

Redes psicológicas: efectos de la fiabilidad y del tamaño muestral en la recuperación de parámetros

Adrián Moll Vinuesa

Máster en Metodología de las Ciencias del
Comportamiento y de la Salud



MÁSTERES
DE LA UAM
2018 – 2019

Facultad de Psicología

UNIVERSIDAD AUTÓNOMA DE MADRID Y
UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE PSICOLOGÍA



**REDES PSICOLÓGICAS: EFECTOS DE LA FIABILIDAD Y DEL TAMAÑO
MUESTRAL EN LA RECUPERACIÓN DE PARÁMETROS**

Por:

ADRIÁN MOLL VINUESA

Tutor:

FRANCISCO JOSÉ ABAD GARCÍA

Trabajo de Fin de Máster

Máster en Metodología de las Ciencias del Comportamiento y de la Salud

Fecha de presentación: 6/2019

Índice

Resumen.....	3
Introducción	4
Tipos de redes: redes directas e indirectas	5
Estimación de redes: Regularización de los ejes	7
Índices de resumen de las redes: Índices de centralidad	8
Algunos problemas en la estimación e interpretación de las redes psicológicas	9
Propósito del estudio	12
Método	13
Estructura simulada.....	13
Generación de las matrices de varianza-covarianza y estimación de las redes.....	14
Criterios de evaluación de la recuperación de la red y de los índices de centralidad	19
Resultados.....	19
Recuperación de los ejes	19
Recuperación de los índices de centralidad	22
Discusión y conclusiones	25
Resultados principales	25
Limitaciones del estudio.....	28
Implicaciones prácticas.....	29
Referencias bibliográficas.....	30
Anexos.....	33

Resumen

El presente estudio trata de evaluar el efecto que tienen el error de medida y el tamaño muestral sobre la estabilidad de las redes psicológicas. Se esperaba que conforme se redujeran el tamaño muestral y la fiabilidad de los indicadores, la recuperación de las conexiones en las redes, así como de los índices de centralidad, fuese peor. Para contrastar estas hipótesis, se simuló datos de una red psicológica, asumiendo un modelo causal conocido, manipulando ambos factores. Se encontraron efectos principales de ambos factores en la recuperación de la mayoría de los índices de estabilidad y de centralidad ($\eta^2_p \geq .14$). Además, se encontró efecto de interacción para la correlación entre la red real y la red estimada y para el índice de cercanía. Los índices más estables resultaron ser los relativos a los ejes espurios, así como el de especificidad, lo que es consistente con la “naturaleza conservadora” de los procedimientos de estimación de las redes. Ningún método funcionaba correctamente si el tamaño muestral era inferior a 200, o si las fiabilidades de los indicadores eran menores a 0.7, lo que permite establecer algunas recomendaciones en relación a la aplicación de la técnica.

Introducción

En la última década, la investigación de los sistemas dinámicos presentes en psicología se ha visto impulsada por el uso de técnicas tales como las *redes psicológicas*, que representan una alternativa al modelo factorial clásico (Borsboom & Cramer, 2013). En este contexto, las redes psicológicas se conceptualizan como una herramienta estadística para representar la estructura de interacciones entre las distintas variables psicológicas (Epskamp, Borsboom & Fried, 2018). Estas estructuras o redes están formadas por nodos que simbolizan las variables psicológicas observadas y que se conectan mediante ejes que representan una relación estadística (usualmente correlaciones parciales). Sus principales campos de aplicación se encuentran en la psicología clínica (McNally et al., 2015), en el estudio de la personalidad (Costantini et al., 2015) y de la calidad de vida (Kossakowski et al., 2016) entre otros.

Las redes contrastan con la aproximación tradicional, el análisis factorial, que trata de explicar la estructura de covarianza de las variables observadas, asumiendo la existencia de variables latentes. En la aproximación clásica las variables se interpretan como indicadores, que son producto causal directo de una serie de dimensiones latentes (Borsboom, Mellenbergh & van Heerden, 2003). Un claro ejemplo de esta tradición sería el modelo de los cinco factores de personalidad (McCrae & Costa, 2008), en el que los factores resultantes son los responsables de la estructura de covarianza encontrada en los distintos ítems de personalidad (por ejemplo, la extraversión explicaría los comportamientos relacionados con asistir a fiestas). No obstante, esta perspectiva ha sido criticada recientemente, poniendo de relieve que la estructura de covarianza puede ser debida a las interacciones directas que existen entre las diferentes variables medidas. En el caso de los ítems de personalidad, un ejemplo sería que las personas que tienden a asistir a fiestas conocen a más gente, y a cuanta más gente conozcan, a más fiestas tenderán a ser invitados. Los distintos comportamientos se activarían o inhibirían entre sí. De esta forma se podría explicar la estructura de covarianza sin tener que recurrir a variables latentes (Costantini et al., 2015). Las redes tratan de representar las relaciones más importantes entre las variables sin imponer necesariamente la presencia de factores latentes, aunque en principio pueda considerarse la presencia de estos (Cramer et al., 2012). Mientras que el foco de atención del análisis factorial se centra casi de forma exclusiva en los elementos compartidos entre los indicadores (i.e., los factores comunes),

el análisis de redes psicológicas se fija directamente en las relaciones directas e indirectas entre las variables observadas.

Hasta el momento existe información muy limitada respecto al papel que puede tener la fiabilidad en las variables medidas para la posterior estimación de la red pertinente a la estructura de covarianza de estas. Si bien se asume que el tamaño muestral necesario para estimar correctamente los parámetros asociados a las redes psicológicas suele ser típicamente mayor al que se utiliza normalmente en el ámbito de la psicología (Epskamp et al., 2018), tras una revisión exhaustiva no se ha encontrado mención alguna al efecto que supone tener indicadores con diferentes niveles de fiabilidad en las redes. Dado que la estructura de la red depende directamente de la matriz de covarianzas entre las variables, y que esta última se verá alterada por la presencia de errores de medida, hay suficientes razones para creer que esto podría suponer un problema grave para la técnica que es menester explorar.

El presente trabajo tiene como objetivo estudiar el efecto de los errores de medida y del tamaño muestral en el funcionamiento de las redes psicológicas, así como establecer implicaciones prácticas y recomendaciones para futuras aplicaciones de la técnica.

A continuación, se presentará una breve revisión sobre las propiedades típicas de las redes, así como de los principales métodos de estimación de estas y de los principales índices utilizados en la literatura.

Tipos de redes: redes directas e indirectas

Podemos distinguir, por un lado, las redes directas, en las cuales las relaciones entre los nodos son asimétricas (A influencia a B, pero B no puede influenciar a A), estas son comúnmente llamadas redes DAG (Directed Acyclic Graph). La ventaja aparente de este tipo de redes es que en ellas se explicita la dirección causal de las relaciones, lo que puede ser en realidad un problema, pues difícilmente podemos justificar este tipo de aseveraciones. Por tanto, estas redes solo se utilizan si uno puede asumir que las relaciones no son cíclicas (Constantini et al., 2015).

Para lidiar con este problema, se han diseñado las redes indirectas, en las que todas las relaciones se interpretan como simétricas. En estas redes, los ejes típicamente representan correlaciones parciales, que indican la covarianza entre nodos que excluye la covarianza común debida al resto de variables, permitiendo así interpretar que la relación

que queda entre dos variables sea exclusiva a estas y no debida a la presencia del resto. El coeficiente obtenido se interpretaría pues como un peso de correlación parcializado. Este es el tipo de red más comúnmente utilizado ya que se ha venido recomendando utilizarlas si: a) no se dispone de suficientes pruebas que nos permitan confirmar la direccionalidad de las relaciones típicas de los modelos DAG; b) si en alguno de los nodos se sospecha que la relación es recíproca; o c) si la investigación que se lleva a cabo es de carácter exploratorio o destinada mayormente a visualizar las relaciones más destacadas entre los nodos de la red (Constantini et al., 2015).

Para obtener la matriz a partir de la cual se estima la estructura de las redes indirectas, se parte de la matriz de covarianzas (Σ), la cual se invierte para obtener la matriz de precisión ($K = \Sigma^{-1}$), para posteriormente estandarizarse, obteniendo así la matriz de covarianzas parcializada:

$$Cor(Y_i, Y_j | Y_{-(i,j)}) = -\frac{\kappa_{ij}}{\sqrt{\kappa_{ii}}\sqrt{\kappa_{jj}}}$$

donde κ_{ij} se refiere a cada elemento de K, y $Y_{-(i,j)}$ indica el conjunto de variables restante eliminando a i y j (Epskamp, Waldorp, Möttus & Borsboom, 2018). Los coeficientes de correlación parciales obtenidos son luego utilizados como pesos de los ejes, y cada variable Y_i es representada como un nodo. En la representación gráfica de la red, los ejes obtenidos se representan mediante una línea continua que une los nodos (que se representan como círculos). Mediante el color se marca la dirección de la relación (i.e., en color verde si la relación es positiva, o rojo en caso de ser negativa) y mediante el grosor se simboliza la intensidad de la relación (i.e., mayor grosor cuanto más se acerque su valor a 1 en valor absoluto). Los ejes no estarán presentes únicamente si su valor es exactamente de cero.

Un ejemplo de red simple es la mostrada por la Figura 1, donde la Depresión, tendría un índice de correlación parcial de 0.63 con la Ansiedad, y de -0.54 con la Calidad de Vida, mientras que la relación entre Ansiedad y Calidad de Vida sería de 0.

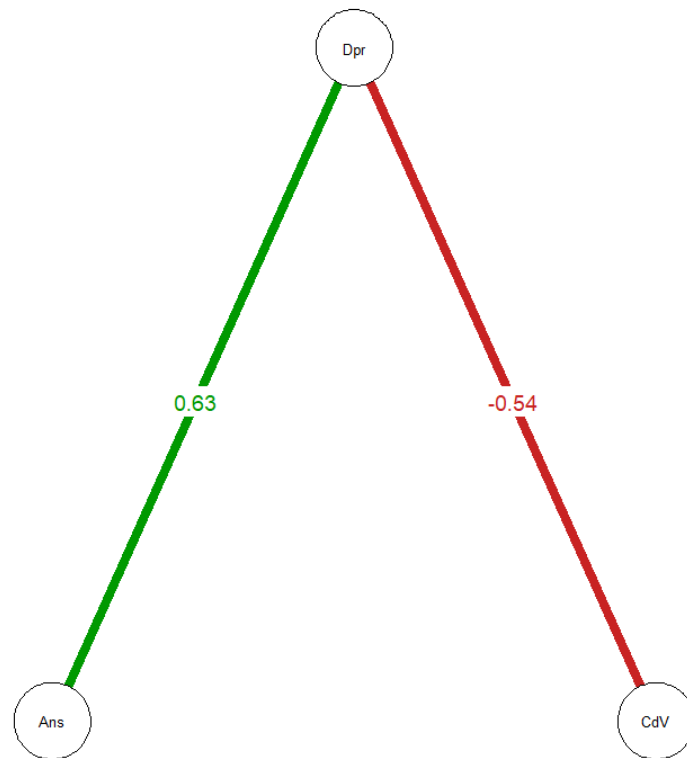


Figura 1: Ejemplo hipotético de una red simple. Dpr = Depresión, Ans = Ansiedad, CdV = Calidad de Vida. Representado por el paquete *qgraph* (Epskamp, Cramer, Waldorp, Schmittmann, & Borsboom, 2012) en R (R Core Team, 2018).

En una red indirecta los ejes pueden interpretarse de distintas formas (Epskamp, 2017):

- Como posibles relaciones causales del tipo $A \rightarrow B$, o bien $A \leftarrow B$ (Pearl, 2000).
- Como efectos predictivos, similares a los típicos coeficientes de regresión (Meinshausen & Bühlmann, 2006).
- Como interacciones mutuas, en las cuales alterar a un nodo, cambia al otro y viceversa (van Borkulo et al., 2014).

Estimación de redes: Regularización de los ejes

Debido a la variación muestral, resulta altamente improbable que el valor de un eje sea exactamente cero. Esto hace que aparezcan redes densas en conexiones, con ejes espurios. Para prevenir esto se han desarrollado métodos que permiten regularizar las correlaciones parciales y obtener una representación más parsimoniosa de las conexiones.

Originalmente se ha intentado corregir la presencia de ejes espurios eliminando aquellos que no fueran estadísticamente significativos después de aplicar una corrección por comparaciones múltiples (e.g., Bonferroni). El problema de esta estrategia es que conlleva una pérdida importante de potencia (Constanti et al., 2015). Esto implica que se fijaran a cero muchas de las conexiones existentes en la realidad. Por esto, en la práctica, se ha recomendado utilizar el método LASSO (least square shrinkage and selection operator; Tibshirani, 1996).

En el procedimiento LASSO se atenúan los coeficientes de correlación parcial obtenidos a través de un parámetro de ajuste llamado lambda (λ). Cuando lambda toma un valor de 0, ningún eje es eliminado; no obstante, conforme más alto sea el valor más ejes se irán removiendo de la red (Foygel & Drton, 2015). Típicamente se estiman un rango de valores de lambda a partir de $\lambda_{min} = R$ hasta λ_{max} , siendo λ_{max} el valor dónde ningún eje es retenido y R una ratio que ayuda a decidir el valor mínimo de lambda. Finalmente, de entre todos los valores contrastados, se selecciona aquel que minimiza al criterio de información EBIC (Extended Bayesian Information Criteria; Chen & Chen, 2008). A su vez, el criterio EBIC hace uso de un hiperparámetro llamado gamma (γ), que no debe ser confundido con lambda. El parámetro γ , al que se le suele asignar un valor entre 0 y 0.5, indica el nivel de preferencia hacia los modelos simples, por lo que un valor mayor implicará escoger modelos con menos ejes. Este procedimiento (LASSO con EBIC) ha mostrado tener un alto nivel de especificidad (i.e., elimina los ejes que no deben estar presentes), pero niveles variables de sensibilidad (i.e., mantiene los ejes que deben aparecer) que dependen del tamaño muestral y de la densidad de la red real (Epskamp & Fried, 2018).

Índices de resumen de las redes: Índices de centralidad

Una vez obtenida la estructura de la red, es habitual estimar diferentes índices que permiten resumir la información de la estructura obtenida. Los índices más utilizados han sido los llamados índices de centralidad, que pretenden operativizar la importancia de los nodos dentro de la red (e.g., cantidad de información que fluye a través del nodo, tolerancia de la red a la extracción del nodo, etc.; Borgatti, 2005). Los más relevantes han sido los siguientes:

- Índice de fuerza (*strength*): se basa en la suma (en valor absoluto) del valor de todos los ejes que conectan a un nodo concreto. Este índice informa de la

capacidad que tiene un nodo para influenciar de forma directa al resto de nodos presentes en la red sin considerar ningún tipo de efecto de mediación (Borgatti, 2005). Un ítem con un índice de fuerza alto sería aquel que puede influenciar directamente a muchos otros ítems, o a su vez, ser influenciado por muchos. Por ejemplo, el nivel de ansiedad podría afectar al insomnio, a la depresión, a la calidad de vida, al estrés percibido, etc. y a su vez este se vería alterado por cambios en los mencionados.

- Índice de cercanía (*closeness centrality*): informa de como cambios en el nodo de interés afectan con rapidez al resto de la red (Borgatti, 2005). Este índice depende del concepto de distancia, definido como el recorrido más corto posible entre dos nodos de interés (este recorrido más corto se llama geodésico). La distancia se calcula como la suma del inverso de los pesos (en valor absoluto) que intervienen en el recorrido. Por ejemplo, si A se conecta con C atravesando B, debería calcularse el inverso de los pesos A – B y B – C para luego sumarse, obteniendo como resultado la distancia entre A y C).
- Índice de intermediación (*betweenness centrality*): indica la cantidad de veces que un nodo concreto actúa como mediador entre las rutas de conexión de otros dos nodos distintos, que necesariamente necesitan atravesarlo para poder conectarse por la red. Esto se calcula sumando el número de geodésicos en las que interviene un nodo (Brandes, 2001). Si un nodo que puntúa alto en este índice fuera eliminado de la red, las distancias entre el resto de los nodos aumentarían.

Algunos problemas en la estimación e interpretación de las redes psicológicas

Interpretación de los índices de centralidad. Estos índices, no obstante, no están exentos de críticas (Bringmann et al., 2018). Por una parte, se argumenta el que estos índices fueron desarrollados para el contexto de las redes sociales en sociología, que se distingue en gran medida de las redes psicológicas, al que se suma el problema de que en el primer contexto ya aparecían problemas con su interpretabilidad y conceptualización (Borgatti, 2005). Por otro lado, estos índices tienden a ser poco estables y tener intervalos de confianza demasiado amplios, especialmente el índice de cercanía y de intermediación (Bulteel, Tuerlinckx, Brose, & Ceulemans, 2016). Además de estos problemas, no se ha desarrollado suficiente investigación acerca del valor predictivo de estos (no parece que

un síntoma central en depresión prediga mejor el trastorno de depresión mayor que otros síntomas menos centrales; Boschloo, van Borkulo, Borsboom, & Schoevers, 2016).

Interpretación causal de los ejes. A pesar de que se recomienda no interpretar de forma causal las relaciones encontradas en las redes, sino considerarlas como “fuentes prometedoras” de causalidad (Pearl, 2000), no son pocos los estudios que hacen uso de un lenguaje completamente causal. Por ejemplo, Jayawickreme et al. (2017), realizó un estudio en el que analizaba las relaciones dinámicas entre eventos traumáticos, eventos estresantes y síntomas típicos de la psicopatología en supervivientes de guerra en Sri Lanka; en su trabajo, si bien recomendaba más investigación en términos generales, proponía centrar la atención futura en los nodos más centrales, tanto en el ámbito de la investigación básica, como en la aplicada (p.e., desarrollo de futuras terapias). Sin embargo, Bringmann et al. (2018) critican ese tipo de inferencias, ya que un nodo puede ser central meramente por tratarse un efecto común; si esto fuera así, intervenir en este tipo de síntomas no alteraría al resto de los nodos presentes.

Además de lo anterior, la presencia de efectos comunes supone un problema para la interpretación de las correlaciones parciales de las que se nutren las redes indirectas. En concreto, es bien conocido el problema de condicionar a variables de efecto común, pues hacer esto supone la generación de una relación espuria negativa entre dos nodos que en realidad son independientes (Pearl, 2000). Koller y Friedman (2009), ejemplifican esto al hipotetizar qué sucedería si se parcializan las correlaciones entre la motivación y la facilidad de un examen con la nota obtenida en un test. Sabemos que la nota obtenida se verá afectada por la motivación y la facilidad del test, pero, aunque conozcamos la motivación de un estudiante, no podemos intuir la facilidad de un examen (por lo que son independientes). No obstante, si nos fijamos en un estudiante que obtuvo notas altas en el test, y sabemos que no estaba motivado, podemos intuir que el examen era fácil. Condicionar estas variables creará una relación negativa espuria. Otro caso en el que podría suceder esto sería al seleccionar submuestras de la población (como grupos que puntúan alto en cierto factor). Gráficamente se vería como en la Figura 2:

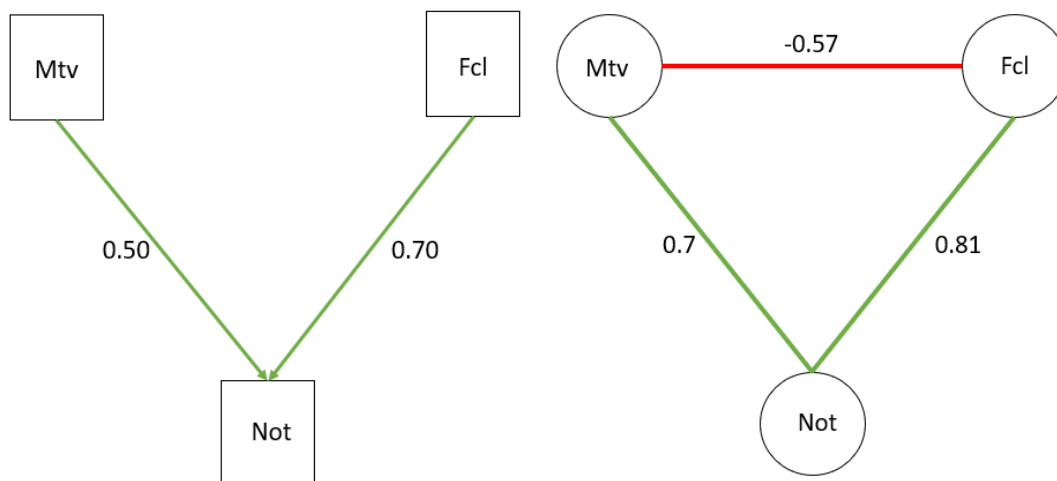


Figura 2: A la izquierda modelo SEM sobre la estructura real de la relación entre Motivación (Mtv), Facilidad del examen (Fcl), y Nota obtenida en el test (Not). A la derecha modelo de redes (parcializado) sobre la relación entre las mismas variables.

No tener en cuenta este problema llevaría a interpretar que existe una relación causal negativa importante entre la motivación y la facilidad del examen, cuando es claramente espuria. En este caso, la estructura de la red es simple, pero es de esperar que, en estructuras con mayor número de nodos y relaciones, la presencia de este tipo de ejes será mucho mayor.

Recuperación de las redes. Otras críticas se han centrado en señalar que las redes podrían no ser estables a través de las muestras; por ejemplo, Epskamp et al. (2018) señalando esta problemática propuso utilizar procedimientos para calcular los intervalos de confianza tanto para los índices de centralidad, como para los pesos obtenidos en los ejes de las redes. Estos autores encontraron que los pesos de los ejes presentaban una variabilidad demasiado amplia, especialmente en muestras pequeñas. Por otro lado, en relación con los índices de centralidad, solo el índice de fuerza parecía ser moderadamente estable.

Mientras que se ha estudiado el efecto del tamaño muestral en la recuperación de las redes, no parece haberse estudiado el efecto que pueda tener la mayor o menor fiabilidad de los indicadores con los que se estima la red. Dado que la fiabilidad de los ítems medidos afecta de manera directa a la estructura de covarianza, y la estimación de la red depende exclusivamente de esta, es de esperar que se encuentre un efecto importante de la precisión de las medidas. Por ejemplo, si se cumple el supuesto de que los errores de medida de X e Y son independientes, la correlación entre las medidas se

verá atenuada por la falta de fiabilidad de estas: $r_{XY} = R_{XY} \sqrt{r_{XX} r_{YY}}$ donde r_{XY} se refiere a la correlación observada, R_{XY} a la correlación verdadera, y r_{XX} e r_{YY} representan a los coeficientes de fiabilidad de las variables que intervienen (Trafimow, 2015). Lo esperado, por tanto, es que los pesos estimados se vean atenuados en mayor medida conforme decrezca la fiabilidad de los ítems. Se debería presumir pues una mejor recuperación de las redes y de los índices de centralidad cuando los ítems presentan fiabilidades más altas.

En la Figura 3 se ejemplifica el efecto de la fiabilidad en una red. Puede observarse que la presencia de ejes espurios aumenta al presentar ítems con fiabilidades menores y tamaños de la muestra más pequeños.

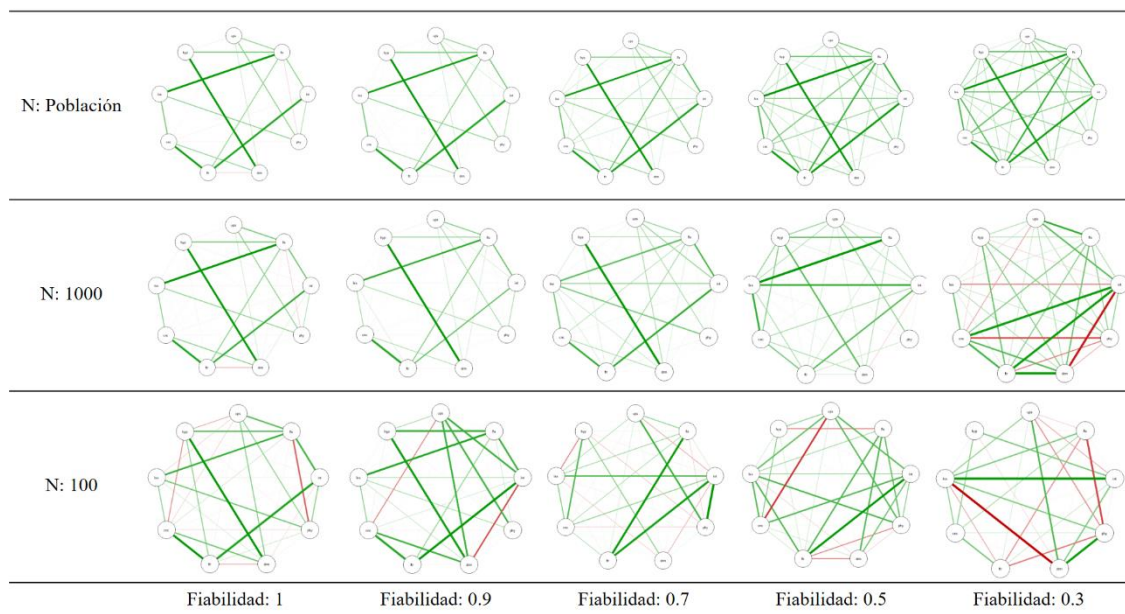


Figura 3: Modelos de redes psicológicas sobre estructura teórica de trastorno de estrés postraumático (PTSD); N = Tamaño de la muestra.

Propósito del estudio

El presente estudio tiene como principal objetivo evaluar el efecto de la fiabilidad de los indicadores en la recuperación de las redes psicológicas y de los índices de centralidad propios de la aplicación de la técnica.

Para esto se llevará a cabo un estudio de simulación en el que se generarán distintas estructuras de covarianzas, con una red real subyacente conocida, pero variando el valor de fiabilidad de los ítems incluidos y el tamaño muestral. Se utilizará como modelo generador la estructura de un conjunto de variables relacionadas con el trastorno de estrés postraumático (PTSD), analizada por Armour, Fried, Deserno, Tsai y Pietrzak

(2017), con el objetivo de valorar una situación compleja y realista (mayor número de nodos y relaciones) típica de las aplicaciones de redes psicológicas.

Finalmente, a partir de los resultados, se proporcionarán algunas recomendaciones para futuras aplicaciones de la técnica, así como sugerencias sobre posibles investigaciones que convendría desarrollar para profundizar en la cuestión.

Método

El presente trabajo se abordó desde una estructura formulada mediante el análisis de las variables recogidas por Armour et al. (2017) sobre PTSD, dónde el número de nodos y ejes presentes es sustancialmente mayor a los anteriores análisis, por lo que se puede valorar la recuperación en el contexto de una red compleja (típico de las aplicaciones prácticas reales en la literatura).

En relación a esa estructura se generaron datos en los que se manipulaba el coeficiente de fiabilidad de las variables incluidas (0.5, 0.7, 0.9 y 1) y el tamaño muestral (100, 200, 500 y 1000).

Estructura simulada

El modelo que se utiliza para comprobar la estabilidad de las redes psicológicas trata de representar un caso realista, típico de las aplicaciones de la técnica, en los que se suelen analizar estructuras complejas (elevado número de nodos y de ejes). Este está basado en los datos publicados por Armour et al. (2017), quienes analizan un modelo de PTSD aplicado a veteranos militares. El modelo contiene 20 ítems que miden 4 dimensiones diferentes propias del trastorno (pensamientos intrusivos, comportamientos de evitación, alteraciones del ánimo y la cognición y alteraciones de la reactividad y activación).

Para conocer la estructura de causalidad que podría subyacer a estos datos se utilizó como base el análisis de redes psicológicas aplicado por McNally et al. (2015) ya que se hacía uso de la misma escala. Dado que esta era una red demasiado densa y la valoración resultaría demasiado compleja, se seleccionaron 9 de las variables analizadas. Se aplicó un análisis confirmatorio a los datos con dicha estructura (ver Figura 4) y tras comprobar que el modelo ajustaba ($CFI = 0.96$, $RMSEA = 0.07$, $SRMR = 0.07$), se tomó como un posible modelo poblacional realista.

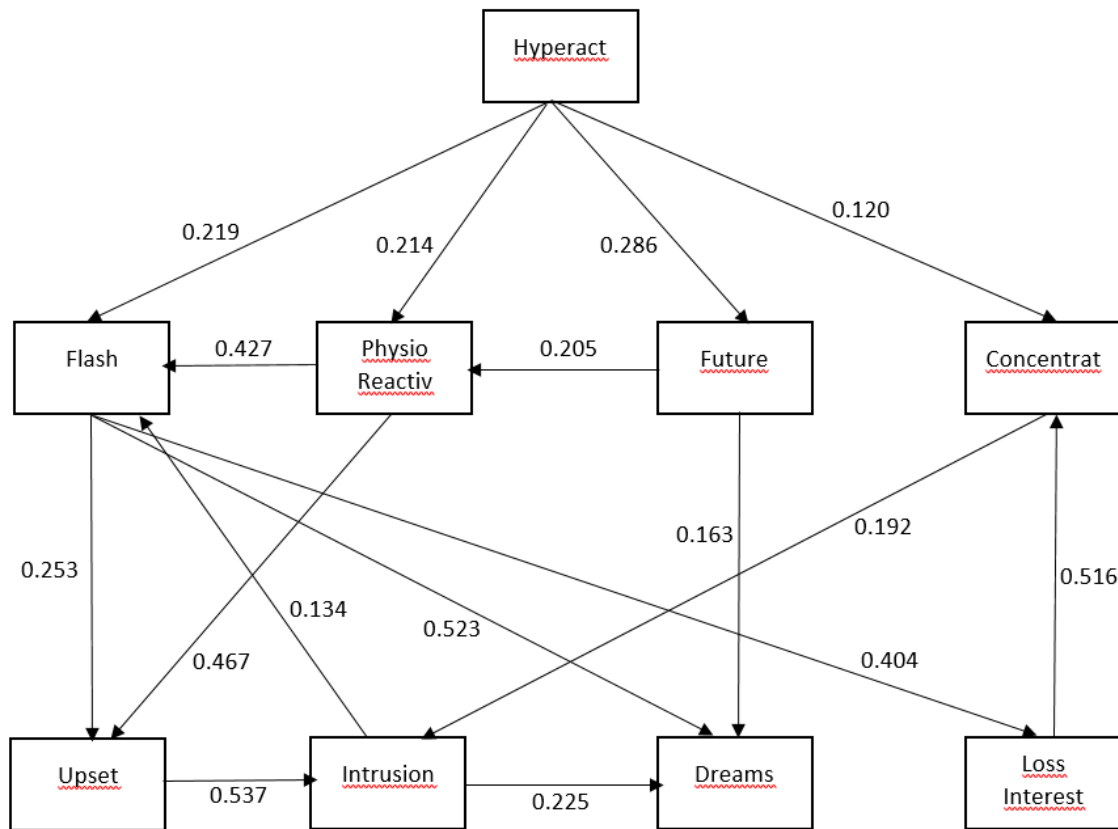


Figura 4: Modelo de PTSD simulado; Hyperact = Hiperactividad; Flash = Recuerdos repentinos; Physio Reactiv = Reactividad fisiológica; Future = Pensamientos acerca del futuro; Concentrat = Problemas de concentración, Upset = Preocupaciones, Intrusion = Pensamientos intrusivos; Dreams = Ensoñaciones; Loss interest = Pérdida de interés.

Generación de las matrices de varianza-covarianza y estimación de las redes

Partiendo del modelo confirmatorio anterior, se generaron distintas matrices de varianzas-covarianzas poblacionales sobre las que aplicar los modelos de redes. Para poder simular datos con distinta fiabilidad, se asumió una estructura subyacente a las variables observadas del modelo. Para ello, se generó una función en R (R Core Team, 2018; disponible en anexos).

Para explicar el procedimiento y la notación, en la Figura 5 se representa una parte reducida del modelo, a modo de ejemplo simplificado.

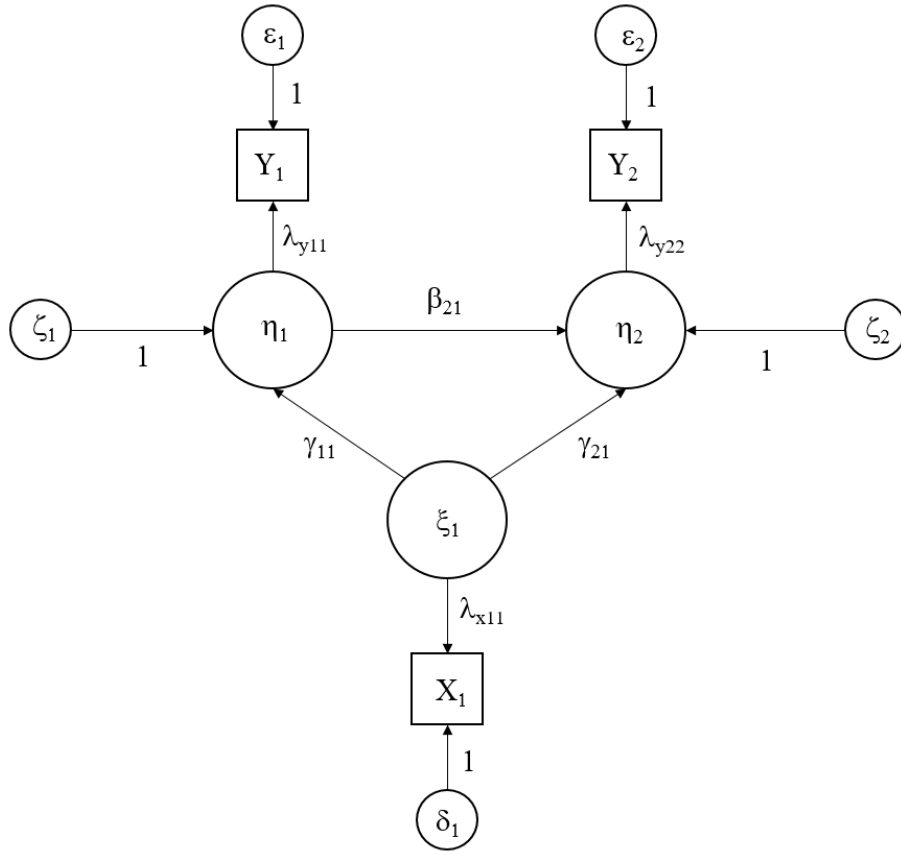


Figura 5: Ejemplo de modelo estructural con parámetros representados. X_1 = Hyperact; Y_1 = Physio Reactiv; Y_2 = Flash.

La matriz de covarianzas reproducida a partir de los parámetros poblacionales del modelo estructural se obtenía como:

$$\Sigma(\theta) = \begin{pmatrix} \Sigma_x(\theta) & \Sigma_{xy}(\theta) \\ \Sigma_{yx}(\theta) & \Sigma_y(\theta) \end{pmatrix} = \begin{pmatrix} \Lambda_x \Phi \Lambda_x' + \Theta_\delta & \Lambda_x \Phi \Gamma' B^{*'} \Lambda_y' + \Theta_{\delta\epsilon} \\ \Lambda_y B^* \Gamma \Phi \Lambda_x' + \Theta_{\epsilon\delta} & \Lambda_y B^* (\Gamma \Phi \Gamma' + \Psi) B^{*'} \Lambda_y' + \Theta_\epsilon \end{pmatrix}$$

dónde: (1) $\Sigma_x(\theta)$ y $\Sigma_y(\theta)$ representan las matrices de varianzas-covarianzas entre las variables observadas asociadas a los factores exógenos (X) y las variables observadas asociadas a los factores endógenos (Y), respectivamente; (2) $\Sigma_{xy}(\theta)$ representa la matriz de covarianzas entre las variables X y las variables Y (siendo $\Sigma_{yx}(\theta)$ la transpuesta).

Los parámetros que aparecen en las matrices son los siguientes:

- Las matrices Λ (Λ_x e Λ_y) contienen los coeficientes de regresión de las variables observadas (X e Y) sobre las variables latentes (ζ y η). Los elementos de esas matrices indican el peso factorial de las variables sobre los factores.

- Las matrices de varianzas-covarianzas de factores latentes exógenos (ζ) y los errores (ζ) de los factores latentes endógenos (η) se denominan Φ y Ψ , respectivamente. Cuando están estandarizadas se interpretan como matrices de correlaciones. En el presente estudio, se fijaba la diagonal de la matriz Φ a 1, de forma que se asume que los factores ζ están estandarizados. También se asumió una matriz Ψ diagonal, lo que implica que las covarianzas entre las variables endógenas están explicadas por el efecto de otras variables. Los valores de la diagonal de Ψ se fijaron para que los factores η estuvieran en métrica estandarizada.
- La matriz Γ contiene los coeficientes de regresión de η sobre ζ , interpretándose como los pesos de los factores exógenos sobre los endógenos.
- La matriz $\mathbf{B}^*$ es la inversa de una matriz que depende de la matriz identidad ($\mathbf{I}$) y de la matriz $\mathbf{B}$, que contiene los coeficientes de regresión de η sobre η , es decir, los pesos de los factores endógenos sobre otros factores endógenos.
- Las matrices Θ_δ y Θ_ϵ representan las matrices de varianza-covarianza de los errores de las variables X e Y . $\Theta_{\delta\epsilon}$ (y su transpuesta, $\Theta_{\epsilon\delta}$) contiene las covarianzas entre los términos de error de X e Y . En el presente estudio se asume que todas las covarianzas entre errores son cero, lo que implica que las covarianzas entre las variables observadas son explicadas por el efecto de los factores latentes.

En primer lugar, se generaba la matriz de covarianzas para los factores (η , ζ) mediante la especificación de las matrices Φ , $\mathbf{B}$ y Γ . Estas matrices definían la direccionalidad de las relaciones. Las covarianzas entre factores exógenos (i.e., variables que no recibían flechas de otras variables del modelo) se definían en la matriz Φ , mientras que las relaciones causales se definían en las matrices $\mathbf{B}$ y Γ .

En segundo lugar, se definía una variable observada por factor y se manipulaban las matrices Λ_x y Λ_y , para representar situaciones con distinta fiabilidad. Esto es, cambiando la intensidad de la relación entre X e ζ (i.e., Λ_x) se manipulaba la fiabilidad del indicador X . Nótese, por ejemplo, que siendo X e ζ variables estandarizadas, un valor λ_x de 0.7 implicaría un coeficiente de fiabilidad de X de 0.7.

Finalmente, a partir de los valores de Φ , $\mathbf{B}$, Γ , Λ_x y Λ_y , se derivaba, a nivel poblacional, la matriz $\Sigma(\theta)$ o estructura de covarianzas entre variables observadas. A partir de dicha matriz poblacional se generaban muestras con distinto tamaño muestral,

que eran utilizada a continuación para calcular la red mediante el paquete *qgraph* (Epskamp et al., 2012).

Nótese que el modelo de la Figura 4 contiene tres tipos posibles de relaciones causales (ver Figura 6): (1) cadena, (2) causa común, (3) efecto común (se espera encontrar errores en este último tipo de relación por la problemática mencionada anteriormente).

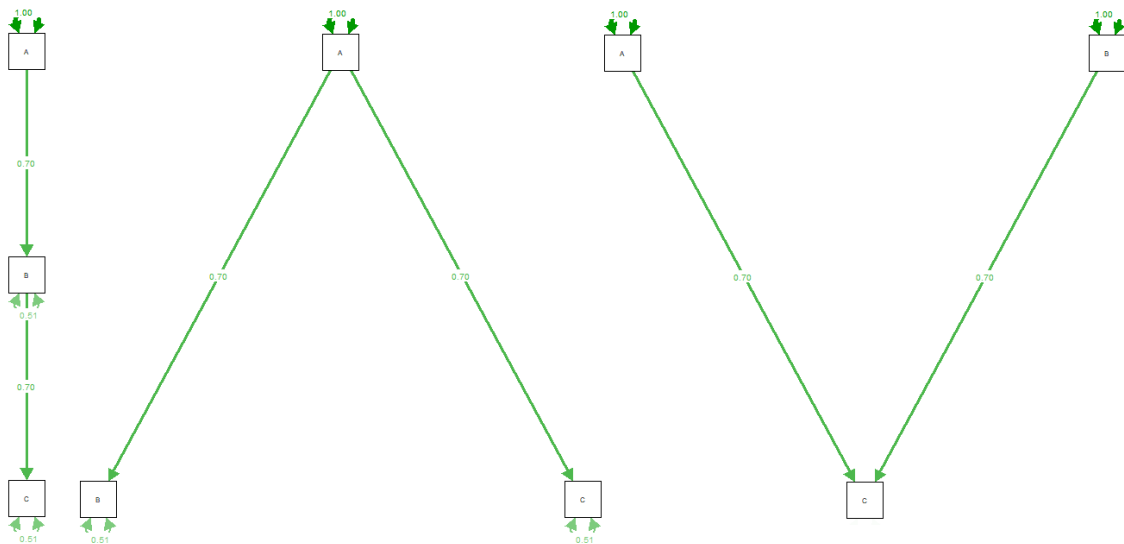


Figura 6: De izquierda a derecha: modelo de cadena, modelo de causa común y modelo de efecto común.

En total se generaron 16 condiciones (4 tamaños muestrales $\times$ 4 niveles de fiabilidad), para las que se generaron 100 réplicas. Para generar muestras de distinto tamaño y distintas réplicas, así como para obtener los índices que se describen en el siguiente apartado, se utilizó la función *netSimulator* del paquete *bootnet* (Epskamp et al., 2018), levemente modificada. La red poblacional sin error de medida con la que se comparaban las diferentes redes estimadas era la siguiente:

Criterios de evaluación de la recuperación de la red y de los índices de centralidad

Siguiendo las recomendaciones establecidas por de Ron, Fried, y Epskamp (2019), se estimó la especificidad (proporción de ejes no presentes en la red verdadera que fueron estimados como no presentes también en las redes estimadas), la sensibilidad (proporción de ejes presentes en la red verdadera que fueron estimados como presentes en las redes estimadas), la correlación (entre los pesos de los ejes de la red verdadera y los pesos estimados en las redes simuladas), la proporción de ejes espurios negativos (cantidad de ejes espurios negativos dividida entre cantidad total posible de estos) y la proporción de ejes espurios positivos (cantidad de ejes espurios positivos dividida entre cantidad total posible de estos). Paralelamente se obtuvieron coeficientes de estabilidad para los índices de centralidad (fuerza, cercanía e intermediación), valorados por la correlación resultante entre el índice estimado y el índice real.

Finalmente, con el propósito de entender mejor los efectos de los factores recogidos, se realizaron diferentes pruebas ANOVA para cada uno de los coeficientes de estabilidad (variables dependientes) teniendo como factores al tamaño muestral y al nivel de fiabilidad de los indicadores (variables independientes). Para delimitar el número de resultados mostrados, únicamente aquellas interacciones que alcanzaron un tamaño del efecto grande ($\eta_p^2 \geq .14$) se representaron gráficamente.

Resultados

Recuperación de los ejes

En la Tabla 1 se presenta una evaluación general de la precisión y estabilidad del modelo de redes psicológicas, la cual incluye la media de los coeficientes recogidos a través de los diferentes niveles de las variables independientes, así como su desviación típica. Todos los resultados referentes a las pruebas ANOVA eran estadísticamente significativos. Los tamaños del efecto resultantes se muestran en la Tabla 2.

Tabla 1: Estabilidad de la red en el modelo PTSD

Grupo	Correlación		Sensibilidad		Especificidad		Espurio Pos		Espurio Neg	
	Media	DT	Media	DT	Media	DT	Media	DT	Media	DT
Fiabilidad:1										
N=100	.84	.05	.47	.06	.94	.07	.02	.03	.00	.01
N=200	.91	.03	.56	.07	.96	.06	.02	.00	.01	.01
N=500	.96	.02	.67	.07	.97	.04	.01	.00	.01	.01
N=1000	.98	.01	.76	.05	.98	.04	.01	.00	.00	.01
Fiabilidad:0.9										
N=100	.74	.09	.47	.06	.91	.08	.05	.03	.00	.01
N=200	.85	.06	.56	.07	.94	.06	.03	.03	.00	.01
N=500	.94	.03	.67	.07	.93	.06	.03	.03	.00	.00
N=1000	.96	.01	.76	.05	.92	.06	.04	.03	.00	.00
Fiabilidad:0.7										
N=100	.47	.15	.28	.08	.89	.08	.06	.04	.00	.01
N=200	.63	.13	.37	.06	.88	.09	.06	.04	.00	.00
N=500	.78	.07	.52	.06	.86	.07	.07	.04	.00	.00
N=1000	.85	.05	.62	.05	.80	.08	.10	.04	.00	.00
Fiabilidad:0.5										
N=100	.18	.20	.11	.07	.93	.07	.04	.03	.00	.01
N=200	.31	.20	.18	.07	.93	.07	.04	.04	.00	.01
N=500	.51	.16	.30	.07	.90	.08	.06	.04	.00	.00
N=1000	.62	.12	.43	.08	.83	.08	.09	.04	.00	.00

Nota. Número de réplicas por condición = 100; DT = Desviación típica; Espurio Pos = Ejes espurios positivos; Espurio Neg = Ejes espurios negativos.

Tabla 2: Tamaños del efecto para estabilidad de la red

	Correlación	Sensibilidad	Especificidad	Espurio Pos	Espurio Neg
Muestra	<u>0.54</u>	<u>0.76</u>	0.05	0.07	0.01
Fiabilidad	<u>0.79</u>	<u>0.82</u>	<u>0.23</u>	<u>0.37</u>	0.07
Interacción	<u>0.17</u>	0.05	0.11	0.13	0.02

Nota. Medida del tamaño del efecto = eta cuadrado parcial; En subrayado efectos grandes ($\eta_p^2 \geq .14$); Todos los efectos son significativos; Espurio Pos= Ejes espurios positivos; Espurio Neg = Ejes espurios negativos.

Los resultados de la Tabla 1 muestran que los índices más precisos y estables a través de las diferentes condiciones fueron los relativos a la presencia de ejes espurios.

La media y desviación típica de estos dos índices fue cercana a 0 en los diferentes niveles de los factores; a pesar de esto, el índice de ejes espurios positivos comprendió valores ligeramente superiores al de ejes espurios negativos, aunque siempre por debajo del 10%. El índice de ejes espurios positivos tendía a aumentar según decrecía la fiabilidad de los indicadores ($\eta_p^2 = .37$) mientras que el índice de ejes espurios negativos no se vio afectado por ninguna de las dos variables independientes.

Para valorar el resto de los índices, se optó por establecer un punto de corte de .70 como valor aceptable. Partiendo de este criterio el índice de estabilidad más robusto fue el de especificidad dado que sus valores son superiores a .80 en las distintas condiciones. No obstante, sí se observó un efecto de la fiabilidad ($\eta_p^2 = .23$). El buen funcionamiento del anterior índice se asoció a un descenso en sensibilidad, que se veía afectada tanto por el tamaño muestral ($\eta_p^2 = .76$), como por la fiabilidad ($\eta_p^2 = .82$). La sensibilidad sólo era aceptable con tamaños de la muestra muy elevados ($N = 1000$) y con indicadores de fiabilidad alta (mayor o igual que .90). Por el contrario, en otras condiciones (e.g., muestras pequeñas) la red no representaba muchos de los ejes que sí deberían aparecer, lo que ilustra la naturaleza conservadora de la técnica.

Por último, respecto al índice de correlación entre los tamaños de los ejes en la red real y la estimada, este se ve afectado en gran medida por el tamaño de la muestra ($\eta_p^2 = .54$), por la fiabilidad ($\eta_p^2 = .79$) y por la interacción de ambos factores ($\eta_p^2 = .17$). Los resultados de esta interacción pueden verse representados en la Figura 8.

Se observó que el funcionamiento del índice mejoraba con fiabilidades superiores, pero que la mejora era superior para tamaños de la muestra más pequeños. La estabilidad de la red con tamaños de la muestra de 100 sujetos era aceptable con fiabilidades de 0.9. Si la fiabilidad de los indicadores era de 0.7, el tamaño de la muestra necesario era de al menos 500 sujetos. Con fiabilidades inferiores, independientemente del tamaño muestral, ningún resultado era aceptable. En definitiva, la correlación era aceptable con al menos 500 sujetos y fiabilidades de .70.

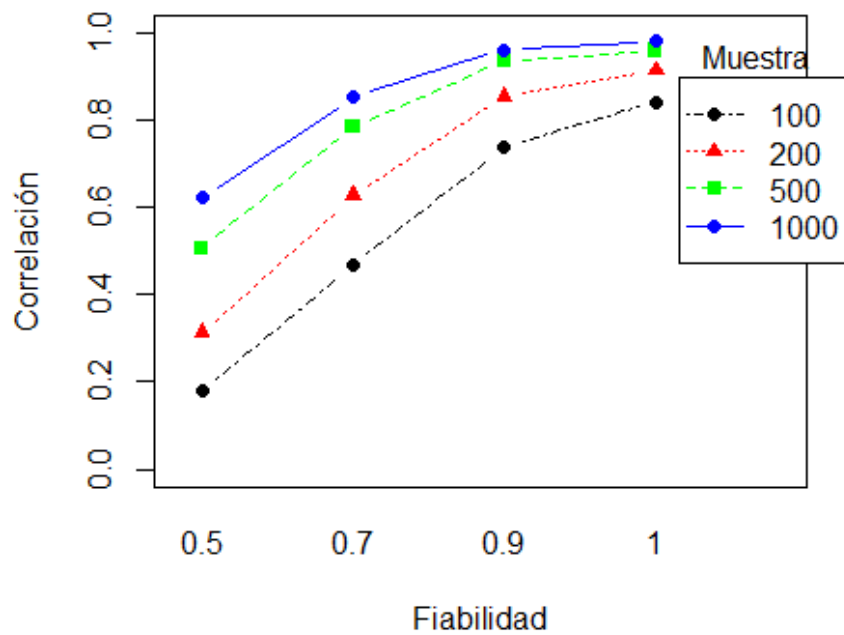


Figura 8: Representación de la interacción para el índice de correlación entre red estimada y red real.

Recuperación de los índices de centralidad

En la Tabla 3 se presentan los resultados consecuentes a la evaluación de la estabilidad en los índices de centralidad. De nuevo, se realizaron las pertinentes pruebas ANOVA para estos índices y se encontró que todos los efectos eran estadísticamente significativos. Los tamaños del efecto se muestran en la Tabla 4.

Los resultados de la Tabla 3 muestran que los índices de centralidad son en general menos estables que los analizados anteriormente. El índice de fuerza resulta ser el más estable, pues incluso en la condición de nivel más bajo de fiabilidad (0.5), si se dispone del suficiente tamaño muestral ($N = 1000$), la recuperación era aceptable (0.81). No obstante, a pesar de obtener la mayor estabilidad, era a su vez afectado en mayor medida tanto por el factor del tamaño muestral ($\eta_p^2 = .35$), así como por la fiabilidad de los indicadores ($\eta_p^2 = .42$).

Tabla 3: Estabilidad de la centralidad en el modelo PTSD

Grupo	Fuerza		Cercanía		Intermediación	
	Media	DT	Media	DT	Media	DT
Fiabilidad:1						
N=100	.78	.11	.76	.21	.68	.27
N=200	.87	.07	.88	.07	.83	.15
N=500	.92	.05	.94	.04	.90	.09
N=1000	.95	.03	.96	.02	.95	.04
Fiabilidad:0.9						
N=100	.77	.14	.69	.30	.58	.32
N=200	.87	.09	.86	.12	.75	.23
N=500	.94	.04	.95	.04	.89	.11
N=1000	.96	.03	.96	.02	.94	.06
Fiabilidad:0.7						
N=100	.59	.24	.19	.31	.40	.37
N=200	.74	.15	.40	.40	.54	.37
N=500	.88	.06	.88	.11	.74	.24
N=1000	.92	.04	.92	.04	.79	.18
Fiabilidad:0.5						
N=100	.28	.35	.00	.00	.14	.39
N=200	.45	.31	.02	.11	.20	.39
N=500	.67	.22	.15	.31	.42	.43
N=1000	.81	.10	.54	.41	.56	.33

Nota. Número de repeticiones por condición = 100; DT = Desviación típica

Tabla 4: Tamaños del efecto para estabilidad de los índices de centralidad de la red

	Fuerza	Cercanía	Intermediación
Muestra	<u>0.35</u>	<u>0.39</u>	<u>0.20</u>
Fiabilidad	<u>0.42</u>	<u>0.65</u>	<u>0.34</u>
Interacción	0.11	<u>0.22</u>	0.02

Nota. Medida del tamaño del efecto = eta cuadrado parcial; En subrayado efectos grandes ($\eta_p^2 \geq .14$); Todos los efectos son significativos.

Por otro lado, el índice menos robusto resultó ser el referente a la intermediación que, de forma similar, se vio afectado en gran medida por ambos factores (muestra: $\eta_p^2 = .20$; fiabilidad: $\eta_p^2 = .34$). Por ejemplo, con coeficientes de fiabilidad de 0.7, se requería de un tamaño muestral elevado (N=500) para alcanzar un valor aceptable. Incluso con el nivel de fiabilidad más alto (i.e., 1), se requería una muestra de al menos 200 sujetos.

Por último, respecto al índice de cercanía, mostró funcionar de forma similar al anterior, aunque con valores ligeramente superiores en las diferentes condiciones. En este caso, se encontraron efectos tanto principales (muestra: $\eta_p^2 = .39$; fiabilidad: $\eta_p^2 = .65$) como de interacción ($\eta_p^2 = .22$). Los resultados de esta interacción pueden verse representados en la Figura 9. Conforme aumenta la fiabilidad, aumenta la estabilidad del índice, pero esto se da en mayor o menor medida en función del tamaño de la muestra. Para muestras grandes ($N \geq 500$), el mayor incremento se produce cuando la fiabilidad aumenta de 0.5 a 0.7, mientras que para muestras más pequeñas se producen incrementos más importantes en niveles más altos de fiabilidad (0.7 a 0.9). La estabilidad del índice era aceptable con muestras de 100 sujetos si la fiabilidad era perfecta. En caso de no ser así se requería de al menos 200 sujetos si se tenía fiabilidades muy altas (0.9). Con un nivel en el factor de fiabilidad de 0.7, era necesario un tamaño muestral de 500. Por último, con el menor nivel del factor fiabilidad, independientemente del tamaño muestral, el resultado era inaceptable.

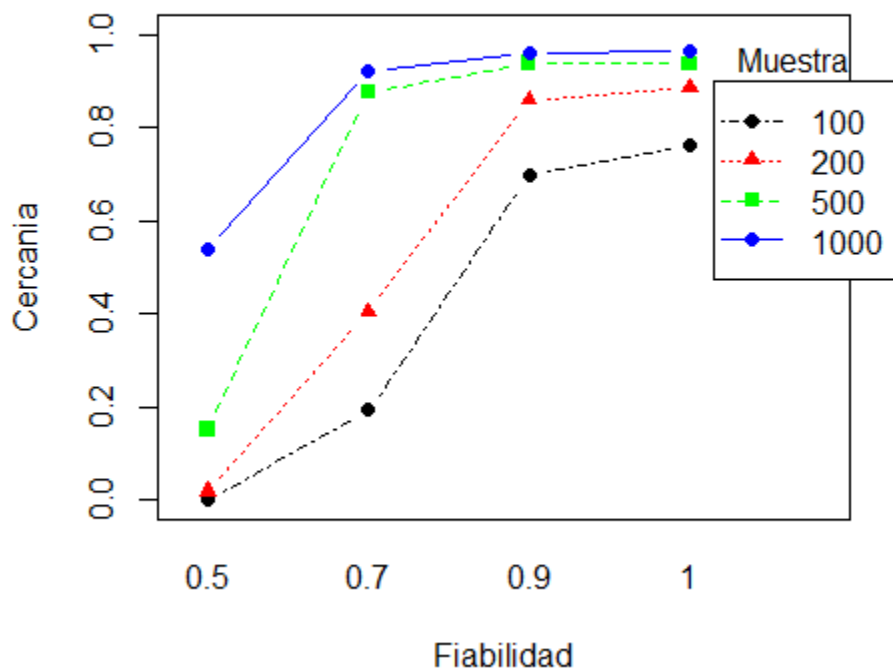


Figura 9: Representación de la interacción para el índice de centralidad de cercanía.

Discusión y conclusiones

En los últimos años, se ha observado un incremento del modelado de redes en el campo de la psicometría (Schmittmann et al., 2013). Si bien estas técnicas permiten indagar sobre la estructura de causalidad entre los indicadores recogidos y analizar la relevancia o centralidad de cada uno de estos, hasta el momento hay información limitada acerca de la estabilidad de estas redes al manipular condiciones como la fiabilidad y su interacción con el tamaño muestral. Además, se desconoce también la medida en que la presencia de efectos comunes (Pearl, 2000), puede dificultar la interpretabilidad de las redes resultantes. Teniendo en cuenta la problemática descrita previamente, se llevó a cabo la simulación de diferentes matrices de covarianzas teóricas donde dichas condiciones se ven reflejadas en diferentes niveles. Para poder medir la precisión de estos métodos se simulaban muestras basadas en dichas matrices, a partir de las cuales se realizaba el cómputo de los diferentes índices de interés, contrastando así la estabilidad y robustez de la técnica además de la información referente a la centralidad. Los resultados más importantes acerca del efecto de la fiabilidad y el tamaño de la muestra de este estudio son discutidos a continuación.

Resultados principales

Los índices de estabilidad revisados en las diferentes condiciones revelan que la estabilidad y precisión de los Ejes espurios negativos es la mayor entre el resto de los índices. Este está seguido de los Ejes espurios positivos y la Especificidad, cuyos resultados siguen siendo muy positivos. En cuanto a la Correlación entre la red real y la red estimada, mostró funcionar relativamente bien, no siendo el caso para la Sensibilidad pues su comportamiento era inaceptable en la mayoría de las condiciones yendo estos resultados en línea con lo observado en anteriores estudios (Epskamp et al., 2018).

Si prestamos atención al efecto que tiene el tamaño de la muestra en la estabilidad de la red se observa que, incluso en muestras pequeñas, se mantienen niveles elevados de Especificidad, siendo escasa la aparición de ejes espurios. En cambio, reducir el tamaño muestral reduce la correlación entre la red estimada y la red real, así como la Sensibilidad. Estos resultados están en línea con las expectativas teóricas, puesto que el procedimiento LASSO tiende a ser conservador en muestras pequeñas (Epskamp et al., 2018). Esto beneficia, por definición, al índice de Especificidad (pues este índice es sensible a la aparición de ejes espurios), pero se asocia a una menor potencia o Sensibilidad. También

afecta, de forma similar, aunque en menor medida, a la correlación entre la red estimada y la red real ya que se acaban fijando a cero algunos ejes originalmente presentes en la estructura verdadera. No obstante, como se ha comentado con anterioridad, el índice funciona relativamente bien con tamaños muestrales superiores a 200.

Al observar el efecto de la fiabilidad observamos un patrón similar al comentado previamente para el tamaño muestral, con algunas diferencias. Por un lado, aunque el índice de Ejes espurios negativos permanece robusto a las variaciones de fiabilidad, estas sí afectan a la aparición de Ejes espurios positivos y, consiguientemente, a la Especificidad. Esto podría deberse a que el modelo simulado presentaba únicamente ejes positivos, por lo que la cantidad de Ejes negativos espurios simulados al atenuarse las correlaciones por el efecto de la fiabilidad no se debe haber alterado en gran medida. Sí se encuentra un efecto para la Especificidad. Podría deberse a que al atenuarse las correlaciones (Trafimow, 2015) costará más identificar qué ejes deben mantenerse y cuáles eliminarse. Si se hipotetiza que existe un efecto de mediación entre 3 variables ($A \rightarrow B \rightarrow C$) y se correlacionan, se generará una relación espuria positiva entre $A \rightarrow C$. Esta debería eliminarse al parcializar las relaciones de las variables, pero probablemente al verse reducida la fuerza de la relación por el efecto de la fiabilidad de los indicadores, el procedimiento de parcialización no sea capaz de determinar si se trata de una relación espuria o de una relación real. No obstante, el índice sigue siendo robusto pues al comprobar las estimaciones concretas están son cercanas a 0 en la mayoría de las condiciones. Para los índices de Sensibilidad y Correlación entre red real y red estimada este efecto es aún más marcado, yendo su explicación en línea con el efecto conservador del procedimiento LASSO, aunque en este caso el efecto es mayor que el observado para el tamaño muestral. Probablemente se deba a la atenuación de las correlaciones por la inclusión de defectos en la fiabilidad de los indicadores, presentando este efecto además una interacción no despreciable con el tamaño de la muestra para el índice pertinente a la Correlación.

Analizando los índices de centralidad se observa que el índice más estable a través de las diferentes condiciones fue el de Fuerza, seguido por el de Cercanía, y siendo el índice relativo a la Intermediación el menos preciso, resultados que van en línea con lo mencionado en otros estudios (Bulteel et al., 2016). Aun así, todos los ítems de centralidad se vieron afectados de forma relevante por los diferentes factores manipulados.

Por un lado, analizando el efecto que ejerce el tamaño muestral sobre la estabilidad del índice Fuerza, observamos que este es bastante elevado. Dado que este índice depende de los pesos estimados (pues es la suma en valor absoluto de cada uno de los ejes que afectan a un nodo), y que estos se ven alterados por el procedimiento LASSO, dependiendo a su vez del tamaño de la muestra como se ha comentado anteriormente, cobra sentido el efecto encontrado. El elevado efecto de la fiabilidad va en línea a lo subrayado por la atenuación de las correlaciones, y por tanto de la variabilidad de los pesos estimados que aportan la información para el cómputo de la Fuerza.

Por otro lado, si prestamos atención al índice de Intermediación, parece ser el menos afectado por las diferentes variables independientes, no obstante, se debe tener en cuenta que el índice sufre de más problemas de estimación que el resto de los índices de centralidad, además el efecto encontrado sigue siendo elevado yendo estos resultados en línea con previas investigaciones (Bulteel et al., 2016). Al tratarse de un índice que informa de la cantidad de veces que un nodo actúa como mediador entre rutas de conexión de otros nodos, y, al eliminar ejes el procedimiento LASSO, parece razonable que al disminuir el tamaño muestral se reduzca el número de relaciones de mediación, dificultando así la correcta estimación del índice. El efecto encontrado en el factor fiabilidad para el propio índice, podría estar en línea con este mismo fenómeno.

Por último, para el índice de cercanía, se encuentra un efecto de interacción entre el tamaño muestral y la fiabilidad. Dado que este último índice se basa en las distancias, y que éstas son resultado directo del inverso de los pesos representados, es consecuente razonar que las distancias calculadas se verán muy alteradas (por el efecto de la fiabilidad) y que, al ser el procedimiento LASSO una técnica conservadora desaparecerán pesos en la red (reflejándose claramente en la baja precisión de la sensibilidad), por lo que estas distancias tampoco se tomarán en cuenta. Estos efectos dan como resultado el patrón obtenido en la representación de la interacción, dónde se puede apreciar como ante bajos efectos de los dos factores comentados, la recuperación es prácticamente nula, siendo necesario que, o bien se disponga de una muestra elevada (500) y una fiabilidad aceptable (0.7), o que en caso contrario dispongamos de unos ítems con fiabilidades altas (0.9) y una muestra aceptable (200).

Limitaciones del estudio

Las limitaciones del estudio son las propias de un estudio de simulación: sus conclusiones son sólo generalizables a las condiciones que fueron analizadas. A continuación, se comentarán algunas limitaciones adicionales, además de recomendaciones para futuras investigaciones. Para empezar, los análisis fueron realizados simulando variables continuas, a pesar de que habitualmente se trabaja con variables ordinales o categóricas, por lo que más investigación es necesaria para comprender mejor el efecto que supondría utilizar ese tipo de medidas. Otra limitación a destacar es el hecho de que todas las variables fueron simuladas con el mismo nivel de fiabilidad, por lo que podría resultar de especial interés comprobar si se modifican las conclusiones cuando los indicadores son heterogéneos en esta propiedad. Hay que destacar también el hecho de que no se consideraron saturaciones cruzadas, para facilitar la generación de matrices de covarianzas además de su interpretabilidad. Últimamente se ha venido trabajando con modelos que combinan la aproximación factorial con las redes psicológicas, lo que podría haber resultado de interés evaluar ya que ofrecen índices de ajuste, no obstante, su estimación es muy costosa dado que su desarrollo es muy temprano (Epskamp, Rhemtulla & Borsboom, 2017). También es de especial interés comprobar en qué grado el tamaño de la red y el tipo de conexiones afecta a la precisión en la recuperación de estas. Es remarcable el hecho de que el modelo base para simular matrices de covarianzas presenta únicamente ejes positivos. Aunque esta elección fue intencional, para abordar el problema en un escenario simple, sería interesante, de cara a futuras investigaciones, plantear modelos que representen relaciones con distinta polaridad. Por último, es importante señalar que, ante diferentes tamaños de los pesos simulados, se podrían ver alteradas algunas de las conclusiones extraídas. Trabajar con modelos que simulen estructuras de pesos grandes, medios, bajos y mixtos permitiría tener una perspectiva más generalizable.

Otras limitaciones tienen que ver con los criterios de precisión utilizados. En el presente estudio se tuvo que seleccionar un valor arbitrario (0.70) para decidir si la estabilidad en los resultados obtenidos para los diferentes índices era aceptable o no. En este sentido, no hay consenso al respecto y otros puntos de corte podrían ser más apropiados. Cabe tener presente el hecho de que en el presente estudio se ha estado comparando la red psicológica teórica con las redes psicológicas simuladas, pero no con un modelo factorial tradicional, por lo que puede que si se tomara como real a este último

y no a la red psicológica, tras calcular la cantidad de ejes espurios resultantes de estimar las redes psicológicas por comparación al modelo factorial los resultados hubiesen sido dispares a los obtenidos en este caso (probablemente se hubiese detectado un mayor número de ejes espurios), ya que si se toma de referencia a la red poblacional como el modelo teórico, al ya estar representadas las relaciones parcializadas en esta red, no permite distinguir que ejes serían espurios en el modelo factorial pues se toman todos como verdaderos.

Implicaciones prácticas

Los resultados de este estudio sugieren que existen diferencias importantes en la estabilidad de las redes psicológicas dependiendo de la fiabilidad de los indicadores y del tamaño muestral. A raíz de estos resultados, creemos que es importante que a los investigadores aplicados sean conocedores de estos efectos, para tratar de evitar posibles interpretaciones erradas, seleccionar el tamaño de la muestra apropiado para su investigación, y comprobar de antemano la fiabilidad de los indicadores para conocer en qué medida la red resultante puede fluctuar con respecto a la realidad.

Dada la naturaleza de la técnica LASSO, los índices de especificidad y ejes espurios negativos mostraron ser robustos a través de las diferentes condiciones, por lo que no debería ser un problema mayor el confiar en la veracidad de la presencia de los ejes representados, si se asume que el modelo de redes psicológicas es el verdadero para los ítems utilizados. No sucede lo mismo con el resto de los indicadores de la red psicológica, por lo que se debería ser especialmente precavido con considerar que no existen más relaciones más allá de las que la red representa, así como con la magnitud de los pesos obtenidos. Por otro lado, el índice de centralidad más estable fue el de fuerza, siendo especialmente robusto frente al factor de fiabilidad, por lo que puede ser una buena medida acerca de lo central que puede ser un indicador, siendo siempre cauto acerca de la utilidad interpretativa de este tipo de índices. No es el caso para el resto de los índices de centralidad, por lo que se debería prestar especial atención y precaución ante estos.

Finalmente, se recomienda que, dado que ningún método parece funcionar excepcionalmente bien con muestras inferiores a 200 sujetos, a no ser que se disponga de ítems con fiabilidades perfectas, lo que es irreal en la práctica, se utilicen muestras de 200 sujetos o más e indicadores con fiabilidades cercanas a 0.7 (lo que suele ser habitual), con el fin de cerciorarse de la precisión de los resultados obtenidos.

8. Referencias bibliográficas

- Armour, C., Fried, E., Deserno, M., Tsai, J., & Pietrzak, R. (2017). A network analysis of DSM-5 posttraumatic stress disorder symptoms and correlates in U.S. military veterans. *Journal Of Anxiety Disorders*, 45, 49-59.
- Borgatti, S. P. (2005). Centrality and network flow. *Social Networks*, 27(1), 55-71.
- Borsboom, D., & Cramer, A. (2013). Newtwork analysis: an integrative approach to the structure of psychopathology. *Annual review of clinical psychology*, 9(1), 91-121.
- Borsboom, D., Mellenbergh, G., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, 110(2), 203-219.
- Boschloo, L., van Borkulo, C., Borsboom, D., & Schoevers, R. (2016). A prospective study on how symptoms in a network predict the onset of depression. *Psychotherapy and Psychosomatics*, 85(3), 183–184.
- Brandes, U. (2001). A faster algorithm for betweenness centrality. *The Journal of Mathematical Sociology*, 25(2), 163-177.
- Bringmann, L., Elmer, T., Epskamp, S., Krause, R., Schoch, D., Wichers, M., Wigman, J., & Snippe, E. (2018). What do centrality measures measure in psychological networks? *Unpublished*.
- Bulteel, K., Tuerlinckx, F., Brose, A., & Ceulemans, E. (2016). Using raw var regression coefficients to build networks can be misleading. *Multivariate Behavioral Research*, 51(2), 330–344.
- Chen, J., & Chen, Z. (2008). Extended bayesian information criteria for model selection with large model spaces. *Biometrika*, 95(3), 759–771.
- Cramer, A., van der Sluis, S., Noordhof, A., Wichers, M., Geschwind, N., & Aggen, S. et al. (2012). Measurable Like Temperature or Mereological Like Flocking? On the Nature of Personality Traits. *European Journal of Personality*, 26(4), 451-459.
- Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Möttus, R., Waldorp, L., & Cramer, A. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, 54, 13-29.

- de Ron, J., Fried, E., & Epskamp, S. (2019). Psychological Networks in Clinical Populations: A tutorial on the consequences of Berkson's Bias. *Unpublished*.
- Epskamp, S. (2017). *Network psychometrics* (tesis doctoral). Universidad de Ámsterdam, Ámsterdam, Países bajos.
- Epskamp, S., Borsboom, D., & Fried, E. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, 50(1), 195-212.
- Epskamp, S., Cramer, A., Waldorp, L., Schmittmann, V., & Borsboom, D. (2012). qgraph: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software*, 48(4), 1-18.
- Epskamp, S., & Fried, E. (2018). A tutorial on regularized partial correlation networks. *Psychological Methods*, 23(4), 617-634.
- Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized Network Psychometrics: Combining Network and Latent Variable Models. *Psychometrika*, 82(4), 904-927.
- Epskamp, S., Waldorp, L., Möttus, R., & Borsboom, D. (2018). The Gaussian Graphical Model in Cross-Sectional and Time-Series Data. *Multivariate Behavioral Research*, 53(4), 453-480.
- Foygel, R., & Drton, M. (2015). High-dimensional Ising model selection with bayesian information criteria. *Electronic Journal of Statistics*, 9 (1), 567–607.
- Jayawickreme, N., Mootoo, C., Fountain, C., Rasmussen, A., Jayawickreme, E., & Bertuccio, R. (2017). Post-conflict struggles as networks of problems: A network analysis of trauma, daily stressors and psychological distress among Sri Lankan war survivors. *Social Science & Medicine*, 190, 119-132.
- Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: principles and techniques*. Cambridge, MA, USA: MIT press.
- Kossakowski, J., Epskamp, S., Kieffer, J., van Borkulo, C., Rhemtulla, M., & Borsboom, D. (2016). The application of a network approach to Health-Related Quality of Life (HRQoL): introducing a new method for assessing HRQoL in healthy adults and cancer patients. *Quality of Life Research*, 25(4), 781-792.

- Meinshausen, N., & Bühlmann, P. (2006). High-dimensional graphs and variable selection with the Lasso. *The Annals Of Statistics*, 34(3), 1436-1462.
- McCrae, R., & Costa, P. (2008). Empirical and theoretical status of the five-factor model of personality traits. In *The SAGE Handbook of Personality Theory and Assessment: Volume 1 - Personality Theories and Models* (pp. 273-294). SAGE Publications Inc.
- McNally, R., Robinaugh, D., Wu, G., Wang, L., Deserno, M., & Borsboom, D. (2015). Mental disorders as causal systems: A network approach to posttraumatic stress disorder. *Clinical Psychological Science*, 3(6), 836-849.
- Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press.
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Robinaugh, D. J., Millner, A. J., & McNally, R. L. (2016). Identifying highly influential nodes in the complicated grief network. *Journal of Abnormal Psychology*, 125, 747-757.
- Schmittmann, V., Cramer, A., Waldorp, L., Epskamp, S., Kievit, R., & Borsboom, D. (2013). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas In Psychology*, 31(1), 43-53.
- Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal Of The Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.
- van Borkulo, C., Borsboom, D., Epskamp, S., Blanken, T., Boschloo, L., Schoevers, R., & Waldorp, L. (2014). A new method for constructing networks from binary data. *Scientific reports*, 4 (5918), 1–10.

9. Anexos

```
library(matlib)
library(BBmisc)
modelo <- function(LambdaX, LambdaY, BetaMatrix, GammaMatrix) {
  sink(tempfile())
  on.exit(sink())
  LAMBDAY <- LambdaY
  LAMBDAX <- LambdaX
  LAMBDAYVECTOR <- as.vector(LAMBDAY)
  LAMBDAYVECTOR <- LAMBDAYVECTOR[LAMBDAYVECTOR!=0]
  nEndo <- length(LAMBDAYVECTOR);nEndo
  LAMBDAXVECTOR <- as.vector(LAMBDAX)
  LAMBDAXVECTOR <- LAMBDAXVECTOR[LAMBDAXVECTOR!=0]
  nExo <- length(LAMBDAXVECTOR);nExo
  GAMMA <- GammaMatrix
  BETA <- BetaMatrix
  PHI <- diag(ncol(GAMMA))
  PSI <- diag(ncol(BETA))
  IDENTIDAD <- diag(nrow(BETA))
  BETAINV <- solve(IDENTIDAD - BETA)
  PESOS <- BETAINV %*% GAMMA
  BETAINV2<-BETAINV^2
  PSI_no_escalada1 <- PESOS %*% PHI %*% t(PESOS)
  PSI_no_escalada <- diag(diag(nrow(PSI_no_escalada1)) - PSI_no_escalada1)
  A <- BETAINV2
  b <- PSI_no_escalada
  options(warn=-1)
  if (is.error(try(Solve(A, b, fractions = FALSE),silent=T)) == T) {
    result <- showEqn(A, b)
    result <-unlist(lapply(strsplit(result, " "),as.numeric))
    result <-result[!is.na(result)]
  }
```

```

diag(Psi)<-result}
else {
  result<-Solve(A, b, fractions = FALSE)
  result<-unlist(lapply(strsplit(result, " "),as.numeric))
  result<-result[!is.na(result)]
  diag(Psi)<-result}
options(warn=0)
THETAEPS <- diag(nEndo)
THETADELTA <- diag(nExo)
for (i in (1:length(LAMBDAYVECTOR))) {
  THETAEPS[i,i] <- 1-LAMBDAYVECTOR[i]^2 }
for (i in (1:length(LAMBDAXVECTOR))) {
  THETADELTA[i,i] <- 1-LAMBDAXVECTOR[i]^2 }
THETADELTAEPS <- matrix(0,nrow=nExo,ncol=nEndo)
THETAEPSDELTA <- t(THETADELTAEPS)
SIGMAX <- LAMBDAX %*% PHI %*% t(LAMBDAX) + THETADELTA
SIGMAY <- LAMBDAY %*% BETAINV %*%
(GAMMA%*%PHI%*%t(GAMMA) + PSI ) %*% t(BETAINV) %*% t(LAMBDAY)
+ THETAEPS
SIGMAXY <- LAMBDAX %*% PHI %*% t(GAMMA) %*% t(BETAINV) %*%
t(LAMBDAY) + THETADELTAEPS
SIGMAYX <- LAMBDAY %*% BETAINV %*% GAMMA %*% PHI %*%
t(LAMBDAX) + THETAEPSDELTA
X <- cbind(SIGMAX, SIGMAXY)
Y <- cbind(SIGMAYX, SIGMAY)
COVPOB <- rbind(X,Y)
return(COVPOB)
invisible(force()) }

```