Published in final edited form as:

J Abnorm Child Psychol. 2015 August ; 43(6): 1011–1023. doi:10.1007/s10802-014-9950-1.

Punishment insensitivity in early childhood: A developmental, dimensional approach

Sara R. Nichols, PhD¹, Margaret Briggs-Gowan, PhD², Ryne Estabrook, PhD¹, James Burns, MS¹, Jacqueline Kestler, MPH¹, Grace Berman, BA¹, David Henry, PhD³, and Lauren Wakschlag, PhD^{1,4}

¹Department of Medical Social Sciences, Northwestern University, Chicago, IL

²Department of Psychiatry, University of Connecticut, Farmington, CT

³Institute for Health Research and Policy, University of Illinois at Chicago, Chicago, IL

⁴Institute for Policy Research, Northwestern University, Evanston, IL

Abstract

Impairment in learning from punishment ("punishment insensitivity") is an established feature of severe antisocial behavior in adults and youth but it has not been well studied as a developmental phenomenon. In early childhood, differentiating a normal:abnormal spectrum of punishment insensitivity is key for distinguishing normative misbehavior from atypical manifestations. This study employed a novel measure, the Multidimensional Assessment Profile of Disruptive Behavior (MAPDB), to examine the distribution, dimensionality, and external validity of punishment insensitivity in a large, demographically diverse community sample of preschoolers (three-five years) recruited from pediatric clinics (N=1,855). Caregivers completed surveys from which a seven-item Punishment Insensitivity scale was derived. Findings indicated that Punishment Insensitivity behaviors are relatively common in young children, with at least 50% of preschoolers exhibiting them sometimes. Item response theory analyses revealed a Punishment Insensitivity spectrum. Items varied along a severity continuum: most items needed to occur "Often" in order to be severe and behaviors that were qualitatively atypical or intense were more severe. Although there were item-level differences across sociodemographic groups, these were small. Construct, convergent, and divergent validity were demonstrated via association to low concern for others and noncompliance, motivational regulation, and a disruptive family context. Incremental clinical utility was demonstrated in relation to impairment. Early childhood punishment insensitivity varies along a severity continuum and is atypical when it predominates. Implications for understanding the phenomenology of emergent disruptive behavior are discussed.

Punishment insensitivity has been linked to severe antisocial behavior, particularly psychopathy, but has received scant attention as a developmental phenomenon in relation to emergent psychopathology (Dadds & Salmon, 2003; Lykken, 1957). In part, this is because the extreme (criminal) behaviors associated with psychopathy are not easily translated to developmentally-meaningful terms for young children. Within a developmental framework,

we have suggested that the marked deficits in empathy, internalization of rules, and socio-moral function that mark psychopathy can be understood developmentally as *reduced responsiveness to socialization* with two components: *low concern for others* and *punishment insensitivity* (Briggs-Gowan et al., 2013; Wakschlag et al., 2014). Low concern for others reflects callous disregard of others' needs and feelings and has been well studied in pediatric populations (Frick 2012; Wakschlag et al., 2014). Recently, reliability and validity has been demonstrated in preschoolers, suggesting that behaviors previously considered in their extreme form and/or in adults or adolescents have earlier developmental expression (Ezpeleta, de la Osa, Granero, Penelo, & Domenech, 2013; Hyde, 2013; Kimonis et al., 2006; Willoughby, Waschbusch, Moore, & Proper, 2011). Punishment insensitivity has been defined as lack of behavioral response to the presentation of a punishment or aversive stimulus designed to change behavior (Dadds & Salmon, 2003). It reflects failure to learn from punishment (rather than lack of empathy or insensitivity to feelings per se). It is theorized as decrements in internalization of rules and ability to inhibit prohibited behavior (Dadds & Salmon, 2003). As with low concern, deficits found in punishment insensitivity may reflect problems processing parental socialization cues, which rely heavily on expression of disappointment and anger when children do not comply (Kochanska & Aksan, 2004). Thus, these two components of reduced responsiveness to socialization are theoretically linked via failures to adaptively respond to the emotions and cues of other. In contrast to callous disregard of others' feelings, developmentally-based studies of punishment insensitivity have been lacking. This is the focus of the present paper.

Applying a developmental lens to punishment insensitivity is important for two reasons: (1) *There is increasing evidence that severe antisocial behavior has roots in early childhood* (Moffitt & Caspi, 2001; Blair, 2001). If punishment insensitivity plays a prominent role in developmental pathways to antisocial behavior, identifying its early expression would allow for more effective early identification. Moreover, difficulty learning from punishment weakens the effectiveness of standard disruptive behavior treatments. In particular, individuals with punishment insensitivity are less responsive to standardized parenting training interventions because they are poorer at learning from alteration of environmental contingencies for reward and punishment (Matthys, Vanderschuren, Schutter, & Lochman, 2012). Thus, pairing early identification with targeted treatments specific to this subgroup may be crucial for altering chronic trajectories. (2) *Punishment insensitivity has differentiated neurocognitive correlates*. In particular, it has been differentially associated with deficits in reinforcement learning (Matthys, Vanderschuren, & Schutter, 2013; Finger et al., 2011; Newman & Kosson, 1986; White et al., 2013). This suggests its specificity and the importance of considering its unique contribution to the development of severe antisocial behavior.

Examining punishment insensitivity as a developmental phenomenon

There has been substantial progress in characterization of disruptive behavior disorders and syndromes (DBDs) in preschool children over the past decade (Wakschlag, Tolan, & Leventhal, 2010). Methods specifically designed to differentiate the normative misbehavior of this age period from emergence of clinical problems have been developed and validated (Egger & Angold, 2004; Wakschlag et al., 2008; Wakschlag et al., 2014). This has enabled

the application of dimensional approaches that model behavior along a normal to abnormal spectrum, improving the ability to capture patterns of atypicality in their nascent stages (Wakschlag et al., 2012). In particular, multidimensional approaches to DBDs have been (Wakschlag et al., 2012). In our own prior work, this has focused on validation and replication of a four dimension model of established features of disruptive behavior (aggression, noncompliance, temper loss and low concern for others; Wakschlag et al., 2012). In a prior phase of study, we applied Item Response Theory to model the severity continuum of each of the dimensions (Wakschlag et al., 2014). Item Response Theory (IRT) is a statistical method that can be leveraged as a means of examining the empirical boundary between typical and atypical behavior because it evaluates response patterns along a latent severity continuum (Reise & Waller, 2009). The present study extends this IRT approach by applying it to punishment insensitivity, a behavioral feature that has not been examined in prior dimensional work with young children.

Developmental correlates of punishment insensitivity

Although punishment insensitivity per se has not been studied in early childhood, Dadds and Salmon (2003) theorize behavioral, regulatory, and family correlates within a developmental framework. Related behavioral constructs. Callous/unemotional behavior (insensitivity) and noncompliance with rules and directions (poor internalization) are conceptually linked to punishment insensitivity. Developmental methods designed to differentiate normative vs. atypical expressions of these behaviors are important validators of the punishment insensitivity construct in early childhood. Motivational regulation of behavior. Motivational systems that underlie regulation of behavior and affect may be impaired. This could reflect an imbalance in the behavioral inhibition and behavioral activation systems (BIS/BAS) theorized by (Gray, 1987; 1990), and Zuckerman's impulsive-sensation seeking personality dimension (Zuckerman, 2012). In particular, the BIS is theorized to be sensitive to signals of punishment and novelty and to inhibit behavior that may lead to negative outcomes, whereas the BAS motivates activation towards goals. Young children high on punishment insensitivity may be impaired in their capacity to learn from environmental cues that indicate when behavior should be inhibited (e.g., low fear, self-control), and to exhibit strong reactions when goals are blocked (e.g., temper loss, aggression). Learning to refrain from misbehavior is based on aversive conditioning (e.g., associating misbehavior with fear of punishment or causing another's distress; Matthys et al., 2013). Further, punishment insensitivity is closely linked to Cloninger's reward dependence construct which has been assessed reliably in early childhood (Constantino, Cloninger, Clarke, Hashemi, & Przybeck, 2002). Social attachment and interest in others' approval is a substrate of internalization (Kochanska & Aksan, 2006), and early childhood reward dependence predicts later delinquency (Tremblay, Pihl, Vitaro, & Dobkin, 1994). Neurocognitively, adults and older youth with psychopathic traits have shown systematic deficits in responsiveness to reward and punishment. Multiple studies have demonstrated that those with callous/unemotional traits have decrements in passive avoidance (i.e., the ability to avoid stimuli that have been associated with punishment; Matthys et al., 2013; Frick & White, 2008). In prior work in this sample, we have demonstrated these same neurocognitive decrements in preschoolers high on *Punishment Insensitivity* (Briggs-Gowan et al., 2013). Thus, further developmental

examination of the spectrum of punishment insensitivity, particularly distinguishing its typical from atypical features, may yield insights that are important for continued refinement of effective treatments and early identification of early-emerging DBDs. Disruptive and chaotic family context. The family environment, particularly the extent to which rules and discipline are delivered in a clear, consistent and contingent manner, critically influences the effectiveness of punishment (Dadds & Salmon, 2003). A key aspect of this is regularity and contingency of environmental cues (Chase-Lansdale, Wakschlag, & Brooks-Gunn, 1995). Young children in angry, chaotic, and inconsistent family environments are likely to be insensitive to punishment, since responses to their behavior or misbehavior are unpredictable and harsh (Dadds & Salmon, 2003; Hyde et al., 2010; Waller et al., 2012; Waller, Gardner, & Hyde, 2013)

To validate the punishment insensitivity construct in early childhood, the present work utilizes a developmental psychopathology framework. Specifically, atypical behaviors are conceptualized as severe manifestations of a spectrum that also includes misbehaviors that are normative within a developmental period (Wakschlag et al., 2010). Behaviors that are common but not frequent and/or occur in mild form are conceptualized as normative. In contrast, behaviors that are recalcitrant to environmental contingencies, pervasive, and highly resistant to change are conceptualized as atypical. Prior work has identified these spectra in other aspects of early disruptive behavior, but not punishment insensitivity (Wakschlag et al., 2014). Here we examine the developmental spectrum of punishment insensitivity in a community sample of preschoolers utilizing a new *Punishment Insensitivity* scale of the Multidimensional Assessment Profile of Disruptive Behavior (MAP-DB^a)¹; (Wakschlag et al., 2014). In particular, we test whether *Punishment Insensitivity* in preschoolers manifests along a severity continuum.

Specific Aims

Aim I: Model the normal:abnormal spectrum of *Punishment Insensitivity*

Hypothesis IA: Item-level frequency distributions will distinguish normative from atypical punishment insensitivity behaviors.

Hypothesis IB: Punishment Insensitivity will demonstrate a unidimensional structure.

Aim II: Test for differences in *Punishment Insensitivity* by child age, sex, ethnicity, and poverty status

Hypothesis II. Although there will be some item-level variation, model fit for Punishment Insensitivity as a whole will be invariant across subgroups.

Aim III: Demonstrate the external validity of *Punishment Insensitivity*

Hypothesis III. Punishment Insensitivity will demonstrate construct and convergent/divergent validity and incremental clinical utility.

^aThis scale was originally titled the “Multidimensional Assessment of *Preschool* Disruptive Behavior” but has been since renamed to reflect its use and validation across a broader age range.

Methods

Participants

The Multidimensional Assessment of Preschoolers (MAPS) Study is comprised of a large, diverse sample of preschoolers recruited from the waiting rooms of multiple pediatric clinics in a large, U.S. urban area. Two samples of children were recruited. Phase I was designed to calibrate the MAP-DB questionnaire (N=1,516). Phase II was a similar but independent sample. Parents completed a “replication” survey of the MAP-DB (N=1,855). The current study reports on the Phase II sample because its replication survey provided greater coverage of the *Punishment Insensitivity* construct via an expanded set of items.

Survey Sample

Participant eligibility was established through a brief set of screening questions asking about their child’s age (3–5 years old), being the child’s legal guardian, parental ability to participate in English or Spanish, and not having already participated in the prior phase with the target child or a sibling. All parents accompanied by young children were approached (N = 4,329). Of these, 2,331 were eligible (ineligibility was primarily due to being out of the target age range, n = 1,031), 2,056 (88.2%) consented to participate and 1,903 (81.6% of all eligible) completed surveys.

Forty-five children were excluded from analyses due to significant developmental delays reported by the parent in the survey: 33 with autism spectrum disorders and 12 with general developmental delays. Three children were excluded due to incomplete data. The resultant analytic sample on which the IRT modeling of the Punishment Insensitivity scale was 1,855. Virtually all (97.5%) participants were biological parents and 92.3% were mothers. Participants were fairly evenly distributed by child sex (51.3% girls, 48.7% boys), age (.5% 2-year olds, 40.0% 3-year olds, 37.1% 4-year olds, 22.4% 5-year olds), race/ethnicity (42.4% African American, 32.0% Hispanic, 23.9% Non-Hispanic White, 1.7% Other), and poverty status determined using federal poverty guidelines based on annual household income and household size (44.7% poor, 55.3% non-poor).

Validation Sub-Sample

A stratified, random sample was drawn from the Phase II sample to participate in an intensive laboratory-based protocol to assess correlates and mechanisms of emergent psychopathology pathways. Families were eligible if the parent who completed the initial survey was the child’s biological mother and was able to participate in English. To exclude children with developmental delays who would be unlikely to complete laboratory-based assessments, children whose mothers reported significant delays were not eligible. Delays were defined as autism spectrum disorder, currently receiving services for global cognitive delays, or language skill in the two-year-old range or lower on a brief language screener and receiving services. By design, the sub-study was oversampled for disruptive behavior and intimate partner violence. A total of 746 preschoolers were sampled: 54 high disruptive behavior and violence exposed, 243 high on disruptive behavior only, 104 violence exposed only, and 345 with neither. Of the 746, 504 participated (67.6% response). Seven additional children were later excluded due to information obtained at the visit, indicating delays (e.g.,

autism, spina bifida) and one child due to language barriers. The resultant analytic sample for the validation portion of the study was 496 (36 disruptive and violence exposed, 170 disruptive only, 73 violence exposed only, 217 neither).

Participants in the sub-sample versus those who were eligible did not participate were similar on MAP-DB scores, violence exposure, child age and sex, and family structure (t s ranged from -1.87 to 1.88 , X^2 from 0.003 to 1.98 , all $ps > .05$). However, participants were more likely to have completed high school (75.2% vs. 65.6%, $\chi^2 = 7.36$, $p < .01$) and more likely to be living in poverty (46.4% vs. 34.9%, $X^2 = 57.91$, $p < .0001$) than non-participants. Participants and nonparticipants also differed in ethnicity (47% vs. 33.9% African American, 30% vs. 22.7% Hispanic, 21.4% vs. 41.3% Caucasian, and 1.6% vs. 2.1% Other, $\chi^2 = 33.06$, $p < .0001$). Finally, test-retest data were available from a demographically comparable subsample of 107 sub-study participants who completed the lab packet including the MAP-DB within 90 days of the original replication survey.

Procedures

Parental consent was obtained for all participants and all study procedures were approved by institutional review boards at three universities. *Punishment Insensitivity* and some of the correlates (inconsistent parenting, MAP-DB dimensions, ITSEA scales) were assessed within the replication survey. The remaining correlates were assessed by maternal report on questionnaires completed during the intensive sub-study visit.

Measures

Mothers completed the MAP-DB, a parent-report questionnaire designed to differentiate normative misbehavior from atypical patterns along dimensions reflecting core features of disruptive behavior (Wakschlag et al., 2014). *Punishment Insensitivity* was assessed as a component of a broader MAP-DB construct reflecting reduced responsiveness to socialization. Due to its different response scale, *Punishment Insensitivity* was not included in previous multidimensional modeling of the MAP-DB (Wakschlag et al., 2014). *Punishment Insensitivity* was conceptualized along a continuum reflecting the emerging capacities of young children to respond to discipline scale development was based on an iterative process that included literature review of extant measures at different age periods, clinical experience with young children, piloting, and expert opinion (for details of MAP-DB development, see Wakschlag et al., 2014). *Punishment Insensitivity* was assessed with 13 questions about how the child responded when “punished or disciplined after misbehaving” in the past month. Items were rated in terms of the proportion of time the child behaved the way described when punishment was administered, using a six-point scale (0=Never, 1=Hardly Ever, 2=Sometimes, 3=Often, 4=Most of the Time, 5=All of the Time). Sample items include “Refuse to apologize, not care when punished,” and “Act like rules didn’t matter.” In preliminary data from this sample, a briefer version of the Punishment Insensitivity scale demonstrated associations to preschoolers’ performance on a computerized passive avoidance learning task, with high Punishment Insensitivity associated with poorer learning from punishment (Briggs-Gowan et al., 2013). The item set was enhanced for the present study.

Related Behavioral Constructs

Dimensions of the MAP-DB were used as construct validators to tap into callous disregard for others' feelings and lack of internalization of rules. As an additional validator, the Inventory of Callous/Unemotional Traits (ICU; Frick & White, 2008), the most widely used pediatric measure of callous behavior, was administered to a subset of children in a longitudinal follow-up visit approximately seven months later ($n=161$) (see below).

Callousness and low concern for others—*Low concern for others* was assessed with the MAP-DB. The MAP-DB *Low Concern for Others* dimension assesses disregard for others' feelings and pleasure in others' distress (e.g., "enjoy making others mad," "act like didn't care when someone felt bad or sad"; nine items, $\alpha=.92$). Ratings are based on past month frequencies: 0=*Never*; 1=*Rarely* (less than once per week); 2=*Some* (1–3) days of the week; 3=*Most* (4–6) days of the week; 4=*Every* day of the week; 5=*Many times each day*.

Callousness was assessed with the ICU. The 11-item *Callousness* subscale assesses indifference to others (e.g., "does not care who s/he hurts to get what wants," "feelings of others are unimportant to him/her," $\alpha=.71$). Behaviors are rated on a four-point scale (i.e., not at all true to definitely true).

Internalization of rules—As an indicator of convergent validity, the MAP-DB *Noncompliance* dimension was used. This assesses defiance and intransigent disobedience (ranging from normative refusal to follow directions to provocative disobedience), and inflexible behavior, (ranging from a "reflexive no" to argumentativeness) (22 items, $\alpha=.96$). These scales have been previously described in detail (Wakschlag et al., 2014). *Sneakiness* is an 11-item MAP-DB scale tapping into lack of internalization of rules manifest in covert behavior (e.g., disobedience when no one is looking, hiding misbehavior, and cheating, $\alpha=.91$).

Divergent validity—To establish divergent validity, constructs theorized to be negatively related to *Punishment Insensitivity* were assessed. The eight-item ICU *Caring* subscale assessed sensitivity to others and internalized standards as manifest by concern for others' feelings and motivation to "do his/her best" ($\alpha=.84$). *Social attachment* was assessed with an eight-item subscale from the Devereux Early Childhood Assessment (DECA; LeBuffe & Naglieri, 1999), which assesses social initiative and attachment over the past month on a four-point scale from never – very frequently (e.g., "is comforted by adult," "shows affection"; $\alpha=.82$). Two subscales were used from a preschool version of the Infant-Toddler Social Emotional Assessment (ITSEA; Carter & Briggs-Gowan, 2006), which was adapted for use with three to five-year-olds. ITSEA

Empathy (e.g., "tries to make up for misbehaving," "is aware of others feelings"; seven items, $\alpha=.77$) and ITSEA *Prosocial Peer* (e.g., "plays nicely," "takes turns"; six items, $\alpha=.78$). Ratings on the ITSEA are made on a three-point scale (not true/rarely to very true/often).

Motivational Regulation Constructs

Convergent validity—The MAP-DB *Aggression* and *Temper Loss* scales were used. The 25-item *Aggression* dimension assesses physical, verbal, and relational aggression ranging in severity (e.g., “aggressive when frustrated,” “hurts someone on purpose,” $\alpha=.95$). The 22-item *Temper Loss* dimension assesses irritable mood and tantrums (e.g., “has tantrums during daily routines,” “stays angry for a long time,” $\alpha=.97$). The preschool version of the ITSEA (Carter & Briggs-Gowan, 2006) was used to assess *Activity/Impulsivity* (e.g., “is restless and can’t sit still”; $\alpha=.79$).

Divergent validity—*Self-control* was assessed with a seven-item subscale of the Social Skills Improvement System (SSIS; Gresham & Elliot, 2008), which assesses the child’s ability to inhibit behavior in high motivation contexts (e.g., “resolves disagreements with you calmly, responds appropriately when pushed or hit”) on a four-point scale from never to almost always ($\alpha=.84$). *Fearfulness* was assessed via the ITSEA Inhibition to Novelty subscales (e.g., “is quiet or less active in new situations,” “is shy”; $\alpha=.79$).

Disruptive and chaotic family context—Three measures were used to assess domains of the family environment theorized to be linked to punishment insensitivity (Dadds & Salmon, 2003). The six-item Confusion, Hubbub, and Order Scale (CHAOS; Coldwell, Pike, & Dunn, 2006; Matheny, Wachs, Ludwig, & Philips, 1995) was used to assess the lack of household regularity and structured routines (e.g., “it’s a real zoo in our home,” “we have regular routines”). Items are rated on a five-point scale from definitely untrue to true ($\alpha=.56$). Maternal *inconsistency* was assessed via a four-item subscale of the Parenting Practices Scale (e.g., “threaten but not punish,” “not punish because child talked you out of it”), rated along a six-point scale from never to always ($\alpha=.59$) (Gorman-Smith, Tolan, Zelli, & Huesmann, 1996). Maternal *irritability* was assessed with the PROMIS Anger scale (Cella et al., 2007). Mothers indicated how often over the past week they experienced an irritable and angry mood (e.g., “ready to explode,” “stayed angry for a long time”) rated on a five-point scale from never – to always ($\alpha=.93$).

Measures for Incremental Clinical Utility Analyses

Impairment—Four indicators of child impairment were assessed: (1) Parental worry about the child, which has been shown to be a strong predictor of help-seeking for social-emotional/behavioral problems in young children (Ellingson, Briggs-Gowan, Carter, & Horwitz, 2004). Parents were asked to rate: (1) how worried they were about their “child’s behavior, for example, temper tantrums, aggressive behavior, or being disobedient?” on a four-point scale from not at all to very worried. This item was dichotomized to compare little or no worry (1 or 2) vs. some or very worried (3 or 4) (13.7%); (2) Whether the child had ever been evaluated or received services for behavioral or emotional problems (5.6%); (3) Expulsion from preschool or day care due to problem behaviors (5.6%); and (4) functional impairment on the Children’s Global Assessment Scale (CGAS; Shaffer et al., 1983). Scores < 60 on the CGAS are considered impaired (3.8%).

Disruptive behavior—A summary disruptive behavior score was generated from the MAPDB *Aggression*, *Noncompliance*, *Temper Loss* and *Low Concern* dimensions (78 items,

$\alpha=.98$) (Wakschlag et al., 2014). This score was used as a covariate for incremental utility regressions.

Analytic Approach

The primary analytic tool used in this paper is the Graded Response Model (GRM; Samejima, 1969). This is one of several common item response models used to assess the continuous dimensions underlying ordinal or categorical data (Lord, 1980; Baker, 2001; Embretson & Reise, 2000; Hambleton, Swaminathan, & Rogers, 1991). GRM and related approaches estimate the difficulty and discrimination of each item and place items and people on comparable scales. Thus, those with a given score on the punishment insensitivity trait are more likely to endorse symptoms with difficulties below their score and less likely to endorse symptoms above their trait score. GRM fit using IRTPRO (Cai, du Toit & Thissen, 2011) were used to examine Aim I, testing whether *Punishment Insensitivity* fit a unidimensional severity continuum and examining the relative severity of items along that hypothesized dimension.

Aim II questions concerning model fit across sociodemographic strata were examined by comparing the items as set using multiple group structural equation models, and item-level effects assessed through tests of differential item functioning (DIF). For Aim III, external validity was examined via correlation to a range of similar and dissimilar constructs as well as incremental utility for explaining variance in impairment over and above established features of disruptive behavior. Convergent validity was evaluated by examining correlations between *Punishment Insensitivity* and constructs hypothesized to be related to it. To establish divergent validity, constructs theorized to be negatively related or unrelated to *Punishment Insensitivity* were assessed. We assessed incremental clinical utility by entering *Punishment Insensitivity* in stepwise logistic regression models predicting measures of child impairment with disruptive behavior already in the model.

Results

The full pool of 13 *Punishment Insensitivity* items was reviewed through IRTPRO to identify items with problems of local dependency (defined as dependency among items that is not related to the construct being measured and assessed by a standardized local dependency chi-square ≥ 20) or high residual correlations (Chen & Thissen, 1997; Cai, Yang, & du Toit, 2011).

Six items were dropped due to problems of local dependence: *Not seem sorry after getting into trouble*; *Act sorry after misbehaving*; *Change his or her behavior after being punished*; *Blame someone else for something s/he did wrong*; *Keep doing what s/he was doing after being told to stop*; and *Act innocent when 'caught in the act*. None of the remaining items had high residuals, resulting in a final seven-item model (see Table 1; $\alpha=.89$; test-retest ICC=.78). These seven items showed strong evidence of unidimensionality based on an eigenvalue decomposition of their covariance structure, with a strong first eigenvalue ($\lambda_1=4.626$), all other eigenvalues below zero ($\lambda_2=.661$), and a ratio of the first to the second eigenvalues exceeding Lord's recommendation of 4:1 ($\lambda_1/\lambda_2=6.998$); (Embretson & Reise, 2000; Lord, 1980). In a sample of 107 mothers who completed the MAP-DB twice in 90

days, *Punishment Insensitivity* demonstrated good test-retest reliability (intra-class correlation; ICC=.76). In a sample of 236 mothers who completed the MAP-DB 3 times over a two year period ($M=21.3$ months; $SD=9.5$ months), *Punishment Insensitivity* demonstrated good longitudinal stability (ICC=.68).

Hypothesis 1A

Frequency distributions will distinguish normative vs atypical behaviors. Item distributions supported this hypothesis (Table 1). Most of the behaviors (71%) were exhibited by more than half of the children at some time in the previous month. More specifically, approximately 70% of children were reported to sometimes *Deny doing something not allowed*, and *Act like s/he didn't hear you when you said no*. In comparison, items such as *Not care when punished* and *Act like rules didn't matter* were comparatively rare, with fewer than 31% of children reported to ever have engaged in these behaviors in the past month. Notably, across all sociodemographic groups, very few caregivers (5%) indicated that their child displayed punishment insensitivity behaviors *Most* or *All of the time*.

Hypothesis 1B

A unidimensional severity structure will be demonstrated. We modeled dimensionality to construct a scale measuring the underlying *PI* trait using IRT, using GRM (Samejima, 1969) with IRTPRO. Individual trait scores were obtained by applying parameter estimates to item responses. Scores were generated from IRTPRO with a mean of zero and a standard deviation of one. Item slopes indicate how strongly individual items are related to the *Punishment Insensitivity* dimension. As shown in Table 2 and Figure 1, item slopes ranged substantially. Figure 1 graphically represents the amount of information about the punishment insensitivity latent trait that each item provides. As can be seen, Item three provided the maximum information (“Act like rules didn’t matter,” slope=4.29) and Item two provided the least information (Deny he or she did something that was not allowed, slope=1.43). Scores for each person on the trait are generated as a part of fitting this model and are used as our measure of *Punishment Insensitivity* going forward. The test information curve for this scale (Online Resource, Figure S1) reveals that *Punishment Insensitivity* provides more information at higher levels of the trait, meaning it is most differentiating for more severe levels of behavior.

As shown in Table 1, many of the *Punishment Insensitivity* behaviors are “normative” (i.e., displayed by more than half of preschool children at least once in the preceding month). This highlights the importance of disentangling normative manifestations from more unusual presentations via establishing frequency thresholds at which they become severe. For example, some behaviors may be normative when they occur occasionally, but indicative of more severe *Punishment Insensitivity* when they are the child’s predominant response to discipline. The category thresholds (b_1 – b_5) indicate the underlying severity level at which the transition from one response category to the next is likely to take place (e.g., from “often” to “most of the time”). These thresholds help to pinpoint the response categories at which an item tends to be associated with underlying severity on the overall latent *Punishment Insensitivity* dimension, defined as scores 95th percentile or t-score of 65 (category thresholds >1.51). These scores are bolded in Table 2. For each item, the first

bolded response category indicates the point at which the behavior starts to be associated with high underlying severity. When this point occurs at the higher end of the response scale, an item tends to be associated with severity only when it occurs frequently. For example, *Act like he or she didn't hear you when you said 'no'* is associated with high severity when it occurs *Most of the time* or *All the time*. In contrast, *Keep on misbehaving no matter what you do* is associated with severity when it occurs *Often*.

IRT also provides insight into the ordering of items along a dimensional severity continuum. The item location (theta) is a summary statistic that represents the average difficulty across all of the category thresholds (b_1 – b_5) for a given item, providing insight into the relative severity of individual items in relation to one another. As shown in Table 2, the *Punishment Insensitivity* items provide coverage across the full spectrum of behavior. Two out of seven (28.6%) items had thetas below the 1.51 severity threshold in the sample, suggesting they are normatively occurring manifestations of *Punishment Insensitivity*. The remaining five items tend to be more difficult, with item locations > 1.51 . The ordering of item locations in Table 2 demonstrates the severity continuum. Behaviors that were qualitatively atypical, i.e., intense (e.g., “Not care when punished”) or intransigent (e.g., “*Keep on misbehaving no matter what you do*”) were more likely to have thetas $> 95^{\text{th}}$ percentile.

Hypothesis II

PI will demonstrate some item-level variation but there will be invariance in model fit. Item-level analyses were conducted to examine patterns of *PI* across child age, sex, poverty, and ethnicity groups.

Sociodemographic differences—Frequency distributions of each item were generally similar across sex, age, poverty, and ethnicity status (for details see Online Resource Table S1). Significant sex and ethnic differences were each identified only for a single item (“Keep on misbehaving no matter what you do;” “Act like rules didn't matter,” respectively). In contrast, differences by age and poverty status were common. More than half of the items exhibited age decreases across the preschool years, and 4/7 items showed poverty status differences, with parents living in poverty endorsing more frequent occurrence of *Punishment Insensitivity*.

Model invariance—To further explore potential sociodemographic variation in *Punishment Insensitivity*, measurement invariance was examined across child age, sex, poverty, and ethnicity status in two different ways. The invariance of all item parameters was first assessed simultaneously by comparing a fully invariant model, where only latent means and variances vary across groups, to one where all loading and threshold parameters vary across groups. Item-level results followed up through tests of differential item functioning (DIF). Robust maximum likelihood estimation with logit link indicated that factor variant models were a better fit for the data across all sociodemographic subgroups (for details see Online Resource Table S2). Tests of differential item functioning via the lordif package (Choi, Gibbons, & Crane, 2011) were then used to ascertain which items were showing non-invariance by race and ascertain the size of these effects. These tests showed that five of the seven items indicated some level of DIF, but these effects were

remarkably small. Race accounted for 0.3%–1.0% of the variation in all items by pseudo- R^2 , well below the cutoffs distinguishing “small or negligible” DIF from moderate (3.5%) and large (7%) DIF (Jodoin & Gierl, 2001). The effects of sex (two items, .2% – .6%), age (one item, .7%) and poverty status (two items, .2% – .5%) were even smaller.

Hypothesis III

Punishment Insensitivity will demonstrate construct validity and incremental clinical utility. We next examined external validity of *Punishment Insensitivity* in the validation subsample. Bivariate correlations demonstrating construct and convergent/divergent validity are shown in Table 3a.

Related Constructs

Punishment Insensitivity was positively associated with callous and low concern behaviors and poorer internalization of rules as expressed in noncompliant and sneaky behaviors. Punishment Insensitivity was negatively related to concern for others as manifest in behaviors such as caring, social attachment, empathy, and prosocial behavior.

Motivational Regulation

Punishment Insensitivity was generally associated in expected directions with problems in motivational regulation: positively to aggression, temper loss and impulsivity, and negatively to self-control. However, contrary to expectations, *Punishment Insensitivity* was positively related to fearfulness.

Family Context

Punishment Insensitivity was also associated with a chaotic, inconsistent, and angry family context.

Incremental clinical utility

A series of stepwise regressions was conducted to test the incremental utility of *Punishment Insensitivity* for predicting impairment above and beyond established features of disruptive behavior (Table 3b). *Punishment Insensitivity* was associated with significant incremental risk of problems related to the child’s behavior for most impairment indicators, (i.e., parental worry, expulsion from preschool or day care), and CGAS functional impairment. In fact, for expulsion and CGAS impairment, *Punishment Insensitivity* was uniquely associated with impairment, whereas disruptive behavior was not. Its impact on impairment can best be understood via the significantly elevated odds ratios (ORs) for three of the four impairment outcomes. ORs are a common tool for assessing the change in odds of an event given a predictor. For example, for every one point increase in *Punishment Insensitivity*, the odds of functional impairment on the CGAS increased by 3.3 [Confidence interval=1.25–8.88].

Discussion

While evidence has been growing for the importance of punishment insensitivity in childhood-onset pathways to severe antisocial behavior, the developmental spectrum of these behaviors has not been established. In this study, we examined this question in a

diverse community sample of preschool-aged children. We have demonstrated that, even in young children, a psychometrically robust severity spectrum of punishment insensitivity can be demonstrated. *Punishment Insensitivity* items increased in difficulty from ignoring prohibitions, to misbehaving regardless of consequences, to imperviousness to punishment. Although the majority of children exhibited some *Punishment Insensitivity*, severe manifestations were rare (<10–11% of children). At the severe end, these behaviors suggest atypical resistance to socialization.

Within this developmental framework, we have conceptualized *Punishment Insensitivity* as indifference to discipline, ranging from more normative lack of responsiveness to callous disregard of punishment. We have validated this construct psychometrically via IRT analyses, and in relation to established correlates. At the severe end, the dimension suggests early problems in moral regulation (e.g., “did not seem to know right from wrong”) and the acquisition of moral (e.g., rule compatible, internalized) conduct (Kochanska, Koenig, Barry, Kim, & Yoon, 2010). Validity of *Punishment Insensitivity* was demonstrated to a nomological net of related constructs and via incremental contribution to impairment, beyond traditional disruptive behavior features. The construct validity of this approach was supported via association to the related constructs of callousness and noncompliance. The association to callousness was detected both in relation to the most widely used measure of pediatric callous/unemotional traits (Frick & White, 2008), and to a developmentally-based approach to assessing low concern for others and noncompliance in a manner specifically designed to distinguish normative from atypical misbehavior in young children (Wakschlag et al., 2010). Convergent and divergent validity were demonstrated largely as theorized. Preschoolers high on *Punishment Insensitivity* were rated as less likely to exhibit attachment behaviors (e.g., show affection, trust) and to act empathically (e.g., trying to make others feel better when upset). This is consistent with the theory that internalization of rules is promoted by close attachment to parental figures and discomfort with their displeasure (Kochanska, Gross, Lin, & Nichols, 2002). Punishments are also more likely to change children’s behavior when they are consistent, contingent, warm and predictable. Thus, positive associations of *Punishment Inconsistency* to family chaos, anger and inconsistency are theoretically coherent.

Counter to prediction, however, *Punishment Insensitivity* was positively related to fearfulness, which is inconsistent with findings from studies of older youth with psychopathic traits (Frick & White, 2008) and observations of young children’s bold, risk-taking behavior in low risk samples (Kochanska et al., 2002). These prior studies link *fearlessness* to punishment insensitivity. This counter-intuitive finding may be a measurement artifact: as our measure assessed inhibition to novelty (better capturing the anxious-fearful end of the spectrum) rather than display of fearless behavior per se. Conversely, this may reflect a difference stemming from the high risk nature of the early childhood sample. That is, for young children who have trouble regulating their emotions and behavior and live in violent or unpredictable environments, *Punishment Insensitivity* may reflect the lack of environmental scaffolding and contingencies that promote learning from punishment cues, rather than a fearless indifference to socialization. A third developmental alternative explanation is suggested by intriguing findings from a recent

study — perhaps the first study of its kind to investigate temperamental substrates of callousness beginning in infancy. Mills-Koonce et al., also found that observed high intensity *fearful* behavior in toddlers (& corollary physiologic indicators of fearfulness) predicted callous traits at age seven (Mills-Koonce et al., 2014). They theorize a developmental transformation by which high fearfulness in early life is indicative of increased susceptibility to environmental context. Within this framework, it is postulated that exposure of young high fearful children to harsh and unpredictable contexts (notably both were correlated with *Punishment Insensitivity* here) may lead to severe and chronic dysregulation. Over time this is hypothesized to result in the biobehavioral stress regulation system transitioning from hyper- to hypo-reactivity as development unfolds (Mills-Koonce et al., 2014). As evidence accrues that the substrates of callousness are reliably identifiable in early childhood (Willoughby, Mills-Koonce, Gottfredson, & Wagner 2014), these findings collectively point to the need for in-depth, multi-level, longitudinal investigation (beginning in the first years of life) to pinpoint the underpinnings and developmental unfolding of these pathways (Frick & White, 2008). Explicating this is a key area for future longitudinal investigation.

Modeling *Punishment Insensitivity* via the GRM allowed for a greater and more nuanced understanding of the scale. Unlike a simple sum score and basic measures of reliability, this approach allowed for analysis of reliability and precision across the range of the scale (as shown in the Figures) and focuses on each item rather than the scale in aggregate. Item difficulties and their component category thresholds tell us about the frequency of endorsement of each item, while slopes or discriminations tell us which items are better at discriminating between varying levels of *Punishment Insensitivity*. As was found in prior work with other MAP-DB scales (Wakschlag et al., 2014), *Punishment Insensitivity* was better at differentiating children at the more severe end of the scale. This is best illustrated via the test information curve (Figure S1) which highlights that the *Punishment Insensitivity* scale is effective at assessing information from about the mean through approximately three standard deviations above the mean. As these findings are from an unselected sample, they should be representative of the general population. This differentiation is crucial for identifying children in need of evaluation or services, relative to distinctions at the normative end of the scale (e.g., distinction between “hardly ever” and “sometimes”).

Contrary to our hypothesis, *Punishment Insensitivity* items were not invariant across demographic sub-groups, with DIF particularly evidence across racial/ethnic groups. These differences were significant but small. They may be artifactual (e.g., differential interpretation of item wording). Multi-method approaches (e.g., direct observations of punishment insensitivity) as well as qualitative examination of this construct in diverse samples will be useful in teasing this out. While the non-invariance of item parameters is not ideal, the effects found in this scale are so small that they should have negligible impact, and were detectable only due to the use of nuanced methods such as GRM, and our large and diverse sample.

The present study is limited by sole reliance on parental, predominantly maternal, report for the primary measure of punishment insensitivity. Future investigations should incorporate reports by other caregivers and teachers. Correlations may also be inflated due to shared

method variance derived from the fact that all were assessed via parental report. Longitudinal, multi-method investigations are an important next step and will be crucial to determining the extent to which this trait is an early developmental marker of the callous/unemotional pathway that demarcates severe antisocial pathway (Frick & White, 2008). *Punishment Insensitivity* at this age is hypothesized to be a precursor to later callous pathways. However, the cross-sectional nature of the present study does not allow us to determine whether it is a developmental form of callousness (i.e., with same etiology and course) or a prodromal pattern that may be linked to later callousness but may also reflect developmental lags in internalization of rules, learning, and/or self-control. Our prior findings in this sample linking *Punishment Insensitivity* at preschool age to task-based problems in passive avoidance learning seen in adult psychopaths support the former (Briggs-Gowan et al., 2013). In particular, identifying individual differences, such as which preschool children high on *Punishment Insensitivity* will display chronic and severe mental health problems and which will not, will be fruitful.

As with any measurement method, the metric for severity is shaped by the rating approach. *Punishment Insensitivity* was assessed based on proportionate frequency (i.e., the proportion of disciplinary events in which these behaviors manifest). Within the multidimensional conceptual framework of the MAP-DB, we believe that this approach best captures the phenomenology of this narrow-band dimension because its severity derives from the extent to which it predominates in disciplinary encounters. Other aspects of severity are captured based on manifestations of corollary behaviors across the MAP-DB dimensional spectra (i.e., *Temper Loss*, *Aggression*, *Noncompliance* and *Low Concern for Others*). For example, when *Punishment Insensitivity* is accompanied by intense, dysregulated tantrums, this is captured in severity of *Temper Loss*, whereas when *Punishment Insensitivity* is associated with *Aggression* that occurs “out of the blue,” severity may be captured in terms of the developmental unexpectability of behaviors that occur out of context. Examining *Punishment Insensitivity* in concert with these other behaviors will provide a comprehensive picture of the severity of the child’s behavior across multiple facets. The present findings are just the first phase of using psychometric modeling to inform a developmental understanding of punishment insensitivity and its implications for clinical decision-making in treatment. For example, items dropped due to local dependence in this initial phase modeling, should be retained for future work that extends the MAP-DB across age periods as they be better differentiated in older samples. Methods such as cognitive interviewing may also be used to get a deeper understanding of whether DIF related to sub-group status (e.g., race/ethnicity) reflects differences in item interpretation. Clinical applications will require further modeling with in-depth measurements of clinical status in order to determine meaningful thresholds for treatment indication, recommendations for targeted treatment, and whether or not the scale is useful as a measure of change. The development of computer adaptive test (CAT) versions will also be useful for efficient assessment of these behaviors within clinical settings (e.g., pediatric waiting rooms).

The MAP-DB provides a psychometrically rigorous, dimensional approach to characterizing traits such as *Punishment Insensitivity* in early development. This is increasingly important for a more nuanced, dimensional approach that conceptualizes psychopathologic tendencies

along a continuum and links them to specific biological mechanisms (Insel, 2009). With the increased recent interest in identifying early precursors to lifespan mental disorders, the exploration of early insensitivity to punishment may provide modifiable targets to prevent more severe sequelae of these tendencies when they manifest as psychopathic tendencies in older youth and adults. Clinically, if a subgroup of children can be identified with *Punishment Insensitivity* at a young age, tailoring treatment may be useful. Gold standard, empirically - validated approaches to the treatment of disruptive behavior involves parent training emphasizing consistent discipline and rewarding positive behaviors while ignoring or extinguishing negative ones. However, individual differences in sensitivity to punishment and reward may result in mixed or weak effects for these treatments. Indeed, there is some evidence that suggests that children with callous-unemotional traits may be less responsive to traditional parent training interventions than those without (Hawes & Dadds, 2005; Waller et al., 2013). More generally, rates of responsiveness to intervention are moderate and the more severe the child's symptoms, the less responsive they may be to parent intervention (Beauchaine, Webster-Stratton, & Reid, 2005). The incremental variance in impairment explained by *Punishment Insensitivity*, above and beyond established disruptive behavior features, suggests that it should be separately considered in the development of tailored interventions. This developmental, dimensional approach also provides a preliminary empirical foundation for delineating the boundaries of normative and problematic punishment insensitivity behaviors as they unfold.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Funding: Lauren Wakschlag, Margaret Briggs-Gowan, and David Henry were supported by National Institute of Mental Health grants R01MH082830 and 2U01MH082830-06. Lauren Wakschlag was also supported by the Walden & Jean Young Shaw Foundation. We thank Alice Carter, Barbara Danis, Carri Hill and Patrick Tolan for their valuable contributions to the development of the MAP-DB. We also thank the pediatric clinics and participants from Rush University, the University of Illinois at Chicago, NorthShore University HealthSystem, North Suburban Pediatrics and the following pediatric research group practices for their participation: Healthline in Valparaiso, IN, Healthline in Michigan City, IN, and Associated Pediatricians in Valparaiso, IN. We gratefully acknowledge David Cella's formative insights on the art of measurement science, and his ongoing scientific leadership.

References

- Baker, FB. The basics of item response theory. 2nd ed.. College Park, MD: ERIC Clearinghouse on Assessment and Evaluation; 2001.
- Beauchaine TP, Webster-Stratton C, Reid MJ. Mediators, moderators, and predictors of 1-year outcomes among children treated for early-onset conduct problems: A latent growth curve analysis. *Journal of Consulting & Clinical Psychology*. 2005; 73:371-388.
- Blair RJR. Neurocognitive models of aggression, the antisocial personality disorders and psychopathy. *Journal of Neurology, Neurosurgery, and Psychiatry*. 2001; 71:727-731.
- Briggs-Gowan MJ, Nichols SR, Voss J, Zobel E, Carter AS, McCarthy KJ, Wakschlag L. Punishment insensitivity and impaired reinforcement learning in preschoolers. *Journal of Child Psychology and Psychiatry*. 2013; 55:154-161. [PubMed: 24033313]
- Cai, L.; du Toit, SHC.; Thissen, D. IRTPRO: Flexible, multidimensional, multiple categorical IRT modeling. Lincolnwood, IL: Scientific Software International; 2011a.

- Cai, L.; Yang, JS.; du Toit, SHC. IRTPRO for Windows: User's guide. Lincolnwood, IL: Scientific Software International; 2011b.
- Carter, AS.; Briggs-Gowan, MJ. ITSEA Infant-Toddler Social and Emotional Assessment Examiner's Manual. San Antonio, TX: Harcourt Assessment, Inc.; 2006.
- Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, Rose M. The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap Cooperative Group during its first two years. *Medical Care*. 2007; 45:S3–S11. [PubMed: 17443116]
- Choi SW, Gibbons LK, Crane PK. Lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and monte carlo simulations. *Journal of Statistical Software*. 2011; 39:1–30. [PubMed: 21572908]
- Coldwell J, Pike A, Dunn J. Household chaos: Links with parenting and child behaviour. *Journal of Child Psychology and Psychiatry*. 2006; 47:116–1122.
- Constantino J, Cloninger CR, Clarke N, Hashemi B, Przybeck T. Application of the seven-factor model of personality to early childhood. *Psychiatry Research*. 2002; 109:229243.
- Dadds MR, Salmon K. Punishment insensitivity and parenting: Temperament and learning as interacting risks for antisocial behavior. *Clinical Child & Family Psychology Review*. 2003; 6:69–86. [PubMed: 12836578]
- Egger, H.; Angold, A. The Preschool Age Psychiatric Assessment (PAPA): A structured parent interview for diagnosing psychiatric disorders in preschool children. In: DelCarmen-Wiggins, R.; Carter, A., editors. *Handbook of infant, toddler and preschool mental health assessment*. New York: Oxford; 2004. p. 223-246.
- Ellingson KD, Briggs-Gowan M, Carter AS, Horwitz SM. Parent identification of early emerging child behavior problems: Predictors of sharing parental concern with health providers. *Archives of Pediatrics & Adolescent Medicine*. 2004; 158:766. [PubMed: 15289249]
- Embretson, SE.; Reise, SP. *Item response theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.
- Ezpeleta L, de la Osa N, Granero R, Penelo E, Domenech JM. Inventory of callous-unemotional traits in a community sample of preschoolers. *Journal of Clinical Child and Adolescent Psychology*. 2013; 42:91–105. [PubMed: 23095075]
- Ezpeleta L, Granero R, de la Osa N, Penelo E, Domenech JM. Dimensions of oppositional defiant disorder in 3-year-old preschoolers. *Journal of Child Psychology & Psychiatry*. 2012; 53:1128–1138. [PubMed: 22409287]
- Finger EC, Marsh AA, Blair KS, Reid ME, Sims C, Ng P, Blair RJ. Disrupted reinforcement signaling in the orbitofrontal cortex and caudate in youths with conduct disorder or oppositional defiant disorder and a high level of psychopathic traits. *The American Journal of Psychiatry*. 2011; 168:152–162. [PubMed: 21078707]
- Frick PJ. Developmental pathways to conduct disorder: Implications for future directions in research, assessment and treatment. *Journal of Clinical Child and Adolescent Psychology*. 2012; 41:378–389. [PubMed: 22475202]
- Frick PJ, White S. Research review: The importance of callous-unemotional traits for developmental models of aggressive and antisocial behavior. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 2008; 49:359–375.
- Gorman-Smith D, Tolan P, Zelli A, Huesmann L. The relation of family functioning to violence among inner city minority youths. *Journal of Family Psychology*. 1996; 10:115–129.
- Gray, JA. *The psychology of fear and stress*. 2 nd ed.. Cambridge: Cambridge University Press; 1987.
- Gray JA. Brain systems that mediate both emotion and cognition. *Cognition & Emotion*. 1990; 4:269–288.
- Gresham, F.; Elliot, S. *Social Skills Improvement System: Rating scales manual*. Minneapolis, MN: Pearson; 2008.
- Hambleton, RK.; Swaminathan, H.; Rogers, HJ. *Fundamentals of item response theory*. Newbury Park, CA: Sage; 1991.
- Hawes D, Dadds MR. The treatment of conduct problems in children with callous-unemotional traits. *Journal of Consulting & Clinical Psychology*. 2005; 73:737–741. [PubMed: 16173862]

- Hyde L, Shaw D, Gardner F, Cheong J, Dishion T, Wilson M. Dimensions of callousness in early childhood: Links to problem behavior and family intervention effectiveness. *Development and Psychopathology*. 2013; 25:347–363. [PubMed: 23627949]
- Hyde L, Shaw D, Moilanen KL. Developmental precursors of moral disengagement and the role of moral disengagement in the development of antisocial behavior. *Journal of Abnormal Child Psychology*. 2010; 38:197–209. [PubMed: 19777337]
- Insel T. Disruptive insights in Psychiatry: Transforming a clinical discipline. *Journal of Clinical Investigation*. 2009; 119:700–705. [PubMed: 19339761]
- Jodoin MG, Gierl MJ. Evaluating Type I error and power using an effect size measure with logistic regression procedure for DIF detection. *Applied Measurement in Education*. 2001; 14:329–349.
- Kimonis E, Frick PJ, Boris N, Smyke A, Cornell A, Farrell J, Zeanah C. Callous-unemotional features, behavioral inhibition, and parenting: Independent predictors of aggression in a high-risk preschool sample. *Journal of Child and Family Studies*. 2006; 15:745–746.
- Kochanska G, Aksan N. Conscience in childhood: Past, present and future. *Merrill Palmer Quarterly*. 2004; 50:299–310.
- Kochanska G, Aksan N. Children's conscience and self-regulation. *Journal of Personality*. 2006; 74:1587–1617. [PubMed: 17083659]
- Kochanska G, Gross J, Lin M, Nichols K. Guilt in young children: Development, determinants and relations with a broader system of standards. *Child Development*. 2002; 73:461–482. [PubMed: 11949903]
- Kochanska G, Koenig J, Barry R, Kim S, Yoon J. Children's conscience during toddler and preschool years, moral self, and a competent, adaptive developmental trajectory. *Developmental Psychology*. 2010; 46:1320–1332. [PubMed: 20822241]
- LeBuffe, PA.; Naglieri, JA. *The Devereux Early Childhood Assessment*. Lewisville, NC: Kaplan Press Publishing; 1999.
- Lord, F. *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1980.
- Lykken DT. A study of anxiety in the sociopathic personality. *The Journal of Abnormal and Social Psychology*. 1957; 55:6.
- Matheny AP, Wachs TD, Ludwig JL, Philips K. Bringing order out of chaos: Psychometric characteristics of the Confusion, Hubbub, and Order Scale. *Journal of Applied Developmental Psychology*. 1995; 16:429–444.
- Matthys W, Vanderschuren LJ, Schutter DJ. The neurobiology of oppositional defiant disorder and conduct disorder: Altered functioning in three mental domains. *Development and Psychopathology*. 2013; 25:193–207. [PubMed: 22800761]
- Matthys W, Vanderschuren LJ, Schutter DJ, Lochman JE. Impaired neurocognitive functions affect social learning processes in oppositional defiant disorder and conduct disorder: Implications for interventions. *Clinical Child and Family Psychology Review*. 2012; 15:234–246. [PubMed: 22790712]
- Mills-Koonce W, Wagner N, Willoughby M, Stifter C, Blair C, Granger D. Greater fear reactivity and psychophysiological hyperactivity among infants with later conduct problems and callous-unemotional traits. *Journal of Child Psychology and Psychiatry*. 2014 epub ahead of print.
- Moffitt TE, Caspi A. Childhood predictors differentiate life-course persistent and adolescence-limited antisocial pathways among males and females. *Development & Psychopathology*. 2001; 13:355–376. [PubMed: 11393651]
- Newman JP, Kosson DS. Passive avoidance learning in psychopathic and nonpsychopathic offenders. *Journal of Abnormal Psychology*. 1986; 96:252–256. [PubMed: 3745647]
- Reise S, Waller N. Item response theory and clinical measurement. *Annual Review of Clinical Psychology*. 2009; 5:25–46.
- Samejima F. Estimation of a latent ability using a response pattern of graded scores. *Psychometrika*. 1969; 35:139–139.
- Shaffer D, Gould MS, Brasic J, Ambrosini P, Fisher P, Bird H, Aluwahlia S. A children's global assessment scale (CGAS). *Archives of General Psychiatry*. 1983; 40:1228–1231.

- Tremblay RE, Pihl RO, Vitaro F, Dobkin PL. Predicting early onset of male antisocial behavior from preschool behavior. *Archives of General Psychiatry*. 1994; 51:732739.
- Wakschlag L, Briggs-Gowan M, Choi S, Nichols S, Kestler J, Burns J, Henry D. Advancing a multidimensional, developmental spectrum approach to preschool disruptive behavior. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2014; 53:8296.
- Wakschlag L, Henry DB, Tolan PH, Carter AS, Burns JL, Briggs-Gowan M. Putting theory to the test: Modeling a multidimensional, developmentally-based approach to preschool disruptive behavior. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2012; 51:593–604. [PubMed: 22632619]
- Wakschlag L, Hill C, Carter AS, Danis B, Egger HL, Keenan K, Briggs-Gowan M. Observational assessment of preschool disruptive behavior, part I: Reliability of the disruptive behavior diagnostic observation schedule (DB-DOS). *Journal of the American Academy of Child and Adolescent Psychiatry*. 2008; 47:622–631. [PubMed: 18434926]
- Wakschlag L, Tolan P, Leventhal B. "Ain't misbehavin": Towards a developmentally-specified nosology for preschool disruptive behavior. *Journal of Child Psychology and Psychiatry*. 2010; 51:3–22. [PubMed: 19874427]
- Waller R, Gardner F, Hyde L. What are the associations between parenting, callous-unemotional traits, and antisocial behavior in youth? A systematic review of evidence. *Clinical Psychology Review*. 2013; 33:593–608. [PubMed: 23583974]
- Waller R, Gardner F, Hyde LW, Shaw DS, Dishion TJ, Wilson MN. Do harsh and positive parenting predict parent reports of deceitful-callous behavior in early childhood? *Journal of Child Psychology & Psychiatry*. 2012; 53:946–953. [PubMed: 22490064]
- White SF, Pope K, Sinclair S, Fowler KA, Brisline SJ, Williams WC, Blair JR. Disrupted expected value and prediction error signaling in youths with disruptive behavior disorders during a passive avoidance task. *American Journal of Psychiatry*. 2013; 170:315–323. [PubMed: 23450288]
- Willoughby M, Mills-Koonce W, Gottfredson N, Wagner N. Measuring callous unemotional behaviors in early childhood: Factor structure and the prediction of stable in middle childhood. *Journal of Psychopathology and Behavioral Assessment*. 2014; 36:30–42. [PubMed: 24729655]
- Willoughby MT, Waschbusch DA, Moore GA, Proper CB. Using the ASEBA to screen for callous unemotional traits in early childhood: Factor structure, temporal stability, and utility. *Journal of Psychopathology and Behavioral Assessment*. 2011; 33:19–30. [PubMed: 21483647]
- Zuckerman, M. Models of adult temperament. In: Zentner, M.; Shiner, RL., editors. *Handbook of Temperament*. New York: Guilford; 2012. p. 41-66.

Figure 1.
Punishment Insensitivity Item Information Curves

Table 1

MAP-DB/ Punishment Insensitivity Frequency Distributions (%) (N=1,855)¹

		Never	Hardly Ever	Sometimes	Often	Most of the Time	All the Time
1.	Act like he or she didn't hear you when you said "no"	30.6	23.3	33.1	8.1	2.8	2.2
2.	Deny he or she did something that was not allowed	28.4	28.8	31.6	6.7	2.9	1.7
3.	Act like rules didn't matter	60.8	21.2	14.1	1.7	1.2	1.0
4.	"no Keep on misbehaving matter what you do"	49.5	30.6	15.2	2.3	1.3	1.1
5.	"right Act like he or she did not know from wrong"	46.6	28.1	19.7	2.6	1.3	1.6
6.	Not care when punished	63.6	21.0	10.4	2.4	1.3	1.3
7.	Refuse to apologize	47.6	27.5	18.6	3.0	2.0	1.2

¹ MAP-DB-Multidimensional Assessment Profile of Preschool Disruptive Behavior

Table 2

MAP-DB Punishment Insensitivity Severity Thresholds¹ (N=1,855)

Category Thresholds								
	Slope	Hardly Ever or higher	Sometimes or higher	Often or higher	Most of the time or higher	All		
						b5	b4	Item Location (Theta)
Punishment insensitivity (95 th % threshold=1.51)								
1.	Act like he or she didn't hear you when you said "no"	2.20	-0.61	0.14	1.40	2.11	2.68	1.14
2.	Deny he or she did something that was not allowed	1.43	-0.87	0.27	1.85	2.66	3.53	1.49
3.	Act like rules didn't matter	4.29	0.29	0.94	1.87	2.21	2.69	1.60
4.	Keep on misbehaving "no matter what you do"	2.90	0.00	0.94	1.92	2.34	2.84	1.61
5.	Act like he or she did not know "right from wrong"	2.19	-0.09	0.84	2.04	2.49	2.92	1.64
6.	Not care when punished	2.94	0.40	1.16	1.92	2.32	2.73	1.71
7.	Refuse to apologize	1.82	-0.06	0.92	2.08	2.58	3.32	1.77

Note.

¹ Data are derived from IRT Graded Response Models.² Bolded numbers indicate category thresholds or item locations above the 95th percentile.

Table 3

External Validity of the MAP-DB Punishment Insensitivity Scale

3a. Bivariate Association of Punishment Insensitivity with Theorized Correlates					
	r_1	mean	std. dev.	minimum	maximum
I. Related Constructs					
<i>Convergent Validity</i>					
MAP-DB Low concern	0.69***	4.50	6.13	0	45
ICU Callousness	0.35***	4.49	3.24	0	24
MAP-DB Noncompliance	0.77***	29.18	19.75	0	110
MAP-DB Sneaky	0.70***	10.59	8.79	0	55
<i>Divergent Validity</i>					
ICU Caring	-0.45***	15.28	4.65	5	24
DECA Attachment	-0.31***	27.43	3.90	7	32
ITSEA Empathy	-0.17***	1.45	0.37	0	2
ITSEA Prosocial Peer	-0.37***	1.39	0.39	0	2
II. Motivational Regulation					
<i>Convergent Validity</i>					
MAP-DB Aggression	0.68***	15.16	16.57	0	120
MAP-DB Temper Loss	0.72***	22.47	20.16	0	110
ITSEA Activity/Impulsivity	0.58***	0.89	0.43	0	2
<i>Divergent Validity</i>					
SSIS Self Control	-0.42***	10.50	3.88	0	21
ITSEA Inhibition to Novelty	0.09*	0.87	0.52	0	2
III. Family Context					
Household CHAOS	0.25***	2.30	0.62	1	4.5
Parenting Inconsistency	0.44***	3.49	3.08	0	17
PROMIS Anger	0.29***	7.05	6.34	0	32

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

3b. Incremental Clinical Utility of Punishment Insensitivity²

Worry	Odds Ratio (Confidence Interval)			
	Expulsion	Services	C-GAS	
Disruptive Behavior	1.02 *** (1.01–1.03)	1.01 * (1.00–1.02)	1.01 (0.99–1.02)	
Punishment insensitivity	2.08 * (1.15–3.76)	0.97 (0.48–1.93)	3.33 * (1.25–8.88)	

² Data are from logistic regression analyses; ns range from 448–495

Note.

- * = $p < .05$;
- ** = $p < .01$;
- *** = $p < .001$