Measuring addiction propensity and severity: The need for a new instrument∗

Kevin P. Conwaya,*, Janet Levyb, Michael Vanyukovc, Redonna Chandlera,
Joni Rutterm, Gary E. Swane, Michael Nealef

a Division of Epidemiology, Services, and Prevention Research, National Institute on Drug Abuse, National Institutes of Health, Department of Health and Human Services, 6001 Executive Blvd., Room 5153, MSC 9589, Bethesda, MD 20892-9589, USA

b Duke University School of Nursing, SON New Building, Room 1019, 307 Trent Drive, Durham, NC 27710, USA

c Departments of Pharmaceutical Sciences, Psychiatry and Human Genetics, University of Pittsburgh, SALK 528, Pittsburgh, PA 15260, USA

d Division of Basic Neuroscience and Behavioral Research, National Institute on Drug Abuse, National Institutes of Health, Department of Health and Human Services, 6001 Executive Blvd., Room 4282, MSC 9589, Rockville, MD 20852, USA

e Center for Health Sciences, SRI International, 333 Ravenswood Avenue, Menlo Park, CA 94025, USA

f Department of Psychiatry, Virginia Commonwealth University, 1200 East Broad Street, PO Box 980710, Richmond, VA 23298-0710, USA

Corresponding author at: Division of Epidemiology, Services, and Prevention Research (DESPR), National Institute on Drug Abuse, 6001 Executive Blvd., Suite 5185, Bethesda, MD 20892-9589, USA. Tel.: +1 301 443 6504; fax: +1 301 443 2636.

E-mail address: kconway@nida.nih.gov (K.P. Conway).

∗ Disclaimer: The views and opinions expressed in this paper are those of the authors and do not necessarily represent the views of the National Institute on Drug Abuse, National Institutes of Health, or any other governmental agency.

Drug addiction research requires but lacks a valid and reliable way to measure both the risk (propensity) to develop addiction and the severity of manifest addiction. This paper argues for a new measurement approach and instrument to quantify propensity to and severity of addiction, based on the testable assumption that these constructs can be mapped onto the same dimension of liability to addiction. The case for this new direction becomes clear from a critical review of empirical data and the current instrumentation. The many assessment instruments in use today have proven utility, reliability, and validity, but they are of limited use for evaluating individual differences in propensity and severity. The conceptual and methodological shortcomings of instruments currently used in research and clinical practice can be overcome through the use of new technologies to develop a reliable, valid, and standardized assessment instrument(s) to measure and distinguish individual variations in expression of the underlying latent trait(s) that comprises propensity to and severity of drug addiction. Such instrumentation would enhance our capacity for drug addiction research on linkages and interactions among familial, genetic, psychosocial, and neurobiological factors associated with variations in propensity and severity. It would lead to new opportunities in substance abuse prevention, treatment, and services research, as well as in interventions and implementation science for drug addiction.

Drug and Alcohol Dependence 111 (2010) 4–12

Published by Elsevier Ireland Ltd.

111 (2010) 4–12

Published by Elsevier Ireland Ltd.
1. Introduction

Drug addiction research spans the gamut from neuroscience and genetics to prevention, treatment and services. As in other biobehavioral disciplines, this research requires a valid and reliable measure of both the propensity to develop addiction and the severity of manifest addiction. This paper examines the need for such a measure as well as the assumption on which it is based, namely that these constructs share the same dimension. It provides a critical review of theory, data, and current instrumentation to show how a new measurement tool could impact the field of addiction research. Finally, it considers the next steps required for the development and fruition of a new instrument for measuring addiction.

Fueled by technological innovations and multidisciplinary scientific collaborations, addiction science has advanced rapidly over the past 20 years. The notion that addiction “runs in families” (Merikangas and Conway, 2009; Bierut et al., 1998; Merikangas et al., 1998: Reich et al., 1988; Schuckit et al., 1972; Winokur et al., 1970) is now attributed in large part to additive genetic factors (Kendler et al., 2003a; Tsuang et al., 1996). This has lead scientists to search for genes that influence variation in the risk for addiction, with a number of replication studies now emerging on candidate genes (Full et al., 2009; Kreek et al., 2005; Li and Burmeister, 2009; Saxon et al., 2005; Schuckit, 2009; Uhl, 2004; Vanyukov and Tarter, 2000). Scientific advances in neuroscience have shown that addiction is a chronic and often relapsing disease that is linked to pathological changes in neural circuitry, including those involved in reward and motivation, learning and memory, cognitive control, decision making, mood, and interoceptive awareness (Adinoff, 2004; Hyman et al., 2006; Kalivas and Volkow, 2005; Kalivas and O’Brien, 2007; Koob and Le, 1997; O’Brien, 2003; Volkow et al., 2003, 2004). Importantly, these changes involve the same structures and processes that contribute to the mechanisms of behavior regulation and its deviations, predating drug use (Vanyukov et al., 2003b). Impairment of the addicted brain provides a neurological basis for the cardinal behavioral manifestation of drug addiction—persistent drug use despite serious adverse consequences. Longitudinal observations of addicts demonstrate the chronic relapsing nature of addiction and the need for long-term treatment strategies (Dennis and Scott, 2007; Dennis et al., 2005). However, the malleability of the liability phenotype, including “maturing-out” of conditions that meet diagnostic criteria for dependence (Dawson et al., 2006; Newcomb et al., 2001), underscores the nonspecific character and construct validity of liability while still holding promise for effective intervention. In contrast to such static traits as stature in adults, liability to addiction is dynamic over time and development. It can be viewed as forming an ontogenetic trajectory that can be tracked and measured, and given its inherent dynamism, potentially changed (Tarter and Vanyukov, 1991, 1994).

Advances in these and other important areas of addiction science have outpaced progress in the measurement of addiction, and the ability to quantify and measure this construct accurately remains a looming methodological challenge (Conway et al., 2006; Craddock et al., 2008; Merikangas and Avenevoli, 2000; Merikangas and Conway, 2008; Neale et al., 2006). Currently, research usually measures drug addiction in the broadest sense by classifying and contrasting “addicts” to “non-addicts” according to diagnostic criteria. Genetic risk for addiction is similarly categorized based on a parent’s diagnosis or family history. This approach implicitly assumes categorical distinctions between groups and categorical similarities among the individuals within each group regarding their symptoms and group-identified features of addiction. Few studies have focused on variations among individuals within each group, despite the common knowledge that meaningful individual differences exist within any given cohort of addicts and controls. Such heterogeneity is reflected in the hundreds of different combinations of addiction symptoms that meet diagnostic criteria (Vanyukov et al., 2003b). Although needed for clinical practice, the diagnosis per se is less than optimal when it comes to genetic and other etiology research, or for informing efforts in the primary prevention of addiction.

A precise method and instrument for measuring individual variation in liability to addiction is therefore needed to build on and advance the science and understanding of the multifactorial nature of drug abuse and addiction. A term introduced in human genetics by Falconer (1965, p. 52), liability is a latent (unobservable) quantitative trait that, if measured, “would give us a graded scale of the degree of affectedness or of normality”, with these two categories divided by a threshold. The “gradations of normality” (the subthreshold liability phenotypes) correspond to variation in the risk (propensity), whereas “gradations of affectedness” (the suprathereshold phenotypes, likely to be assigned a clinical diagnosis) correspond to variation in severity, comprising the two portions of the liability distribution. Applied to addiction, severity refers to the degree of maladaptive compulsive drug-seeking and using behavior displayed by an individual, corresponding to variation in liability above the diagnostic threshold. Propensity refers to the probability of the disorder’s onset, corresponding to liability variation below the diagnostic threshold. It follows that individual differences in severity manifest as different degrees of maladaptive drug-seeking and drug-using behavior. Similarly, variation in propensity manifest as behavioral precursors of addiction. The ultimate desirable goal of a new instrument of addiction would be to provide a single scale (see Fig. 1) by which an individual’s liability to addiction (propensity or severity) can be quantified using a numeric score.

The plausibility of such a single common (versus drug-specific) liability dimension (latent trait) and the feasibility of its measurement are supported by clinical, neurobiological, genetic, and statistical findings (Vanyukov et al., 2003a). Liabilities to addictions to specific drugs are both phenotypically and genetically highly correlated, with minimal specific genetic variance; most of the variance in addiction liability is associated with common (thus, likely, brain-behavioral) mechanisms rather than with drug action per se (Kendler et al., 2007; Tsuang et al., 1998). Indeed, neurobiological data pertaining to drug reward suggest that many drug effects
between an index of behavior disinhibition (locating symptoms of conduct and attention-deficit hyperactivity disorder, novelty seeking, and substance use on the same dimension) and an index of performance on neuropsychological tasks measuring attentional control and response inhibition (Young et al., 2009). A similar brain-behavior connection is noted for disinhibition/externalizing and the reduced amplitude of the P300 event-related potential (Iacono et al., 2002; Justus et al., 2001), particularly its main time-frequency components, theta and delta, indexing, respectively, memory encoding and attention processes, and signal matching, decision making, and memory updating processes (Gilmore et al., 2009). The association between P300 amplitude and an index of the latent externalizing trait (accounting for variance shared between liabilities to alcohol, nicotine and drug dependence, conduct disorder and adult antisocial behavior) is of genetic origin (Hicks et al., 2007). As shown by genetic studies of phenotypes based on neuroimaging, genetic associations of behavioral response regulation and personality characteristics comprising the behavior dysregulation construct are mediated by the function of neural circuitry involving targets of drug action (Harriri et al., 2008; Hariri, 2009).

Statistically, substance use disorders and related symptoms have been shown to best fit a model whereby their covariances are substantially accounted for by a single dominant factor (Iacono et al., 2002, 2006; Compton et al., 2009; Saha et al., 2007; Wu et al., 2009). Unidimensionality was also established for items that comprise the TLI, the instrument (described above) that measures transmissible liability to illicit drug-related substance use disorder in children. Interestingly, the TLI items are largely related to a set of externalizing problems (e.g., conduct disorder, antisocial personality disorder) and temperament-like indicators of nonspecific addiction risk (e.g., behavioral disinhibition, constraint, sensation seeking, and novelty seeking) collectively referred to as behavioral undercontrol, dysregulation, or externalizing psychopathology (Cadoret et al., 1995; Chassin et al., 1991; Conway et al., 2002, 2003; Elkins et al., 2004; Iacono et al., 1999, 2008; Krueger et al., 2002, 2007; Sher et al., 2000). Evidence suggests a common genetic variance between these traits and addiction (Mustanski et al., 2003). Some studies suggest that the broad externalizing factor is more heritable than the constituent individual disorders (80–85% versus 40–70%) (Hicks et al., 2004; Moffitt, 2005; Young et al., 2000), as well as being more useful for gene identification (Dick et al., 2008).

Variation in common (nonspecific) liability to addiction could help explain the typical pattern of progression “softer” to “harder” types of substances used (Tarter et al., 2006). Variation in externalizing traits could also mediate heritability of liability to addiction (Button et al., 2006; Kirillova et al., 2008; Slutske et al., 2002; Swendsen et al., 2002; Tarter et al., 2004; Vanyukov et al., 2007). From a genetic-risk perspective, familial antisocial addiction may be particularly potent (or severe). Studies of adopted-away offspring of fathers who are antisocial addicts (compared to either antisocial or addicted), for example, have found the offspring to be at greatest risk for substance abuse themselves (Langbehn et al., 2003). Prevention research has demonstrated how interventions that directly or indirectly target externalizing problems can effectively prevent drug abuse (Hawkins et al., 2008; Kellam et al., 2008; Poduska et al., 2008), perhaps through impacting underlying neuroregulation systems (Romer and Walker, 2007) and familial processes that enhance behavioral regulation (Brody et al., 2009).

2. Severity measurement instruments

The following briefly reviews instruments that measure addiction propensity and severity (see Table 1). The instruments in this review were identified through a literature search of sev-
Instruments designed to measure addiction severity.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>What it measures</th>
<th>Operationalization of severity</th>
<th>Item-selection methodology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addiction Severity Instrument (ASI)</td>
<td>Severity of alcohol and drug use</td>
<td>Need for treatment across 6 domains</td>
<td>Clinical judgment</td>
</tr>
<tr>
<td>Alcohol Dependence Scale (ADS)</td>
<td>Severity of alcohol dependence</td>
<td>DSM symptomatology, loss of control, obsessive drinking style, two aspects of withdrawal</td>
<td>Clinical judgment, item/factor correlation</td>
</tr>
<tr>
<td>Benzodiazepine Questionnaire (BDEQ)</td>
<td>Severity of benzodiazepine dependence</td>
<td>DSM symptomatology, pleasurable effects of drug, perceived need for drug in order to function properly</td>
<td>Clinical judgment, item/factor correlation</td>
</tr>
<tr>
<td>Chemical Use, Abuse, and Dependence Scale (CUAD)</td>
<td>Severity of alcohol and drug use</td>
<td>DSM symptomatology</td>
<td>Clinical judgment</td>
</tr>
<tr>
<td>Drug Use Screening Inventory (DUSI)</td>
<td>Severity of alcohol and drug use</td>
<td>Consequences of drug use</td>
<td>Clinical judgment</td>
</tr>
<tr>
<td>Global Appraisal of Individual Needs (GAIN)</td>
<td>Severity of alcohol and drug use</td>
<td>DSM symptomatology, substance use frequency, behavioral complexity</td>
<td>Clinical judgment</td>
</tr>
<tr>
<td>Severity of Alcohol Dependence Questionnaire (SADQ)</td>
<td>Severity of alcohol dependence</td>
<td>DSM symptomatology, three aspects of withdrawal, rapidity of reinstatement after abstinence</td>
<td>Clinical judgment, item/factor correlation</td>
</tr>
<tr>
<td>Severity of Amphetamine Dependence Questionnaire (SamDQ)</td>
<td>Severity of amphetamine dependence</td>
<td>DSM symptomatology, three aspects of withdrawal, rapidity of reinstatement after abstinence, depression, lethargy</td>
<td>Clinical judgment, item/factor correlation</td>
</tr>
<tr>
<td>Severity of Dependence Scale (SDS)</td>
<td>Severity of drug dependence</td>
<td>DSM symptomatology, compulsivity of drug use</td>
<td>Clinical judgment, item/factor correlation</td>
</tr>
<tr>
<td>Severity of Opiate Dependence Questionnaire (SODQ)</td>
<td>Severity of opiate dependence</td>
<td>DSM symptomatology, three aspects of withdrawal, rapidity of reinstatement after abstinence</td>
<td>Clinical judgment, item/factor correlation</td>
</tr>
<tr>
<td>Substance Dependence Severity Scale (SDSS)</td>
<td>Severity of drug dependence</td>
<td>DSM symptomatology</td>
<td>Clinical judgment</td>
</tr>
<tr>
<td>Substance Use Involvement Index (SUI)</td>
<td>Severity of alcohol and drug use</td>
<td>DSM symptomatology, liability of substance use</td>
<td>Clinical judgment, item/factor correlation, item-response theory</td>
</tr>
</tbody>
</table>

2.1. Addiction Severity Index (ASI)

The ASI (McLellan et al., 1980) is a widely used instrument that assesses the need for alcohol or drug treatment across six domains of functioning (Chemical Abuse, Medical, Psychological, Legal, Family/Social, and Employment/Support). Within each domain, the interviewer provides global ratings of severity (referred to as Interviewer Severity Ratings, or ISRs) based on the patient's responses to objective items as well as the patient's assessment of how bothered he/she is by problems in each domain. The ISRs, which range from 0 to 9, form the basis for the patient's profile and treatment plan. Revised versions of the ASI (McDermott et al., 1996) use more advanced psychometric methods, partially in response to studies that failed to replicate high inter-rater reliabilities of the original version of the ASI (McLellan et al., 1985; Hodgens and Elguebally, 1992). The ASI is one of the most frequently used instruments in addiction research and practice, and it serves as one of the core measures in the NIDA Clinical Trials Network. The reliability and validity of the ASI is well documented, but less is known about its dimensionality or whether its items reflect differences in severity.

2.2. Drug Use Screening Inventory (DUSI)

The DUSI (Kirisci et al., 1995; Tarter et al., 1990) was developed as a diagnostic instrument for the treatment of alcohol or drug problems in adolescents. Despite the word “screening” in its name, this is a phenotypic assessment instrument comprising a series of scales measuring severity of substance abuse as reflected in various life domains: Substance Use, Behavior Patterns, Health Status, Psychiatric Disorder, Social Skills, Family System, School Adjustment, Work, Peer Relationships, Leisure/Recreation. Kirisci et al. (1995) found that the items are scalable as indicators of a unidimensional latent trait, that the scores among individuals with high or moderate substance use were measured with greater precision than those at lower levels, and that scores accurately predicted the level of substance use and consequences at 6-month follow-up.

2.3. Global Appraisal of Individual Needs (GAIN)

The GAIN (Dennis et al., 2003) is a large set of questionnaires designed to provide diagnostic and severity information for both adolescents and adults seeking treatment for drug or alcohol problems. The GAIN’s 16-item Substance Problem Scale (SPS) scale consists of the DSM-IV criteria for substance use diagnosis plus five screener items that measure variation in severity (i.e., weekly use, family and friends complaining about use, continued use despite fights, and use being time consuming). Acceptable reliability and validity estimates have been reported (1999 manual and Dennis et al., 2006). Riley et al. (2007) found that the five screener items had relatively low location parameters, indicating that the items reflect variation at the lower (or “milder”) end of a severity continuum.

2.4. Severity of Alcohol Dependence Questionnaire (SADQ) and Alcohol Dependence Scale (ADS)

The SADQ (Stockwell et al., 1979) and the ADS (Skinner and Allen, 1982) were explicitly designed to extend the Edwards and

---

1 Considerable research attention has been focused on the severity of nicotine addiction, and results suggest a latent trait for nicotine dependence (Strong et al., 2003a,b, 2007, 2009). Less is known about the severity of alcohol and/or illicit drug addiction, however, which underscores the need for additional research of this kind.
Gross (1976) conceptualization of alcoholism to severity of addiction to illicit drugs. The SADQ includes four subscales: physical withdrawal, affective withdrawal, drinking to relieve withdrawal, and rapidity of reinstatement after abstinence. The ADS includes four slightly broader subscales: psychophysical withdrawal, psychomotor withdrawal, loss of behavioral control, and obsessive drinking style. Reliability and validity estimates were high for both instruments. Kahler et al. (2003a,b) reported that the ADS items primarily measure a single dimension of alcoholism severity, with some items corresponding to greater severity than others.

2.5. Severity of Opioid Dependence Questionnaire (SODQ), the Severity of Amphetamine Questionnaire (SamDQ), and the Benzodiazepine Dependence Questionnaire (BDEPQ)

The SODQ and the SamDQ adapted the SADQ for opioids and amphetamines, respectively. These instruments consist of subscales designed to measure the severity of physical withdrawal symptoms, affective withdrawal symptoms, the extent to which drugs are used to relieve withdrawal symptoms, and the rapidity of reinstatement after periods of abstinence. The BDEPQ (Baillie and Mattick, 1996) assesses aspects of benzodiazepine dependence including the extent to which pleasurable effects were anticipated as a result of drug use, the extent to which drug use is needed to complete daily life activities, and a general dependence factor. The reliability and validity of these instruments are well documented (Baillie and Mattick, 1996; Churchill et al., 1993; Sutherland et al., 1986; Topp and Mattick, 1997), but information is limited regarding dimensionality and whether some items reflect greater severity than others.

2.6. Chemical Use Abuse and Dependence (CUAD)

The CUAD scale (Mcgovern and Morrison, 1992) is a semi-structured interview designed to provide DSM-III-R substance use diagnoses as well as severity indicators in both clinical and research settings. For each substance used, weights based on clinical judgment are assigned for each symptom to reflect increasing clinical severity, and total severity scores are computed by summing the weights across symptoms. The CUAD has demonstrated reliability and validity (Mcgovern and Morrison, 1992), but it is unknown whether it is unidimensional or whether some items reflect greater severity than others.

2.7. Substance Dependence Severity Scale (SDSS)

The SDSS (Miele et al., 2000) is a semi-structured interview designed to provide DSM-IV diagnoses and severity indicators in both clinical and research settings. For each substance used, the SDSS provides DSM-IV diagnoses and four different 30-day severity indicators; two for intensity of symptoms (SEV, WORST SEV), and two for frequency of symptoms (DAYS, WORST DAYS). The reliability and validity of the SDSS is documented, but information is lacking about its dimensionality and ability to assess severity.

2.8. Severity of Dependence Scale (SDS)

The SDS (Gossop et al., 1995) was developed as a brief scale to measure the degree of dependence experienced by different types of drug users. In contrast to the other instruments discussed here, the five-question SDS “is primarily a measure of compulsive use and it does not include items to measure tolerance, withdrawal, or reinstatement” (p. 612). Responses are self-rated on a 4-point Likert scale, and summed to form a total score. Scale reliability and validity has been reported and factor analytic methods have confirmed unidimensionality of the five items.

2.9. Substance Use Involvement Index (SUII)

The Substance Use Involvement Index (SUII; Kirisci et al., 2002), derived using item-response theory (IRT; described below), is based on the respondent’s endorsement of lifetime use (yes/no) of 10 categories of drugs: alcohol, cannabis, cocaine/crack, opiates, amphetamines, methylphenidate, sedatives, tobacco, hallucinogens, PCP and inhalants. One study reported that a unidimensional factor model adequately fit the data for both adult males and adult females, and the factor loadings were not significantly different between gender groups (Kirisci et al., 2002). Interestingly, Kirisci and colleagues also reported that equated item-location parameters were higher in females than males, suggesting that endorsement of substance use by a male is (by itself) indicative of a lower level of severity than the same endorsement by a female. These results are consistent with data showing that females have a higher liability threshold (Kendler et al., 2003b), i.e., they require higher factor scores in order to manifest a disorder.

3. Propensity measurement instruments

Measurement of propensity to addiction has rarely been attempted and is considerably more difficult than measurement of addiction severity, especially given the absence of face-value indicators of risk with high construct validity. Obviously, only premorbid characteristics may be safely used as indicators of propensity because drug use and addiction have behavioral and psychological effects. Despite knowledge of many “risk” and “protection” factors (psychological and psychopathological variables associated with addiction risk) (Glantz et al., 2005; Hawkins et al., 1992), it is not evident on an a priori basis that premorbid indicators can be used, or how, in constructing an index of propensity to addiction.

3.1. Transmissible Liability Index (TLI)

One approach to the selection of propensity indicators has been used in the derivation of the index of transmissible liability to addiction related to illicit drugs (the Transmission Liability Index, TLI), based on a high-risk/family design and IRT in the NIDA-funded Center for Education and Drug Abuse Research (CEDAR). Inasmuch as addiction risk is transmissible within families (mostly due to its high heritability), characteristics that discriminate groups of children with affected versus unaffected parents (high- and low-average-risk groups, HAR and LAR, respectively) are likely to be indicators of the transmissible component of children’s addiction liability/propensity (the variance component correlated between relatives/generations).

The TLI derivation method involves using a large set of items from numerous psychological and psychiatric instruments that were originally selected based on their potential for measuring variables related to addiction risk and psychopathology. Through a variety of statistical methodologies (e.g., factor analysis, IRT), the items were selected and transformed into a set of unidimensional constructs characterizing individual behavior/personality (e.g., antisociality, attention, mood) (see measurement model and other details in Vanyukov et al., 2003a). The resulting 45-item set selected for sons of the probands at age 10–12 was used to assess the quality of items and estimate TLI. The findings show that the TLI is a valid and reliable scale, and highly predictive (e.g., O.R. = 1.81, 95% C.I.: 1.12–2.30) of substance use disorder (Vanyukov et al., 2009; Kirisci et al., 2009). Twin studies have also shown that TLI is highly heritable ($H^2 = 0.79$) (Vanyukov et al., 2009; Hicks et al., in press), further supporting its derivation as a measure of a transmissible trait.
The longitudinal design of the CEDAR study permits further refinement of the method used in the creation of the TLI, such that the liability/propensity indicator constructs and items will be referenced to the offspring’s own, rather than parental, outcomes such as drug use and drug addiction. Such referencing, aside from being based entirely on the offspring’s own phenotype, directly connects indicators to a given manifest trait, which will enable indexing of the resultant total, rather than only transmissible, liability. This extended liability index methodology has potential as a prototype for development of an index to cover the entire range of phenotypic distribution (i.e., both propensity and severity) and identify candidate biologic and genetic mechanisms of drug use and its precursors and antecedents.

4. Summary and future directions

4.1. Limitations of existing instruments

The assessment instruments in use today for measuring addiction severity (described above and appearing in Table 1) have proven utility, reliability, and validity, but they are of limited use for evaluating individual differences in propensity/severity of the hallmark characteristic(s) of drug addiction, i.e., compulsivity in seeking and using drugs despite harmful consequences. First, existing instruments are based on a variety of related but different constructs of addiction severity including behavioral and social consequences, quantity or type of DSM symptoms endorsed, use patterns within and across substances, and number of different DSM diagnoses. Only one, the SDS, focuses on compulsive drug taking and seeking. Second, the content of existing measures is unlikely to reflect the full range of addiction severity. This limitation can be partly attributed to the fact that the content of many instruments derives from the DSM, which virtually by design, is not optimal for measuring variability in the severity of addiction. Indeed, clinical and epidemiologic studies applying IRT models show that DSM symptoms are redundant and fail to capture variability across the full range of the addiction continuum (Compton et al., 2009; Gillespie et al., 2007; Langenbucher et al., 1995; Saha et al., 2007; Wu et al., 2009). Many existing instruments also rely on clinical judgment in the absence of item-selection techniques, thereby increasing the likelihood of including items that are redundant, collinear, or weakly related to the construct of interest.

Third, only one extant instrument (TLI) is currently available to measure propensity to addiction, let alone both the propensity and the severity of addiction on the same metric. It should be noted that the latent liability phenotype assessed by the TLI, while pertaining to the risk in the nonaffected individuals and thereby to prospective prediction, is nevertheless a cross-sectional measure. Substance use or lack thereof, or consumption with multiple graded categories (e.g., none, low, heavy) are observed phenotypic outcomes that may or may not be used as indicators of latent liability. Using items immediately related to drug use may be problematic, of course, because that would preclude the index’s use in studying respective behaviors as outcomes. Although substance use is a necessary condition for addiction development, it seems likely that the psychological/behavioral components of liability such as personality largely influence substance use initiation and/or, more importantly, development of addiction. Substance use can thus be viewed as a manifestation of liability forming an ontogenetic trajectory that can be tracked and measured.

Fourth, most extant instruments were constructed using classical test theory methodology, which has its own psychometric limitations that directly bear on severity measurement. The use of factor analysis, for example, does not utilize the full power of modern scaling techniques (Embretson and Reise, 2000; Hambleton and Swaminathan, 1985; Lord and Novick, 1968) that take item properties into account (e.g., whether one item indicates greater “severity” than another item). While producing scales with high levels of internal consistency, the use of factor analytic methods alone does not allow one to build instruments to discriminate individuals along selected ranges of an underlying trait.

4.2. Advantages of item-response theory (IRT) for measuring propensity and severity

Whereas various methods are available to analyze latent traits, IRT appears particularly suited for the derivation of an index of liability to addiction. Briefly, IRT (e.g., Embretson and Reise, 2000) is a psychometric test theory that relates the performance of an examinee on a test item to a latent trait (ability) that the test is intended to measure. This relationship (e.g., in a simple case, between the trait and the probability of a correct response) is described by an item-response function (IRF). While ability level is a characteristic of the examinee, the examinee’s performance will also depend on parameters that characterize the test items themselves. The widely used two-parameter model includes location (difficulty) or $b$, the trait value at which the probability of a correct response exceeds 0.5, and discrimination or $a$, which is proportional to the slope of the IRF at point $b$ on the trait scale. In this model, the parameters provide for gradients in item difficulty and capacity to discriminate different values of the trait. In contrast to the classical psychometric test theory, which was used to develop most of the instruments reviewed herein, IRT provides testable models. A data-fitting IRT model yields estimates that have unique value for trait measurement; i.e., item parameters are invariant across samples (subpopulations) of subjects (the trait distribution does not influence the estimates), and trait estimates are invariant across items used.

The many advantages of IRT have been underutilized in severity instrument development, though it has successfully been used in the measurement of transmissible addiction risk (Vanyukov et al., 2003a,b, 2009; Kirisci et al., 2006, 2009). Selection of items based on the information they contribute is a key advantage of item-response theory (Hambleton et al., 1991), which is likely to inform the development of the liability index. The most difficult question centers on the problem of selecting and constructing items that are highly informative for measuring low propensity values. Nevertheless, some such items (e.g., “I don’t move around much at all in my sleep”, “I usually eat the same amount each day”, and “Changes in plans make my child restless”) have been identified and applied in an existing propensity measurement instrument, the TLI (Kirisci et al., 2009; Vanyukov et al., 2009). These and other items are potential candidates for an instrument that measures both propensity and severity on the same scale.

How and to what extent the propensity and severity phenotypes can be characterized by a common, underlying, and continuous unidimensional trait of liability to addiction (Chan et al., 2008) remain large, empirical challenges suitable for IRT analyses. Dimension-sharing between propensity and severity need not be perfect for the derivation of a unidimensional index to quantify sub- and suprathreshold liability; showing that a unidimensional model offers the best fit for the data is sufficient. The unidimensionality of a trait essentially refers to the structure of covariance between items used in its measurement (and testable for their dominance in accounting for the covariance) rather than to a single causal influence or the total number of influences potentially determining the shared variance. Even if there is specificity to the unaffected and affected phenotypic distributions (i.e., those of propensity and severity), the variance they share with the common liability dimension, based on genetic data, will likely be large and sufficiently informative to be targeted for measurement. If such tests fail and the two constructs do not share the same dimension to an appre-
ciable degree, continuous indices will need to be developed that quantitate these constructs separately and are applied to the nonaffected (or asymptomatic) and affected (symptomatic) populations. Measurement of common rather than drug-specific liability is also desirable from a practical standpoint, given its universal application.

4.3. Advantages of an instrument measuring propensity to and severity of addiction

The conceptual and methodological shortcomings of instruments currently used in research and clinical practice can be overcome through the use of emerging technologies to develop a reliable, valid, and standardized assessment instrument to measure and distinguish variations in phenotypic expression of the underlying latent trait that comprises propensity and severity of drug addiction. The distribution of the hypothesized latent trait in the population is depicted in Fig. 1. Key requirements of this assessment instrument include the ability to accurately and comprehensively measure gradations along the propensity/severity continuum using broad and discriminating content that captures the essence of addiction: to detect meaningful variation between, within, and across individuals over time that is scalable along the underlying dimension; and to allow for efficient assessment of the construct with minimal burden for administration, training, and cost to the researcher, clinician, research participant, or patient. Coupling modern psychometric methodology with sophisticated computer technology (e.g., IRT-based computerized adaptive testing [CAT]) enables measurement covering the full phenotypic scale, with maximum flexibility, accuracy, and efficiency (Dragswog and Olson-Buchanan, 1999; Kirisci and Hsu, 1992). CAT-based instruments, for example, can administer different sets of items to different people optimal for their liability phenotypes, can provide analyses to explore the relative difficulty of items, and can generate a single score to represent an individual’s location along an underlying trait (the liability phenotype). This technique is routinely applied in the education field (e.g., the administration of the Graduate Record Examination applies IRT analyses), and can certainly be adapted for use in the field of addiction.

Whether propensity to and severity of addiction are locatable on essentially the same dimension remains to be determined, yet the cumulative evidence on neurobiological mechanisms of addiction, combined with psychometric evidence, gives merit to a common construct with minimal burden for administration, training, and cost to the researcher, clinician, research participant, or patient. Coupling modern psychometric methodology with sophisticated computer technology (e.g., IRT-based computerized adaptive testing [CAT]) enables measurement covering the full phenotypic scale, with maximum flexibility, accuracy, and efficiency (Dragswog and Olson-Buchanan, 1999; Kirisci and Hsu, 1992). CAT-based instruments, for example, can administer different sets of items to different people optimal for their liability phenotypes, can provide analyses to explore the relative difficulty of items, and can generate a single score to represent an individual’s location along an underlying trait (the liability phenotype). This technique is routinely applied in the education field (e.g., the administration of the Graduate Record Examination applies IRT analyses), and can certainly be adapted for use in the field of addiction.

Whether propensity to and severity of addiction are locatable on essentially the same dimension remains to be determined, yet the cumulative evidence on neurobiological mechanisms of addiction, combined with psychometric evidence, gives merit to a common construct with minimal burden for administration, training, and cost to the researcher, clinician, research participant, or patient. Coupling modern psychometric methodology with sophisticated computer technology (e.g., IRT-based computerized adaptive testing [CAT]) enables measurement covering the full phenotypic scale, with maximum flexibility, accuracy, and efficiency (Dragswog and Olson-Buchanan, 1999; Kirisci and Hsu, 1992). CAT-based instruments, for example, can administer different sets of items to different people optimal for their liability phenotypes, can provide analyses to explore the relative difficulty of items, and can generate a single score to represent an individual’s location along an underlying trait (the liability phenotype). This technique is routinely applied in the education field (e.g., the administration of the Graduate Record Examination applies IRT analyses), and can certainly be adapted for use in the field of addiction.

Conflict of interest

Author Swan served on the Pfizer National Advisory Board (1 day, 2008) to consult on issues related to varenicline.

Acknowledgements

Role of funding source: NIDA had no role in the analysis, preparation, and writing of the report, or in the decision to submit the paper for publication.

Contributors: Authors Conway took primary responsibility for writing the manuscript drafts. Authors Conway and Levy managed the literature searches and summaries of related work. Authors Levy, Vanyukov, Chandler, Rutter, Swan, and Neale contributed to the manuscript design, writing, and editing of manuscript drafts. All authors contributed to and have approved the final manuscript.

References


