Supplemental Materials

Inclusion criteria and data collection procedures for the five samples

Sample 1

Participants were instructed to install the MetricWire (MetricWire Inc., 2019) application onto their personal smartphones. Participants were assigned to either the event- or signal contingent recording condition as part of a larger study (Himmelstein et al., 2019).

Those assigned to the signal-contingent condition ($n = 148$) were sent a random prompt six times per day over seven days between the hours of 10 AM and 10 PM. The prompts asked participants to reflect on their most recent social interaction of at least five minutes that had occurred since the last assessment. Prompts were spaced at a minimum of 90 minutes apart. Participants were sent a reminder prompt if they did not respond to the initial prompt within 15 minutes. The survey was made unavailable after 30 minutes of no response. Individuals in the event-contingent condition ($n = 140$) were instructed to complete a survey after each social interaction they had that lasted at least five minutes.

In the event-contingent recording condition, participants were asked to access the MetricWire app and complete a survey each time they had a social interaction. The event-contingent condition social interaction survey was identical to that used in the signal-contingent condition, with the exception of the question of whether or not the participant had experienced a social interaction since the previous survey. Participants in both conditions received course credit if they completed the baseline assessment and an average of four EMA surveys per study day. Analyses of contingency groups revealed no systematic differences in means or variances of the VIAS or affect (Himmelstein et al., 2019), and consequently data across the two conditions was

merged for the present study. In total, 10,345 responses were collected, with participants responding to an average of 35.92 ($SD = 14.62$) surveys each.

Sample 2

The EMA protocol in this sample was identical to that of the signal-contingent condition group in Undergraduate Sample 1 (i.e., there was no event-contingent condition). Participants in this sample were incentivized to participate with a chance to win one of several prizes, with likelihood of winning a prize increasing proportionate to the number of EMA surveys submitted. In total, 8,768 responses were collected from included participants, averaging 22.14 ($SD = 14.16$) surveys each.

Sample 3

Participants were recruited online and through flyers. For inclusion, participants had to be between 18 and 40 years of age. Participants also had to be users of a smartphone running iOS or Android software. To recruit a distinct community sample, individuals were not eligible if they were enrolled in a full-time undergraduate program. Participants were pre-screened to ensure a gender-balanced sample as well as adequate representation of personality traits of interest to the larger study of narcissism. Namely, modesty was assessed in the pre-screen using the NEO Personality Inventory – Revised (Costa & McCrae, 1992). Low modesty, a core feature of narcissism, was oversampled such that a 2:1:1 ratio of low, moderate, and high levels of modesty within each gender were recruited.

In Community Sample 3, participants viewed a brief training video on how to install and use the MetricWire application (MetricWire Inc., 2019). The signal-contingent protocol from Undergraduate Samples 1 and 2 was also used in Community Sample 3. Participants were incentivized to participate through random drawings for Amazon.com gift cards upon completion

of the baseline questionnaires and guaranteed gift certificates for completion of the EMA surveys. In total, 17,735 responses were collected, with participants responding to an average of 51.86 ($SD = 17.62$) surveys each.

Sample 4

The three days of baseline EMA data collection was described in Chin et al., 2019, but is described again below. EMA surveys were administered using participants' personal smartphones using Qualtrics surveys delivered through SurveySignal text links and MetricWire surveys. Participants were prompted to complete five EMA surveys daily at quasi-random times each day (30 surveys total across the baseline and postintervention periods). Text links were sent during each of five 2-hr blocks distributed between 9:00 a.m. and 7:00 p.m., with links expiring after 45 min. Participants were also prompted to complete daily diary assessments at 8:30 p.m. each day; links were sent at exactly 8:30 p.m. and remained active until 11:30 p.m. Participants were trained to complete all EMA assessments during the baseline study appointment. The EMA assessment began on a Wednesday and concluded on a Friday. In total, 2,849 observations were analyzed with an average of 19.25 ($SD = 8.67$) valid submissions per participant.

Sample 5

Individuals with active substance dependence or a history of amnestic disorder, bipolar I disorder, cognitive disorder, delirium, delusional disorder, dementia, schizophrenia, or schizoaffective disorder were excluded from the study.

Participants were issued a smartphone to use during the 21-day EMA protocol. The smartphones featured a pre-installed, custom application that provided random prompts to participants over 12 hours, with the beginning of the start time chosen by the participants. The software divided the 12-hour period into 6 equal intervals and randomly prompted participants

once during each interval. In addition, participants had the ability to initiate surveys from within

the software and were asked to use this feature after each social interaction (i.e., event contingent

recording; Ellison et al., 2020). In total, 4,419 responses were analyzed, with participants

completing an average of 81.83 ($SD = 53.42$) surveys.

Table A1

Response scales and anchors used in the five samples

Construct	Reference	Undergraduate Sample 1	Undergraduate Sample 2	Community Sample 3	Community Sample 4	Outpatient Sample 5	Scale	Anchors
Affect	Watson & Clark (1999)	X	X	X			100-point	"Not at all" to "Extremely"
					X		7-point	"Not at all" to "Extremely"
						X	100-point	"Not at all" to "Extremely"
Momentary Stress	*Ad hoc*	X		X			100-point	"Not at all" to "Extremely"
	Ad hoc					X	100-point	"Low" to "High"
Trait dominance and warmth	Markey & Markey (2009)	X					5-point	"Very inaccurate" to "Very accurate"
Interpersonal Problems	Soldz et al. (1995)		X	X			5-point	"Not at all" to "Extremely"
	Boudreaux et al. (2018)			X			4-point	"Not a problem" to "Serious problem"
Interpersonal Values	Locke et al. (2012)		X				5-point	"Not important to me" to "Extremely important to me"
Big-5 Personality	Soto & John (2017)	X		X			5-point	"Disagree strongly" to "Agree strongly"
Narcissism	Pincus et al. (2009)	X	X				5-point[†]	"Very untrue of me" to "Very true of me"
	Schoenleber et al. (2015)		X	X			5-point	"Very untrue of me" to "Very true of me"
	Sherman et al. (2015)		X	X			5-point	"Very untrue of me" to "Very true of me"

Construct	Citation					Scale	Anchors
Perceived stress	Cohen et al. (1983)			X		5-point	"Never" to "Very often"
Loneliness	Russell (1996)			X		4-point	"Never" to "Always"
Borderline Personality Disorder	Zanarini et al. (2003)				X	2-point	"Yes" or "No"
Anxiety	Wright & Simms (2014)	X				5-point	"Very untrue of me" to "Very true of me"
	Beck & Steer (1990)		X			4-point	"Not at all" to "It bothered me a lot"
	Newman et al. (2002)				X	Varies	Varies
Depression	Wright & Simms (2014)	X				5-point	"Very untrue of me" to "Very true of me"
	Beck et al. (1996)		X	X	X	4-point	Varies

Note. Due to an administration error during data collection, the PNI was assessed using a 5-point scale in both samples rather than the traditional 6-point scale.

Table A2

Anchors of ad hoc momentary items in Community Sample 5

Item	Anchors
Ability to perform important tasks	"Low" to "High"
Ability to relate to others	"Poor" to "Excellent"
Anxiety	"Mild" to "Severe"
Depression	"Mild" to "Severe"
How well do you know this person?	"Not at all" to "Very well"
Current feelings of "Pleasantness"	"Unpleasant" to "Pleasant"
Mental health symptoms	"Mild" to "Severe"
Physical health symptoms	"Mild" to "Severe"
Self-esteem	"Low" to "High"

References

Beck, A. T., & Steer, R. A. (1990). *Manual for the revised Beck Anxiety Inventory*. The

    Psychological Corporation.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory–II.*

    (2nd ed.). The Psychological Corporation.

Boudreaux, M. J., Ozer, D. J., Oltmanns, T. F., & Wright, A. G. C. (2018). Development and

    validation of the circumplex scales of interpersonal problems. *Psychological Assessment*,

    30(5), 594-609.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress.

    Journal of Health and Social Behavior, *24*(4), 385–396.

Locke, K. D., Craig, T., Baik, K. D., & Gohil, K. (2012). Binds and bounds of communion:

    effects of interpersonal values on assumed similarity of self and others. *Journal of*

    Personality and Social Psychology, *103*(5), 879-897.

Markey, P. M., & Markey, C. N. (2009). A brief assessment of the interpersonal circumplex: The

    IPIP-IPC. *Assessment*, *16*(4), 352-361.

MetricWire Inc. (2019). MetricWire (Version 4.2.8). https://metricwire.com

Newman, M. G., Zuellig, A. R., Kachin, K. E., Constantino, M. J., Przeworski, A., Erickson, T.,

    & Cashman-McGrath, L. (2002). Preliminary reliability and validity of the Generalized

    Anxiety Disorder Questionnaire-IV: A revised self-report diagnostic measure of

    generalized anxiety disorder. *Behavior Therapy*, *33*(2), 215-233.

Pincus, A. L., Ansell, E. B., Pimentel, C. A., Cain, N. M., Wright, A. G. C., & Levy, K. N.

    (2009). Initial construction and validation of the Pathological Narcissism Inventory.

    Psychological Assessment, *21*(3), 365-379.

Russell, D. W. (1996). UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. *Journal of Personality Assessment*, *66*(1), 20-40.

Schoenleber, M., Roche, M. J., Wetzel, E., Pincus, A. L., & Roberts, B. W. (2015). Development of a brief version of the Pathological Narcissism Inventory. *Psychological Assessment*, *27*(4), 1520-1526.

Sherman, E. D., Miller, J. D., Few, L. R., Campbell, W. K., Widiger, T. A., Crego, C., & Lynam, D. R. (2015). Development of a short form of the Five-Factor Narcissism Inventory: The FFNI-SF. *Psychological Assessment*, *27*(3), 1110-1116.

Soldz, S., Budman, S., Demby, A., & Merry, J. (1995). A short form of the Inventory of Interpersonal Problems Circumplex Scales. *Assessment*, *2*(1), 53-63.

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117-143.

Watson, D., & Clark, L. A. (1999). *The PANAS-X: Manual for the positive and negative affect schedule-expanded form*. The University of Iowa.

Wright, A. G. C., & Simms, L. J. (2014). On the structure of personality disorder traits: Conjoint analyses of the CAT-PD, PID-5, and NEO-PI-3 trait models. *Personality Disorders: Theory, Research, and Treatment*, *5*(1), 43-54.

Zanarini, M. C., Vujanovic, A. A., Parachini, E. A., Boulanger, J. L., Frankenburg, F. R., & Hennen, J. (2003). A screening measure for BPD: The McLean Screening Instrument for Borderline Personality Disorder (MSI-BPD). *Journal of Personality Disorders*, *17*(6), 568-573.