


An integrated multimodal model of alcohol use disorder generated by data-driven causal discovery analysis

Eric Rawls ^{1✉}, Erich Kummerfeld² & Anna Zilverstand¹

Alcohol use disorder (AUD) has high prevalence and adverse societal impacts, but our understanding of the factors driving AUD is hampered by a lack of studies that describe the complex neurobehavioral mechanisms driving AUD. We analyzed causal pathways to AUD severity using Causal Discovery Analysis (CDA) with data from the Human Connectome Project (HCP; $n = 926$ [54% female], 22% AUD [37% female]). We applied exploratory factor analysis to parse the wide HCP phenotypic space (100 measures) into 18 underlying domains, and we assessed functional connectivity within 12 resting-state brain networks. We then employed data-driven CDA to generate a causal model relating phenotypic factors, fMRI network connectivity, and AUD symptom severity, which highlighted a limited set of causes of AUD. The model proposed a hierarchy with causal influence propagating from brain connectivity to cognition (fluid/crystalized cognition, language/math ability, & working memory) to social (agreeableness/social support) to affective/psychiatric function (negative affect, low conscientiousness/attention, externalizing symptoms) and ultimately AUD severity. Our data-driven model confirmed hypothesized influences of cognitive and affective factors on AUD, while underscoring that addiction models need to be expanded to highlight the importance of social factors, amongst others.

¹Department of Psychiatry and Behavioral Sciences, University of Minnesota, Minneapolis, MN, USA. ²Institute for Health Informatics, University of Minnesota, Minneapolis, MN, USA. ✉email: erawls89@gmail.com

Lifetime incidence of alcohol use disorder (AUD) is as high as 29–30%^{1,2}, with alcohol use being a leading cause of death [3 million worldwide in 2016 or 5.3% of all deaths worldwide³]. Success rates for quitting drinking in AUD are low [30–40%^{4,5}], which has been attributed to the multi-causality of the mechanisms underlying AUD and the need for more targeted treatments⁶. However, the development of targeted interventions for AUD is hampered by a lack of studies investigating multifactorial mechanisms driving AUD.

Early theories of addiction maintenance proposed single key mechanisms, such as allostasis⁷, or hedonic signaling⁸. These early theories have given way to multifactorial models of addiction, such as the “three-stage cycle” model⁹, which proposes that negative affect, incentive salience, and executive function are functional domains involved in addiction. There is a great deal of empirical support for the involvement of these three domains. The three domains have been mapped onto corresponding personality profiles that confer addiction risk¹⁰ and have been used to develop a set of proposed neuroclinical assessment tools^{11,12} that were successfully applied to AUD¹³. However, a three-domain model is far from encompassing the entire phenotypic space that contributes to AUD. The NIMH RDoC^{14,15} proposed 23 functional domains underlying psychopathology, recognizing a need for multivariate models that incorporate increasing knowledge of the many functional domains contributing to psychiatric dysfunction. A recent consensus paper on a multivariate assessment approach for addiction identified another seven “addiction-specific” domains in addition to the RDoC domains¹⁶. Critically, in all of these approaches, the proposed functional domains were identified by expert consensus and therefore might not exactly match the true underlying domains that exist in the data. For example, in an exploratory analysis of a large public dataset, Van Dam et al.¹⁷ derived seven phenotypic factors that only partially mapped onto RDoC domains, but predicted psychiatric distress. A more recent addiction theory¹⁸ identified ten domains contributing to maladaptive decision-making in addiction. A systematic review of neuroimaging studies in addicted populations implicated the involvement of at least six different neurobiological mechanisms in AUD¹⁹. These recent developments underscore the need for data-driven, multivariate analysis methods capable of fully examining and describing the large phenotypic space underlying addiction, if we are to understand the central question of how multi-causal factors underlie the maintenance and escalation of alcohol use.

In the current study, we leveraged the deep behavioral and psychiatric phenotyping²⁰ and high-resolution neuroimaging data²¹ from the Human Connectome Project (HCP)²². Using data from nearly 1000 participants, we first derived a set of data-driven domains underlying the full range of phenotypic functioning measured in the HCP dataset. We extracted whole-brain connectivity metrics from 12 data-derived resting-state functional magnetic resonance imaging (fMRI) networks²³ to measure individual neurobiological differences. To examine the relationships between fMRI network connectivity, phenotypic domains, and AUD symptom severity, we applied Causal Discovery Analysis (CDA), a class of machine learning techniques that learns causal models from input data. These methods search the enormous set of possible structural models and return a graph representing estimated causal relationships in the data. The particular method we applied, Greedy Fast Causal Inference (GFCI)²⁴, uses conditional dependence relations to discover when unmeasured variables confound the relationships between measured variables, making this method particularly powerful for real-world data sets that cannot possibly capture every variable of interest. By (1) deriving data-driven domains encompassing the whole phenotypic space measured in HCP, (2) extracting whole-

brain network connectivity profiles, and (3) applying CDA to the resulting phenotypic and neurobiological domains, we generated an integrated, multimodal causal model of neurobehavioral factors contributing to AUD symptom severity.

Results

Exploratory factor analysis: decomposing the phenotypic space measured in the HCP. To reduce the phenotypic space measured in the HCP to a set of underlying domains, we conducted an exploratory factor analysis (EFA) in the entire HCP sample that had complete phenotypic data ($n = 933$, 53.5% females). Based on the results of Monte Carlo simulation we extracted 18 factors ($p < 0.05$) from the 100-variable phenotypic space measured in the HCP dataset, which collectively accounted for 47% of common variance. Results of the Monte Carlo permutation test for eigenvalue significance are presented in Supplementary Fig. S1 (eigenvalue significance), Supplementary Data 1 (observed, random, and resampled eigenvalues), and Supplementary Table S1 (percent variance explained per factor, eigenvalues, and cumulative variance). EFA model fit indices indicated good factor separation (RMSEA = 0.03, Tucker–Lewis Index = 0.86).

Factors, in order of common variance accounted for, were associated with: (1) Somaticism (high DSM/ASR somaticism, high DSM depression, low PSQI sleep quality), (2) Fluid Cognition (high Raven’s progressive matrices performance), (3) Internalizing (high DSM/ASR anxiety, high DSM depression, high NEO-FFI neuroticism), (4) Gambling Task Reaction Time (slow gambling task reaction time), (5) Conscientiousness/Attention (low DSM attention deficit hyperactivity disorder, low ASR attention problems, and high NEO-FFI conscientiousness), (6) Visuospatial Processing (high Penn short line orientation task performance), (7) Social Support (high NIH toolbox friendship, low loneliness, low perceived rejection and perceived hostility, high emotional and instrumental support), (8) Processing Speed (high NIH Toolbox Flanker Total Score, fast fMRI emotion task RT), (9) Externalizing (high ASR aggression and rule-breaking, high DSM antisocial, high NIH toolbox aggression), (10) Social Withdrawal (high ASR withdrawal, high DSM avoidance, low NEO-FFI extraversion), (11) Language Task Performance (high fMRI language task story average difficulty, and high math problem accuracy), (12) Relational Task Reaction Time (slow fMRI relational task reaction time [RT]), (13) Delay Discounting (high delay discounting AUC for \$200 and \$40k), (14) Working memory (fMRI N-Back task fast reaction time [RT] and high accuracy), (15) Negative Affect (high NIH toolbox anger, fear, sadness and stress), (16) Crystallized IQ (high NIH toolbox English reading and picture vocabulary, high education, and high NEO-FFI openness), (17) Positive Affect (high NIH toolbox life satisfaction, positive affect, and meaning and purpose, and NEO-FFI extraversion), and (18) Agreeableness (low aggression and high NEO-FFI agreeableness). Factor loadings for each item are available in Supplementary Data 2 (raw variables) and Supplementary Data 3 (standardized variables), and the factor correlation structure is available in Supplementary Data 4.

Causal discovery of the neurobehavioral underpinnings of AUD

Model fit and quality metrics. To build a multidomain model of the causal neurobehavioral underpinnings of AUD, we used a recently developed causal discovery machine learning algorithm called GFCI²⁴, which is particularly powerful for use with observational data because it has the capacity to determine when causal relationships are impacted by unobserved confounding variables. We submitted the 18 factors extracted by EFA,

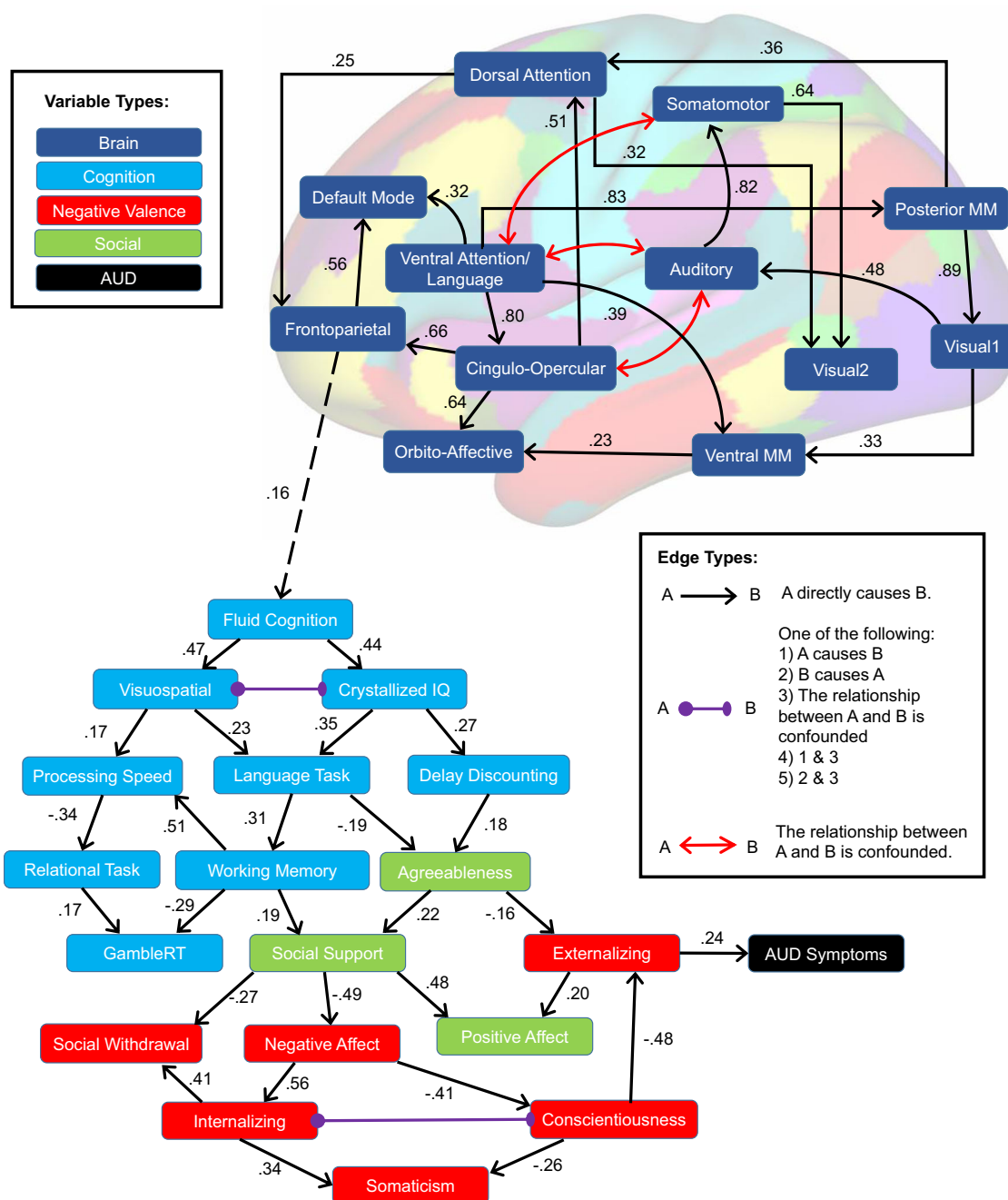


Fig. 1 Causal discovery of the neurobehavioral underpinnings of alcohol use disorder. Causal discovery analysis of the neurobehavioral determinants of AUD symptom severity in the HCP dataset was done using Greedy Fast Causal Inference (GFCI). GFCI returns a partial ancestral graph (PAG) depicting causal relationships between a set of variables, while assessing for unmeasured third variables in relationships (confounders). Standardized edge weights recovered via structural equation modeling (SEM) are displayed in text next to each edge in the graph. The overall SEM fit was good, RMSEA = 0.06, Tucker-Lewis Index = 0.91.

within-network brain resting-state fMRI connectivity from 12 previously defined resting-state networks²³, and a measure of AUD severity (AUD symptom count) to a unified causal modeling framework using GFCI. The output of GFCI is presented in Fig. 1.

We extracted measures of model fit and effect sizes for causal relationships using structural equation modeling (SEM). The SEM fit to this model indicated a good fit, RMSEA = 0.06, Tucker-Lewis Index = 0.91, and every edge in the causal model was significant in the corresponding SEM ($p < 0.001$). Recovered edge weights from SEM were presented overlaid on the GFCI graph. Stability testing (jackknife analysis) demonstrated the high

stability of the model (Supplementary Fig. S2 and Supplementary Data 5). Furthermore, the model was highly replicable when using a separately defined brain network parcellation (RMSEA = 0.06, Tucker-Lewis Index = 0.87, every edge $p < 0.001$) (Supplementary Fig. S3). As such, the results described here do not appear to depend on the specific parcellation used to measure fMRI resting-state network connectivity.

Causal model results interpretation. First, we found that brain network connectivity measures and phenotypic factors largely separated into two interconnected separate clusters. The brain

network subgraph indicated several salient points. We found that high connectivity within the ventral attention/language network during rest caused high connectivity within default mode (DMN), cingulo-opercular, and multimodal sensory association networks—networks that play a central role in self-reflective brain processes. We also found effects indicating causal influences of cingulo-opercular connectivity on attentional processing (cingulo-opercular → dorsal attention). Finally, we found converging causal influences onto the frontoparietal network (cingulo-opercular/posterior multimodal → frontoparietal, dorsal attention → frontoparietal).

Brain connectivity intersected with behavioral phenotypic variables in a link between high frontoparietal connectivity and high Fluid Cognition. From this point, causal influences propagated from Fluid Cognition to Visuospatial Processing and Crystallized IQ, replicating a well-studied effect that individuals high in fluid cognitive ability will also be high in crystallized intelligence^{25,26}. From there causal influences proceeded to more specific cognitive measures, including Working Memory, Language Task Performance (verbal and math ability), and Delay Discounting. We found a direct link between Crystallized IQ and Delay Discounting, such that individuals higher in Crystallized IQ also exhibited lower (less impulsive) discounting rates. These cognitive measures were then in turn causally linked to affective, social, and psychiatric factors. High Language Task (verbal and math) performance and less impulsive Delay Discounting caused Agreeableness. Low Working Memory performance, and low Agreeableness caused lowered Social Support, and decreased Social Support contributed to increased Negative Affect, increased Social Withdrawal, and decreased Positive Affect. High Negative Affect in turn contributed to higher Internalizing symptoms, and lower Conscientiousness/Attention. Low Conscientiousness/Attention, and low Agreeableness caused high Externalizing psychopathology, while high Externalizing psychopathology directly caused increased AUD symptom severity.

Previous hypotheses have particularly focused on the influences of negative affect, incentive salience, and executive function in AUD. Our results support a causal role for cognitive and affective influences on AUD, while expanding our understanding of the complex multifactorial space contributing to AUD.

Discussion

Early addiction models posited that addiction was due to single key mechanisms^{7,8}, but modern addiction models have begun to emphasize the multifactorial mechanisms underlying addiction^{9,11,16,18,19}. In this study, we used a data-driven approach to characterize phenotypic domains in a large community sample, and examined whole-brain network connectivity at a large scale using a data-driven network analysis and parcellation²³. We then modeled large-scale brain and behavioral influences on AUD symptom severity using CDA. Our results shed light on the relationship between brain network connectivity and phenotypic domains in general, as well as providing specific information on how brain and behavioral factors contribute to the severity of AUD, and which could be targeted in treatment.

Our analysis significantly expands the multifactorial space of current addiction models. Our factor analysis uncovered a variety of factors that map relatively well onto domains elaborated in RDoC (Table 1). For example, we found factors that mapped well onto aspects of the RDoC Cognitive Systems domain, the RDoC Negative Valence Systems domain, and the RDoC Social domain. To assist in interpreting the large-scale domains that our data-driven factors mapped onto, we grouped factors based on their correlations. Interestingly, we found that the Conscientiousness/

Attention and the Social Withdrawal factors correlated with other factors in the Negative Valence Systems domain, rather than the RDoC-assigned grouping of these factors (Cognitive Systems: Attention and Social Systems: Affiliation & Attachment, respectively). A previous review found that inattention and anxiety are tightly linked²⁷, but our results provide evidence of the direct link between inattention and negative affect. We also found that the Delay Discounting factor, while considered part of the RDoC Positive Valence System: Reward Valuation subconstruct²⁸, correlated instead with Cognitive Systems factors, suggesting delay discounting is more related to Cognitive Control/Impulsivity domains than to Reward Valuation^{29–31}.

Here we summarize several key points from our mapping of causal influences on AUD symptom severity onto the RDoC framework. First, our analysis uncovered strong evidence for the direct causal effect of the Negative Valence domain in AUD. This specifically included a causal influence of the Negative Affect factor on AUD, mediated through Conscientiousness/Attention (which correlated with other Negative Valence Systems factors). While many neurobiological models of addiction agree on the importance of negative affect in AUD^{9,19}, this is not unanimously agreed upon by experts in the addiction sciences¹⁶. The presented empirical data hence provides important empirical evidence implicating the broader Negative Affect Domain (as defined in RDoC) as an important treatment target in AUD.

Second, our analysis uncovered strong evidence for a mediating/buffering role of the Social Systems domain in AUD. Low Social Support and low Agreeableness were indirect causes of AUD severity and fully mediated the effect of cognition on the negative valence domain, providing strong empirical evidence that addiction models should incorporate measures of social function^{32,33}. Epidemiological research has established a solid link between social affiliation and drug addiction³⁴, and increased social affiliation is associated with decreased risk of relapse in drug users who are seeking treatment³⁵. Despite the considerable evidence research has uncovered for the importance of social affiliation as a protective factor in addiction³⁶, current neurobiological models of addiction generally fail to consider social factors³² and their close relationship to cognitive/affective factors.

We generally found weak evidence for the involvement of the Positive Valence Systems domain in this analysis, although this is likely due to a limitation of the dataset employed. Positive Valence Systems subdomains, including reward-based domains that are particularly important in addiction¹⁶ are relatively neglected in the HCP dataset³⁷. We did find a causal influence of Delay Discounting on AUD severity, but as noted previously the Delay Discounting factor correlated with other Cognitive factors, suggesting that Delay Discounting (as measured in the HCP study) might be related to Cognitive Control/Impulsivity more so than Reward Valuation^{29–31}.

Finally, our analysis provides strong evidence that prefrontal cortex (PFC) brain networks, and associated cognitive factors, are situated at the top of the causal hierarchy of influences on AUD severity. The role of PFC and associated high-level cognitive factors in addiction is often referenced and is a part of major current theories of addiction^{19,38}, but our results are among the first to empirically demonstrate this hierarchical influence on AUD. Importantly, our causal model indicates that cognitive influences on AUD severity may extend far beyond the traditional consensus that inhibitory control is the most important cognitive influence on addiction^{16,38,39}, as we also highlighted influences of fluid and Crystallized IQ, working memory, and language/math ability on AUD.

Our results also shed light on theorized externalizing and internalizing pathways to AUD. Previous research has shown that externalizing symptomatology predicts AUD in young adult

Table 1 Discovered factors (EFA) using 100 phenotypic measures.

Domains	Factors grouped according to factor correlations ^a	RDoC subdomains ^b
Negative Valence	Externalizing (high ASR aggression and rule-breaking, high DSM antisocial, high NIH aggression)	Frustrative Non-reward
	Conscientiousness/Attention (low DSM ADHD, low ASR attention problems, and high NEO-FFI conscientiousness)	<i>Attention^c</i>
	Somaticism (high DSM/ASR somaticism, high DSM depression, low PSQI sleep quality)	Sustained threat
	Internalizing (high DSM/ASR anxiety, high DSM depression, high NEO-FFI neuroticism)	Potential threat, Sustained threat
	Negative Affect (high NIH anger, fear, sadness and stress)	Acute threat, loss, sustained threat
	Social Withdrawal (high ASR withdrawal, high DSM avoidance, low NEO-FFI extraversion)	<i>Affiliation & Attachment^c</i>
	Visuospatial Processing (high Penn short line orientation task performance)	Visual
	Delay Discounting (high delay discounting AUC for \$200 and \$40k)	<i>Reward valuation^c</i>
	Language Task Performance (high fMRI language task story average difficulty, and high math problem accuracy)	Language behavior
	Crystalized IQ (high NIH English reading and picture vocabulary, high education, and high NEO-FFI openness)	Declarative memory
Cognition	Fluid Cognition (high Raven's progressive matrices performance)	Working memory
	Gambling Task Reaction Time (slow gambling task reaction time)	
	Working memory (fMRI N-Back task fast reaction time (RT) and high accuracy, fast Penn word memory RT)	Declarative/working memory
	Processing speed (high NIH flanker total score, fast fMRI emotion task RT)	
	Relational Task Reaction Time (slow fMRI relational task RT)	
	Social Support (high NIH friendship, low loneliness, low perceived rejection and perceived hostility, high emotional and instrumental support)	Affiliation & Attachment
	Positive Affect (high NIH life satisfaction, positive affect, and meaning and purpose, and NEO-FFI extraversion)	Perception and Self
	Agreeableness (low aggression and high NEO-FFI agreeableness)	Affiliation & Attachment

^aFactors are grouped according to correlations between the factors.

^bThe right column indicates the RDoC domain each factor most closely approximated.

^cFactors whose correlation structure did not match the RDoC domain assignment for that factor ($n = 3$) are displayed in italics.

samples⁴⁰. Our data-driven causal model revealed that externalizing fully mediated the impact of all other (measured) causes of AUD; that is, AUD is unrelated to other phenotypic or brain network factors when externalizing is controlled for. Note that our externalizing factor consisted of ASR rule-breaking, aggression, and antisocial scales, and NIH toolbox aggression. The ASR rule-breaking scale contains an item that assesses whether a subject “gets drunk,” but this does not appear to have influenced the current analysis. First, this is only one out of 14 items on the rule-breaking scale (and 40 items in total that contributed to the Externalizing factor). Second, all four scales that formed the Externalizing factor independently correlated with AUD symptom severity (all $p < 0.001$), such that each individual aspect of Externalizing appears to be related to AUD severity. Overall the model hence supports that externalizing symptoms in general mediate the causal influence of other factors on AUD.

Research has also focused on the high coincidence of AUD and internalizing disorders^{41–43}. Previous causal modeling research found a causal path from internalizing disorder to drinking behavior in AUD (mediated through drinking-to-cope)⁴⁴. The current model contains a confounded relationship between Internalizing and Conscientiousness/Attention, indicating an inability of the model to determine the relationship between these two factors, possibly due to underlying constructs (e.g. drinking motives) that were not captured in this dataset. Therefore, it remains to be further described by future research if negative affect is a common underlying cause of internalizing and AUD symptoms⁴¹, or if there is an independent causal influence of internalizing psychopathology on AUD symptoms. It is possible that this relationship could be better examined through a longitudinal study, as pathodevelopmental perspectives on AUD have proposed that early stages of addiction are characterized by low levels of internalizing, but later stages of addiction are characterized by increasing levels of internalizing⁴¹.

Interestingly, our results suggest an important role of fluid cognition in AUD, but this is seldom addressed in current neurobiological models of addiction. However, executive function is central to neurobiological models of addiction^{9,19,38,39}. Fluid cognition (here, measured through performance on an abbreviated form of Raven's progressive matrices)⁴⁵, or a person's ability to reason and think abstractly and flexibly, has an intuitive relationship with the concept of executive function. Authors have often considered working memory to be either indicative of executive function⁴⁶, or of fluid cognition⁴⁷, and executive function and fluid cognition are similarly impacted by brain lesions⁴⁸. Our results generate the hypothesis that fluid cognition and Crystalized IQ, including problem-solving and abstract reasoning, lie at the beginning of a causal hierarchy eventually influencing AUD severity. This fits previous empirical evidence demonstrating that AUD is associated with deficits attributed to various fluid cognitive abilities or executive function such as working memory^{49,50} and planning and goal maintenance⁵¹, but expands on this by indicating that these factors have a causal influence on AUD symptom severity. Previous research has also demonstrated that high-level cognition predicts initiation of substance use in adolescence⁵², lifetime drug use and abuse⁵³, and addiction treatment outcomes⁵⁴. Our model thus adds to the growing body of empirical evidence that proposes a causal role of cognition as a primary resilience factor and potential treatment target in AUD.

Our results also provide critical empirical evidence for the role of hierarchical brain network interactions in AUD. We found a direct brain-phenotype link from frontoparietal (executive) network connectivity to fluid cognition, corroborating previous evidence of this link in healthy populations^{55–59}. In addition, frontoparietal network dysfunction has been implicated in impaired inhibitory control^{19,60,61} and self-regulation^{27,61}. Individuals with AUD show decreased recruitment of frontoparietal

network during social–emotional processing^{62,63}, decision-making⁶⁴, and cognitive control⁶⁵, as well as decreased frontoparietal connectivity during rest⁶⁶. Frontoparietal disengagement during social–emotional processing predicts relapse in AUD⁶⁷, and decreased frontoparietal activation during inhibitory control predicts later drinking in adolescents^{68,69}. Our results indicate that causal influences of frontoparietal network connectivity on AUD are mediated through deficits in overall cognitive ability and its downstream effects on the broader Negative Affect domain. A recent systematic review⁷⁰ also showed that targeting dlPFC (part of the frontoparietal network)²³ can improve cognitive deficits in addiction, including executive functions. This is particularly relevant, as our data-driven model also indicates that neuromodulation of frontoparietal network could improve executive functioning, with downstream effects on AUD severity.

We also found direct effects of cingulo-opercular and dorsal attention network connectivity onto the frontoparietal network. The cingulo-opercular network reacts to salient stimuli regardless of positive/negative valence^{71,72}, while the dorsal attention network supports the external focusing of attention⁷³ and encodes top-down control and working memory load⁷⁴. Individuals with AUD also show decreased cingulo-opercular activation during social–emotional processing^{62,63}, cognitive control^{65,75,76}, and decision-making⁶⁴. Our results suggest that dysfunctional connectivity in salience and attentional networks can contribute to cognitive dysfunction in AUD, with these effects being mediated through executive network connectivity. The presented causal model hence provides direct evidence for brain-directed treatment approaches targeted at the frontoparietal network, such as cognition-enhancing therapy^{61,77,78}, pharmacological interventions (cognitive enhancers)⁷⁷ or neuromodulation treatment (e.g., by external devices)^{70,79} or neurofeedback interventions²⁷.

A key result generated from our CDA is the role of ventral attention/language network connectivity as a central “hub” in the brain during resting-state. The ventral attention/language network has been characterized in many brain network partitions^{80–82}, and while it historically has been implicated in language processing²³, more recent research has described a role in orienting attention to unexpected events^{83,84}. Ventral attention/language network connectivity caused cingulo-opercular network connectivity directly, and indirectly caused dorsal attention network connectivity (mediated through cingulo-opercular connectivity, and posterior multimodal association network connectivity). Therefore, ventral attention/language network connectivity exerts influences on frontoparietal network connectivity through multiple different pathways, and might have long-range impacts on cognition and eventually on AUD severity. The potential involvement of ventral attention/language network in AUD appears to have been scarcely investigated. This network encompasses vlPFC regions that are close to left dlPFC regions that are often targeted in neuromodulation interventions for addiction⁷⁹, and therefore neuromodulation targeted at left dlPFC might also stimulate ventral attention/language networks. Left vlPFC regions are also implicated in cognitive interventions for addiction⁶¹. Future analysis should examine the relationship between brain ventral attention/language networks and cognitive dysfunction in AUD, and the implications for treatment.

The analysis method used in the current manuscript is not free of limitations, and other limitations are also imposed by the nature of the dataset we used. Notably, the fact that we did not find any cycles (i.e. variable *X* causes variable *Y*, and variable *Y* causes variable *X*) in the current data does not mean that they do not exist. The causal discovery algorithm used in this analysis cannot discover recurring cycles in cross-sectional data, but is capable of discovering recurrent cycles when more than one time

point is measured and the cycles unfold over time. Future analysis should incorporate longitudinal data to specifically test the possibility that recurrent cycles might contribute to AUD⁹. This limitation extends to the brain network subgraph we recovered as well; the causal discovery algorithm we used cannot recover bidirectional relationships in cross-sectional data, so some brain network links that are actually bidirectional processing streams may instead be represented as the predominant causal relationship between two networks. Finally, the CDA algorithm also uses a penalized likelihood score, therefore potentially missing weak causal links present in the data; however, this practice also serves to increase the confidence in the causal relationships the algorithm does find.

An important limitation of the dataset is the extremely limited assessment of the Positive Valence domain in the HCP dataset. Current perspectives in addiction emphasize the role of Positive Valence domains^{9,16,19,38,39}, but the HCP dataset does not contain many measures in this domain. The data did contain a measure of Delay Discounting, which had causal influences on AUD severity, but this factor appeared to be grouped with other cognitive factors and could not be interpreted as a unique measure of Reward Valuation. The HCP dataset also contained a gambling choice fMRI task, but this task did not provide a phenotypic measure of incentive salience processing. Future analysis will need to carefully measure Positive Valence domains, in addition to the domains measured in the HCP dataset, to determine where these domains fit in an overall causal model of AUD. Another limitation of the dataset is that the cross-sectional design employed by the HCP is also unable to assess certain predictions of pathodevelopmental perspectives on addiction, such as the possibility that different causal factors are involved in early and late stages of AUD⁴¹.

To conclude, this study is the first to conduct a machine learning search for causal influences of AUD symptoms over a wide phenotypic and neurobiological space. We found phenotypic factors related to several RDoC domains, and confirmed hypothesized influences of a Negative Valence (Negative affect > Conscientiousness/Inattention > Externalizing > AUD symptom severity) and Cognitive Systems (Fluid Cognition > Crystallized IQ > Working Memory/Language/Math > Social/Affective/Psychiatric factors > AUD symptom severity) on AUD. The model proposed a hierarchy with causal influence propagating from brain function to cognition (Fluid/Crystallized Cognition, Language/Math & Working Memory) to social (Agreeableness/Social Support) to affective/psychiatric function (Negative Affect, low Conscientiousness/Attention, Externalizing symptoms) and ultimately AUD symptoms. These results underscore (a) a strong causal link between prefrontal brain function/cognition and affective/psychiatric factors and (b) an important buffer function of social factors (Social Support, Agreeableness). Our data-driven model hence confirmed hypothesized influences of cognitive and affective factors on AUD, while underscoring that traditional addiction models need to be expanded to highlight the importance of social factors, among others. Results further demonstrated that it is possible to reduce a broad phenotypic space (100 measures) to a limited set of causal factors of AUD, which can inform future research. We argue that the presented causal model of AUD provides evidence for exploring two different kinds of treatment approaches, specifically for investigating (a) “top-down” interventions aimed at enhancing high-level cognition, including brain-directed interventions targeting the executive network and (b) “integrative” interventions that take the interplay between brain/cognitive, affective/psychiatric factors, and social factors into account. We note that we did not investigate the individual heterogeneity of the causal factors involved in this model, but only provided a static causal model of an “average”

Table 2 Demographic characteristics of the final sample (n = 926).

Demographics	Options	Total N	AUD	Control	AUD – Control difference
Gender	M	428	128	300	$\chi^2 = 28.74, p < 0.001$
	F	498	76	498	
Race	White	700	166	534	$\chi^2 = 19.56, p = 0.002$
	Black/African-American	130	15	115	
	Asian/Nat. Hawaiian/Other Pacific Islander	57	8	49	
	Other	39	15	24	
Age	Mean	28.84	28.65	28.88	$t(924) = 0.80, p = 0.43$
	Standard deviation	3.69	3.38	3.74	
Education	Mean	14.98	14.95	14.99	$t(924) = 0.31, p = 0.76$
	Standard deviation	1.77	1.75	1.77	
Income	Mean	5.10	5.00	5.13	$t(924) = 0.78, p = 0.44$
	Standard deviation	2.13	2.15	2.13	
AUD symptoms	0	538			
	1	184			
	2	98			
	3	46			
	4	43			
	5+	17			
AUD diagnosis	Yes	204			
	No	722			

Table 3 List of AUD symptoms as described in the DSM.

DSM-IV-TR	Symptoms	DSM-5
Alcohol Abuse	1. Recurrent use of alcohol resulting in a failure to fulfill major role obligations at work, school, or home (e.g., repeated absences or poor work performance related to alcohol use; alcohol-related absences, suspensions, or expulsions from school; neglect of children or household)	Y
	2. Recurrent alcohol use in situations in which it is physically hazardous (e.g., driving an automobile or operating a machine when impaired by alcohol use).	Y
	3. Recurrent alcohol-related legal problems (e.g., arrests for alcohol-related disorderly conduct).	N
	4. Continued alcohol use despite having persistent or recurrent social or interpersonal problems caused or exacerbated by the effects of alcohol (e.g., arguments with spouse about consequences of intoxication)	Y
Alcohol Dependence	1. Need for markedly increased amounts of alcohol to achieve intoxication or desired effect; or markedly diminished effect with continued use of the same amount of alcohol.	Y
	2. The characteristic withdrawal syndrome for alcohol; or drinking (or using a closely related substance) to relieve or avoid withdrawal symptoms.	Y
	3. Drinking in larger amounts or over a longer period than intended.	Y
	4. Persistent desire or one or more unsuccessful efforts to cut down or control drinking	Y
	5. Important social, occupational, or recreational activities given up or reduced because of drinking	Y
	6. A great deal of time spent in activities necessary to obtain, to use, or to recover from the effects of drinking	Y
	7. Continued drinking despite knowledge of having a persistent or recurrent physical or psychological problem that is likely to be caused or exacerbated by drinking.	Y

individual with AUD symptoms, as a first step. We believe that this initial step of describing a comprehensive, integrated, multimodal but also reduced model (in a principled data-driven way) is crucial. We see the provided causal model as a working model, which can be further expanded (e.g. by the RDoC Positive Valence factors), explored with regard to individual heterogeneity and used in predictive modeling studies on alcohol use trajectories in active users, as well as in individuals with AUD in treatment.

Methods

Subjects. We analyzed data from the final release of the WU-Minn Human Connectome Project (n = 1206, aged 22–35, 54% females). All subjects provided written informed consent at Washington University. The CDA included all subjects who had complete data from all modalities (phenotypic n = 933, resting-state fMRI n = 1085, final n = 926). Subjects with AUD comprised 22% of the included sample, and 37% of subjects with AUD were females. See Table 2 for demographic characteristics of the included sample.

Outcome measure: AUD symptom severity. Subjects were assessed for symptoms of alcohol abuse and dependence using the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA). Symptom count data were provided for DSM-IV-TR alcohol abuse and alcohol dependence (Table 3). Symptom counts for alcohol abuse were provided as 0, 1, or 2+, and symptom counts for alcohol dependence were provided as 0, 1, 2, or 3+ (i.e. truncated symptom counts were provided). Given the low number of subjects with the highest symptom counts, it is unlikely that including a more fine-grained symptom count would have provided much additional information. It is therefore also likely that participants mostly had mild/moderate AUD severity, and that as the sample was young adults (age 22–35), this sample likely represents an early stage of AUD.

DSM-5 re-categorized alcohol abuse and alcohol dependence into a single disorder (AUD) using the criteria of both alcohol abuse and dependence, with one symptom of alcohol abuse removed (legal problems) and one symptom added (craving). We reconstructed this change by adding alcohol abuse and dependence symptom counts for each subject. Given recent interest in dimensional rather than categorical psychiatric dysfunction, we used the AUD symptom count (severity) as our primary outcome variable.

Behavioral and self-report measures. The HCP dataset contains a wide array of self-report, diagnostic and behavioral measures assessing domains of cognition,

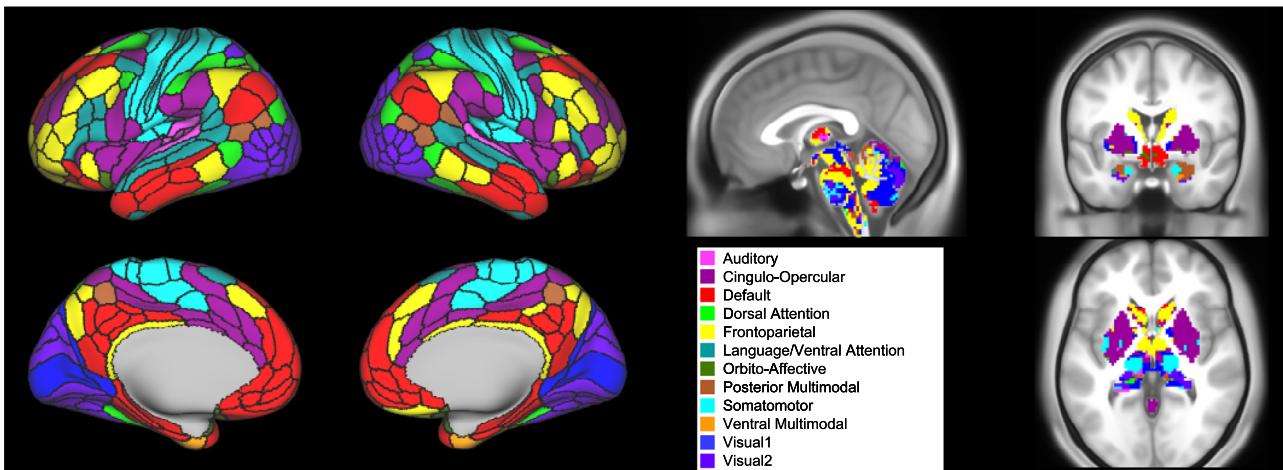


Fig. 2 Cole-Anticevic brain-wide network partition (CAB-NP). We conducted a whole-brain parcellation and assigned brain parcels to 12 large-scale networks according to the Cole-Anticevic brain-wide network partition (CAB-NP) parcellation²³. This parcellation built on the Glasser multimodal cortical parcellation by including subcortical and cerebellar parcels, and assigned each of the 718 parcels to a large-scale brain network using Louvain community detection.

emotion, social function, psychiatric dysfunction, and personality. We selected all available measures from these domains for further analysis (100 in total). We included all provided measures, but when provided, we used summary scores rather than item-level or minor scores as long as the summary score encompassed the task construct of interest⁸⁵. For example, for the Short Continuous Penn Test, we included the summary statistics of sensitivity and specificity but not more specific scores, and for delay discounting we included the area under the curve for \$200 and \$40k but not the individual discounting levels. For behavioral tasks completed in the scanner, we included separate behavioral measures for each of the major conditions. Note that in the interest of more completely characterizing the phenotypic space measured in the HCP data, we included behavior from tasks completed in the scanner but not fMRI measures of task-related brain activation.

If both age/gender-adjusted and unadjusted scores were provided, we included only the adjusted scores. For example, data from the Achenbach self-report inventory included raw scale scores, and scale scores adjusted for age and gender; given the strong relationship between psychiatric symptomology and age/gender, we included the age/gender-adjusted measures from this self-report inventory. This age/gender adjustment was thus applied to 14/100 (14%) of variables. Data from the NIH Toolbox cognition battery were likewise adjusted for participant age, given the strong dependency between age and cognitive abilities. Finally, the sensory sensitivity battery provided by the HCP consortium included measures of noise, taste, smell, and pain intensity, likewise controlled for the influence of age. Thus, age adjustment was applied to 11/100 (11%) of variables. Therefore in sum, 25% of variables were adjusted in some manner for participant characteristics (age: 11%, age/gender: 14%), while the remaining 75% of variables were not adjusted for any participant characteristics. We also excluded all items that were linear combinations of other data. For a list of included variables and excluded variables see Supplementary Data 6; for descriptive statistics of the included variables see Supplementary Data 7.

Factor analysis of phenotypic data. To reduce the phenotypic space measured in the HCP to a set of underlying domains, we conducted an EFA in the entire HCP sample that had complete phenotypic data ($n = 933$, 53.5% females). Factors were extracted using maximum likelihood as calculated by the expectation-maximization algorithm, and we used an oblimin rotation to allow for correlated factors. EFA was calculated in R using the “psych” package⁸⁶. The choice of EFA over similar data reduction schemes such as PCA was made because EFA explicitly accounts for error due to unreliability in measurement⁸⁷ unlike PCA⁸⁸. Furthermore, maximum likelihood extraction is robust against violation of distributional assumptions [as discussed in refs. ^{89,90}], providing a powerful technique for reduction of questionnaire and other behavioral data. As factor analysis conducts an eigenvalue decomposition of the correlation matrix among items, this analysis also does not assume that items submitted to factor analysis are measured on similar scales. Oblimin rotation allows for correlated factors, which is critical to data reduction over a large phenotypic variable space as we expect many factors to be closely related but separable (for example, negative affect and internalizing psychopathology). We used Monte Carlo permutation analysis (parallel analysis)⁹¹ to determine how many factors were statistically significant at $p < .05$ ⁹². Monte Carlo simulation was also calculated using the “psych” package for R⁸⁶. While some methods for choosing the number of factors in a solution are not robust against violation of distributional assumptions, the use of permutation analysis to choose the number of

extracted factors ensures that this selection does not depend on any assumptions regarding underlying distributional qualities.

Resting-state fMRI acquisition and preprocessing. High-resolution structural and functional MRI data were collected on a Siemens 3T Connectome Skyra scanner with a 32-channel head coil at Washington University. See Uğurbil et al.²¹ for a full description of the acquisition parameters for rfMRI in the HCP database. Resting-state fMRI (rfMRI) was collected over 2 days in four runs of 14:33 each. Structural data were preprocessed as described in Glasser et al.⁹³, using the most recent version of the HCP preprocessing pipeline (4.1). Briefly, anatomical image preprocessing consisted of bias field and gradient distortion correction, coregistration of T1w and T2w images, and linear and nonlinear registration to MNI space. Cortical surfaces were constructed using FreeSurfer. Surface files were transformed to MNI space, registered to the individual’s native-mesh surfaces, and downsampled.

Functional MRI preprocessing is fully described in Glasser et al.⁹³. Briefly, volumetric fMRI were subjected to gradient distortion correction, motion correction, and referencing to T1w. All transforms were concatenated and run in a single nonlinear resampling to MNI space. Data were then masked by the PostFreeSurfer brain mask and normalized. This volumetric timeseries was then mapped to a combined cortical surface and subcortical voxel space (“grayordinates”) and smoothed with a 2 mm FWHM Gaussian.

Finally, fMRI data were high-pass filtered (FWHM = 2355 s) and cleaned of artifacts using ICA-FIX^{94,95}. Artifact components and 24 motion regressors⁹⁶ were regressed out of the data in a single step. This produced the final ICA-FIX denoised versions of the data in both volumetric and CIFTI (combined grayordinates and subcortical/cerebellar voxels) space (<https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release>). CIFTI data were used for all primary analyses, while a replication analysis used the volumetric data to ensure the results replicated with multiple different brain network measures.

rfMRI parcellation and network assignment. Our analysis of rfMRI network connectivity only included subjects who completed at least one full day of rfMRI acquisition (two runs of 14:33), and 94% of subjects had four full runs of rfMRI data. We parcellated the whole brain, including cortex, subcortex, and cerebellum, into 718 parcels using the Cole-Anticevic brain-wide network partition (CAB-NP)²³, a parcellation scheme that builds on the Glasser et al.⁹⁷ multimodal cortex parcellation (360 parcels). We chose the CAB-NP parcellation because while the Glasser parcellation used multiple measures including myelination, rfMRI activity, and anatomical landmarks to delineate a fine-grained map of cortical space, it did not include any subcortical voxels, and did not explicitly assign parcels to large-scale networks using principled statistical methods. The CAB-NP parcellation built on the Glasser parcellation by (1) assigning subcortical and cerebellar voxels to parcels and (2) by using Louvain community detection to delineate 12 large-scale networks consisting of cortical, subcortical, and cerebellar regions (Fig. 2).

As the choice of parcellation can influence network connectivity profiles⁹⁸, we furthermore investigated whether the results of this part of the analysis would replicate when using an independently derived set of resting-state networks. We used the set of 300 regions-of-interest (ROIs) and associated network assignments described in ref. ⁸², which included the 264 ROIs described in ref. ⁸¹ with added subcortical and cerebellar ROIs (Supplementary Fig. S4). This parcellation used a

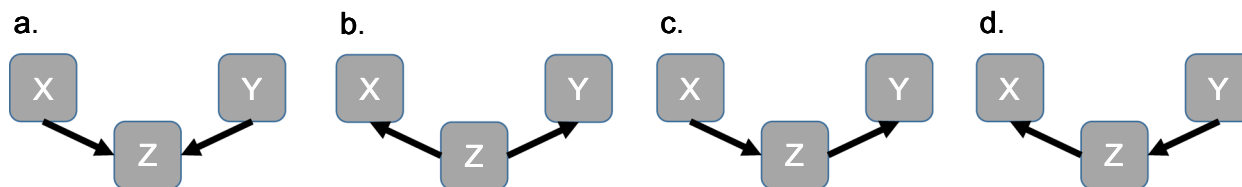


Fig. 3 Discovery of causal orientations using conditional independence relationships. Four different ways that three variables X, Y, and Z could be causally related. **a** is a structure known as a “collider,” in which X and Y both cause Z, but X and Y are not related. In this structure, X and Z are dependent, and Y and Z are dependent, while X and Y are independent. However, when Z is conditioned on (controlled for), X and Y are dependent. Meanwhile, in panel **b**, Z causes both X and Y. In this structure, X and Y are dependent because of their common cause, and are independent when Z is conditioned on. In panels **c**, **d**, X and Y are dependent because one causes Z, which then causes the other. In both of these panels, X and Y are independent when Z is conditioned on, as Z is the only link from X to Y. GFCI utilizes conditional independence tests to determine causal direction in graph edges, specifically by identifying “collider” cases in the graph (since these cases imply different conditional dependencies than the other three cases).

set of functionally defined spherical ROIs in volumetric space, and employed the Infomap algorithm to compute network assignments for each ROI. As this analysis was intended as a replication of our primary results, we only selected the networks that aligned closely with the CAB-NP networks (Supplementary Table S2).

Calculation of rfMRI network connectivity. For each network, we computed pairwise Pearson correlations between each pair of parcels in the network. Pearson correlations were transformed to approximate a normal distribution using Fisher’s z-transform. Within each of the networks, we then took the average of the parcel-to-parcel correlations to obtain a summary statistic for within-network connectivity^{99–103}. This procedure therefore summarizes how tightly connected (coherent) the regions comprising each of the 12 networks are with each other. This resulted in a set of average network-level connectivity values for each subject.

CDA: Greedy Fast Causal Inference. Causal models represent, often graphically, the set of cause-and-effect relationships that are present within a set of data¹⁰⁴. As the number of variables in a dataset increases, so too does the space of possible causal models that could give rise to the observed data, making the problem of identifying which of the potential causal models best fits the observed data very difficult. CDA uses machine learning to determine which causal models are best supported by the data¹⁰⁵. There are many CDA algorithms that make a wide variety of assumptions and have varying performance characteristics; for review, see Glymour et al.¹⁰⁶.

In the current study, we applied GFCI²⁴, an accurate and fast algorithm for establishing causal relationships from data even in the presence of unmeasured confounds. GFCI operates in two phases. GFCI begins by searching the space of possible graphs to create a preliminary graph that minimizes a penalized likelihood score, in this case the Bayesian Information Criteria¹⁰⁷. This initial search phase is done using the Fast Greedy Equivalence Search method¹⁰⁸. After the initial search phase, the algorithm refines the discovered graph by conducting a series of conditional independence tests. This phase rules out any edges that imply conditional dependencies not borne out by the data (for an example, see Fig. 3). Specifically, this phase capitalizes on the fact that “collider” structures (Fig. 3a; a case where two variables both cause a third variable) imply a separate set of conditional dependencies than any other ways those three variables could be causally related (Fig. 3b–d). The most important distinction of GFCI compared to other causal discovery methods is that GFCI can detect confounding factors, and as such is particularly suited to analysis of real-world data, where there is no guarantee that every relevant variable has been measured. The output of the GFCI algorithm is a partial ancestral graph with edge types indicating causal relationships, uncertain relationships, and the presence of unmeasured confounding variables. GFCI analysis was implemented using Tetrad. Analysis was run with default parameters; that is, using $\alpha = 0.01$, maximum degree of the graph = 100, and a penalty discount of two. Penalized likelihoods for models were calculated using the Bayesian Information Criteria¹⁰⁷, which is the default model fit index in Tetrad and the most commonly used model fit index in CDA.

To recover effect sizes of the edges in the model, we fit a structural equation model (SEM) to the graph structure using the “lavaan” package for R¹⁰⁹. We present the graph GFCI learned from the full dataset with SEM effect sizes for each edge. As an additional analysis of model stability in smaller samples, we also conducted a stability analysis by resampling 90% of the sample¹¹⁰ without replacement with 1000 repetitions (jackknifing). Finally, to test the replicability of the brain model, we used the R package “lavaan” to fit a SEM representing the discovered causal model (CAB-NP)²³ to the alternately derived brain network data (Greene-300)⁸².

Statistics and reproducibility. Factor analysis (parallel analysis and factor extraction) was done using the “psych” package (version 2.0.12)⁸⁶, and factor rotation used the “GPARotation” package (version 2014.11-1) for R (version 3.6.3). fMRI network connectivity analysis was done using Connectome Workbench

(version 1.4.2) and MATLAB (version R2018b). Causal discovery analysis (GFCI) was done using Tetrad (version 6.7). Stability analysis (jackknifing) was done in Tetrad (version 6.7) using 90% subsampling and 1000 replications. Effect size calculation was done using the “lavaan” package (version 0.6-6) running in R (version 3.6.3). We used a very large, well-documented, and publicly available dataset (HCP; $n = 926$) to ensure the representativeness of the analysis, and we additionally validated the stability of the neurobiological analyses using a secondary, independently derived brain parcellation, and the stability of the causal discovery analyses using a jackknifing procedure.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data analyzed for this study are available from the WU-Minn Human Connectome Project Consortium (<https://www.humanconnectome.org/study/hcp-young-adult/document/1200-subjects-data-release>). These data are available for public use with a data use agreement. Raw data sufficient to reproduce the GFCI and SEM results is available (Supplementary Data 8).

Code availability

All code and analysis packages used in this manuscript are freely available with the exception of MATLAB. MATLAB was used for convenience (fMRI network correlations) but is not required to replicate analyses. Connectome Workbench is freely available from <https://www.humanconnectome.org/software/get-connectome-workbench>. R is freely available from <https://www.r-project.org/>. The R packages “psych”, “GPARotation”, and “lavaan” can be freely installed within R. Tetrad is freely available from <https://www.ccd.pitt.edu/tools/>.

Received: 29 September 2020; Accepted: 24 February 2021;

Published online: 31 March 2021

References

- Grant, B. F. et al. Epidemiology of DSM-5 alcohol use disorder: results from the National Epidemiologic Survey on Alcohol and Related Conditions III. *JAMA Psychiatry* **72**, 757–766 (2015).
- Hasin, D. S., Stinson, F. S., Ogburn, E. & Grant, B. F. Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Arch. Gen. Psychiatry* **64**, 830–842 (2007).
- WHO. *Global Status Report on Alcohol and Health 2018*. (2018).
- Chiappetta, V., Garcia-Rodriguez, O., Jin, C. J., Secades-Villa, R. & Blanco, C. Predictors of quit attempts and successful quit attempts among individuals with alcohol use disorders in a nationally representative sample. *Drug Alcohol Depend.* **141**, 138–144 (2014).
- Kranzler, H. R. & Soyka, M. Diagnosis and pharmacotherapy of alcohol use disorder: a review. *JAMA* **320**, 815 (2018).
- Litten, R. Z. et al. Heterogeneity of alcohol use disorder: understanding mechanisms to advance personalized treatment. *Alcohol. Clin. Exp. Res.* **39**, 579–584 (2015).
- Koob, G. F. & Le Moal, M. Drug addiction, dysregulation of reward, and allostasis. *Neuropsychopharmacology* **24**, 97–129 (2001).
- Koob, G. F. Drug addiction: the Yin and Yang of hedonic homeostasis. *Neuron* **16**, 893–896 (1996).

9. Koob, G. F. & Volkow, N. D. Neurobiology of addiction: a neurocircuitry analysis. *Lancet Psychiatry* **3**, 760–773 (2016).
10. Belcher, A. M., Volkow, N. D., Moeller, F. G. & Ferré, S. Personality traits and vulnerability or resilience to substance use disorders. *Trends Cogn. Sci.* **18**, 211–217 (2014).
11. Kwako, L. E., Momenan, R., Litten, R. Z., Koob, G. F. & Goldman, D. Addictions neuroclinical assessment: a neuroscience-based framework for addictive disorders. *Biol. Psychiatry* **80**, 179–189 (2016).
12. Kwako, L. E. et al. Addictions neuroclinical assessment: a reverse translational approach. *Neuropharmacology* **122**, 254–264 (2017).
13. Kwako, L. E. et al. Neurofunctional domains derived from deep behavioral phenotyping in alcohol use disorder. *Am. J. Psychiatry* **176**, 744–753 (2019).
14. Cuthbert, B. N. & Insel, T. R. Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Med.* **11**, 126 (2013).
15. Cuthbert, B. N. & Kozak, M. J. Constructing constructs for psychopathology: The NIMH research domain criteria. *J. Abnorm. Psychol.* **122**, 928 (2013).
16. Yücel, M. et al. A transdiagnostic dimensional approach towards a neuropsychological assessment for addiction: an International Delphi Consensus Study. *Addiction* **114**, 1095–1109 (2019).
17. Van Dam, N. T. et al. Data-driven phenotypic categorization for neurobiological analyses: beyond DSM-5 labels. *Biol. Psychiatry* **81**, 484–494 (2017).
18. Redish, A. D., Jensen, S. & Johnson, A. A unified framework for addiction: vulnerabilities in the decision process. *Behav. Brain Sci.* **31**, 415–487 (2008).
19. Zilverstand, A., Huang, A. S., Alia-Klein, N. & Goldstein, R. Z. Neuroimaging impaired response inhibition and salience attribution in human drug addiction: a systematic review. *Neuron* **98**, 886–903 (2018).
20. Barch, D. M. et al. Function in the human connectome: Task-fMRI and individual differences in behavior. *NeuroImage* **80**, 169–189 (2013).
21. Uğurbil, K. et al. Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *NeuroImage* **80**, 80–104 (2013).
22. Van Essen, D. C. et al. The WU-Minn Human Connectome Project: an overview. *NeuroImage* **80**, 62–79 (2013).
23. Ji, J. L. et al. Mapping the human brain's cortical-subcortical functional network organization. *NeuroImage* **185**, 35–57 (2019).
24. Ogarrio, J. M., Spirtes, P. & Ramsey, J. A hybrid causal search algorithm for latent variable models. *Conference on Probabilistic Graphical Models, PMLR* **52**, 368–379 (2016).
25. Cattell, R. B. Theory of fluid and crystallized intelligence: a critical experiment. *J. Educ. Psychol.* **54**, 1–22 (1963).
26. Cattell, R. B. *Intelligence: Its Structure, Growth and Action*. xxii, 694 (North-Holland, 1987).
27. Zilverstand, A., Parvaz, M. A. & Goldstein, R. Z. Neuroimaging cognitive reappraisal in clinical populations to define neural targets for enhancing emotion regulation. A systematic review. *NeuroImage* **151**, 105–116 (2017).
28. Lempert, K. M., Steinglass, J. E., Pinto, A., Kable, J. W. & Simpson, H. B. Can delay discounting deliver on the promise of RDoC? *Psychol. Med.* **49**, 190–199 (2019).
29. Dalley, J. W., Everitt, B. J. & Robbins, T. W. Impulsivity, compulsivity, and top-down cognitive control. *Neuron* **69**, 680–694 (2011).
30. Koffarnus, M. N., Jarmolowicz, D. P., Mueller, E. T. & Bickel, W. K. Changing delay discounting in the light of the competing neurobehavioral decision systems theory: a review. *J. Exp. Anal. Behav.* **99**, 32–57 (2013).
31. Schel, M. A., Scheres, A. & Crone, E. A. New perspectives on self-control development: highlighting the role of intentional inhibition. *Neuropsychologia* **65**, 236–246 (2014).
32. Heilig, M., Epstein, D. H., Nader, M. A. & Shaham, Y. Time to connect: bringing social context into addiction neuroscience. *Nat. Rev. Neurosci.* **17**, 592–599 (2016).
33. Tomek, S. E. & Olive, M. F. Social influences in animal models of opiate addiction. In *International Review of Neurobiology* Vol. 140 (eds. Olive, M. F. & Tomek, S. E.) Ch. 4, 81–107 (Academic Press, 2018).
34. Berkman, L. F., Kawachi, I. & Glymour, M. M. *Social Epidemiology* (Oxford University Press, 2014).
35. Havassy, B. E., Hall, S. M. & Wasserman, D. A. Social support and relapse: commonalities among alcoholics, opiate users, and cigarette smokers. *Addict. Behav.* **16**, 235–246 (1991).
36. Leach, D. & Kranzler, H. R. An interpersonal model of addiction relapse. *Addict. Disord. Treat.* **12**, 183–192 (2013).
37. Manza, P., Shokri-Kojori, E. & Volkow, N. D. Reduced segregation between cognitive and emotional processes in cannabis dependence. *Cereb. Cortex* **30**, 628–639 (2020).
38. Goldstein, R. Z. & Volkow, N. D. Dysfunction of the prefrontal cortex in addiction: neuroimaging findings and clinical implications. *Nat. Rev. Neurosci.* **12**, 652–669 (2011).
39. Goldstein, R. Z. & Volkow, N. D. Drug addiction and its underlying neurobiological basis: neuroimaging evidence for the involvement of the frontal cortex. *Am. J. Psychiatry* **159**, 1642–1652 (2002).
40. Farmer, R. F. et al. Internalizing and externalizing disorders as predictors of alcohol use disorder onset during three developmental periods. *Drug Alcohol Depend.* **164**, 38–46 (2016).
41. Anker, J. J. & Kushner, M. G. Co-occurring alcohol use disorder and anxiety. *Alcohol Res. Curr. Rev.* **40** (2019).
42. Kushner, M. G. Seventy-five years of comorbidity research. *J. Stud. Alcohol Drugs. Suppl.* **75** (Suppl 17) 50–58 (2014).
43. Pacek, L. R. et al. Comorbid alcohol dependence and anxiety disorders: a national survey. *J. Dual Diagn.* **9**, 271–280 (2013).
44. Anker, J. J., Kummerfeld, E., Rix, A., Burwell, S. J. & Kushner, M. G. Causal network modeling of the determinants of drinking behavior in comorbid alcohol use and anxiety disorder. *Alcohol. Clin. Exp. Res.* **43**, 91–97 (2019).
45. Bilker, W. B. et al. Development of abbreviated nine-item forms of the Raven's Standard Progressive Matrices Test. *Assessment* **19**, 354–369 (2012).
46. McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A. & Hambrick, D. Z. The relationship between working memory capacity and executive functioning: evidence for a common executive attention construct. *Neuropsychology* **24**, 222–243 (2010).
47. Shelton, J. T., Elliott, E. M., Matthews, R. A., Hill, B. D. & Gouvier, Wm. D. The relationships of working memory, secondary memory, and general fluid intelligence: working memory is special. *J. Exp. Psychol. Learn. Mem. Cogn.* **36**, 813–820 (2010).
48. Barbey, A. K. et al. An integrative architecture for general intelligence and executive function revealed by lesion mapping. *Brain* **135**, 1154–1164 (2012).
49. Bechara, A. & Martin, E. M. Impaired decision making related to working memory deficits in individuals with substance addictions. *Neuropsychology* **18**, 152–162 (2004).
50. Goudriaan, A. E., Oosterlaan, J., de Beurs, E. & van den Brink, W. Decision making in pathological gambling: a comparison between pathological gamblers, alcohol dependents, persons with Tourette syndrome, and normal controls. *Cogn. Brain Res.* **23**, 137–151 (2005).
51. Verdejo-Garcia, A. Goal-based interventions for executive dysfunction in addiction treatment. In *Cognition and Addiction*, (ed. Verdejo-Garcia, A.) 277–282 (Elsevier, 2020).
52. Squeglia, L. M. et al. Neural predictors of initiating alcohol use during adolescence. *Am. J. Psychiatry* **174**, 172–185 (2016).
53. Keyes, K. M., Platt, J., Kaufman, A. S. & McLaughlin, K. A. Association of fluid intelligence and psychiatric disorders in a population-representative sample of US adolescents. *JAMA Psychiatry* **74**, 179–188 (2017).
54. Domínguez-Salas, S., Díaz-Batanero, C., Lozano-Rojas, O. M. & Verdejo-Garcia, A. Impact of general cognition and executive function deficits on addiction treatment outcomes: systematic review and discussion of neurocognitive pathways. *Neurosci. Biobehav. Rev.* **71**, 772–801 (2016).
55. Santarnecchi, E., Emmendorfer, A. & Pascual-Leone, A. Dissecting the parieto-frontal correlates of fluid intelligence: a comprehensive ALE meta-analysis study. *Intelligence* **63**, 9–28 (2017).
56. Spreng, R. N., Stevens, W. D., Chamberlain, J. P., Gilmore, A. W. & Schacter, D. L. Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *NeuroImage* **53**, 303–317 (2010).
57. Tschentscher, N., Mitchell, D. & Duncan, J. Fluid intelligence predicts novel rule implementation in a distributed frontoparietal control network. *J. Neurosci.* **37**, 4841–4847 (2017).
58. Watson, C. E. & Chatterjee, A. A bilateral frontoparietal network underlies visuospatial analogical reasoning. *NeuroImage* **59**, 2831–2838 (2012).
59. Woolgar, A., Duncan, J., Manes, F. & Fedorenko, E. Fluid intelligence is supported by the multiple-demand system not the language system. *Nat. Hum. Behav.* **2**, 200–204 (2018).
60. Luijten, M. et al. Systematic review of ERP and fMRI studies investigating inhibitory control and error processing in people with substance dependence and behavioural addictions. *J. Psychiatry Neurosci.* **39**, 149–169 (2014).
61. Zilverstand, A., Parvaz, M. A., Moeller, S. J. & Goldstein, R. Z. Cognitive interventions for addiction medicine: understanding the underlying neurobiological mechanisms. *Prog. Brain Res.* **224**, 285–304 (2016).
62. Kim, S. M., Han, D. H., Min, K. J., Kim, B.-N. & Cheong, J. H. Brain activation in response to craving- and aversion-inducing cues related to alcohol in patients with alcohol dependence. *Drug Alcohol Depend.* **141**, 124–131 (2014).
63. Maurage, P. et al. Disrupted regulation of social exclusion in alcohol-dependence: an fMRI study. *Neuropsychopharmacology* **37**, 2067–2075 (2012).
64. Beylergil, S. B. et al. Dorsolateral prefrontal cortex contributes to the impaired behavioral adaptation in alcohol dependence. *NeuroImage Clin.* **15**, 80–94 (2017).
65. Hu, S., Ide, J. S., Zhang, S., Sinha, R. & Li, C. R. Conflict anticipation in alcohol dependence—a model-based fMRI study of stop signal task. *NeuroImage Clin.* **8**, 39–50 (2015).

66. Shokri-Kojori, E., Tomasi, D., Wiers, C. E., Wang, G.-J. & Volkow, N. D. Alcohol affects brain functional connectivity and its coupling with behavior: greater effects in male heavy drinkers. *Mol. Psychiatry* **22**, 1185–1195 (2017).
67. Seo, D. et al. Disrupted ventromedial prefrontal function, alcohol craving, and subsequent relapse risk. *JAMA Psychiatry* **70**, 727 (2013).
68. Norman, A. L. et al. Neural activation during inhibition predicts initiation of substance use in adolescence. *Drug Alcohol Depend.* **119**, 216–223 (2011).
69. Wetherill, R. R., Squeglia, L. M., Yang, T. T. & Tapert, S. F. A longitudinal examination of adolescent response inhibition: neural differences before and after the initiation of heavy drinking. *Psychopharmacology* **230**, 663–671 (2013).
70. Naish, K. R., Vedelago, L., MacKillop, J. & Amlung, M. Effects of neuromodulation on cognitive performance in individuals exhibiting addictive behaviors: a systematic review. *Drug Alcohol Depend.* **192**, 338–351 (2018).
71. Bartra, O., McGuire, J. T. & Kable, J. W. The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage* **76**, 412–427 (2013).
72. Corbetta, M., Patel, G. & Shulman, G. L. The reorienting system of the human brain: from environment to theory of mind. *Neuron* **58**, 306–324 (2008).
73. Riggall, A. C. & Postle, B. R. The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. *J. Neurosci.* **32**, 12990–12998 (2012).
74. Majerus, S., Pétters, F., Bouffier, M., Cowan, N. & Phillips, C. The dorsal attention network reflects both encoding load and top-down control during working memory. *J. Cogn. Neurosci.* **30**, 144–159 (2018).
75. Czapla, M. et al. Do alcohol-dependent patients show different neural activation during response inhibition than healthy controls in an alcohol-related fMRI go/no-go-task? *Psychopharmacology* **234**, 1001–1015 (2017).
76. Fryer, S. L. et al. Differential brain response to alcohol cue distractors across stages of alcohol dependence. *Biol. Psychol.* **92**, 282–291 (2013).
77. Sofuoglu, M., DeVito, E. E., Waters, A. J. & Carroll, K. M. Cognitive enhancement as a treatment for drug addictions. *Neuropharmacology* **64**, 452–463 (2013).
78. Sofuoglu, M., DeVito, E. E., Waters, A. J. & Carroll, K. M. Cognitive function as a trans-diagnostic treatment target in stimulant use disorders. *J. Dual Diagn.* **12**, 90–106 (2016).
79. Song, S., Zilverstand, A., Gui, W., Li, H. & Zhou, X. Effects of single-session versus multi-session non-invasive brain stimulation on craving and consumption in individuals with drug addiction, eating disorders or obesity: a meta-analysis. *Brain Stimul.* **12**, 606–618 (2019).
80. Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S. & Petersen, S. E. Intrinsic and task-evoked network architectures of the human brain. *Neuron* **83**, 238–251 (2014).
81. Power, J. D. et al. Functional network organization of the human brain. *Neuron* **72**, 665–678 (2011).
82. Seitzman, B. A. et al. A set of functionally-defined brain regions with improved representation of the subcortex and cerebellum. *NeuroImage* **206**, 116290 (2020).
83. Vossel, S., Geng, J. J. & Fink, G. R. Dorsal and ventral attention systems: distinct neural circuits but collaborative roles. *Neuroscientist* **20**, 150–159 (2014).
84. Kim, H. Involvement of the dorsal and ventral attention networks in oddball stimulus processing: a meta-analysis. *Hum. Brain Mapp.* **35**, 2265–2284 (2014).
85. Smith, S. M. et al. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nat. Neurosci.* **18**, 1565–1567 (2015).
86. Revelle, W. *Psych: Procedures for Psychological, Psychometric, and Personality Research*, Northwestern University, Evanston, Illinois, USA, <https://CRAN.R-project.org/package=psych> Version 2.0.12. (2017).
87. Suhr, D. Exploratory or confirmatory factor analysis? *SAS SUGI 31 Proceedings: Statistics, Data Analysis, and Data Mining* **200–31**, 66–76 (2006).
88. Bryant, F. B. & Yarnold, P. R. Principal-components analysis and exploratory and confirmatory factor analysis. In *Reading and understanding multivariate statistics* (eds. L. G. Grimm & P. R. Yarnold) 99–136 (American Psychological Association, 1995).
89. Fuller, E. L. & Hemmerle, W. J. Robustness of the maximum-likelihood estimation procedure in factor analysis. *Psychometrika* **31**, 255–266 (1966).
90. Muthén, B. O. & Kaplan, D. A comparison of some methodologies for the factor analysis of non-normal likert variables: a note on the size of the model. *Br. J. Math. Stat. Psychol.* **45**, 19–30 (1992).
91. Horn, J. L. A rationale and test for the number of factors in factor analysis. *Psychometrika* **30**, 179–185 (1965).
92. Glorfeld, L. W. An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educ. Psychol. Meas.* <https://doi.org/10.1177/0013164495055003002> (1995).
93. Glasser, M. F. et al. The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105–124 (2013).
94. Griffanti, L. et al. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *NeuroImage* **95**, 232–247 (2014).
95. Salimi-Khorshidi, G. et al. Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *NeuroImage* **90**, 449–468 (2014).
96. Satterthwaite, T. D. et al. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *NeuroImage* **64**, 240–256 (2013).
97. Glasser, M. F. et al. A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).
98. Messé, A. Parcellation influence on the connectivity-based structure–function relationship in the human brain. *Hum. Brain Mapp.* **41**, 1167–1180 (2020).
99. Hagler, D. J. et al. Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *NeuroImage* **202**, 116091 (2019).
100. Mamah, D., Barch, D. M. & Repovš, G. Resting state functional connectivity of five neural networks in bipolar disorder and schizophrenia. *J. Affect. Disord.* **150**, 601–609 (2013).
101. Repovš, G., Csernansky, J. G. & Barch, D. M. Brain network connectivity in individuals with schizophrenia and their siblings. *Biol. Psychiatry* **69**, 967–973 (2011).
102. Repovš, G. & Barch, D. M. Working memory related brain network connectivity in individuals with schizophrenia and their siblings. *Front. Hum. Neurosci.* **6**, 137 (2012).
103. Van Dijk, K. R. A. et al. Intrinsic Functional connectivity as a tool for human connectomics: theory, properties, and optimization. *J. Neurophysiol.* **103**, 297–321 (2010).
104. Pearl, J. *Causality: Models, Reasoning, and Inference*. xvi, 384 (Cambridge University Press, 2000).
105. Spirtes, P., Glymour, C. N. & Scheines, R. *Causation, Prediction, and Search* (MIT Press, 2000).
106. Glymour, C., Zhang, K. & Spirtes, P. Review of causal discovery methods based on graphical models. *Front. Genet.* **10**, 524 (2019).
107. Schwarz, G. Estimating the dimension of a model. *Annals of Statistics*, **6**, 461–464 (1978).
108. Ramsey, J., Glymour, M., Sanchez-Romero, R. & Glymour, C. A million variables and more: the Fast Greedy Equivalence Search algorithm for learning high-dimensional graphical causal models, with an application to functional magnetic resonance images. *Int. J. Data. Sci. Anal.* **3**, 121–129 (2017).
109. Rosseel, Y. lavaan: an R package for structural equation modeling. *J. Stat. Softw.* **48**, 1–36 (2012).
110. Kummerfeld, E. & Rix, A. Simulations evaluating resampling methods for causal discovery: ensemble performance and calibration. In *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* 2586–2593 (2019).

Acknowledgements

E.R. is supported by National Institutes of Mental Health (NIMH) grant T32-MH115866. E.K. received support for this work from the National Center for Advancing Translational Sciences of the National Institutes of Health Award Number UL1TR000114. A.Z. is supported by the P30 DA048742-01A1 of the National Institute on Drug Abuse. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Institutes of Mental Health. The authors thank Matthew Kushner for valuable input and editorial assistance.

Author contributions

E.R.: conceptualization, methodology, software, validation, formal analysis, investigation, writing—original draft, writing—review & editing, visualization. E.K.: conceptualization, methodology, software, validation, writing—review & editing, visualization, supervision. A.Z.: conceptualization, methodology, validation, resources, writing—review and editing, supervision, project administration.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-021-01955-z>.

Correspondence and requests for materials should be addressed to E.R.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021