–554.
- Spirtes, P., & Glymour, C. (1991). An Algorithm for Fast Recovery of Sparse Causal Graphs. *Social Science Computer Review*, 9(1), 62–72.
- Verma, T. S., & Pearl, J. (1990). Equivalence and Synthesis of Causal Models. *Proceedings of the 6th Conference on Uncertainty in Artificial Intelligence*, 220–227.
- Shimizu, S., Inazumi, T., Sogawa, Y., Hyvärinen, A., Kawahara, Y., Washio, T., Hoyer, P.O., & Bollen, K.A. (2011). DirectLiNGAM: A Direct Method for Learning a Linear Non-Gaussian Structural Equation Model. *J. Mach. Learn. Res.*, 12, 1225–1248.
- Lacerda, G., Spirtes, P.L., Ramsey, J., & Hoyer, P.O. (2008). Discovering Cyclic Causal Models by Independent Components Analysis. *Conference on Uncertainty in Artificial Intelligence*.
- Hoyer, P.O., Shimizu, S., Kerminen, A.J., & Palviainen, M. (2008). Estimation of causal effects using linear non-Gaussian causal models with hidden variables. *Int. J. Approx. Reason.*, 49, 362–378.
- Cotta, G. A., Runge, J., & Drouard, M. (2023). Selecting robust features for machine-learning applications using multivariate causal discovery. *Environmental Data Science*, 2. DOI: 10.1017/eds.2023.7

