ABSTRACT—A long-standing and diverse body of evidence has documented the importance of externalizing characteristics as very early etiologic predictors of a pathway to severe alcohol and other drug problems and substance use disorder (SUD). At the same time, much remains unclear about the mechanistic structure of this pathway, including understanding what the defining characteristics are that encompass the diverse behaviors included in the externalizing domain. This article proposes that the core risk phenotype unifying this domain is behavioral undercontrol–disinhibition. It describes the defining features of this phenotype and outlines the mediators, moderators, and developmental course that characterize the pathway from early risk to a SUD endpoint. A brief summary of the neurocognitive and brain functional response systems that underlie the behavioral phenotype emphasizes the operation of two systems in dynamic tension, one an effortful control system, the other an incentive reactivity system.

KEYWORDS—externalizing behavior; impulsivity; risk phenotype; effortful control; incentive reactivity

EVIDENCE FOR AN EXTERNALIZING PATHWAY

The view that alcoholism—or in modern terminology, alcohol use disorder (AUD)—is actually a heterogeneous group of disorders has had a long history, but a systematic description of number of types, differentiating characteristics, and developmental course has occurred only within the past quarter century. That work, primarily based on retrospective clinical and epidemiologic survey data, suggested that clinical manifestations, etiology, and prognosis differentiated at least two different “types,” one with externalizing comorbid traits and sometimes referred to as “antisocial alcoholism,” the other without these traits (Babor et al., 1992; Cloninger, Sigvardsson, & Bohman, 1988; Zucker, 1987). Longitudinal studies in more recent years have been able to prospectively confirm the importance of externalizing characteristics as strong etiologic predictors of the first type (see Zucker, 2006, 2008, for reviews of the older studies as well as of more recent ones).

This longitudinal work has fleshed out a long list of factors associated with the pathway, of this type of AUD, including a high-risk genetic background marked by a positive family history for AUD (Goodwin, Schulsinger, Knop, Mednick, & Guze, 1973); antisocial comorbidity in at least one of the parents (Zucker, Ellis, Fitzgerald, Bingham, & Sanford, 1996); a rearing environment characterized by poor parenting (Olson, Bates, & Bayles, 1990; Valiente, Lemery-Chalfant, & Reiser, 2007); exposure to abuse, stress, conflict, and violence (Hussong, Curran, Moffitt, & Caspi, 2008; Sanford, Bingham, & Zucker, 1999); low social competence in early childhood (Garnier & Stein, 2002; Pitkänen, Kokko, Lyra, & Pulkkinen, 2008); involvement with deviant peers (Dielman, Butchart, & Shope, 1993; Dishion, Patterson, Stoolmiller, & Skinner, 1991); and earlier use of alcohol and other drugs (Grant, 1998), with heavier and more problematic use thereafter (Clark et al., 2005; Dubow, Boxer, & Huesmann, 2008; Pitkänen et al., 2008) and increasing likelihood of moving into early onset AUD or some other drug use disorder (Grant & Dawson, 1997, 1998; Hawkins et al., 1997). The robustness of the developmental relations is indicated by the fact that the great majority of these longitudinal studies began in early or middle childhood, and two have even followed the associations out to the mid-40s.
In this article, we summarize the multidomain evidence for the role of externalizing behaviors in the etiology of substance use disorders (SUDs). The focus is on two issues: (a) the mechanistic structure of the externalizing pathway as a complex multilevel developmental system encompassing behavior, neurocognitive and brain functional response systems, and social environment, and (b) the critical mediators and moderators of developmental course. The discussion of these issues is organized into three sections: (a) the core characteristics of the risk phenotype, (b) the epigenesis of risk and heterogeneity of developmental course, and (c) the probable neurocognitive and brain functional response systems and mechanisms underlying the risk phenotype. Although the majority of work covered pertains to risk for AUD, because of the evidence of a common underlying liability for all disorders of drug involvement (Kendler, Prescott, Myers, & Neale, 2003; Krueger et al., 2002), we have retained the generic SUD terminology when possible.

**CORE CHARACTERISTICS OF THE EXTERNALIZING RISK PHENOTYPE FOR SUD**

The heavy research focus on the externalizing behavior cluster has been driven by an interest in identifying the core vulnerability trait, or risk phenotype, that precedes the antisocial behavior and substance abuse associated with adult SUD. To qualify as a candidate for the risk phenotype, the cluster would need to be relatively stable across time and contexts (i.e., it would have trait-like characteristics) and be mechanistically connected to the emergence of the adult characteristics. The externalizing cluster of behaviors in large measure satisfies these criteria. Its content involves both aggressive and delinquent behavior (Achenbach, Howell, Quay, & Conners, 1991; Krueger et al., 2002); the common element is difficulty in control of behavior, not affect. Aggressive behavior includes acts of verbal and physical aggression. Delinquent behavior includes conduct problems such as lying, cheating, stealing, and truancy; that is, it involves rule breaking and acting in opposition to social norms rather than people. The correlation between the two behaviors in childhood and in adolescence is high \(r \approx .7\), reflecting strong but far from perfect overlap (Achenbach & Rescorla, 2003). The externalizing construct is itself a dimensional translation of much of antisocial personality disorder (ASPD) symptomatology, and studies have used both ASPD and SUD symptom measures or diagnoses in combination as an index of the risk phenotype (Kendler et al., 2003; Krueger et al., 2002). At the same time, if difficulty in control of behavior is the critical feature, then the construct descriptor probably should be behavioral undercontrol or behavioral disinhibition.

The issue of labeling is more than a simple dispute about terminology. A careful parsing of these constructs indicates that their subcomponental attributes operate at several levels. Behavioral undercontrol, or behavioral disinhibition, refers to a vulnerability of disinhibitory processes that involves the inability or unwillingness or failure to inhibit behavior even in the face of anticipated or already received negative consequences (Hawkins, Catalano, & Miller, 1992; Kandel, 1978; Zucker, 2006). Antisocial behavior, delinquency, and rule breaking fall at the level of observable behavior. However, the construct also subsumes inferential traits, such as impulsivity, low behavioral constraint, impaired impulse control, as well as lack of control over cravings for food, sex, or drugs. Undercontrolled individuals are more likely to carry out activities that are normatively inappropriate or socially disapproved, such as aggression or open conflict with authority; they also have greater proneness to engage in risky behaviors without weighing the consequences of their actions. At the same time, this trait is distinct from other, closely related constructs. For example, inhibition is “a temperament or style of reacting with fear or withdrawal when confronted with novelty, including both novel situations as well as unfamiliar adults or peers” (Fox, Henderson, Marshall, Nichols, & Ghera, 2005; Kagan & Snidman, 2004). Inhibition is not the obverse of undercontrol.

For these reasons, although the externalizing label is still being utilized by others, we turn to the finer-grained undercontrol or disinhibition labels, which are more descriptive of the underlying trait (McGue, Iacono, Legrand, Malone, & Elkins, 2001; Sher & Trull, 1994; Wong et al., 2006). From this point on, we refer to the core disinhibitory construct as one of behavioral undercontrol–disinhibition. Although it is premature to regard one descriptor as more primary than the other, to differentiate them, we use the term **undercontrol** when the referent is to expression at the behavioral level, and the term **disinhibition** when the referent is to the neurophysiological or neurocognitive indicator (also see Kagan & Snidman, 2004).

Although the undercontrol–disinhibition characterization makes the best sense, the umbrella it covers is a large one. The jury is still out about whether this will remain the best superordinate characterization, whether the domain should include all of the component pieces described earlier, and whether such a superordinate level of analysis is most effective for identifying liability. A number of other groups are actively exploring these issues: Two (Dick et al., 2010; Smith et al., 2007) are focused on parsing the impulsivity construct, albeit in different ways; a third proposes a composite indicator of neurobehavioral disinhibition that cuts across behavioral, neuropsychological, and affective domains (Tarter et al., 2003); a fourth proposes that the core liability trait is one of behavioral disinhibition, indexed behaviorally by externalizing symptomatology, but also neurophysiologically via brain P300 response (see Iacono & Malone, 2011; Iacono, Malone, & McGue, 2008); a fifth, which proposes that the core liability is one of behavioral undercontrol, delineates a personality structure virtually identical to that in Iacono et al.’s (2008) behavioral disinhibition construct but focuses more on personality structure and genetic basis without elaboration of neurophysiological structure (Slutske et al., 2002); and a sixth, which is closest to our work here, posits a hierarchic structural
model involving a behavioral disinhibition higher order factor superordinate to three lower factors of impulsive sensation-seeking, antisociality–unconventionality, and externalizing symptomatology (see Bogg & Finn, 2010; Wills et al., 2001).

**EPIGENESIS: THE DEVELOPMENTAL CASCADE OF RISK, HETEROGENEITY OF DEVELOPMENTAL COURSE, AND POSSIBLE INDIRECT GENETIC EFFECTS ON THE DISINHIBITION PATHWAY**

The brief etiologic summary provided at the beginning of this article enumerates some of the major mediators and moderators of the disinhibitory pathway between early childhood risk and adult disorder. A risk-cumulation additive model would predict that the greater the number of risk factors present, the more likely the SUD endpoint. However, recent work based on long-term, multiwave prospective studies and a dynamic cascade model of risk (cf. Dodge et al., 2009) shows that simple additivity of risk only provides a crude model of risk flow. Timing and sequencing of risk aggregation, density of risk pathway, developmental role demands at outcome, and intermediate risk-offset opportunities are all essential to the probability of a risky outcome (Dubow et al., 2008; Merline, Jager, & Schulenberg, 2008; Pitkänen et al., 2008).

Another major source of imprecision in characterizing developmental course is driven by the preponderant use of methods of analysis that are variable centered as opposed to person centered. If the process of risk development is heterogeneous across the population, then variable-centered analysis will, at best, only crudely characterize intraindividual etiology and developmental course, and, at worst, will provide an erroneous picture. A classic study by Schulenberg, O’Malley, Bachman, Wadsworth, and Johnston (1996) that utilized data on binge-drinking frequency in a nationally representative sample of college-age youth illustrates this point. Their trajectory analysis identified six trajectory classes (never, rare, chronic, decreased, increased, and “fling”) over the 18–24 age interval and showed that the mean trajectory of binge drinking was not reflective of the developmental trajectories of use for any of the classes; furthermore, different predictor patterns identified membership in each of the classes. A similar pattern of developmental heterogeneity has been shown for past-year frequency of marijuana use (Schulenberg et al., 2005). Over the past decade, virtually all of the numerous trajectory class studies of adolescent-to-young-adult problem alcohol and marijuana use have found a high-problem class, along with predictors of class membership that clearly differentiate it from one or more of the lower trajectory classes. Studies of early- to middle-adulthood samples have similarly observed this differentiation well into middle adulthood (Jacob, Bucholz, Sartor, Howell, & Wood, 2005; Jester et al., 2008).

In all of this work, predictors of high-trajectory class membership are also the identifiers of the disinhibition–undercontrol pathway: positive SUD family history, parental heavy drinking or AUD, poor parental monitoring and support, early age of onset of alcohol or other drug use, and exposure to heavy-drinking peers. The developmentally very early appearance of a high-problem class has also been demonstrated in a trajectory class analysis of disinhibitory behavior over the course of childhood (Jester et al., 2005; Jester et al., 2008). This work, which identified a high-disinhibition and disruptive behavior class in a longitudinal study spanning ages 3–17 years, not only provides developmental continuity with the older studies; it also indicates very early detectability of the high undercontrol chronic pathway.

The evidence for heterogeneity in the course of undercontrol as a function of gender differences is equivocal. Although differences in level of undercontrol are virtually universal in the stereotypic direction, gender differentiation in causal structure and relations with other variables is mixed. Differences have sometimes been found in predictive models involving undercontrolled behavior, but degree of effect often varies with age, sometimes disappearing with increasing age over the course of childhood and adolescence (Hussong et al., 2007), and sometimes increasing (Hicks et al., 2007). In other studies, gender differences have either been completely absent or present only in subsidiary parts of the analysis (e.g., Dodge et al., 2009; Slutske et al., 2002). The diversity of these findings suggests that gender differentiation in level of undercontrol does not assure that the relations among variables will also be different.

Characterizations of the developmental course of undercontrol currently remain heavily focused on continuity and cumulation of risk, not on discontinuity, which would involve transitions from higher to lower risk, or from lower to higher risk, at some point. Change in developmental course that occurs in a non-linear trajectory—involving acceleration in risk or problem use at one time or drop off at another—would be one such example of the phenomenon (see, e.g., the Schulenberg et al., 2005; “fling group” trajectory, which involves both such transitions). Currently, the field lacks a framework and a theory to understand these apparent discontinuities in risk level. Statistical characterization of a curvilinear trajectory or a trajectory with a quadratic component does not meet this criterion because the pathway is in reality continuity of a function with known (or determinable) rate of change. What is missing, however, is a conceptual framework that can incorporate exogenous factors not predicted by prior exposure or trajectory vector. Schulenberg, Maggs, and O’Malley’s (2003) concept of developmental discontinuities and disturbances (Zucker et al., 2006) is an early effort to incorporate these effects into the system. Discontinuities can create a permanent shift in a trajectory (disturbance as turning point) or, alternatively, can create a “ripple effect” (disturbance as perturbation). The work of Hussong and colleagues (Hussong, Curran, Moffitt, Caspi, & Carrig, 2004; Hussong et al., 2005) on developmental snare and launch point effects is another early effort to refine the solely continuity model that has dominated the field.
Potential Sources of Indirect Genetic Effect
Although the body of evidence indicating a mediated trajectory class of disinhibitory risk from early childhood to adult SUD outcome is quite strong, two aspects of these relations have been largely neglected. One is that many of the risk factors identified as mediators are nested together in the highest risk families. Thus, families that have the highest probability for transmission of genetic risk (both parents actively alcoholic, at least one with an ASPD diagnosis) are also households where conflict, violence, low educational achievement, parental psychiatric comorbidity, and low socioeconomic status are found (Clark, Cornelius, Wood, & Vanyukov, 2004; Clark et al., 2005; Hussong et al., 2007; Loukas, Zucker, Fitzgerald, & Krull, 2003). The other risk factor is that active parent selection of environments takes place not only in marital assortment (Cornelius, Kirisci, Reynolds, Homish, & Clark, 2008; Windle, 1997) but even in the selection of neighborhoods where the parents’ alcoholic psychopathology is more likely to be sustained (Buu et al., 2007) and where the development of their children’s risk is enhanced (Buu et al., 2009). These relations are all suggestive of substantial gene–environment correlations, that is, genetically influenced individual differences in exposure to risky or protective environments. Moffitt, Caspi, and Rutter (2005) argue that the ideal place to look for gene–environment interactions is where there is both an environmental main effect and a genetic one. This appears to be the case here, many times over. At the same time, because these interactions are ultimately neurobiological, even though the behavioral risks are statistically correlated does not mean that the interactions are with the same genes, or have sites of action that involve the same neural circuitry. The next section is a beginning effort at specifying what the critical circuitry might be.

NEUROCOGNITIVE AND BRAIN FUNCTIONAL RESPONSE SYSTEMS AND THEIR OPERATIONAL RELATIONS TO THE UNDERCONTROL–DISINHIBITION DEVELOPMENTAL PATHWAY

A thorough developmental systems analysis of the undercontrol–disinhibition pathway requires characterization of bidirectional influences across levels of system ranging from the social contextual to the behavioral, to the neural, and to the genetic (Gottlieb, 2003). Fortunately for our purposes here, characterization of cross-system relations pertaining to the undercontrol–disinhibition domain have been central to the field for some time, and a substantial amount of work, both cross-sectional and developmental, now exists on the linkages between behavior, neurocognition, and brain response. (Studies involving genetic relations are not covered.)

The neural underpinnings of control and (dis)inhibition involve regions of prefrontal cortex and their extensive circuitries including regions of thalamus, basal ganglia, and limbic regions. At least two psychological systems and their likely neural imple-

mentation need to be considered (Eisenberg et al., 1997; Nigg, 2000). The systems are in dynamic tension and modulate each other throughout development. One is an effortful control system (Rothbart & Bates, 1998). Effortful control refers to the control of behavior and attention in the service of goals that are distal in time and represented in memory or working memory rather than by immediate incentives and cues. An example would be a child’s ignoring a whispering classmate to earn some privilege later in the hour. Although executive functioning is a broader construct that encompasses other component abilities, this activity is subsumable in the EF domain. Effortful control involves the ability to regulate behavior to fit contextual demands and maintain a goal set (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000; Pennington & Ozonoff, 1996). Effortful control likely reflects activation in prefrontal cortical regions (particularly lateral prefrontal) corresponding with suppression of activity in limbic regions (particularly ventral striatum/nucleus accumbens and possibly amygdala). Although thalamic nuclei, particularly the subthalamic nucleus, also appear to be important, the system can still be regarded heuristically as a “top-down” system.

The second system is an incentive reactivity system; it is relatively automatic in that it does not require mental resources and operates rapidly. It responds to novelty or incentive cues for potential near-term reward or loss by interrupting behavior. Thus, a child may stop talking and actively study a stranger who has entered the room (a cue of novelty or uncertainty; Kagan & Snidman, 2004) or refrain when they see a warning that they are about to lose a privilege (potential loss of reward; Gray & McNaughton, 2000). The incentive reactivity system is distinguishable from primitive appetitive systems, such as hunger or fear, which respond to actual reward or loss rather than to simply their signaled potential. The system stimulates high arousal panic or excitement (or freezing) rather than inhibition of previous behavior, and also activates attentional redirection and inspection of the novel stimulus (see Gray & McNaughton, 2000). The appetitive systems are related to the psychobiology of drug response and addiction after drug ingestion. Conversely, the incentive reactivity system activates excess incentive cue responding under conditions of experienced failure; thus, it is a liability marker for drug problems.

This second construct is very similar to that proposed as reactive control by Eisenberg and Morris (2002). It involves suppression of ongoing behavior prompted by signals in the environment rather than by goals in working memory. The neural process is presumed to be subcortical signaling of novelty or potential threat (e.g., from the nucleus accumbens or amygdala), interrupting programs operating in prefrontal cortex. Therefore, it is heuristically seen as a “bottom-up” process in brain. When this process fails, undercontrolled behavior occurs and the individual goes forward with the behavior despite a signal of potential problems (see Heitzeg, Nigg, Yau, Zubiceta, & Zucker, 2008). Undercontrol may also occur due to overfunctioning of
excitement or approach in limbic regions. Although bottom-up, the behavior would functionally appear the same as failure of the top-down system (i.e., failure to modulate excitement for the sake of goals or other less salient contextual demands).

Existing work suggests that top-down and bottom-up responses mutually influence one another (Nigg, 2000; Rothbart & Bates, 2006) both dynamically and developmentally. Thus, incentive response also involves decision-making, but in the context of potential reward or its loss. Imaging studies of healthy adults indicate that the selection of larger later rewards involves dorsolateral prefrontal activation, whereas the selection of smaller, more immediate rewards involves ventral striatum and amygdala activation (Ballard & Knutson, 2009; Hariri et al., 2000; McClure, Lailson, Loewenstein, & Cohen, 2004). With drug use, these systems are perturbed by excitation of incentive response circuits, which can overwhelm weakened control functions (Goldstein & Volkow, 2002; Jentsch & Taylor, 1999) and appear phenotypically as behavioral undercontrol (ultimately leading to more drug use).

Developmentally, just when there is major build-up of opportunity for substance use in later adolescence, neural alterations in both frontal “control” system and subcortical “incentive” system are taking place. The dorsolateral prefrontal cortex is one of the last brain regions to mature, with myelogenesis continuing into early adulthood (Benes, 2001; Gogtay et al., 2004). Conversely, limbic and striatal systems are mature and responding to cues during adolescence (Galvan et al., 2006). Overall, the relatively early maturation of subcortical activation systems compared with prefrontal control systems may bias adolescent motivation toward immediate over long-term reward (i.e., disinhibition). How this risk may work during childhood remains to be discovered.

The development of these systems can be perturbed in multiple ways. Early biological insult or genetic risk may interfere with development. Caregiver scaffolding also plays a crucial role in development of both control systems (for reviews, see Eisenberg & Morris, 2002; Nigg, Hinshaw, & Huang-Pollack, 2000; Rothbart & Bates, 1998, 2006). Further, high stress exposure in early development has lasting effects on these brain and neurochemical systems (Braun, Lange, Metzger, & Poergel, 2000; Bremner & Vermetten, 2001), as well as increasing the likelihood of substance abuse and externalizing problems in adolescence (Denno, Dertke, Bordens, Washburn, & Schneiderl, 1988; Harrison, Fulkerson, & Beebe, 1997). Although much of the work in this area has established only one or another facet of these relations, recent evidence indicates that insults to these systems are detectable simultaneously in neural activation patterns, behavior, and SUD risk, Heitzeg et al. (2008) have been able to identify dysregulation of reward-related circuitry as well as prefrontal control in adolescents at high risk for SUD. Moreover, dysregulation was correlated with externalizing behavior problems. Within the context of high-risk populations, these findings suggest that pre-existing dysregulation of this circuitry from a number of different sources is an early part of the behavioral undercontrol risk pathway. Once substance abuse begins, the system is further perturbed (Heitzeg, Nigg, Yau, Zucker, & Zubieta, 2010), presumably leading to further undercontrol.

In summary, there is ample evidence of important development in regulatory abilities and the neural networks that support these abilities during the period from childhood to early adolescence, as well as into early adulthood. Experiential moderators during this period also affect network development. Alcohol consumption in adolescence may also impede development, but current evidence is insufficient to regard this as a firm conclusion. The current behavioral and imaging data also converge in suggesting that frontostratial-thalamic and limbic networks involved in regulatory control and motivation are also involved in key temperament and personality domains bearing on the development of undercontrol and the regulation of SUD risk. This work strongly suggests that there will be individual differences in neural network functionality between high- and low-undercontrol individuals. This hypothesis needs to be evaluated.

SUMMARY

A convergent network of evidence continues to indicate that one of the core risk pathways to SUD involves a vulnerability to disinhibitory processes. The vulnerability is expressed at the behavioral level by high undercontrol. It is exacerbated by social environmental factors likely to be highly aggregated in high-risk families, and it is regulated at the neural level by two systems, one involving effortful control, largely localized in prefrontal cortical circuitry, and the other involving incentive reactivity, localized in subcortical circuitry. These systems mature at different rates over the course of adolescence and early adulthood; they are also affected by environmental exposure to stress, family conflict, and possibly by precocious alcohol and other drug use. Work to date is suggestive of a considerable synergy across levels of system, and over time for the highest risk subset of the population. However, the demonstration of such relations is still in its infancy.

REFERENCES


Hawkins, J., Catalano, R., & Miller, J. (1992). Risk and protective factors for alcohol and other drug problems in adolescence and...


