HHS Public Access
Author manuscript
Assessment. Author manuscript; available in PMC 2021 October 01.

Published in final edited form as:
Assessment. 2021 April ; 28(3): 796–812. doi:10.1177/1073191120939161.

Impulsive states and impulsive traits: A study of the multilevel structure and validity of a multifaceted measure of impulsive states

Max A. Halvorson, M.A.[1], **Sarah L. Pedersen, Ph.D.**[2,3], **Madison C. Feil, M.S.**[1], **Liliana J. Lengua, Ph.D.**[1], **Brooke S.G. Molina, Ph.D.**[3], **Kevin M. King, Ph.D.**[1]

[1]University of Washington, Department of Psychology

[2]University of Pittsburgh, Department of Psychology

[3]University of Pittsburgh, Department of Psychiatry

Abstract

Emerging evidence suggests impulsive states may be reliably measured in the moment using ecological momentary assessment (EMA); however, research has not investigated whether the multifactorial structure of impulsive traits also characterizes impulsive states. In two independent samples spanning adolescence through young adulthood (n = 211, n = 222), we adapted global self-report measures of impulsive traits to EMA and conducted multilevel confirmatory factor analyses to characterize the within- and between-person factor structure of five impulsive traits (negative urgency, planning, persistence, sensation seeking, and positive urgency). Across both studies, factor models with one factor for each UPPS-P facet fit the data well at both levels, though some latent factors were highly correlated. Aggregated impulsive states, especially negative urgency, predicted ODD symptoms, emotional problems, alcohol problems, and ADHD symptoms. Our results suggest that EMA measures can capture a range of impulsive states that mirrors the heterogeneity seen in the trait literature.

Keywords

Ecological momentary assessment; Impulsivity; Measurement; Factor analysis

Introduction

Individual differences in multiple, related traits (described as "impulsive traits") underpin impulsive behaviors and various forms of psychopathology, such as risky drug and alcohol use, self-injurious behaviors, and mood disorders (Berg et al., 2015). Despite growing consensus that impulsive behaviors are determined by multiple distinct traits (Gullo et al., 2014; King et al., 2014; Sharma, Markon, et al., 2013; Whiteside & Lynam, 2001), understanding of impulsive traits is limited by a reliance on global self-reports, which ask

Correspondence concerning this article should be addressed to Max A. Halvorson and Kevin M. King, Department of Psychology, University of Washington, Box 351525, Seattle WA 98195-1525, (206) 685-3571 (f), maximum2@uw.edu or kingkm@uw.edu.
Supplemental study information, including raw correlation matrices and analysis scripts, is available at https://osf.io/p3xzw/

individuals to report their typical behavior, as opposed to behavior in a given situation. Measuring impulsive states in the moment holds promise as a means of addressing research questions about processes more proximal to behavior (e.g., when and under what conditions do individuals act most impulsively) and identifying more temporally specific points of intervention. To address these research questions, it is important to understand whether individuals' moment-to-moment experience of impulsive states might mirror or differ from more global impulsive traits.

Impulsive Traits and Impulsive States

One prominent model of impulsive traits characterizes individual differences in the tendency to think and plan ahead (planning), persist towards goals (persistence), act on impulse in the face of negative emotions (negative urgency) or positive emotions (positive urgency), and engage in thrill-seeking behavior (sensation seeking; Cyders et al., 2007; Sharma, Markon, et al., 2013; Whiteside & Lynam, 2001). Narrative and meta-analytic reviews, covering hundreds of empirical studies, have suggested that nearly all global self-report measures of "impulsivity" reflect some combination of these five impulsive traits (Cross et al., 2011; Cyders & Coskunpinar, 2011; Duckworth & Kern, 2011; King et al., 2014). However, some research suggests that a three-factor solution may describe these five traits more parsimoniously. Of these five traits, planning and persistence tend to be moderately correlated (*r*s as high as .53, .45, Roley et al., 2017; Whiteside & Lynam, 2001) and load highly (.65 to .76) on a single higher-order factor in confirmatory factor analysis (Cyders & Smith, 2007; Sharma, Kohl, et al., 2013). The same is true of negative and positive urgency, which are highly correlated (*r*s as high as .80, .71, or .69; Pedersen et al., 2016; Roley et al., 2017; Settles et al., 2014) and also loaded highly (.62 to .93) on a single higher-order factor (Cyders & Smith, 2007). Sharma, Kohl, Morgan, & Clark (2013) investigated a large battery of global self-report items intended to measure impulsive traits and reported three factors reflecting planning/persistence, urgency, and sensation seeking. Recent narrative and factor analytic reviews have also argued that the optimal representation is as three weakly correlated traits (Duckworth & Kern, 2011; King et al., 2014; Sharma, Markon, et al., 2013; Smith et al., 2007).

Global self-report measures of impulsivity have strong evidence of predictive and discriminant validity for risky behaviors and mental health outcomes (Berg et al., 2015; Smith et al., 2007), but also have several key limitations. Global self-report measures require individuals to retrospectively recall their behavior across a broad range of time and situations, a reconstructive process that can be laden with bias. Recall is influenced by factors related to the encoding, recollection, and emotional salience of memories (Hunt et al., 2003; Robinson & Clore, 2002), as well as heuristics involved in judging one's response to an item (Shiffman et al., 2008). For instance, many negative and positive urgency items were derived from a facet of neuroticism (Sharma, Markon, et al., 2013; Whiteside & Lynam, 2001), and people higher on neuroticism may be biased towards recalling negative information (Eysenck & Mogg, 1992).

Although person-level relations indicate the importance of impulsive traits on psychopathology, the processes by which they unfold over time are impossible to understand without measures that match the time frame of these processes more closely (Curran & Bauer, 2011; Fisher et al., 2018; Kenrick & Funder, 1988). This mismatch between process-level theories and global self-report measures has been advanced as a reason for relatively poor prediction of future behavior by trait measures: though the disparate situations assessed by a survey measure may be experienced as cognitively or affectively similar, they may not lead to strong prediction of behavior *in a situation, for an individual* (Mischel et al., 2002; Mischel & Shoda, 1995). Many within-person questions about impulsivity – such as which situations make impulsive states more likely, or whether impulsive states precede drinking on a given day – will remain unexamined absent reliable and valid measures of impulsive states. Developing reliable and valid measures of impulsive states is essential to answering process-level questions, and given existing research on impulsive traits, it is critical to test whether measures of impulsive states reflect similar dimensions.

Prior Research on Impulsive States

To date, only three studies of which we are aware have reported on the validity of daily self-report measures of impulsive states. Tomko et al. (2014) developed a four-item measure (example items: “I have felt impatient,” “I made a ‘spur of the moment’ decision”) of impulsive states for EMA research, the Momentary Impulsivity Scale (MIS), and validated it in a clinical sample (n = 105). Participants completed brief assessments six times per day for 28 days. Results supported a single factor, which was moderately associated with global self-report of impulsive traits including BIS-11 subscales and negative urgency, planning, and persistence from the UPPS (*r*s = .26 to .45). Moreover, within-person variability in reports of the MIS over time – measured by the mean squared successive difference (MSSD) – was correlated (*r*s = .21 to .45) with global self-reports of the same impulsive traits, suggesting that impulsive traits are not only associated with levels but also with general variability in impulsive states. Wu & Clark (Wu & Clark, 2003) examined once-a-day reports of 20 commonly-occurring impulsive and planful behaviors (example items: “Blew off my homework”, “Made a to-do list”) in college undergraduates (n = 170). Exploratory factor analyses retained 13 items which formed two uncorrelated ($r = -.03$) subscales: failure to plan and carefree/spontaneous. Person-level aggregates of failure to plan behaviors were positively related to global self-reports of lack of planning (*r*s = .24 to .37) as measured by the Schedule for Nonadaptive and Adaptive Personality (Clark, 1993) and the Barratt Impulsivity Scale-Version 11 (Patton et al., 1995). Carefree/spontaneous behaviors were correlated with a measure of urgency and sensation seeking (the BIS-11 Motor Impulsivity subscale). These findings were then replicated in a sample of college students (n = 152; Sharma, Kohl, et al., 2013), finding similar magnitude and direction of correlations with the SNAP measure. Thus, impulsive states may be measured at the daily or momentary level, and when aggregated, are related to global self-reports.

Given strong evidence that multiple impulsive traits exist, measures of impulsive states should also aim to measure several factors related to impulsive acts. Prior work has not sampled a wide range of items, or had the explicit goal of deriving a small number of factors. This approach makes it difficult to know whether the five-factor model of impulsive traits is mirrored in people’s perceptions of their impulsive states. Measuring impulsive traits as they manifest in everyday life could allow for validation of process-level theories of specific individual traits (e.g., that individuals who score high on global self-report of negative urgency are especially likely to lash out at a friend after experiencing negative affect earlier in the day). Measuring specific impulsive states could also allow for specific prediction from individual states to individual behaviors. On the other hand, if trait measures and EMA equivalents have widely differing factor structures, it may be critical to investigate other factors (e.g., situational or contextual factors) that may explain the association between traits and states. Some early evidence supports the idea of distinguishing between impulsive states: in recently published work with the same sample as Study 2 below, Pedersen et al. (2019) found that adults with childhood ADHD had increased variability on some impulsive state measures (negative urgency, positive urgency, and sensation seeking) but not others (planning and persistence). Thus, the primary aim of the current study was to more thoroughly describe the factor structure of state impulsivity using items representing a wide range of impulsive traits.

The Current Study

In the current study, we analyzed the reliability and validity of an EMA adaptation of a broad questionnaire battery of impulsive traits (the UPPS-P; Whiteside & Lynam, 2001; Cyders et al., 2007). Our goals were 1) to test the degree to which the between-person structure of impulsive traits also describes within-person variation in impulsive states, and 2) to report on the reliability and validity of EMA measures of impulsive states.

To address these aims, we used data from two independently conducted studies which adapted global self-report measures of UPPS-P impulsive traits to assess impulsive states. We hypothesized that the five impulsive domains measured would be differentiable, and that a five-factor solution would best fit the data. However, given extant research on measurement of impulsive states (Sharma, Markon, et al., 2013; Tomko et al., 2014; Wu & Clark, 2003), we also examined the possibility of a more parsimonious solution. We further hypothesized that aggregated impulsive states would show evidence of convergent validity and criterion validity. At the trait level, we hypothesized that aggregated impulsive states would relate positively and moderately with retrospective impulsive trait measures (convergent validity), and would relate positively and moderately with externalizing problems, mood disorders, and emotional problems (criterion validity), following meta-analytic findings by Berg et al. (2015) that impulsive traits predict a broad range of emotional and behavioral problems. At the state level, we hypothesized that one or more of the UPPS-P impulsive states would correlate with existing impulsive state measures (convergent validity; Tomko et al., 2014; Wu & Clark, 2003).

In Study 1, a sample of high school and college students, we aimed to understand whether state impulsivity is best conceptualized as a unitary state or as several distinct dimensions by measuring urgency, planning, and persistence in the moment. Study 2 replicates Study 1 in a sample recruited to contain elevated levels of impulsive traits, expands the factor analyses to include measures of sensation seeking and positive urgency, and includes additional indices of criterion validity.

Methods

Overview of Study 1 and Study 2

Study 1 and Study 2 were EMA studies of daily impulsivity with similar protocols and measures. Study 1 included participants from a public high school and a four-year university in the Northwestern United States and measured three impulsive states (negative urgency, planning, and persistence) in 215 participants, assessed three times per day for 10 days. Study 2 measured five impulsive states (negative and positive urgency, planning, persistence and sensation seeking) in 211 young adult participants in a city in the Northeastern United States, measured six times per day for 10 days. Each study assessed impulsive states during a 10-day period comprising two weekends, in order to oversample certain impulsive states and impulsive behaviors (e.g., heavy alcohol and marijuana consumption) which tend to be more likely on weekends than on weekdays. We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

Study 1 Method

Participants.—Participants were high school students at a public high school in the Pacific Northwest (n = 61) and college students at a four-year public university (n = 161; combined n = 222). The high school sample was 64% female, and 54% White, 17% Asian, 5% Black, and 15% multiracial, with students ranging in age from 15 to 18 (M = 16.6). The college student sample was 49% female and 66% White, 21% Asian, and 12% other race or ethnicity, and consisted of students aged 18–22 (mean age 19.4). The college sample was recruited from among participants in the university's Psychology Subject Pool who reported weekly alcohol or marijuana in a screening survey. Of screened university participants, 17% met eligibility criteria.

Procedure.—Study procedures were approved by the local Institutional Review Board. Participants first completed a web-based baseline survey and were trained in the EMA protocol, either during an in-person session (college sample), or a telephone session (high school sample). Next, participants began a 10-day EMA protocol on the following Friday (college sample), or on the following day (high school sample). During the EMA period, participants received three text messages a day on their personal smartphones containing a link to a Qualtrics survey during a random time within three evenly-spaced assessment windows between 9am and 9pm (college sample) or between 12pm and 9:30pm (high school sample). Response windows were altered in the high school sample to avoid interrupting class time, and the first survey was sent during students' lunch break. Assessment windows were separated by at least two hours to avoid temporal overlap in surveys. Reminder text messages were provided to participants one hour after the initial text; 51% of participants completed the survey within 10 minutes, and 86% completed the survey within one hour. College participants received prorated course extra credit for participation which depended on the number of EMA surveys completed. High school students received up to $30, $1 for each survey completed. We considered participants (n = 3 in the high school sample, n = 4 in the college sample) who completed fewer than 3 days' worth of assessments (9 assessments) to be non-responders; they were excluded from further analyses. Excluding non-responders, response rates for the daily assessments were 81% for high school and 87% for college students. In total, the final analysis sample consisted of 58 high school students (1,416 observations of a possible 1,740) and 157 college students (4,107 observations of a possible 4,710).

Measures

Impulsive traits.—We measured impulsive traits during the baseline survey with the full UPPS-P (Lynam et al., 2006; Whiteside & Lynam, 2001) measure: negative urgency (12 items, e.g., "In the heat of an argument, I will often say things that I later regret"), planning (11 items, e.g., "My thinking is usually careful and purposeful"), persistence (10 items, e.g., "I generally like to see things through to the end"), sensation seeking (12 items, e.g., "I generally seek new and exciting experiences and sensations"), and positive urgency (14 items, e.g., "When I am really excited, I tend not to think of the consequences of my actions"). Response options ranged from 1 (agree strongly) to 4 (disagree strongly). All responses were coded so that higher numbers indicated more impulsivity. Cronbach's alpha values were high for each scale (α > .85). Validity analyses were completed with only the impulsive traits that matched the impulsive states measured.

Impulsive states.—Impulsive state items were adapted from the global self-report version of the UPPS-P by project staff based on item loadings, consultation with experts on item content, and adaptability to the momentary context. For example, "I have trouble controlling my impulses" was modified to "Since the last assessment I had trouble controlling my impulses." Table 1 contains original global self-report items and adapted language. The project team adapted 14 items from the UPPS-P (Whiteside & Lynam, 2001) to represent three impulsive states – negative urgency (6 items), planning (4 items), and persistence (4 items), with item stems changed to reflect behavior "since the last assessment" (or "since you woke up this morning" for morning assessments). Mean scores were calculated for each scale. A several-hour response time frame was chosen to balance several factors: consistency and comparability with prior work using this approach (e.g., Tomko et al., 2014), the burden of responding to surveys, the amount of retrospection and recall on the part of participants, and the likelihood of capturing recent impulsive states at a given assessment. In fact, examining distributions of impulsive state items revealed right-skew, suggesting that high-impulsivity states are somewhat uncommon. . Although it is possible that a briefer time frame would limit recall biases even further, this would have increased the burden on participants, potentially reducing response rates and inducing measurement reactivity. Overall reliability was high for all 3 scales, α = .71 – .84. The items had identical anchors as the trait UPPS (from Agree Strongly to Disagree Strongly), but respondents instead used a slider bar (which recorded responses ranged from 0–100) for the impulsive behaviors items.

In contrast to Study 2, Study 1 adapted items from the UPPS negative urgency and positive urgency scales to be free of emotional content (e.g., "I acted without thinking" rather than "When I am upset I often act without thinking") in order to reduce the items' dependency on participants' evaluation of their mood. The larger goal of our research in this domain is to understand the interplay between affect and impulsive states. Within this broader goal, Study 1 represents our attempt to understand the structure of impulsive states. Understanding relations between affect and impulsive states might be complicated by including affect in the impulsive state measures themselves. To avoid this potential confound, and informed by meta-analytic work indicating that urgency shares common variance with affect-free measures of impulsive traits (Sharma, Markon, & Clark, 2013), we adapted urgency items to be free of affect in Study 1, with the acknowledgment that differences in factor structure across studies may be impacted by differences in item content. These items comprised a general urgency scale. Despite these concerns, subsequent analyses indicated that these items behaved similarly to the negative urgency items in Study 2. Thus, these scale items are reported below as negative urgency.

At each assessment, we administered a random selection of half of each item set to reduce response burden. Because these items were missing completely at random (MCAR), they could be imputed without bias, and scale composites (such as means) may also be estimated using the observed items with no increase in bias, and only minor increases in the standard error of the estimates (Graham, 2009; Graham et al., 2006).

State-level convergent validity.—Two existing EMA measures of impulsive behaviors and states were examined for potential relations with our adapted UPPS-P measures. **Impulsive behaviors** were measured each day at the nighttime assessment using the Daily Behavior Questionnaire (DBQ), an inventory of 31 everyday impulsive behaviors (Sharma, Kohl, et al., 2013; Wu & Clark, 2003). Items form two uncorrelated subscales derived from exploratory factor analysis: failure to plan behaviors (20 items; α = 64; e.g., “Made a to-do list “) and carefree/spontaneous behaviors (11 items; α = .56; e.g., “Made an impulse buy”). For comparability with impulsive state items, we reverse coded the failure to plan behaviors and reported them below as planful behaviors. **Momentary impulsivity** was measured at each assessment using the Momentary Impulsivity Scale (MIS), a brief 4-item measure of impulsive state. Example items included “I said things without thinking” and “I have felt impatient.” Between-person reliability for this measure was high (R_{KRN} = 0.99), whereas within-person reliability was moderate (R_{CN} = 0.56). Overall levels and variability in MIS scores are moderately related to baseline measures of impulsive traits.

Trait-level criterion validity.—To measure criterion validity, we used global self-reports of alcohol problems and emotional problems, both of which have been closely linked to impulsive traits (Berg et al., 2015). **Alcohol problems** were measured with the Rutgers Alcohol Problem Index (RAPI; White & Labouvie, 1989). Individuals self-reported the past-year frequency of 23 different alcohol-related consequences on an ordinal scale (0 = None to 3=More than 5 times). Published reliabilities are high for the RAPI (Cronbach’s α = .92, 3-year test-retest reliability = .40; White & Labouvie, 1989). Example items include “Neglected your responsibilities” and “Was told by a friend, neighbor or relative to stop or cut down drinking.” A global mean score was computed and used in analyses. **Emotional problems** were measured with anxiety, depression, and anger scales from the Patient-Reported Outcomes Measurement Information System (Pilkonis et al., 2011). The three scales have high internal consistencies (αs = .90–.95). Example items for anxiety (28 items) included “I felt anxious;” for depression (28 items) included “I felt hopeless;” and for anger (22 items) included “I was grouchy.” Responses were coded on a Likert scale (1=“never” to 5=“always”). The PROMIS scores are normed on a large national dataset (N=5,239; Pilkonis et al., 2011), and thus T-scores (M=50, SD=10) are available. A subset of the full item set (8 anxiety items, 8 depression items, and 5 anger items) was used to calculate T scores based on national norms. Mean T-scores for the three scales in the Study 1 sample were: anxiety (M=55.02, SD=8.85), depression (M=49.95, SD=10.77), and anger (M=51.10, SD=9.50).

Study 2 Method

Participants.—Participants were 231 moderate-to-heavy current drinkers (n = 117 with childhood ADHD; n = 114 without childhood ADHD). Participants were recruited from the Pittsburgh ADHD Longitudinal Study (PALS, *n* = 88; Molina et al., 2007, 2017) or from the surrounding community (*n* = 143; Pedersen et al., 2019). The sample was 76% male and had a mean age of 28.0 (range 21–35 years old, SD = 4.06). The majority of participants identified as White/European American (67.3%) and 31.8% identified as Black/African American; 0.9% identified as Asian or another race. All participants were current drinkers (drank alcohol in the last month). The larger study from which these data were drawn included an in-lab alcohol administration (described in Hasler et al., 2019), thus additional exclusion criteria were used. Participants were matched across race and ADHD status on self-reported past 30-day drinking behavior.

Procedure.—Study procedures were approved by the local Institutional Review Board, and are described in full in Pedersen et al. (2019). Following participation in a baseline interview and a laboratory protocol, the majority of participants started the EMA protocol on the following Friday (14 participants had a delayed start date due to extenuating circumstances). Participants used their personal smartphone or a study-provided smartphone. All participants were provided in-person instruction on how to complete the prompts (e.g., wait until done driving, do not give out your password). The first assessment was sent 15 minutes after self-reported expected wake time and the last assessment was sent to participants 15 minutes prior to reported expected bedtime. An additional 4 assessments were randomly sent throughout the day. Participants were instructed that they had 10 minutes to complete the prompt and received a reminder text 5 minutes after the first prompt was sent. Prompts were sent six times a day for 10 days. Participants could earn up to $110 for completing at least 80% of the EMA assessments. Those who completed less than 80% of the EMA prompts received corresponding compensation (e.g., 60% completion received $66, 60% of $110). Twenty participants were excluded for providing less than 3 full days' worth of EMA surveys. Excluded participants did not differ from included participants on demographic measures (age, gender), ADHD diagnostic status, or recruitment method. Excluding non-responders, overall compliance for the daily assessments was 72% and the total number of surveys completed comprised 9,154 of 12,660 possible observations.

Measures

Impulsive traits.—During a baseline survey, five impulsive traits were assessed using the 59-item UPPS-P impulsive behavior scale, as described in Study 1.

Impulsive states.—As in Study 1, items (19) were adapted from the UPPS-P to the momentary level to assess negative urgency (4 items; α = .83), (lack of) planning (3 items; α = .87), (lack of) persistence (4 items; α = .81), sensation seeking (4 items; α = .81), and positive urgency (4 items; α = .91) to assess impulsive states. Participants were asked to report on their experiences since the last EMA prompt. Response options ranged from 1 (agree strongly) to 4 (disagree strongly). Composite scores were calculated as in Study 1.

Trait-level criterion validity.—To measure criterion validity, we assessed alcohol problems, adult ADHD symptoms, and adult oppositional defiant disorder (ODD) symptoms. These problems represent a combination of trait-like propensities toward problematic behaviors (ADHD, ODD symptoms) and acute episodes of psychopathology. **Alcohol problems** were measured with a modified version of the Young Adult Alcohol Problems Screening Test (YAAPST; Hurlburt & Sher, 1992) to assess a range of alcohol problems in adulthood. This 36-item questionnaire (response options: 0 = "none or not in the past year" through 3 = "more than 5 times") assesses how frequently problems occurred while drinking or as a result of drinking within the past year. The YAAPST displays good reliability in prior literature (Hurlburt & Sher, 1992) and in our sample (α = .92). For the current study, the sum of frequency scores for each problem reported in the past year (range 0–108; M = 22.75, SD = 19.57) was the outcome variable. **Attention Deficit/Hyperactivity Disorder (ADHD) symptoms** were assessed for the past 6 months with the 18-item Adult ADHD Self Report Scale (ASRS; Kessler et al., 2007). These items are adapted from DSM-IV criteria to mirror ADHD symptoms as assessed in a clinical interview. Respondents report on the frequency of feelings/behaviors in the past 6 months using an ordinal scale ("never", "rarely", "sometimes", "often", "very often"). ASRS scores indicating ADHD show 82%−87% concordance with diagnoses based on a full clinical interview. Cronbach's alpha for the ASRS is in the range of .63–.72. A global mean score was computed and analyzed. **ODD symptoms** in the past 12 months were assessed with an 8-item measure of ODD symptoms adapted from a longer battery of executive function deficit-related behaviors (Barkley et al., 2010; Biederman et al., 2008). Though the form of adult ODD symptoms may differ developmentally from youth ODD symptoms, these problems are related to impulsive traits and show continuity across the lifespan (Babinski, Hartsough, & Lambert, 1999; Satterfield et al., 2007). Each item requested participants report the frequency of a behavior in the last 12 months on an ordinal scale (0=Never or Rarely, 1=Sometimes, 2=Often, 3=Very Often). Items included "Lose my temper," "Argue with others," and "Deliberately annoy people."

Analysis Plan

Descriptive statistics and scale reliability.—Prior to model fitting, we examined the extent of between- and within-person variance in individual items by examining intraclass correlations for items. To estimate the stability of daily measures – that is, how reliable measurements are when averaged across time and scale item – we estimated variance components as recommended by Shrout & Lane (2014) using the VCA software package in R. This coefficient R_{KR} represents the stability of rank-ordering of individuals in their levels of impulsive states across time. Because participants were assessed at the same frequency, but at pseudorandom time points within assessment windows, this quantity was computed treating time as a random effect. We also included coefficient ω at the between-person and within-person levels to measure internal consistency reliability at each level, calculated using the approach advocated by Geldhof et al. (2014). ω is similar in interpretation to

Cronbach's α in that it indexes the ratio of true score variance to total variance at the given level, but does not make the assumption of essential tau equivalence and provides more precise reliability estimates.

Missing data.—Prior to model estimation, we imputed data for impulsive state items that were missing using a fully conditional multiple imputation procedure which accounts for clustering in multilevel data (Enders, Keller, & Levy, 2017), using Blimp 1.1 (Keller & Enders, 2017). Excluding data that were missing by design and fully missing EMA surveys, missing data for impulsive traits or states was rare (under 4% missing for each variable). We imputed data separately for the adolescent and the young adult datasets in Study 1 so as not to impute similarity to the potentially different factor structures of the two. Each imputation was conducted with two chains and created 20 imputed datasets for each of the three datasets (Study 1 adolescent, Study 1 young adult, Study 2). Maximum potential scale reduction factors near 1.00 across two imputation chains (1.00–1.05 in Study 1, 1.03–1.10 in Study 2) suggested that the imputation process converged on stable estimates. Estimates presented below are mean estimates across 20 imputed datasets. Across all estimates, variability due to imputation was at least an order of magnitude smaller than estimates.

Multilevel confirmatory factor analysis.—We conducted multilevel confirmatory factor analysis (ML-CFA) of EMA-assessed impulsive states in MPlus 8.0 (L. K. Muthen & Muthen, 2017). ML-CFA allows for simultaneous examination of factor structure at the between-subject and within-subject levels in one model. We conducted the ML-CFA in a stepwise fashion, following an approach for optimizing model estimation proposed by Muthen (Dyer et al., 2005; B. O. Muthen, 1994). First, we decomposed the total covariance matrix for all impulsive state items into two separate covariance matrices: one estimating between-cluster variation and one estimating pooled within-cluster variation. Next, we conducted traditional CFAs separately on the between-cluster covariance matrix and on the within-cluster covariance matrix. Models were specified with each item loading on its corresponding UPPS-P facet at both the between- and within-person level. Finally, we used the estimates from those models as starting parameter estimates for estimating the final two-level model, which was specified in MPlus in full.

Model fit was assessed using chi-square as an indicator of exact fit. Where exact fit was not achieved (as chi-square is sensitive to violations of normality and sample size; Hu & Bentler, 1999), we used relative fit indices, including the comparative fit index (CFI), standardized root-mean-square residual (SRMR), and root-mean square error of approximation (RMSEA). Using these indices, we judged model fit with reference to standards provided by Hu and Bentler (1999), Kenny & McCoach (2003) and the cautions of Marsh et al. (2004), and attended to modification indices as a means of balancing indicators of model fit (Jackson et al., 2009).

Model selection followed five primary steps. First, we performed item selection, examining items with overlapping content and any items with weak correlations with all others. Second, we attempted to fit the full two-level model in all 20 imputed datasets, evaluating summary fit statistics. Third, we explored modification indices in five randomly-selected imputed datasets to identify potential post-hoc sources of model misfit or misestimation (negative variance estimates; items from the same subscale with highly correlated residual variances). Fourth, we evaluated these adjusted models and interpreted findings. Fifth, we examined the feasibility of reducing the number of latent factors explaining variance in the impulsive states by combining those that were highly correlated.

Validity.—To examine trait-level convergent validity, we computed correlations between EMA-aggregated measures of impulsive domains and corresponding global self-report measures. To examine trait-level criterion validity, we computed correlations between EMA-aggregated measures of impulsive domains and behavioral and emotional problems. We also examined correlations between overall within-person variability in each domain, measured with the MSSD (Jahng et al., 2008; von Neumann et al., 1941), and outcomes. In Study 1, we also assessed state-level convergent validity using EMA measures of daily impulsive behaviors and broad momentary impulsivity.

Study 1 Results

Between- and within-person variability in impulsive states

One-fourth to one-third of the variance in the 12 impulsive state items (measured 3 times per day for 10 days) was explained by person-level differences in Study 1 (ICCs .23–.36; 30 observations per participant). Thus, accounting for between-person differences in impulsive states, substantial within-person variance remained, which encompasses both predictable moment-to-moment variance and within-person error variance. The stability of rank-ordering of individuals on each impulsive state across time was moderate ($R_{KR, urg}$ = .41, $R_{KR, plan}$ = .65, $R_{KR, pers}$ = .63), and scales showed internal consistency at the between-person level ($\omega_{b, urg}$ = .95, $\omega_{b, plan}$ = .96, $\omega_{b, pers}$ = .95) and the within-person level ($\omega_{w, urg}$ = .71, $\omega_{w, plan}$ = .66, $\omega_{w, pers}$ = .66).

ML-CFA Results

Model estimation.—After item selection, the full 3-factor model fit the data well (X^2(103) = 917, CFI = 0.93, $SRMR_{between}$ = 0.05, $SRMR_{within}$ = 0.04, RMSEA = 0.04). We dropped two items: one due to content overlap (the adapted urgency item "I acted without thinking" was identical to the inverse of a planning item in the original UPPS-P) and one due to weak correlations with all other items (the persistence item "I gave up easily" was weakly correlated, all *r*s < .3, with other persistence items). The subsequent full 3-factor model required two post-hoc model specifications to address isolated sources of model misfit. First, we fixed residual variances to zero for three potential Heywood cases whose residual variances were not significantly different from zero at the between-person level to avoid negative variance estimates (Heywood, 1931). Second, high modification indices (above 50) suggested we allow residuals to correlate for two urgency items ("I had trouble controlling my impulses" and "It was hard to resist acting on my feelings") at both the between- and within-person levels.

Model fit and factor loadings.

No cross-loadings were necessary to achieve acceptable model fit. Fit criteria did not vary widely across imputations, and standard deviations were an order of magnitude smaller than estimates, suggesting the imputation was reliable. After the two re-specifications described above, the three-factor model (only [negative] urgency, planning, and persistence were assessed) showed acceptable fit to the data, $X^2(103) = 917$, CFI=.93, $SRMR_{between} = .05$, $SRMR_{within} = .04$, RMSEA = .04. Factor loadings are presented in Table 2.

After fixing the highest factor loading for each factor to 1.00, the median item factor loading was .90 (all loadings over .74) at the between-person level. The models explained a median of 80% of the between-person variance in items. At the within-person level, the median item factor loading was .95 (all loadings over .65), with latent factors accounting for a median of 40% of item variance.

Planning and persistence were highly correlated both between person ($r = .77$) and within person ($r = 81$). Latent factor correlations are reported in Table 3.

Exploratory analysis of correlated latent factors.

In light of strong correlations between planning and persistence, we tested factor models combining correlated impulsive states. First, we combined planning and persistence. Despite fit criteria at or near acceptable levels, nested model comparison (X^2 difference tests) and relative fit criteria (AIC and BIC) indicated inferior fit to the initially hypothesized models. Combining planning and persistence ($X^2_{diff}(4) = 1089$, $p > .05$, ΔAIC = 1260, ΔBIC = 1039), or estimating a one factor model ($X^2_{diff}(12) = 11646$, $p > .05$, ΔAIC = 12361, ΔBIC = 12280) produced significantly poorer fit.

Validity

Trait-level convergent validity.—Correlations between impulsive state measures and global self-report UPPS-P measures are summarized in Table 4. Correlations between the same construct measured via EMA and global self-report were moderate to strong and generally stronger than correlations with dissimilar constructs..

Trait-level criterion validity.—Table 5 reports relations between aggregated impulsive states and emotional and behavioral problems. Again, higher negative urgency predicted higher levels of alcohol problems, while higher planning predicted fewer alcohol problems, and persistence was unrelated. Correlations of EMA-aggregated states with retrospective report of anger, anxiety, and depression followed similar patterns, with negative and positive urgency positively associated with higher levels of problems, and higher planning and persistence associated with lower levels. Variation (MSSD) in impulsive states again related to emotional and behavioral problems with comparable strength and direction to mean levels in each domain.

State-level convergent validity.—Correlations between within-person variation in impulsive states and within-person variation in the MIS (at the momentary level) and DBQ (at the daily level) were calculated to assess state-level convergent validity. Measures were also aggregated to the person level and correlated with one another to assess trait-level convergent validity. The MIS was associated strongly and positively with urgency at both levels (r_{within} =.68, $r_{between}$ = .88). The MIS also related to planning and persistence with small negative correlations, functioning similarly to the urgency measure. DBQ measures of impulsive behaviors correlated positively with state urgency, but not with state planning and state persistence. DBQ measures of planful behavior correlated positively with state planning and state persistence, but negatively with state urgency. See Table 6 for correlations.

Study 2 Results

Study 2 replicated several findings from Study 1 for (negative) urgency, planning, and persistence, and extended these findings for sensation seeking and positive urgency. Differences in results are discussed below.

Between- and within-person variability in impulsive states

Over half the variance in the 18 impulsive state items (measured 6 times per day for 10 days) was explained by between-person differences in Study 2 (ICCs .55–.60; 60 observations per participant). Reliability of measurement, reflecting rank-ordering of individuals on average impulsive states across time, for each impulsive state was higher than in Study 1 ($R_{KR,\ urg}$ = .88, $R_{KR,\ plan}$ = .91, $R_{KR,\ pers}$ = .87, $R_{KR,\ ss}$ = .89, $R_{KR,\ purg}$ = .88). Scales showed comparable internal consistency to Study 1 scales at the between-person level ($\omega_{b,\ urg}$ = .96, $\omega_{b,\ plan}$ = .97, $\omega_{b,\ pers}$ = .98, $\omega_{b,\ ss}$ = .84, $\omega_{b,\ purg}$ = .98) and the within-person level ($\omega_{w,\ urg}$ = .65, $\omega_{w,\ plan}$ = .71, $\omega_{w,\ pers}$ = .70, $\omega_{w,\ ss}$ = .62, $\omega_{w,\ purg}$ = .74).

ML-CFA Results

Estimation.—Following item selection, the full 5-factor model fit the data well ($X^2(218)$ = 1181, CFI = 0.96, $SRMR_{between}$ = 0.05, $SRMR_{within}$ = 0.03, RMSEA = 0.02). In Study 2, one item was dropped due to high content overlap ("I kept my feelings under control" was dropped due to cross-loadings on both negative and positive urgency, and shared method variance with "I lost control when I was in a great mood") and one persistence item was dropped due to weak correlations with all other persistence items ("I gave up easily" was weakly correlated, *r*s < .35, with other persistence items). No negative variance estimates or high modification indices were observed in the 5 imputed datasets examined.

Fit and factor loadings.—No cross-loadings were necessary to achieve acceptable model fit. Fit criteria did not vary widely across imputations, and standard deviations were an order of magnitude smaller than estimates, suggesting imputation was reliable. The five-factor model (negative urgency, planning, persistence, sensation seeking, and positive urgency) showed acceptable fit to the data, $X^2(218)$ = 1181, CFI=.96, $SRMR_{between}$ = .05, $SRMR_{within}$ = .03, RMSEA = .02.

After fixing the highest factor loading for each factor to 1.00, the median item factor loading was .98 (all loadings over .79) at the between-person level. The models explained a median of 92% of the between-person variance in items. At the within-person level, the median item factor loading was .96 (all loadings over .49), with latent factors accounting for a median of 37% of item variance. This lower variance suggests that after taking into account variance due to each impulsive state scale, substantial variance in responses to EMA items (63% of total item variance) remained.

Latent factors were correlated within each level, reflecting shared variance between distinct impulsive traits (between-person level) and between distinct impulsive states (within-person level). At both levels, planning and persistence were highly correlated ($r_{between}$ = .94, r_{within} = .82), while negative urgency was highly correlated with both positive urgency ($r_{between}$ = .97, r_{within} = .71) and sensation seeking ($r_{between}$ = .69, r_{within} = .67). Positive urgency and sensation seeking were also highly correlated with one another ($r_{between}$ = .70, r_{within} = .71). Latent factor correlations are reported in Table 3.

Exploratory analysis of correlated latent factors.—As with Study 1, combining planning and persistence led to poorer model fit ($X^2_{diff}(8) = 112$, $p > .05$, ΔAIC = 190, ΔBIC = 130) than the 5-factor model. Combining negative and positive urgency into a single urgency factor also led to poorer model fit ($X^2_{diff}(8) = 555$, $p > .05$, , ΔAIC = 1006, ΔBIC = 946). Finally, we tested a one-factor model; as expected, a one-factor model did not approach acceptable levels of fit ($X^2_{diff}(20) = 14375$, $p > .05$, ΔAIC = 22703, ΔBIC = 22554).

Validity

Trait-level convergent validity.—Table 4 summarizes findings on person-level convergent validity. Correlations between the same construct measured with EMA and global self-reports were moderate to strong, ranging from .28 (planning) to .65 (positive urgency). Patterns of correlation amongst the five impulsive traits (e.g., trait planning with trait persistence) generally followed the same pattern in both global self-report and EMA-aggregated measures, with some exceptions. Sensation seeking in particular correlated differently with other impulsive domains, depending on the assessment method. Global self-report of sensation seeking was unrelated to other measures (rs −.06–.13). EMA-aggregated sensation seeking levels, however, were positively related to measures of negative and positive urgency (rs = .63 to .65) and negatively related to planning and persistence (rs = −.20 to −.22).

Trait-level criterion validity.—Relations between impulsive states and person-level outcomes are summarized in Table 5. Generally, criterion validity findings replicated and extended findings of Study 1, with slightly higher correlations in Study 2. Higher levels of EMA-aggregated negative and positive urgency were associated with higher reported alcohol problems, ADHD symptoms, and ODD symptoms, while EMA-aggregated planning and persistence were associated with lower levels of all three symptom types. EMA-aggregated sensation seeking was only weakly associated with alcohol problems. Within-person variation (MSSD) in impulsive domains related to alcohol problems, ADHD symptom severity, and ODD symptoms with comparable strength and direction to mean levels in each domain. Because variability scores are only positive, associations for planning and persistence were of the opposite sign.

Discussion

Despite widespread recognition of the potential of EMA to further the study of impulsive behaviors, almost no attention had been paid to how impulsive states might be expressed or modeled in ecologically valid settings. Across two studies spanning adolescence through early adulthood, we found that impulsive states vary along similar domains as impulsive traits and we presented evidence for convergent and criterion validity of these measures. A substantial amount of variance in EMA items reflected between-person differences (around one-third Study 1 and over half in Study 2), meaning that the remainder of variance reflected moment-to-moment fluctuations within individuals. Person-level aggregates of impulsive states corresponded moderately with global self-report measures of the same trait, and correlations with emotional and behavioral problems were similar to those of trait measures. Overall, our results suggest that variation in impulsive domains at the state level reflect several related dimensions of variation rather than a unitary impulsive state, and that these impulsive states are broadly consistent with conceptualizations of impulsive traits.

Building on prior efforts by Tomko et al. (2014), Wu & Clark (2003), and Sharma, Kohl, et al. (2013), we found that impulsive states can be reliably measured and are characterized along multiple dimensions that mirror those captured in global self-reports. The five-factor structure of the UPPS-P model of impulsive traits replicated at both the between-person level and the within-person level in both studies, with three factors tested in Study 1. However, both between- and within-person correlations amongst latent factors of impulsive states showed less differentiation amongst the two pairs of correlated factors of the UPPS-P (planning and persistence, and negative and positive urgency) than typically seen in trait measures. This suggests that in moments when participants reported less planning, they also perceived themselves to be less persistent, and when they perceived themselves to be higher on negative urgency, they also reported higher positive urgency. High correlations between planning and persistence mirror recent findings suggesting that planning and persistence load onto a broad trait reflecting conscientiousness (Sharma, Kohl, et al., 2013), while high correlations between negative and positive urgency echo recent evidence that negative and positive urgency are not always easily differentiated, and may represent a more general impulsive reactivity to emotional states (Carver et al., 2008; Smith & Cyders, 2016). Both within-person variance and between-person variance in urgency and the MIS were highly correlated in Study 1, suggesting a central role of urgency in momentary assessment. Though these two sets of latent factors were highly correlated, exploratory analyses indicated that best-fitting models were still those with separate factors for each impulsive trait, even when selecting models based on fit criteria which penalize for model complexity. Thus, the current study suggests that at least three (and perhaps five, depending on substantive interest) facets of impulsivity can be measured when considering impulsive states.

EMA measures of impulsive states demonstrated strong convergent, divergent, and person-level criterion validity. Supporting convergent and divergent validity, individuals' aggregated levels of impulsive states correlated moderately with their corresponding full-scale trait measures administered at baseline. Impulsive states also correlated more highly with the impulsive traits they were meant to capture than with non-overlapping impulsive traits.

These on-diagonal correlations and validity relations follow a pattern suggested in multitrait-multimethod studies (Campbell & Fiske, 1959) and suggest overlap in construct representation between EMA state scales and global self-report trait scales. Regarding criterion validity, person-level aggregates of impulsive states predicted levels of emotional and behavioral problems. In parallel with results from Berg et al.'s (2015) meta-analysis of impulsive traits, aggregated impulsive states predicted ADHD symptom severity, conduct and alcohol problems, anger problems, anxiety symptoms, and depression symptoms. Also reflecting Berg et al.'s (2015) findings, negative urgency was the strongest, most consistent predictor of problems overall. Finally, we observed momentary convergent validity via strong positive within-person correlations between urgency and a one-factor impulsive state measure (MIS; Tomko et al., 2014) measure, supporting a central role of urgency. Altogether, our findings support the validity of an EMA measure of impulsive states.

Although it has long been noted that sensation seeking is distinct from other impulsive traits (Cross et al., 2011; Duckworth & Kern, 2011), results from Study 2 suggested that sensation seeking was less distinct when measured as a state. Specifically, data from Study 2 showed that sensation seeking states were highly correlated with negative and positive urgency states (rs = .63 to .65). In other words, in moments where participants reported that they "sought out new and exciting experiences" or "enjoyed taking risks", they also reported more acting on impulse and trouble controlling their feelings. One possibility is that this may reflect situational effects (such as responses to emotionally arousing states) that produce increases in both urgency and sensation seeking. Another contributing factor is that the battery of sensation seeking items selected for EMA adaptation did not include items on low-frequency events like skiing down a steep slope or skydiving, which may carry information more distinct from other impulsive domains and which often require advance planning. We also note relatively weak correlations in our data between EMA sensation seeking and alcohol problems. Sensation seeking has broadly been linked to the quantity and frequency of drinking in the trait literature (Coskunpinar et al., 2013; Stautz & Cooper, 2013), and a recent EMA study has provided initial evidence that sensation seeking on a given day predicts whether one drinks on that day (Lydon-Staley et al., 2019). Our findings are consistent with three large-scale meta-analyses finding weaker effect sizes for sensation seeking (as compared to other UPPS-P facets) in predicting alcohol dependence and alcohol problems (Berg et al., 2015; Coskunpinar et al., 2013; Stautz & Cooper, 2013). Future research should continue to distinguish between alcohol use and alcohol problems when investigating links impulsive states and alcohol.

It is critical to develop evidence-based research guidelines and recommendations that balance measurement frequency and validity with respondent burden. Despite nearly identical measures, Study 2 (which measured individuals 6 times per day) measures demonstrated higher stability within-person and greater trait-level convergent validity than Study 1 measures (sampled 3 times per day). However, at the within-person level, the pattern of correlations (and factor loadings) was nearly identical across studies, suggesting this differential reliability was limited to between-person associations of the aggregated measures. This discrepancy highlights the fact that scale reliability in an EMA context is a function of both the number of items in a scale and the number of assessments per individual (Shrout & Lane, 2014). While Study 1 response rates were higher than for Study 2, it is likely that the number of assessments and the method of responding (0 to 100 slider bar in Study 1; 1 to 4 ordinal scale in Study 2) influenced the reliability of measuring between-person variance in impulsive states. It is also possible that both reliability and response rates were likely also influenced by planned missingness in the Study 1 design. Though randomly administering items does not introduce bias and leads to a missing completely at random (MCAR) missing data mechanism, each combination of items is observed less frequently, leading to reduced power for estimating covariances and reduced reliability. When missing data are common (for example in planned missingness designs), data matrices quickly become difficult to impute using traditional methods as convergence is not guaranteed. One limitation of our approach is that it is not yet clear on what timescale impulsive states vary within-person. It is possible that impulsive states fluctuate at a finer-grained timescale than what was assessed in our current study, which may mean that our assessments smooth across multiple episodes of impulsivity. Future methodological work will be critical in guiding researchers to a useful balancing point between comprehensiveness of measurement and response rates, and to better understand the timescale on which states fluctuate.

The current study has several limitations. Though the UPPS-P is a broad, multidimensional, and widely-used measure of multiple impulsive domains, it is possible that additional aspects of impulsive behaviors were not captured by the UPPS-P. Information from the MIS and the DBQ helped to address the concern that the UPPS-P may miss important sources of variation, as MIS scores were closely related to urgency scores and DBQ scores correlated modestly with several UPPS-P impulsive states. Another limitation is that Study 1 only collected measures on 3 of the 5 impulsive traits. EMA self-report measures have inherent limitations; other forms of behavior tracking (e.g., location tracking, biometric measures, informant reports) could supplement knowledge gained through self-report measures. Information on divergent validity was not available – future studies should examine the extent to which different impulsive states differentially relate to a broad range of risky behaviors (e.g., gambling, risky drug use). Moreover, we only examined the between-person criterion validity of our impulsive state measures. Future research should test the association of impulsive states with within-person variation in thoughts, feelings and behaviors relevant to psychopathology. Regarding method variance, it is possible that high latent factor correlations are due to shared method effects (e.g., overall mood effects explaining covariance between impulsive states) rather than true overlap in impulsive states. Though our study did not address this limitation directly, it sets the foundation for future work building out the nomological network of these impulsive states which may help to disentangle method variance from true shared construct variance.

The study samples included high school students, college students who drink alcohol frequently, adults with ADHD histories who currently drink, and adult drinkers with no history of ADHD. Regarding age, our finding that the factor structure of impulsive states in adults and in youth is nearly identical provides some support for the use of these measures across a broad age range. Oversampling for individuals who are likely to experience impulsive states more frequently (students who drink, adults with ADHD) represents both a strength (the potential for increased construct coverage at the high end of impulsive states) and a threat to generalizability (these individuals may not represent the broader population). Though the sample size of the current study is sufficient to characterize within-subject variation (n=14,677 observations), longitudinal research on a broader participant pool would help to establish measurement invariance and generalizability across a wide range of participant characteristics. Regarding statistical limitations, ML-CFA and the methods described above for calculating reliability in multilevel models are relatively novel; thus, comparability to prior samples and benchmarks for "good" psychometrics and statistical power are still lacking.

Contemporary research on transdiagnostic risk factors has placed impulsive traits in a central position, as these traits relate to a broad range of internalizing psychopathology and externalizing psychopathology (Berg et al. 2015). To continue to deepen understanding of the day-to-day process by which impulsive traits relate to high-risk behaviors, more fine-grained assessment is necessary. Combining nomothetic principles with idiographic assessment has the potential to advance our understanding of how impulsive traits predispose impulsive individuals to particular problematic behaviors. Our results suggest that impulsive traits correspond to specific impulsive states, and that these states can be measured distinctly in a daily context. By assessing multiple types of impulsive behaviors, our ability as a field to link personality traits and contextual factors to problematic behaviors can be deepened and broadened. The adapted measures of impulsive states we present here hold promise for linking the "who" questions of current trait theories to the "when" and "under what conditions" questions of state theories.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

We thank the many students and research assistants who assisted in collecting the data for all samples. Dr. Pedersen's work on Study 1 was supported by K01-AA021135 (PI: Pedersen) and the ABMRF/The Foundation for Alcohol Research (PI: Pedersen). Data collection for Study 2 was supported in part by a gift to the University of Washington Center for Child and Family Well-Being from the Maritz Family Foundation.

References

Barkley RA, Murphy KR, & Fischer M (2010). ADHD in adults: What the science says. Guilford Press.

Berg JM, Latzman RD, Bliwise NG, & Lilienfeld SO (2015). Parsing the heterogeneity of impulsivity: A meta-analytic review of the behavioral implications of the UPPS for psychopathology. Psychological Assessment. 10.1037/pas0000111

Biederman J, Petty CR, Fried R, Doyle AE, Mick E, Aleardi M, Monuteaux MC, Seidman LJ, Spencer T, Faneuil AR, Holmes L, & Faraone SV (2008). Utility of an abbreviated questionnaire to identify individuals with ADHD at risk for functional impairments. Journal of Psychiatric Research, 42(4), 304–310. 10.1001/jamainternmed.2014.4010.Influence [PubMed: 17335849]

Campbell DT, & Fiske DW (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56(2), 81. [PubMed: 13634291]

Carver CS, Johnson SL, & Joormann J (2008). Serotonergic function, two-mode models of self-regulation, and vulnerability to depression: What depression has in common with impulsive aggression. Psychological Bulletin, 134(6), 912–943. 10.1037/a0013740 [PubMed: 18954161]

Clark LA (1993). Manual for the schedule for nonadaptive and adaptive personality. Minneapolis: University of Minnesota Press.

Coskunpinar A, Dir AL, & Cyders MA (2013). Multidimensionality in impulsivity and alcohol use: A meta-analysis using the UPPS model of impulsivity. Alcoholism: Clinical and Experimental Research, 37(9), 1441–1450. 10.1111/acer.12131

Cross CP, Copping LT, & Campbell A (2011). Sex differences in impulsivity: A Meta-analysis. Psychological Bulletin, 137(1), 97–130. 10.1037/a0021591 [PubMed: 21219058]

Curran PJ, & Bauer DJ (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. Annual Review of Psychology, 62, 583–619. 10.1146/annurev.psych.093008.100356

Cyders MA, & Coskunpinar A (2011). Measurement of constructs using self-report and behavioral lab tasks: Is there overlap in nomothetic span and construct representation for impulsivity? Clinical Psychology Review, 31(6), 965–982. 10.1016/j.cpr.2011.06.001 [PubMed: 21733491]

Cyders MA, & Smith GT (2007). Mood-based rash action and its components: Positive and negative urgency. Personality and Individual Differences, 43(4), 839–850. 10.1016/j.paid.2007.02.008

Cyders MA, Smith GT, Spillane NS, Fischer S, Annus AM, & Peterson C (2007). Integration of impulsivity and positive mood to predict risky behavior: Development and validation of a measure of positive urgency. Psychological Assessment, 19(1), 107–118. 10.1037/1040-3590.19.1.107 [PubMed: 17371126]

Duckworth AL, & Kern ML (2011). A meta-analysis of the convergent validity of self-control measures. Journal of Research in Personality, 45(3), 259–268. 10.1016/j.jrp.2011.02.004 [PubMed: 21643479]

Dyer NG, Hanges PJ, & Hall RJ (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. Leadership Quarterly, 16(1), 149–167. 10.1016/j.leaqua.2004.09.009

Eysenck MW, & Mogg K (1992). Clinical anxiety, trait anxiety, and memory bias. The Handbook of Emotion and Memory: Research and Theory, 429–450.

Fisher AJ, Medaglia JD, & Jeronimus BF (2018). Lack of group-to-individual generalizability is a threat to human subjects research. Proceedings of the National Academy of Sciences, 115(27), E6106–E6115. 10.1073/pnas.1711978115

Geldhof GJ, Preacher KJ, & Zyphur MJ (2014). Reliability estimation in a multilevel confirmatory factor analysis framework. Psychological Methods, 19(1), 72–91. 10.1037/a0032138 [PubMed: 23646988]

Graham JW (2009). Missing data analysis: Making it work in the real world. Annual Review of Psychology, 60, 549–576. 10.1146/annurev.psych.58.110405.085530

Graham JW, Taylor BJ, Olchowski AE, & Cumsille PE (2006). Planned missing data designs in psychological research. Psychological Methods, 11(4), 323–343. 10.1037/1082-989X.11.4.323 [PubMed: 17154750]

Gullo MJ, Loxton NJ, & Dawe S (2014). Impulsivity: Four ways five factors are not basic to addiction. Addictive Behaviors, 39, 1547–1556. 10.1016/j.addbeh.2014.01.002 [PubMed: 24576666]

Hasler BP, Wallace ML, White SJ, Molina BSG, & Pedersen SL (2019). Preliminary evidence that real world sleep timing and duration are associated with laboratory‐assessed alcohol response. Alcoholism: Clinical and Experimental Research, 43(7), 1575–1584. 10.1111/acer.14076

Heywood HB (1931). On Finite Sequences of Real Numbers. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 134(824), 486–501. 10.1098/rspa.1931.0209

Hu LT, & Bentler PM (1999). Cutoff Criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6, 1–55. 10.1080/10705519909540118

Hunt M, Auriemma J, & Cashaw ACA (2003). Self-report bias and underreporting of depression on the BDI-II. Journal of Personality Assessment, 80(1), 26–30. 10.1207/S15327752JPA8001 [PubMed: 12584064]

Hurlburt SC, & Sher KJ (1992). Assessing negative alcohol consequences in college students. J Am Coll Health, 41, 700–709.

Jackson DL, Gillaspy JA Jr, & Purc-Stephenson R (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. Psychological Methods, 14(1), 6. [PubMed: 19271845]

Jahng S, Wood PK, & Trull TJ (2008). Analysis of affective instability in ecological momentary assessment: Indices using successive difference and group comparison via multilevel modeling. Psychological Methods, 13(4), 354–375. 10.1037/a0014173 [PubMed: 19071999]

Kenny DA, & McCoach DB (2003). Effect of the number of variables on measures of fit in structural equation modeling. Structural Equation Modeling, 10(3), 333–351.

Kenrick DT, & Funder DC (1988). Profiting from controversy. Lessons from the person-situation debate. The American Psychologist, 43(1), 23–34. 10.1037/0003-066X.43.1.23 [PubMed: 3279875]

King KM, Patock-Peckham JA, Dager AD, Thimm K, & Gates JR (2014). On the mismeasurement of impulsivity: Trait, behavioral, and neural models in alcohol research among adolescents and young adults. Current Addiction Reports, 1, 19–32. 10.1007/s40429-013-0005-4

Lydon-Staley DM, Falk EB, & Bassett DS (2019). Within-person variability in sensation-seeking during daily life: Positive associations with alcohol use and self-defined risky behaviors. Psychology of Addictive Behaviors.

Lynam DR, Smith GT, Whiteside SP, & Cyders MA (2006). The UPPS-P: Assessing five personality pathways to impulsive behavior. West Lafayette, IN: Purdue University.

Marsh HW, Hau K-T, & Wen Z (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. Structural Equation Modeling: A Multidisciplinary Journal, 11(3), 320–341. 10.1207/s15328007sem1103_2

Mischel W, & Shoda Y (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. Psychological Review, 102(2), 246. [PubMed: 7740090]

Mischel W, Shoda Y, & Mendoza-Denton R (2002). Situation-behavior profiles as a locus of consistency in personality. Current Directions in Psychological Science, 11(2), 50–54.

Molina BSG, Pelham WE, Gnagy EM, Thompson AL, & Marshal MP (2007). Attention-deficit/hyperactivity disorder risk for heavy dDrinking and alcohol use disorder is age specific. Alcoholism: Clinical and Experimental Research, 31(4), 643–654. 10.1111/j.1530-0277.2007.00349.x

Molina BSG, Sibley MH, Pedersen SL, & Pelham WE (2017). The Pittsburgh ADHD Longitudinal Study. In Attention-deficit hyperactivity disorder: Adult outcomes and its predictors (pp. 105–155). Oxford Universiy Press.

Muthen BO (1994). Multilevel covariance structure analysis. Sociological Methods & Research, 22(3), 376–398. 10.1177/0049124194022003006

Muthen LK, & Muthen BO (2017). Mplus User's Guide (Eighth Edi).

Patton JH, Stanford MS, & Barratt ES (1995). Factor structure of the Barratt impulsiveness scale. Journal of Clinical Psychology, 51(6), 768–774. [PubMed: 8778124]

Pedersen SL, King KM, Louie KA, Fournier JC, & Molina BS (2019). Momentary fluctuations in impulsivity domains: Associations with a history of childhood ADHD, heavy alcohol use, and alcohol problems. Drug and Alcohol Dependence, 205, 107683. [PubMed: 31704385]

Pedersen SL, Walther CAP, Harty SC, Gnagy EM, Pelham WE, & Molina BSG (2016). The indirect effects of childhood ADHD on alcohol problems in adulthood through unique facets of impulsivity. Addiction, 111(9), 1582–1589. 10.1002/jmri.25711.PET/MRI [PubMed: 26999438]

Pilkonis PA, Choi SW, Reise SP, Stover AM, Riley WT, & Cella D (2011). Item banks for measuring emotional distress from the patient-reported outcomes measurement information system (PROMIS®): Depression, anxiety, and anger. Assessment, 18(3), 263–283. 10.1177/1073191111411667 [PubMed: 21697139]

Robinson MD, & Clore GL (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. Psychological Bulletin, 128(6), 934–960. 10.1037/0033-2909.128.6.934 [PubMed: 12405138]

Roley ME, Contractor AA, Weiss NH, Armour C, & Elhai JD (2017). Impulsivity facets' predictive relations with DSM-5 PTSD symptom clusters. Psychological Trauma, 9(1), 76–79. 10.1037/tra0000146.Impulsivity [PubMed: 27243571]

Settles RE, Zapolski TCB, & Smith GT (2014). Longitudinal test of a developmental model of the transition to early drinking. Journal of Abnormal Psychology, 123(1), 141–151. 10.1016/j.neuroimage.2013.08.045.The [PubMed: 24661166]

Sharma L, Kohl K, Morgan TA, & Clark LA (2013). “Impulsivity”: Relations between self-report and behavior. Journal of Personality and Social Psychology, 104(3), 559–575. [PubMed: 23437924]

Sharma L, Markon KE, & Clark LA (2013). Toward a theory of distinct types of “impulsive” behaviors: A meta-analysis of self-report and behavioral measures. Psychological Bulletin, 140(2), 374–408. 10.1037/a0034418 [PubMed: 24099400]

Shiffman S, Stone AA, & Hufford MR (2008). Ecological momentary assessment. Annu. Rev. Clin. Psychol, 4, 1–32. [PubMed: 18509902]

Shrout PE, & Lane SP (2014). Psychometrics. In Mehl MR & Conner TS (Eds.), Handbook of Research Methods for Studying Daily Life (pp. 302–320). Guilford Publications.

Smith GT, & Cyders MA (2016). Integrating affect and impulsivity: The role of positive and negative urgency in substance use risk. Drug and Alcohol Dependence, 163, S3–S12. 10.1016/j.drugalcdep.2015.08.038 [PubMed: 27306729]

Smith GT, Fischer S, Cyders MA, Annus AM, Spillane NS, & McCarthy DM (2007). On the validity and utility of discriminating among impulsivity-like traits. Assessment, 14(2), 155–170. 10.1177/1073191106295527 [PubMed: 17504888]

Stautz K, & Cooper A (2013). Impulsivity-related personality traits and adolescent alcohol use: A meta-analytic review. Clinical Psychology Review, 33(4), 574–592. 10.1016/j.cpr.2013.03.003 [PubMed: 23563081]

Tomko RL, Solhan MB, Carpenter RW, Brown WC, Jahng S, Wood PK, & Trull TJ (2014). Measuring impulsivity in daily life: The Momentary Impulsivity Scale. Psychological Assessment, 26(2), 339–349. 10.1037/a0035083 [PubMed: 24274047]

von Neumann J, Kent RH, Bellinson HR, & Hart BI (1941). The mean square successive difference. The Annals of Mathematical Statistics, 12(2), 153–162. 10.1214/09-STS284

Whiteside SP, & Lynam DR (2001). The five factor model and impulsivity: Using a structural model of personality to understand impulsivity. Personality and Individual Differences, 30(4), 669–689. 10.1016/S0191-8869(00)00064-7

Wu KD, & Clark LA (2003). Relations between personality traits and self-reports of daily behavior. Journal of Research in Personality, 37(4), 231–256. 10.1016/S0092-6566(02)00539-1

Table 1.

Measures of impulsive traits/states

Construct	Original Item	EMA Adaptation	S1	S2	Item
(Negative) Urgency	It is hard for me to resist acting on my feelings.	It was hard for me to resist acting on my feelings.	X	X	resist
	I have trouble controlling my impulses.	I had trouble controlling my impulses.	X	X	icontrol
	When overjoyed, I feel like I can't stop myself from going overboard.	I felt like I couldn't stop myself from going overboard.[a]	X		stop
	I often get involved in things I later wish I could get out of.	I got involved in something I later wished I could get out of.	X		involve
	I always keep my feelings under control.	I kept my feelings under control.[b]	X	X	fcontrol
	Others are shocked or worried about the things I do when I am feeling very excited	Others were shocked or worried about the things I did.[a]	X		shock
Planning	I usually think carefully before doing anything.	I thought carefully before doing anything.	X	X	thought
	I tend to value and follow a rational, "sensible" approach to things.	I followed a rational, "sensible" approach to things.	X	X	rational
	Before making up my mind, I consider all the advantages and disadvantages.	Before making up my mind, I considered all the advantages and disadvantages.	X	X	consider
Persistence	I generally like to see things through to the end.	I saw things through to the end.	X	X	seethrough
	I concentrate easily.	I concentrated easily.	X	X	concentrate
	I finish what I start.	I finished what I started.	X	X	finish
Sensation Seeking	I generally seek new and exciting experiences and sensations.	I sought out new and exciting experiences and sensations.		X	newexp
	I quite enjoy taking risks.	I enjoyed taking risks.		X	risk
	I sometimes like doing things that are a bit frightening.	I wanted do something that was a bit frightening.		X	frighten
	I would enjoy fast driving.	I enjoyed fast driving.		X	fastdrive
Positive Urgency	I tend to lose control when I am in a great mood.	I lost control when I was in a great mood.		X	pcontrol
	Others are shocked or worried about the things I do when I am feeling very excited.	Others were shocked or worried about the things I did when I was feeling very excited.		X	pshock
	When overjoyed, I feel like I can't stop myself from going overboard.	When overjoyed, I felt like I couldn't stop myself from going overboard.		X	pstop
	I tend to act without thinking when I am really excited.	I acted without thinking when I was really excited.		X	pnothink

[+] Stem for items was "Please indicate how much you agree or disagree with the statement."

[a] Item was originally a positive urgency item, but was adapted in Study 1 to be stripped of emotional content;

[b] Item was worded oppositely in Study 2: "I lost control."

Table 2.1.

Between-person factor loadings

	Between-Person Factor Loading									
	(Negative) Urgency		Planning		Persistence		SS	Pos Urg	% Variance Explained	
	Study 1 (n=215; 30 obs/person)	Study 2 (n=211; 60 obs/person)	1	2	1	2	2	2	1	2
resist	1.00 *	0.97							29%	87%
icontrol	0.93	1.00 *							75%	93%
stop	0.97	-							100%	-
involve	0.79	0.89							61%	78%
fcontrol	0.74	-							86%	-
shock	0.80	-							80%	-
thought			1.00 *	1.00 *					100%	98%
rational			0.87	0.96					75%	83%
consider			0.93	1.00					80%	95%
seethrough					1.00 *	0.98			100%	95%
concentrate					0.74	1.00 *			58%	86%
finish					0.84	0.99			86%	99%
newexp							1.00 *		-	92%
risk							0.84		-	76%
frighten							0.79		-	39%
fastdrive							0.92		-	76%
pcontrol								1.00 *	-	97%
pshock								0.96	-	99%
pstop								0.84	-	90%

Note: SS = Sensation Seeking; Pos Urg = Positive Urgency. Items listed in same order as in Table 1.

* Item loading fixed to 1.00. % Variance Explained indicates the % of variance in the item explained by its latent factor.

Table 2.2.

Within-person factor loadings

	Within-Person Factor Loading									
	(Negative) Urgency		Planning		Persistence		SS	Pos Urg	% Variance Explained	
	Study 1 (n=215; 30 obs/person)	Study 2 (n=211; 60 obs/person)	1	2	1	2	1	1	1	2
resist	0.96	0.98							29%	34%
icontrol	0.87	1.00 *							25%	40%
stop	1.00 *	-							44%	-
involve	0.65	0.73							14%	22%
fcontrol	0.96	-							54%	-
shock	0.71	-							32%	-
thought			1.00 *	1.00 *					47%	43%
rational			0.96	0.97					44%	43%
consider			0.88	0.98					37%	43%
seethrough					1.00 *	0.91			50%	38%
concentrate					0.74	0.88			29%	37%
finish					0.99	1.00 *			48%	52%
newexp							1.00 *		-	34%
risk							0.90		-	37%
frighten							0.49		-	12%
fastdrive							0.96		-	30%
pcontrol								0.96	-	37%
pshock								1.00 *	-	49%
pstop								0.81	-	36%

Note: SS = Sensation Seeking; Pos Urg = Positive Urgency. Items listed in same order as in Table 1.

* Item loading fixed to 1.00. % Variance Explained indicates the % of variance in the item explained by its latent factor.

Table 3.

Correlations among latent factors.

	(Negative) Urgency	Planning	Persistence	Sensation Seeking	Positive Urgency
Between-person factor correlations					
(Negative) Urgency	1	−0.24	−0.19		
Planning	−0.54	1	0.77		
Persistence	−0.55	0.94	1		
Sensation Seeking	0.69	−0.26	−0.23	1	
Positive Urgency	0.97	−0.48	−0.49	0.70	1
Within-person factor correlations					
	(Negative) Urgency	**Planning**	**Persistence**	**Sensation Seeking**	**Positive Urgency**
(Negative) Urgency	1	−0.21	−0.10		
Planning	−0.20	1	0.81		
Persistence	−0.11	0.82	1		
Sensation Seeking	0.67	0.06	0.21	1	
Positive Urgency	0.71	−0.18	−0.08	0.71	1

Note: Correlations above the diagonal are from Study 1 (n=215, 30 observations/person). Correlations beneath the diagonal are from Study 2 (n=211, 60 observations/person).

Table 4.

Person-level convergent validity: Raw correlations for baseline and EMA-aggregated measures of impulsive traits.

Study 1

	T NU	T Pr	T Pe	T SS	T PU	MS NU	MS Pr	MS Pe	MS SS	MS PU
T (Negative) Urgency	1									
T Planning	−0.27	1								
T Persistence	−0.42	0.40	1							
T Sensation Seeking										
T Positive Urgency										
MS (Negative) Urgency	0.55	−.0.24	−0.29			1				
MS Planning	−0.27	0.28	0.23			−0.20	1			
MS Persistence	−0.25	0.08	0.32			−0.19	0.77	1		
MS Sensation Seeking										
MS Positive Urgency										

Study 2

	T NU	T Pr	T Pe	T SS	T PU	MS NU	MS Pr	MS Pe	MS SS	MS PU
T (Negative) Urgency	1									
T Planning	−0.51	1								
T Persistence	−0.62	0.65	1							
T Sensation Seeking	0.05	−0.06	0.13	1						
T Positive Urgency	0.74	−0.51	−0.53	0.14	1					
MS (Negative) Urgency	0.58	−0.40	−0.51	0.06	0.68	1				
MS Planning	−0.38	0.47	0.46	0.11	−0.40	−0.51	1			
MS Persistence	−0.39	0.37	0.48	0.15	−0.38	−0.54	0.92	1		
MS Sensation Seeking	0.32	−0.25	−0.19	0.36	0.47	0.63	−0.22	−0.20	1	
MS Positive Urgency	0.48	−0.36	−0.41	0.06	0.65	0.95	−0.46	−0.49	0.65	1

Note: T denotes trait measure, MS denotes mean of state measurements. Column names correspond to row names. Gray shaded cells indicate cross-method correlations. Empty cells are included in Sample 1 for comparability across studies.

Author Manuscript
Author Manuscript
Author Manuscript

Table 5.

Person-level criterion validity: Raw correlations between impulsive domain measures and behavioral/emotional problems.

		Study 1				Study 2		
		Anger	**Anxiety**	**Depression**	**Alcohol Problems**	**Alcohol Problems**	**ODD Symptoms**	**ADHD Symptoms**
(Negative) Urgency	Global self-report	0.62	0.55	0.51	0.31	0.42	0.55	0.62
	EMA-aggregate	0.40	0.37	0.36	0.30	0.34	0.44	0.49
	MSSD	0.21	0.20	0.21	0.15	0.19	0.42	0.46
Planning	Global self-report	−0.09	−0.08	−0.09	−0.30	−0.25	−0.31	−0.47
	EMA-aggregate	−0.16	−0.08	−0.15	−0.12	−0.22	−0.23	−0.32
	MSSD	0.07	−0.05	−0.02	0.00	0.12	0.28	0.22
Persistence	Global self-report	−0.25	−0.34	−0.39	−0.09	−0.30	−0.48	−0.58
	EMA-aggregate	−0.26	−0.24	−0.28	−0.03	−0.25	−0.29	−0.34
	MSSD	0.09	−0.01	0.05	−0.04	0.09	0.24	0.21
Sensation Seeking	Global self-report					0.06	−0.05	0.00
	EMA-aggregate					0.17	0.22	0.20
	MSSD					0.08	0.36	0.22
Positive Urgency	Global self-report					0.35	0.46	0.52
	EMA-aggregate					0.32	0.39	0.37
	MSSD					0.16	0.44	0.32

Note: MSSD = mean squared successive deviation; ODD = oppositional defiant disorder; ADHD = attention deficit/hyperactivity disorder; EMA = ecological momentary assessment

Table 6.

Momentary convergent validity: Raw between- and within-person correlations between Study 1 UPPS-P impulsive state measures,Tomko et al. (2014) Momentary Impulsivity Scale, andSharma, Kohl, Morgan & Clark (2013) Daily Behavior Questionnaire.

	Urgency	Planning	Persistence	MIS	DBQ - Spontaneous	DBQ - Planful
Urgency		−0.19	−0.16	0.68	0.16	−0.12
Planning	−0.25		0.60	−0.10	−0.06	0.19
Persistence	−0.22	0.76		−0.06	−0.05	0.16
MIS	0.88	−0.15	−0.11		0.22	−0.05
DBQ - Spontaneous	0.25	−0.06	−0.03	0.38		0.01
DBQ - Planful	−0.20	0.38	0.31	−0.11	0.17	

Note: Values above diagonal are within-person correlations, values below diagonal are between-person correlations. MIS = Momentary Impulsivity Scale, DBQ = Daily Behavior Questionnaire.