Development, validation, and faking-resistance of an implicit measure of psychopathy in the workplace

Ryan Cook[1], Nicolas Roulin[1], & Keith Joy[2]

Article "in press" in *Human Performance*

Corresponding author: Ryan Cook, MS 323, Saint Mary's University, 923 Robie Street | Halifax, N.S. Canada | B3H 3C3, ryan.cook@smu.ca

[1] Psychology Department, Saint Mary's University, Halifax, Nova Scotia, Canada
[2] Canadian Armed Forces

Abstract

Researchers have called for faking-resistant measures of psychopathic personality that can be self-administered in high-stakes contexts (e.g., hiring). We developed and validated an implicit measure of psychopathy contextualized in workplace situations. We first detail how the measure is framed, conceptualized, and rooted in psychopathy literature. We then describe the item development process, and Study 1 involves expert review and refining the item list. In Study 2 (N = 396), we examine internal consistency and factor structure for a 22-item version of the measure. In Study 3 (N = 251), we demonstrate test-retest reliability, construct-related validity, and provide initial evidence for criterion-related validity through a two-wave study. Study 4 analyzes the measure using item response theory, based on a sample of 6,746 job seekers, demonstrating effectiveness for measuring high levels of psychopathy. In Study 5 (N = 219) we provide evidence of faking-resistance and criterion-related validity with behavioural (two weeks later) and self-report (one year later) outcomes. Finally, Study 6 provides promising evidence of incremental validity using an organizational sample (N = 615). Overall, scores on this new implicit measure are reliable, with acceptable construct-related, criterion-related, and incremental validity, while also being faking-resistant. Implications for use in workplace settings are discussed.

Keywords: psychopathy, personality, selection, conditional reasoning,

Development, Validation, and Faking-resistance of an Implicit Measure of Psychopathy in the Workplace

Manipulative, deceitful, arrogant, insensitive, remorseless, cold-hearted, egocentric, risky, and parasitic; these are some adjectives commonly used to describe people high on trait psychopathy. Psychopathy in the workplace has been of particular interest to the field of psychology over the past two decades (Babiak, 2007; Smith et al., 2014) and is usually referred to as corporate psychopathy, successful psychopathy, professional psychopathy, or workplace psychopathy. Although there may be minor differences among these terms (e.g., "corporate" and "industrial" usually describing white-collar professions specifically), they are often used interchangeably.

Henley (2002) found individuals scoring high on psychopathic personality prefer risky work activities, a solitary work style, a desire to control or lead, and enjoy collecting wealth or prestige, but avoid careers with high dependency on others, objectively boring, or stereotypically female. They are also drawn to law enforcement and military professions, along with business, contact sports, and politics (Smith et al., 2014). Some aspects of psychopathy can be adaptive, and even beneficial, in some contexts. For instance, Lilienfeld et al. (2012) analyzed personality data for 42 U.S. Presidents and found components of corporate psychopathy were correlated with better ratings of performance, persuasiveness, crisis management, and leadership ability.

However, workplace psychopathy is associated with many negative outcomes for organizations (Boddy et al., 2010; Mathieu & Babiak, 2016b), which largely outnumber its potential benefits (Smith & Lilienfeld, 2013). For instance, high levels of psychopathic personality are associated with high degrees of counterproductive work behaviours (CWBs) such as theft and interpersonal abuse (Scherer et al., 2013), unethical decision-making (Stevens et al., 2012), white-collar crime (Pardue et al., 2013) and more (see Lilienfeld et al., 2015). Psychopathy levels are also associated with worse task

performance, such as consistent losses for hedge fund managers (ten Brinke et al., 2018). A meta-analysis by Landay et al. (2019) showed a negative (but curvilinear) relationship between psychopathic personality and leadership effectiveness. Importantly, leaders with psychopathic tendencies can also impact colleagues or subordinates at work: They are more likely to use an abusive and authoritarian leadership style, which results in conflicts, bullying, lower affective well-being and increased CWBs by subordinates (Boddy, 2011; Scherer et al., 2013), as well as decreased job satisfaction and higher turnover intent among colleagues and subordinates (Boddy et al., 2010; Mathieu & Babiak, 2016b). To quantify the negative impacts of potential psychopathic tendencies in leaders, researchers estimated that abusive supervision cost organizations $23.8 billion in the U.S. alone (Tepper et al., 2006), and a single psychopathic leader may cause up to $100 million worth of damage or 40% staff turnover in a single year (Boddy et al., 2015). With regard to white-collar crime, a recent Association of Certified Fraud Examiners (ACFE) report suggests that fraud and embezzlement amount to $1.08 trillion in the U.S. annually (ACFE, 2020).

Given the risks of hiring or promoting individuals with highly psychopathic personality, especially into leadership positions where they have autonomy and control over others, it is surprising that existing hiring tools are limited in their ability to accurately and efficiently screen for these people (Smith & Lilienfeld, 2013). Existing measures of psychopathy generally fall into the following categories: (1) They require a detailed clinical assessment which is very resource-intensive, especially for organizations using it for non-clinical purposes; (2) They require a third-party rater who has experience with the individual (e.g., a co-worker), which comes with the inherent issues of laypeople over-estimating levels of psychopathy in others (Caponecchia et al., 2012); (3) They are overt self-reports and thus highly susceptible to faking and socially desirable responding, which is magnified in incentivized or high-stakes contexts (Kelsey, 2016; Smith & Lilienfeld, 2013).

Implicit personality measures, such as Conditional Reasoning Tests (CRTs), have been suggested as a potential answer to these problems. CRTs are still self-administered, yet their implicit nature makes them more resistant to faking (Bowler & Bowler, 2014; Galić et al., 2014), even when given an incentive (LeBreton et al., 2007; Wiita et al., 2020). In fact, both personality and CRT researchers have specifically called for developing CRTs as the most promising method for assessing psychopathy (and other dark, socially undesirable personality traits) in a workplace setting (James & LeBreton, 2012; O'Boyle et al., 2012; Spain et al., 2014; Wu & Lebreton, 2011).

The current research aims to answer these calls and to develop and validate an implicit measure of psychopathy in the workplace, labelled the Conditional Reasoning Test for Workplace Psychopathy (CRT-WP). As such, we contribute to the workplace psychopathy, personality, and implicit measurement (e.g., CRT) literatures in the following ways. First, by using a hybrid approach combining best-practice in CRT development (James & LeBreton, 2012; LeBreton et al. 2020) with classical test development building from established theories/models of psychopathy (e.g., Hare, 1985; Lilienfeld & Andrews, 1996; Patrick, 2010), we develop justification mechanisms and work-related conditional reasoning problems designed to assess psychopathic tendencies in organizational settings. We then refine our measure based on expert review (Study 1) and factor analysis (Study 2) to ensure content adequacy, theoretical relevance, and practicality. In Study 3, we provide initial evidence of the temporal stability, construct-related validity (via relationships with established measures of psychopathy), and criterion-related validity (via relationships with CWBs) of CRT-WP scores. Study 4 explores the item- and test-level functioning from an Item Response Theory (IRT) standpoint, providing detailed information about the effectiveness of the CRT-WP to measure high levels of psychopathy. Study 5 demonstrates the faking-resistant nature of the CRT-WP in contrast to other relevant overt measures (dark triad, honesty-humility, integrity) and provides additional evidence of

criterion-related validity with both behavioural outcomes (e.g., selfish decisions in a dictator game) and self-reports of CWBs and cyber-loafing one year later. Finally, Study 6 provides evidence for the incremental validity of CRT-WP scores to predict CWBs beyond self-reported personality measures in an organizational sample. Importantly, the CRT-WP is designed as a *workplace-contextualized* measure (thus enhancing its job-relevance) of *sub-clinical* levels of the personality trait of psychopathy, but it is not designed or intended to make clinical assessments or a binary psychopath/non-psychopath decision. Instead, it measures levels of a personality trait on a continuum, similar to most common personality inventories. Overall, we provide an implicit measure of sub-clinical and workplace-contextualized psychopathy with scores that are reliable and valid.

Workplace Psychopathy

Babiak (1995) described psychopaths in the workplace as charismatic, influential, and viewed positively by most co-workers, but also highly manipulative, coercive, oppressive, and viewed more negatively by those who were less "useful" to them. Babiak suggested that these "corporate psychopaths" have the same fundamental personality traits as other psychopaths, but lack the easily detectable anti-social behaviours commonly associated with criminal psychopaths (e.g., physical violence, crime) *or* use ways to avoid getting caught (e.g., by acting behind office doors).

Defining corporate psychopathy is difficult, and researchers are still exploring how and why these successful psychopaths differ from the typical criminal psychopaths who often have trouble with the law from an early age (Hare, 1985). There are three competing models to explain their differences and similarities (Lilienfeld et al., 2015). First, the *moderated-expression* model states that "successful" psychopathy results from some of the core traits and behaviours being mitigated by other factors like intelligence or effective parenting. Second, the *differential-configuration* model posits that successful psychopaths possess different or additional traits compared to criminal psychopaths, for instance high

conscientiousness and boldness (Lilienfeld et al., 2015). Finally, the *differential-severity* model states that corporate psychopathy is simply a milder expression of clinical psychopathy. While other explanations have been provided, for instance using neurobiological differences (Gao & Raine, 2010), there is the least, and limited, empirical support for the third model (Lilienfeld et al., 2015), some for the first (Blickle & Schütte, 2017; Boddy et al., 2010), and the most for the second (Vergauwe et al., 2021).

Organizations often evaluate employees or applicants using personality tests (e.g., conscientiousness, integrity), however, assessing psychopathic personality is rarely considered (Mathieu & Babiak, 2016b) despite the potential to help reduce occurrences of abusive supervision, unethical decision making, and manipulation. Doing so could help prevent people with psychopathic traits from reaching leadership roles or positions of authority over the public (i.e., law enforcement) or clients (i.e., health care). It is important to emphasize that psychopathy is *not* considered a mental disorder (e.g., in the DSM-V), but other personality constructs are (e.g., narcissistic or antisocial personality disorders). Thus, as we describe in more detail below, while designing the CRT-WP we deliberately eliminated dimensions of psychopathy related criminal tendencies or history, which could be connected to mental disorders (e.g., antisocial personality) and thus risks of being legally challenged (e.g., based on the Americans with Disabilities Act in the United States). As we discuss next, the lack of psychopathic personality assessment in hiring is largely because existing measures have various practical challenges.

Existing Psychopathy Measures

In Table 1, we summarize the issues with applying existing measures of psychopathic personality to organizational settings. For instance, what is considered by most as the best assessment of clinical psychopathy, the Psychopathy Checklist-Revised (PCL-R; Hare, 1980), requires a clinical

professional performing a multifaceted assessment, which would be (a) too costly for most organizations, (b) likely lead to negative reactions from employees or applicants, and/or (c) open organizations to legal challenges. The B-Scan 360 (Babiak & Hare, 2012) and the Psychopathy Measure – Management Research Version (PM-MRV; Boddy, 2011) require raters familiar with the targets' behaviours and tendencies (e.g., coworkers) to make a valid assessment, which is unpractical in hiring contexts, for example. The B-Scan also replicates the same structure as (and strongly correlates with) the PCL-R, which could result in similar legal issues.

Perhaps more importantly, all existing self-report measures are susceptible to impression management or faking. These self-report measures include items such as "I sometimes enjoy hurting those who care about me", which can be easily recognized as measuring an undesirable behaviour, and thus easily faked (Mathieu & Babiak, 2016b). This problem inherent to face-valid self-report measures becomes magnified when psychopathic personality is the focus, as the trait itself is partially defined by the ability to lie, deceive, and manipulate (Kelsey, 2016; Kelsey et al., 2015). Psychopathy is positively correlated with both intentions to fake and past faking behaviour in hiring contexts too (Fisher et al., 2018; Roulin & Krings, 2016). There is evidence that faking is an issue for the Self-Report Psychopathy scale (SRP; Paulhus et al., 2012), the Levenson Self-Report Psychopathy scale (LSRP; Levenson et al., 1995), and the Triarchic Psychopathy Measure (TriPM; Patrick, 2010), three popular self-report measures of psychopathy (Kelley et al., 2018; Kelsey, 2016; Kelsey et al., 2015; Rogers et al., 2002). The Psychopathic Personality Inventory Revised (PPI-R; Lilienfeld & Andrews, 1996) is the only one to include "validity scales" that measure virtuous and deviant responding. Yet, it is still vulnerable to some degree of faking (Edens et al., 2001; Kelsey et al., 2015; Kelsey, 2016; Marcus et al., 2018). Moreover, although the virtuous responding scale of the PPI-R can in fact detect those who are responding in the most positive light (Anderson et al., 2013), explicit faking is different from

virtuous responding (for an in-depth explanation, see Burns & Christiansen, 2011). To our knowledge, there is no research on the susceptibility to faking of the B-Scan Self (Mathieu & Babiak, 2016b), though since it is a workplace version of the SRP, it may prove to be equally susceptible.

Self-report measures are still the most commonly used method to assess psychopathy in the workplace (Mathieu & Babiak, 2016b; Verschuere et al., 2014), although researchers have argued that faking seriously limits their validities (Verschuere et al., 2014), that measures without faking detection scales are of questionable utility (Robinson & Rogers, 2015), or perhaps that self-report measures should not be used to measure psychopathy at all (Kelsey, 2016; Kelsey et al., 2015). In response, several researchers have suggested that implicit measures, like CRTs, could be the most promising method for measuring dark personality traits, such as psychopathy, for work-related purposes (O'Boyle et al., 2012; Spain et al., 2014; Wu & LeBreton, 2011). Building on these recommendations, we describe below the development of an implicit (CRT-based) measure of workplace psychopathy that is self-administered (i.e., low-cost and not requiring expert raters), was developed using non-criminal samples, and prevents faking even when presented with an incentive.

Implicit Measurement of Personality through Conditional Reasoning Tests

A conditional reasoning problem gives a set of logical premises in the context of a scenario, and a series of response options for test-takers to select the most logical one. A set of these problems make up a CRT, which measures conditional reasoning and logical decision-making. James (1998) hypothesized that this CRT format could be used to measure implicit personality. He argued that differences in one's personality are reflected in attitudinal tendencies and biases to favour, adopt, and choose certain options in everyday life. For example, consider two individuals having issues understanding some of the material on an upcoming test. The first person, high in trait Achievement Motivation, is likely to have an implicit tendency to believe that hard work and continued effort pays

off, so they continue to study. The second person, low in Achievement Motivation, may implicitly start to distance themselves from the importance of the test and doubt their ability to improve, resulting in them choosing to give up studying. Thus, if we design conditional reasoning problems that present scenarios such as this, with response options that are associated with different attitudinal tendencies and biases, we should be able to implicitly measure the personality trait of interest without the respondent knowing that this is what is being measured.

The first step to creating a CRT which implicitly measures personality is to develop a set of "justification mechanisms" (JMs; James et al., 2004; Schoen et al., 2021). JMs are descriptions of these reasoning processes, tendencies, and biases that are inherent to the personality construct of interest. JMs are established by consulting the literature to consider theories, conceptualizations, and previous measures of the construct. Although some CRTs have designed JMs to capture newly created purely implicit constructs, as we describe in detail below, we used a hybrid approach to build JMs derived from established models of psychopathy. Once JMs have been identified, conditional reasoning problems that measure those JMs are created. Each CRT problem, or item, measures one or more JM. For every conditional reasoning problem there are four possible response options, and the goal is to measure levels of the personality trait with an individual's responses. Essentially, this is achieved by implicitly directing respondents to make a dichotomous choice between two meaningful options. Two of the four response options are designed to be easily recognized as "incorrect" and illogical solutions to the problem, and are thus quickly eliminated by test-takers. The remaining two options are designed to be equally logical and "correct" solutions. However, one relies on relevant biases, tendencies and beliefs which aligns with high levels of the personality construct of interest, while the other aligns with tendencies and beliefs associated with low levels.

When James et al. (2004) originally put forth the CRT-A, they established consistent scale reliabilities and correlations with relevant performance measures. For instance, the CRT-A predicts various CWBs (Berry et al., 2010; Galić, 2016; Galić et al., 2014). Unlike the self-report measures listed earlier, there is ample evidence that CRT-based measures prevent faking through their implicit nature. Even when given an incentive, such as being told that results would influence chances of being hired, CRTs are still resistant to faking (LeBreton et al., 2007; Wiita et al., 2020). However, this requires that the implicit nature remains intact, as research demonstrates that participants can fake if told how the test functions beforehand (Bowler et al., 2013; LeBreton et al., 2007; Wiita et al., 2020).

Development of the CRT-WP

We used a hybrid approach in the development of the CRT-WP, whereby we followed the guidelines from the CRT literature (e.g., James & LeBreton, 2012), but instead of generating JMs to capture a new implicit construct (i.e., implicit psychopathy), we relied on established theories and models of (explicit) psychopathy to guide our framing of the biases, tendencies, and ways of thinking for the implicit side of psychopathic personality. As such, we began with consulting a wealth of literature on psychopathy theories and measurement. Over 60 peer-reviewed articles and manuals related to the psychopathy measures referenced earlier (i.e., PCL-R, SRP, PPI, LSRP, TriPM) were reviewed to ensure that the CRT-WP was not inherently biased toward any particular conceptualization of psychopathy. However, we excluded the clinical elements of psychopathy (i.e., particular to the PCL-R and SRP) to follow the *sub-clinical* route (i.e., similar to the theories behind the PPI-R, LSRP, and TriPM). This, alongside contextualizing the CRT-WP in strictly workplace-relevant situations, makes our implicit measure better-positioned for use in employment settings (Smith & Lilienfeld, 2013). It also clearly disconnects the CRT-WP content from personality traits related to mental disorders (e.g., antisocial personality) and thus offers better legal protection (e.g., against potential

claims based on the Americans with Disabilities Act in the United States or similar legislations elsewhere).

Justification Mechanisms

The JM development process for the CRT-WP was focused on workplace psychopathy, as the JMs for criminal psychopathy may be different for some traits (i.e., fearlessness) in different contexts (Lilienfeld et al., 2015). We generated six JMs to comprehensively cover the factors from the main existing measures of psychopathy: *Externalization, Carefree Impulsivity, Social Superiority, Fearlessness, Ruthless Self-Interest,* and *Insensitivity.* Table 2 presents the descriptions (i.e., the necessary biases in thinking and mental justifications) for the six JMs and highlights how each of them covers core traits/factors from established models and (explicit) measures psychopathy (i.e., PPI/PPI-R, TriPM, and LSRP). As noted above, it was a methodological and theoretical decision for the JMs to *not* cover factors related to non-job-relevant behaviours (in the PCL-R and SRP) such as "revocation of conditional release," "juvenile delinquency," and "short-term marital relationships", since they cover content that may be illegal, or at least unethical, to ask employees or applicants. As a result, the CRT-WP aligns with conceptualizations of psychopathy where criminal or anti-social behaviours are considered an outcome, but not part of the construct (like in the PPI-R and TriPM). Some JMs, such as social superiority, may be related to other dark triad traits (narcissism and Machiavellianism). This is consistent with existing measures of psychopathy that regularly overlap with these two constructs (see Table 2), and the three factors of the dark triad are also empirically inter-correlated (O'Boyle et al., 2012).

Item Generation

Item generation followed a similar process to that of the CRTs for Aggression (James et al., 2004), and Creative Personality (Schoen et al., 2018). In both cases, the authors suggested that CRTs

should ideally include 18 to 25 items. We thus initially generated twice as many items as desired (Hinkin, 1998). We wrote ten items for each of the six JMs, resulting in a total of 60 items. Although CRT items often cover more than one JM, and rarely produce clean factors perfectly aligned with the JMs (James & LeBreton, 2012), we wrote items specifically centred on one JM and thus one factor from established theories/models of psychopathy. All our items were written as work-specific scenarios because we intend to capture work-related psychopathic tendencies, but also to enhance face validity (i.e., to make it more believable that the CRT-WP is measuring problem-solving ability in the workplace and increase respondents' desires to select *the right answer*). Below is an example item from the CRT-WP, two others are available in Appendix A of the online supplement.

> Firing employees is an essential part of operating an organization. Generally, no employees want to be fired. As a result, sometimes when employees are aware that their organization is considering firing people, they become very worried and begin to act differently. In these desperate times, some people will find ways to make co-workers look bad in order to make them a more likely target for the upcoming firings. Many of these co-workers who get targeted do in fact get fired as a result.
>
> What is most true about the co-workers who were fired?
>
> A) They are the most obnoxious employees in company meetings
> **B) They are casualties of a survival-of-the-fittest environment**
> C) They benefit from having strong memory skills in this situation
> *D) They are being targeted by unfair and dirty tactics*

Recall that respondents are told that this is a conditional reasoning test, so they are truly trying to use their best analytical skills to select *the* correct response option. It is evident that options *a* and *c* do not serve as a logical response, guiding respondents to the other two choices with *equally logical* responses: a psychopathic option (*b*) or a non-psychopathic option (*d*). Selecting the psychopathic option is scored +1, the non-psychopathic options -1, and illogical options zero. A total score measuring psychopathic tendencies and biases can then be computed (James et al., 2004). We note that the complete list of items is not included here, nor in supplementary material. This approach is

consistent with previously developed CRTs for implicit personality. Indeed, the validity and faking-resistance of the implicit measure relies on its true nature and content being unknown to the test-takers. Thus, making all items public could compromise the CRT-WP's integrity. However, the authors will provide all items and scoring instructions to researchers interested in using the CRT-WP upon request.

Overview of Studies

We present six separate studies below. Study 1 involved a panel of subject-matter-experts (SMEs) who sorted and reviewed the preliminary 60 items. Study 2 includes an initial data collection where the remaining items were given to an online sample, and the factor structure and reliability of the measure were assessed. In Study 3, the refined version of the CRT-WP was completed by another online sample to examine the psychometric properties of the CRT-WP scores (e.g., construct-related validity, test-retest reliability) via two waves of data collection. Study 4 examined the CRT content using item response theory analyses with a large sample of actual job seekers. In Study 5, the faking-resistance of the CRT-WP was assessed by assigning participants to either an honest or simulated hiring condition, and then examining how their scores were associated with behavioural measures (e.g., selfish decisions) and self-reports (e.g., CWBs) collected at later time points. Finally, in Study 6 we show that CRT-WP scores account for incremental variance in CWBs above and beyond self-reported personality measures in an organizational context.

Study 1 – Item and Scale Revision with Subject Matter Experts

Methods

Fourteen SMEs participated in this initial review study: 13 graduate students in industrial/organizational (n = 8) and forensic (n = 5) psychology programs who were taking courses on psychometric assessment and development, and one professional clinical psychologist. Since the CRT-

WP is meant to measure sub-clinical levels of psychopathic personality, these SMEs were deemed appropriate. SMEs were given a paper-and-pencil document that contained the names and descriptions of each JM (similar to Table 2), and the list of 60 initially developed CRT-WP items in a randomized order. First, SMEs were asked to identify which JM each item was measuring (e.g., Ruthless Self-Interest). Second, they were also asked to identify which of the four response options represented a high/low degree of that JM (e.g., high and low degrees of Ruthless Self-Interest). SMEs were also given an example item which demonstrated how to complete the task and could comment on confusing wording or difficulties in understanding items and response options.

The responses of the SMEs were analyzed using percent agreement, and ability to identify the correct response options for each item. Any item that did not reach 75% agreement on the identification of the JM, as well as both high and low response options, would be either removed from the item list or re-worded based on SMEs' comments, dependent on how many items met the cut-off score for each JM (Hinkin, 1998). These guidelines were set beforehand so that there could be an adequate number of items for each JM included in the final version.

Results and Discussion

Items were classified into three categories based on their percent agreement: (1) items which scored above the 75% cut-off percentage on all three aspects the SMEs had to assess (which JM the item measured, identifying both the low and high JM responses – a total of 26 items); (2) items which scored above the cut-off percentage on the identification of the low and high response options, but below the cut-off score on the JM identification (20 items); (3) items that did not reach the percentage agreement cut-off for at least one of the low/high response identifications (14 items). For a complete breakdown of percent agreement, see Table S1 of the online supplement. Items in the first category were kept "as is" and those in the third category were eliminated. Items in the second category were

reassessed for confusing wording or whether they were best suited to measure a JM other than the one intended. Of the 20 items, 17 were reworked or re-assigned to a different JM and three were eliminated. For items that were reworked, this generally involved slight edits to provide more detail in the item stem, better align the response options with the intended JM, or ensure the psychopathic and non-psychopathic options were equally logical. In total, this phase of the development process resulted in 43 items (6 Externalization, 6 Carefree Impulsivity, 6 Social Superiority, 8 Fearlessness, 9 Ruthless Self-Interest, 8 Insensitivity).[2]

Study 2 – Exploratory Factor Analysis and Initial Reliabilities

Carrying forward the refined list of items from Study 1, in Study 2 we attempted to find a clear and interpretable factor structure within the CRT-WP, while also eliminating poorly functioning items using exploratory factor analyses (EFAs). This study was thus largely exploratory.

Method

Sample

Participants were recruited using Mechanical Turk (MTurk). MTurk samples are more diverse and produce results at least as valid/reliable as undergraduate student samples (Hauser et al., 2019; Landers & Behrend, 2015). The study was restricted to MTurk users who had already completed at least 100 prior surveys with at least a 90% approval rating (by using Cloud Research and their additional paid recruiting resources, as suggested by Hauser et al., 2019), were at least 18 years old, and were from the United States or Canada. We obtained 497 complete responses, but 76 respondents

[2] Seven of the 10 fearlessness items required re-wording or being moved to another JM (carefree impulsivity). Our initial definitions for carefree impulsivity and fearlessness were thus slightly altered based on SMEs' comments to distinguish the two concepts (i.e., spontaneity, lack of long-term planning, and tendency for impulsive decisions vs. preference for high-risk behaviours and the lack of anxiety or fear experienced in those situations).

were removed based on their responses to attention check items (e.g., “I eat concrete daily”). Another common gauge of attention used in previous CRT research is the number of illogical response options selected by respondents. James et al. (2004) and Schoen et al. (2018) suggest to remove any participant who selects at least 25% illogical response options. We followed this guideline and eliminated 25 participants who selected 11 or more of these options. We note that many of the 76 participants who failed the attention checks also endorsed more than 25% of illogical CRT response options, highlighting the overlap between the two approaches. Our final sample was 396 (almost 10 participants per item ratio), which is satisfactory for EFA (Worthington & Whittaker, 2006).

Mean age was 37.36 (*SD* = 10.89). Participants were 60.4% male and 39.4% female, 78.5% White, 9.1% Black, 5.1% Hispanic, 4.8% East Asian, 1.0% South Asian, and 1.6% Native/Aboriginal, Pacific Islander, or Other. Regarding education, 66.9% reported having at least some college/university education, 32.1% were high school graduates, while 1% had less than a high school education. When asked about employment, 87.6% of participants reported working at least part-time (62.6% full-time), 7.1% were unemployed, 2.5% were students, and 1.5% were retired.

Procedure and Scoring

The consent form told participants that they were being given the Conditional Reasoning Test for Workplace *Problems*, which measured problem-solving and logical reasoning in relation to different work-related problems. The use of passive deception here was necessary, given the implicit nature of the CRT-WP. However (here and in all subsequent studies) participants were debriefed about the true nature of the test, and the reason for the deception at the end of the study. Participants completed the 43 CRT-WP items from Study 1, presented in a randomized order, and a demographic questionnaire. For each CRT-WP item, selecting a “psychopathic” or “high JM” response option was scored as a +1, either of the two illogical response options as 0, and the “non-psychopathic” or “low

JM" response option as -1. CRT-WP scores represent the sum across all items. Previous CRT research has used this scoring system, although others have also been used (James & LeBreton, 2012). A short measure (IPIP-20; Goldberg, 1999) was included simply to embed the attention check items. Mean completion time was 28.78 minutes (*SD* = 9.73), compensation was USD $3, and participants were debriefed on the true nature of the CRT-WP.

Results

Response frequencies for all 43 items can be found in Table S2 of the online supplement. In theory, since high levels of psychopathy are less common in the general population (Coid et al., 2009), the psychopathic response options should be selected less frequently than the non-psychopathic ones. However, any one item does not assess whether someone scores high on psychopathy or not, and some JMs (e.g., fearlessness) can be associated with positive outcomes. For instance, one item had 77.5% of respondents choose the high JM (psychopathic) response. Additionally, previous research recommends that the two illogical options combined should not exceed 5% selection for any item, as this is indicative of poor item content (James, 1998; James & McIntyre, 2000). Two items, for which 15.4% and 5% of respondents chose the illogical response options, were thus removed.

We conducted EFAs with categorical factor indicators in Mplus version 7.4 (Muthén & Muthén, 2017). The Geomin rotation method was used with an oblique rotation due to expected intercorrelations between factors. Since multiple EFAs were being conducted to determine the best factor structure, various indicators were compared including eigenvalues, χ^2, CFI, RMSEA, SRMR, interpretability, and reliability coefficients. Given the categorical scoring, the Kuder-Richardson-20 coefficient (KR-20; Kuder & Richardson, 1937) was used to measure reliability as recommended by previous CRT research (see James & LeBreton, 2012; LeBreton et al., 2020).

For each round of EFA, we computed indicators of fit for 1-6 factor models. Items were considered problematic, and removed, if they did not load strongly on any factor (i.e., factor loadings lower than .20; following the guidelines for CRT validation provided by James and LeBreton, 2012) or if they forced an additional factor beyond an already theoretically-meaningful structure (e.g., a three-factor structure with two clearly interpretable factors, but a third factor with seemingly random items). Following five iterations of this EFA process, and removal of items which had problematic response rates (e.g., high endorsement of illogical options), 22 items remained. EFAs of this final 22-item version showed fit indices had levelled off and there were no non-significant factor loadings. Although two items showed some cross-loading on both factors, a subsequent EFA without these items demonstrated poorer fit and there was no longer a discernable best factor structure. As a result, the 22-item version was considered the final version.

The resulting fit indices for 1- to 6-factor models from this final 22-item EFA are presented in Table S3 of the online supplement. Both interpretability and fit indices reaching acceptable thresholds support this model as the best structure (e.g., χ^2/df = 1.00, *CFI* = 0.99, *SRMR* = 0.07). According to χ^2 difference tests, the 3-factor model is superior to the 2-factor model, however it is lacking theoretical meaningfulness. The rotated loadings for the 22 items in this final 2-factor structure are presented in Table S4 of the online supplement. Total scores were computed for both factors and the overall 22-item measure.[3] Correlations between CRT-WP overall and factor scores as well as dichotomously-coded demographic variables are presented in Table 3, along with means, standard deviations, and reliability coefficients.

[3] At this stage, twelve additional participants were removed based on choosing six or more illogical responses, as the 25% cut-off had changed (from 11 to 6) due to the decreased number of items (from 43 to 22).

Discussion

The negative values for the means of the final 22-item CRT-WP (*M* = -8.34, *SD* = 6.49) and its two factors (*M*s -5.68 and -2.68, *SD*s = 3.21 and 4.99, respectively) indicated that most people were choosing more non-psychopathic options than psychopathic ones. This was expected given that only few people are highly psychopathic. Although most of the factor loadings are lower than the normal .50 - .70 threshold range (Conway & Huffcutt, 2003), this is expected because of the categorical items and EFA approach and typical in CRT research (James & LeBreton, 2012; LeBreton et al., 2020). In addition, KR-20 reliabilities for both factors and the overall CRT-WP were high (ranging from .77 to .80).

Factor 1 covers more individual-oriented psychopathy (one's own thoughts and behaviours) and includes all carefree impulsivity and fearlessness items, along with items from externalization and social superiority that focus on one hypothetical character. Factor 2 covers other-oriented psychopathy (thoughts and behaviours directed at others) and includes ruthless-self interest and insensitivity items, along with externalization and social superiority items that do not have one specific person to focus on, but instead provide general scenarios (e.g., talked about how "some employees" or "some people in leadership positions" behave toward others). This two-factor structure makes theoretical sense when compared with other psychopathy measures described in Table 2. For example, the other-oriented psychopathy factor consists of the JMs covered by the primary factor of the LSRP, whereas the individual-oriented psychopathy factor covers almost exclusively the elements of the secondary LSRP factor. Similarly, the other-oriented factor of the CRT-WP aligns with factor 1 of the PCL-R/SRP, which describes psychopaths as remorseless or manipulative. The individual-oriented factor covers factor 2 of the PCL-R/SRP, which is focused on chronic instability, social deviance, and impulsive decision-making. In comparison to the Triarchic model of psychopathy, the individual-oriented factor

of the CRT-WP aligns with disinhibition and boldness, while the other-oriented factor aligns with meanness.

Although we treat this factor structure as providing theoretical support for the content of the CRT-WP, consistent with previous CRT research (James & LeBreton, 2012), we recommend using the overall composite score. This is because the overall measure is designed to cover one construct (i.e., workplace psychopathy), but individual items may cover more than one JM, as mentioned earlier. As a result, we will only discuss overall CRT-WP scores from this point onward, although future research may want to return to confirm the suggested two-factor structure or explore alternative ones.

Study 3 – Two-wave Study with Confirmatory Factor Analysis and Validity Assessment

The goal of Study 3 was to further demonstrate the reliability and validity of the new and refined version of the CRT-WP. In Wave-1, participants completed the revised 22-item CRT-WP in a randomized order. Like in Study 2, the same three attention-check items were embedded within the mini-IPIP for wave one, alongside the same demographic questions. One week later (Wave-2), participants were invited back to complete the CRT-WP again, as well as two overt self-report measures of psychopathy (i.e., the TriPM and SRP-III), a self-report of CWBs, and a measure of socially-desirable responding. Confirmatory factor analyses (CFA) were conducted on Wave-1 of CRT-WP data using Mplus version 7.4.

Correlating scores from Wave-1 and Wave-2 provides an estimate of test-retest reliability for the CRT-WP. Previous CRTs in the literature have demonstrated good test-retest reliability in multiple studies, with values ranging from .74 to .87 (James et al., 2004, James et al., 2005; Schoen et al., 2018). Therefore, scores on the CRT-WP from the two waves should be positively related, with a value in the range of those reported in past CRT validation efforts.

Hypothesis 1: Participants' CRT-WP scores from Wave-1 and Wave-2 will be positively correlated, demonstrating test-retest reliability.

Additionally, scores on an implicit/covert CRT and overt self-report measures of the same construct should be significantly but *not* strongly correlated (James et al., 2004, James et al., 2005; Schoen et al., 2018), with correlations typically around .30 or lower (James & LeBreton, 2012; LeBreton et al., 2020). For instance, James and LeBreton (2012) correlations ranging from .14 to .23 between the CRT-A and self-reports of aggression, whereas Schoen et al. (2018) reported correlations.18 to .33. for CRT-CP and self-reports of creativity. In theory, this occurs due to overt and covert measures assessing different dimensions or components of the same construct (Bing et al., 2007). To reduce common method bias (Podsakoff et al., 2003), CRT-WP scores from Wave-1were correlated with Wave-2 overt measures.

Hypothesis 2: CRT-WP scores and (a) SRP-III and (b) TriPM scores will be positively correlated.

Existing implicit CRTs have predicted relevant outcomes for the construct assessed, such as aggressive behaviours or CWBs (James et al., 2005; James & LeBreton, 2012). Although the current study does not have an objective measure of psychopathic behaviours at work, self-reported perpetration of CWBs was used as a proxy. The positive relationship between psychopathy and increased CWBs was discussed above, and is well-documented in the literature (Boddy, 2011; Scherer et al., 2013). Therefore, it is expected that participants' scores on the CRT-WP will be significantly associated with self-reported CWBs.

Hypothesis 3: CRT-WP scores will be positively correlated with self-reported CWBs.

Finally, the hallmark of the CRT-WP is that it is implicit and covert, which should prevent respondents from engaging in impression management or faking to artificially reduce their scores. As a result, CRT-WP scores should be *unrelated* to socially desirable responding while the overt measures should be significantly *related* to it.

Hypothesis 4: CRT-WP scores will not be correlated with socially desirable responding, while the SRP-III and TriPM will be.

Methods

Sample

For Wave-1, data were collected from 301 MTurk users who did not complete Study 2. The same recruitment/screening criteria were applied. Sample size was decided based on recommendations for CFA analyses (Worthington & Whittaker, 2006), accounting for the expected removal of some participants. Based on responses to attention check items, 37 respondents were removed. Thirteen more participants were removed based on having selected six or more illogical responses, resulting in a final sample of 251 participants for Wave-1.

The sample was 58.2% male, 40.2% female, and 1.6% identifying as "Other." Mean age was 36.71 (*SD* = 10.46). Most participants were White (77.7%), with 8.4% Black, 4.8% Latino, 4% East Asian, 1.2% South Asian, and 4% combined Middle Eastern, Native/Aboriginal, Pacific Islander, or "Other." Regarding education, 68.9% reported having at least some college/university education, while 31.1% reported being a high school graduate. The sample was 88.4% employed (68.5% full-time), 5.2% unemployed, 4% students, 1.6% retired, and 0.8% "Other." Mean completion time for was 25.83 minutes (*SD* = 10.48). Given that Wave-1 only consisted of the CRT-WP and demographics, along

with consent and debriefing forms, it is estimated that mean completion time for the CRT-WP specifically is approximately 18-19 minutes.

One week later, 208 (81.89%) participants returned for Wave-2. Using the same criteria as before, four participants were removed based on attention check items and six participants were removed due to selecting six or more illogical responses on the CRT-WP. Thus, 198 participants made up the final sample for Wave-2. There was no significant difference in Wave-1 CRT-WP scores between participants who returned vs. did not, t (248) = 0.17, d = 0.03 (see Table S5 for details).

Measures

CRT-WP. The 22-item version of the CRT-WP following Study 2 was used. The measure continued to be scored in the same +1, 0, and -1 format, and we computed an overall 22-item score. For both waves, item order was randomized. Reliability coefficients are reported in Table 4.

Self-reported Psychopathy. First, participants were asked to complete the Self-Report Psychopathy scale (SRP-III; Paulhus et al., 2012). We used a 34-item version (Mahmut et al., 2011), with a 5-point Likert-scale response format ranging from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). This version has previously demonstrated good reliability and validity in a community, non-clinical sample. It covers the four sub-scales: callous affect, erratic lifestyle, interpersonal manipulation, and criminal tendencies. An example item is "I sometimes enjoy hurting the people who care for me" (callous affect). The criminal tendencies sub-scale was *not* included in the composite score for the SRP because it assessed an aspect of psychopathy that was purposefully not covered by the CRT-WP.

The second measure of self-report psychopathy was the brief version of the Triarchic Psychopathy Measure (TriPM; Patrick, 2010), which consists of 58 items measuring the three factors of boldness, meanness, and disinhibition. Participants respond using a scale of 1 (*Strongly Disagree*) to 5

(*Strongly Agree*). An example item is "I sometimes insult people on purpose to get a reaction from them" (meanness). Both reliability and validity for the brief TriPM have been established (Patrick, 2010). Both an overall composite score and individual factor scores were calculated.

Counterproductive Work Behaviours. CWBs were measured by the Counterproductive Work Behaviour Checklist (CWB-C; Spector et al., 2006). It is a 33-item measure which consists of five subscales: abuse, production deviance, sabotage, theft, and withdrawal. Participants indicate how often they engage in each CWB on a scale from 1 (*Never*) to 5 (*Daily or Almost Daily*). An example item is "Blamed someone at work for an error you made" (abuse). Respondents were instructed to answer with respect to their current job or (e.g., for unemployed MTurk users) their last job if they did not currently have one. Total composite scores and individual factor scores were calculated.

Socially Desirable Responding. We used the 8-item impression management factor of the Balanced Inventory for Desirable Responding (BIDR-16; Hart et al., 2015), which measures deliberate false presentation. This short version of the BIDR has scale properties similar to the original version and is scored in the same manner (Hart et al., 2015). Respondents rate each item on a 7-point scale from 1 (*Not* True) to 7 (*Very True*), and we used a continuous scoring as recommended by Hart et al. An example is "When I hear people talking privately, I avoid listening".

Results

Wave-1

The frequency of response selections for each item was similar to that of Study 2, with no item having a meaningfully different distribution. We compared a series of theoretically-meaningful models using CFA in Mplus: A 1-factor model, the 2-factor model found in Study 2, a 3-factor model which attempted to mimic the Triarchic conceptualization of psychopathy; a 6-factor model aligned with our

six JMs; and a final model with a 2nd order structure (one first-order factor for each JM and one overarching "psychopathy" factor)[4]. The fit indices for these five models are presented in Table S6 of the online supplement. All models showed good fit, including the 1-factor and 2-factor models.

Wave-2

The means, standard deviations, reliability coefficients, and inter-correlations for all Wave-1/2 variables are presented in Table 4. The KR-20 reliability coefficients for the CRT-WP was larger in Wave-2 than in Wave-1 (or Study 2), indicating good internal consistency. Supporting Hypothesis 1, the correlation between CRT-WP scores from Wave 1 and Wave 2 (r = .72) demonstrated acceptable temporal stability (or test-retest reliability).

As expected, CRT-WP scores from Wave-1 had significant weak positive correlations with both the SRP-III (r = .31, p < .001) and the TriPM (r = .18, p = .01), supporting both Hypotheses 2a and 2b. We examined whether CRT-WP scores from Wave-1 predicted CWB scores from wave two. The correlation was significant (r = .18, p = .01), supporting Hypothesis 3.

Finally, overall CRT-WP scores from Wave-1 or Wave-2 were not significantly correlated with impression management (rs = -.10 and -.11), whereas overt measures of psychopathy were (i.e., r = -.40 for the SRP-III, r = -.34 for the TriPM), supporting Hypothesis 4.[5]

Additional unplanned analyses were performed to explore the so-called "channeling hypothesis" (Bing et al., 2007). According to this hypothesis, and integrative models of personality, explicit/overt and implicit/covert measures assess different elements of the same construct. Similar to

[4] At the request of an anonymous reviewer, we also tested a bifactor model with the two factors from the 2-factor model and one general psychopathy factor. The fit for this model was worse (e.g., RMSEA = .03, CFI = .82) than other models.
[5] All the relationships between Wave-1 CRT-WP scores and Wave-2 variables (i.e., CRT-WP, SRP, TriPM, CWB) were also replicated using a full-information maximum likelihood (FIML) model and results were identical to those using data from returning participants only (see Table S7 in our online supplement).

how an iceberg has a visible portion above the surface, and a hidden portion below the surface, explicit personality is overt and self-reportable, whereas implicit personality is unobservable by conscious thought and only appears in tendencies, biases, and so on. Thus, to get a complete picture of the entire construct (i.e., the whole iceberg) both observable and implicit portions should be considered. As such, we explored whether adding the interaction between the overt (SRP-III or TriPM) and implicit (CRT-WP) measures of psychopathy would explain additional variance in CWBs beyond the main effects of these predictors alone. In both cases, the interaction term was significant. That is, interactions between the SRP-III and CRT-WP ($\Delta F = 10.17$, $p < .01$, $\Delta R^2 = .03$, $ß = .68$), and the TriPM and the CRT-WP ($\Delta F = 3.90$, $p = .05$, $\Delta R^2 = .01$, $ß = .73$), were both helpful in predicting CWBs beyond the two pairs of measures alone (see Table S8 and Figure S1 of the online supplement).

Although Study 3 was not designed to test incremental validity, we also explored whether the CRT-WP scores provided incremental prediction of CWBs above and beyond the SRP or the TriPM scores, using regression analyses. In both cases, the CRT-WP was not a significant predictor when the SRP ($\beta = .02$, $p = .76$) or the TriPM ($\beta = .49$, $p < .001$) were entered first in the regression.

Discussion

The results of this third study provide further evidence for the reliability and factor-structure, and preliminary evidence for the validity of the CRT-WP. Estimates of internal consistency were high, and Wave-1 and Wave-2 scores indicated acceptable test-retest reliability. In line with previous literature showing weak but significant relationships between CRTs and overt measures of the same constructs (James et al., 2004; Schoen et al., 2018), the CRT-WP was weakly, yet significantly, correlated with both the SRP-III and the TriPM. These correlations fell within the hypothesized range for implicit measures, providing support for that the CRT-WP as a measure of psychopathy. Indeed, although overt self-report measures of the same construct should correlate highly, covert and overt

measures assess the construct differently and thus produce weaker intercorrelations (e.g., Bing et al., 2007). In addition, our implicit measure was workplace-specific, whereas other measures were not. CRT-WP scores being unrelated to impression management also provides initial evidence that the measure is working covertly as intended. In contrast, scores on the overt self-report measures of psychopathy were moderately to strongly correlated with impression management. Taken together, these findings highlight the potential value of the CRT-WP to both the psychopathy literature and its assessment in practice.

CRT-WP scores from Wave-1 were significantly related to CWB scores. Although the correlation is relatively weak in magnitude (r = .18), a meta-analysis reported an uncorrected correlation of r = .06 between (overt) psychopathy and CWBs (O'Boyle et al., 2012), suggesting that this correlation can be viewed favourably. In addition, CWBs were also self-reported in an overt manner, and our findings show that they were also susceptible to impression management (r = -.46, p < .001), suggesting that some participants may have under-reported their CWBs. Because timing of measurement, item characteristics, and social desirability are important sources of common method variance bias (e.g., Podsakoff et al., 2003), it partly explains why measures of psychopathy (i.e., SRP, TriPM) that were, measured concurrently, overt, explicit, self-reported, and also prone to impression management were more strongly related to self-reported CWBs than the CRT-WP.[6] This might also explain why CRT-WP scores did not provide incremental prediction of CWBs beyond SRP or TriPM scores. Yet, we acknowledge that the correlations between the CRT-WP and CWBs scores should ideally have been higher, relative to those for self-reports of psychopathy. Similarly, CRT-WP scores should ideally provide incremental validity over self-report measures of psychopathy or dark

[6] Following Lindell and Whitney's (2001) method-variance marker approach, we computed partial correlations between the SRP, Tri-PM, or the CRT-WP (at T1 and T2) and CWBs, controlling for BIDR-IM. In support of our argument, when compared to raw correlations, corrected ones were substantially reduced for the SRP (from .49 to .37) and the TriPM (from .51 to .42), while the reduction is slightly smaller for the CRT-WP (.18 to .15 for T1, .14 to .10 for T2).

personality. This is why we continued to explore CWB as a key outcome (see Studies 5 and 6). Future studies could also include an objective measure of CWBs to assess predictive validity (e.g., recorded reports of counterproductive or deviant behaviors).

Lastly, there was a consistent negative correlation between all three psychopathy measures and age. These correlations are likely evidence of the steady decline of "negative" personality traits, including psychopathy, with age (Olver & Wong, 2015). Our correlations can be viewed as further evidence that all three measures are capturing the same construct. Importantly, the CRT-WP was not correlated with any other demographic variable, while the SRP-III and TriPM were both significantly correlated with participant sex, with men scoring higher in psychopathy. One explanation is that men truly do score higher on psychopathy (Coid et al., 2009), but the CRT-WP fails to capture this tendency. Alternatively, research suggests that the existing conceptualizations of psychopathy (especially the PCL-R and SRP-III) could be biased toward measuring traits and tendencies more common among male psychopaths (Forouzan & Cooke, 2005). Since the PCL-R and SRP frameworks were primarily developed with male samples, past research has suggested that these measures do not cover the female-specific psychopathic traits and tendencies. Women high in psychopathy are likely to have the same underlying personality as it relates to psychopathy, but how it is *expressed* through behaviour is different (for a review, see Nicholls & Petrila, 2005). The lack of correlation with participant sex would suggest that the CRT-WP may be a more sex-neutral measure of psychopathic tendencies which can measure both expressions equally.

Study 4 – Item Response Theory (IRT) Analysis of CRT-WP Items

Both DeSimone and James (2015) and Smith et al. (2020) underline the importance of conducting Item Response Theory (IRT) analyses as the final step in developing, refining, and validating items of implicit CRTs. IRT refers to a set of processes and models which can be used to

understand psychological measures (Embretson & Reise, 2000). Classical test theory assumes that test items make up a larger scale which measures a particular construct equally well regardless of the level of that construct. IRT focuses on scores for each individual item, is not concerned with "overall" scores, and instead assesses how accurate each item is at measuring the latent construct, or Theta (θ), at all levels of that construct (Embretson & Reise, 2000). Therefore, it is important to link CRT items (and indirectly the overall scale) to the underlying latent construct (e.g., psychopathy) using IRT to supplement traditional analyses. Moreover, IRT informs us about whether items are better at measuring low or high levels of the trait, and how well each item can discriminate those levels.

As participants complete an implicit measure of personality, they believe the test is assessing their logical reasoning ability. The response options are designed so that respondents are unknowingly restricted to two equally correct options, and thus which one they choose (or their best "guess" between the two) is guided by their implicit biases or tendencies. For these reasons, 2-PL models have been used in previous IRT research using implicit CRTs (DeSimone & James, 2015; Galić et al., 2014; Theriault, 2019). The fit of different models can be compared using traditional fit indices like BIC and -2 loglikelihood ratio. We thus compare 1-PL, 2-PL, and 3-PL models, and predict that:

> *Hypothesis 4*: The CRT-WP will demonstrate better fit to the 2-PL IRT model when compared to the 1-PL or 3-PL models.

Recall that justification mechanisms (JMs) describe the implicit biases, tendencies, and inclinations of an underlying personality trait. CRT items are written based on those JMs. However, individuals high on the trait or construct may not necessarily possess high levels of all JMs, and multiple JMs often guide thinking or behaviour in any given situation. Therefore, people higher in psychopathic personality should choose the psychopathic (i.e., "correct") option more often and CRT-WP items should show a positive relationship between the latent trait (θ) and discrimination

parameters. DeSimone and James (2015) found that this was true for 21 of the 22 CRT-A items. We thus predict:

> *Hypothesis 5*: There will be a positive relationship between estimates of latent psychopathy and discrimination parameters for each CRT-WP item.

Finally, the Item Characteristic Curves (ICCs) resulting from IRT analyses illustrate how each CRT-WP item functions, how much information it provides (through the difficulty and discrimination parameters), and how well it contributes to the overall measurement. Smith et al. (2020) noted that ICCs help understand which items work well, while flagging poorly functioning items due to a lack of discrimination or a "difficulty" level that is too low or high. ICCs could then be used to further refine the CRT-WP by identifying items in need of additional analyses and consideration.

Research Question 1: Do the ICCs show adequate difficulty or discrimination levels for the CRT-WP items?

Methods

Sample

Data were collected between March 2022 and August 2023 from potential job applicants interested in learning more about the selection process of a branch of the Canadian federal government. Potential job applicants could complete a practice cognitive ability test on the organization's website and were invited to complete the CRT-WP as well for research purposes. A total of 26,489 surveys were started, with 8,583 choosing to complete the optional CRT-WP portion, resulting in 6,746 potential applicants with usable responses (based on attentiveness, completion, and the number of illogical responses on CRT-WP items). Participants were 58.0% male with a mean age of 28.07 (*SD* = 9.60). Participants identified as 50.5% White/Caucasian, 25.4% Visible Minorities, 4.8% Indigenous

persons, and 2.1% identified as a person with a disability. The highest level of education completed was high school for 46.5%, community/professional college for 19.6%, while 31.0% had completed at least some university.

Measures

CRT-WP. The same 22-item version of the CRT-WP following Study 3 was used. The measure was scored using the same format, and we computed an overall 22-item score.

Results & Discussion

All IRT analyses were conducted using IRTPRO version 6.0 for Windows. Hypothesis 4 was tested by running 1-PL, 2-PL, and 3-PL IRT models for the CRT-WP items, and comparing fit indices (e.g., χ^2 difference, BIC). Consistent with Hypothesis 4, the 2-PL model showed superior fit indices to the 1-PL($\Delta\chi 2$ (21) = 642.1, $p < .001$) and the 3-PL model ($\Delta\chi 2$ (21) = 90.4, $p < .001$) [7]. This result also aligns with previous research examining the CRT-A with IRT (DeSimone & James, 2015; Galić et al., 2014; Theriault, 2019)[8]. This supports that the guessing parameter is not incrementally relevant to measuring the underlying trait (e.g., psychopathy) because test-takers are being guided by implicit biases and tendencies when selecting the (psychopathic or non-psychopathic) option. Further analysis of the 2-PL model suggested that four (of the 22) items had poor fit according to S-X^2 statistics (see Table S9 of Online Supplement). However, exploring the frequency tables of these items and comparing them to the rest of the scale did not reveal any discernable differences in response frequency or patterns.

[7] We also explored a 2-dimensional IRT model based on the 2-factor model of Study 2. This 2-dimensional model did fit the data better than the unidimensional 2-PL model ($\Delta\chi 2$ (21) = 84.6, $p < .001$), but the difference was not appreciable, similar to DeSimone and James (2015). Since CRTs are scored unidimensionally, the 2-PL model was retained.

[8] In response to a helpful suggestion from an anonymous reviewer, we also conducted all IRT analyses using the Nominal Response Model. While the fit was much worse compared to the unidimensional models (similar to DeSimone & James, 2015), these analyses did reveal that six items included an illogical response option with a probability of selection that significantly increased (to over 10-20% likelihood) at higher levels of Theta. Future work could re-phrase these six illogical response options to make them even less logical, ideally decreasing the expected response frequency.

Hypothesis 5 was tested by analyzing the relationship between θ (latent psychopathy) and discrimination for each CRT-WP item using the best-fitting model (i.e., 2PL) and examining the ICCs. Supporting Hypothesis 5, all discrimination parameters were positive, indicating that endorsing the psychopathic option for each item did in fact relate to higher levels of the same underlying trait. Discrimination parameters ranged from 0.30 to 1.33, with an average of 0.70. The items with lower discrimination values are less able to differentiate between levels of the underlying trait, and thus provide less information.

For RQ1, 17 of 22 item ICCs displayed the point of median probability between +1 and +3 levels of theta, with one item at +0.75 and the other three at +3.87 or higher. This means that for a test-taker to select the psychopathic option, they need to be on the higher end of trait psychopathy. This is consistent with the CRT-WP's goal to identify and screen the individuals with higher-levels of workplace psychopathic personality, while being less concerned with whether someone is "extremely low" or "very low" on the trait. Accordingly, the total information curve peaks at +1.8 levels of Theta, and individual information curves show that most items are contributing significant information to the overall measurement. These difficulty and discrimination parameters can be found in Table S9, and the ICC and information curves in Figure S2-4, of the Online Supplement.

Overall, our IRT analyses of the CRT-WP was conducted in line with the recommendation of Smith et al. (2020) and following previous implicit personality research (i.e., with the CRT-A). In general, the results dictate that (1) guessing is not an important parameter in why participants select either the psychopathic or non-psychopathic option, suggesting that responses rather capture biases and underlying personality; (2) all items are measuring the same underlying trait; and (3) both individual items and the overall CRT-WP are best at identifying and measuring high levels of that trait (i.e., high psychopathy). Not all items were perfect in terms of fit, discrimination, or the information provided,

but most items met the standards in all three of those areas. Only one item showed poor performance in both fit and information provided, so future validation efforts may explore removal of this item for parsimony. Compared to previous IRT research with the CRT-A (DeSimone & James, 2015), the performance of CRT-WP items in this study were satisfactory and contribute to further evidence that the implicit measure is generally functioning as intended.

Study 5 – Three-wave Study on Faking-Resistance and Criterion-Related Validity

The objective of this fifth study was to assess the faking-resistant nature of the CRT-WP in a simulated hiring scenario, while also providing additional evidence for criterion-related validity by using behavioural and time-separated outcomes.

A key benefit of CRTs is their resistance to impression management or faking because of their implicit and covert nature. It is thus important to demonstrate faking-resistance for the CRT-WP. We therefore compare CRT-WP scores between test-takers instructed to remain completely honest vs. to imagine completing the measures as part of a hiring process for an attractive job. As noted earlier, overt self-report measures of psychopathy (and personality in general) are susceptible to faking and impression management (Kelsey, 2016; Mathieu & Babiak, 2016a), and people high in psychopathy are more likely to fake in hiring contexts (Roulin & Krings, 2016). However, research contends that implicit measures like CRTs prevent faking even when participants are provided with an incentive (LeBreton et al., 2007; Wiita et al., 2020). Thus, it is hypothesized that participants will be able to fake overt personality measures, but not the implicit CRT-WP, in a simulated hiring scenario.

Hypothesis 6: Scores in a simulated hiring (vs. honest) context will be significantly different (in favourable directions) for all *overt* personality measures, but not for the CRT-WP.

Impulsivity and selfish behaviour are two facets central to conceptualizations of psychopathy (e.g., Hare, 1985; Lilienfeld & Andrews, 1996), and are reflected in the JMs at the core of CRT-WP (see Table 2). Additionally, resulting from tendencies of disinhibition and self-interest, cheating behaviours have been related to psychopathic personality in many studies (e.g., Ljubin-Golub et al., 2020). The CRT-WP should thus be associated with impulsivity, cheating, and selfish behaviour.

Hypothesis 7: CRT-WP scores will be positively correlated with behavioural indicators of (a) impulsivity, (b) cheating, and (c) selfish behaviour.

Finally, as mentioned above (Study 3), the relationship between high levels of psychopathy and increased perpetration of CWBs is well-documented in the literature (Boddy, 2011; Scherer et al., 2013). Further, recent research has identified that psychopathy is especially predictive of deviant online behaviours such as cyberloafing, even more so than other dark traits (Moor & Anderson, 2019). Therefore, it is expected that:

Hypotheses 8: CRT-WP scores will be positively correlated with (a) CWBs and (b) cyberloafing.

Method

Sample

We collected complete data from 249 Prolific users living in the United Kingdom and Ireland. We screened participants to have management experience, be 18 years old, fluent in English, and be willing to complete studies involving deception. One participant failed to meet these criteria, and only six participants failed any attention-check items or chose illogical responses on 25% or more of the CRT-WP items, leaving 243 participants. An extra 24 participants failed an additional screening check

at the end of the survey by responding "No" to a question confirming management experience. Thus, the final sample for Wave-1 was $N = 219$. Mean completion time was 27.03 minutes ($SD = 10.58$).

Overall, participants were 53% female, 90% Caucasian, 59.9% university-educated, and 85.9% employed (65.8% full-time), with a mean age of 38.49 years ($SD = 10.78$). When participants were invited back two weeks later, 192 participants (87.67%) returned to complete Wave-2. And, 155 participants (70.77%) returned to complete Wave-3 one year later.[9] Participants were compensated with £4.50 for completing Wave-1, £1.00 base compensation for Wave-2 with the potential to earn up to £4.85 more depending on their performance on the tasks and games, and £1.00 for completing Wave-3. Sample attrition was random and unrelated to demographic characteristics (except age) and, most importantly, CRT-WP scores, $t(217) = 0.51$, $d = 0.11$, $p = .61$ for Wave 2 and $t(217) = 0.08$, $d = 0.01$, $p = .94$ for Wave 3 (see Table S10 of the Supplement for details).

Procedure & Design

In Wave-1, participants were randomly assigned to either an honest or a simulated-hiring condition. Those in the honest condition were told to remain as honest as possible while responding to a series of assessments/measures. Those in the simulated-hiring condition were told to imagine they were completing the measures as part of the hiring process for a very attractive, well-paying job. Such experimental conditions and instructions are typical to personality faking research (Viswesvaran & Ones, 1999).[10] Participants completed the CRT-WP, overt measures of personality (with three attention-check items embedded), and demographic questions.

---

[9] We also conducted analyses with only the $N = 116$ individuals who reported being still employed at the time of the Wave 3 data collection (during the Covid-19 pandemic), but results were equivalent to those with $N = 155$.

[10] Analyses of a manipulation check item confirmed that the manipulation worked: In the honest condition, 99.1% of participants responded with "Strongly Agree" when asked if they completed the measures as honestly as possible. Similarly, 98.3% of participants in the simulated hiring condition responded they either "Agree" or "Strongly Agree" that they completed the measures as seriously as they would if they were applying for a real job.

Participants who successfully completed Wave-1 were eligible to take part in Wave-2 two weeks later. It included three behavioral tasks designed to measure impulsivity, cheating, and selfish decision-making (see *Measures* below) and a debriefing form informing them about the true nature of the study. One year later, Wave-3 included measures of CWBs and cyberloafing, and a debriefing.

Measures

CRT-WP. The same 22-item version of the CRT-WP was used here, with the same instructions and scoring as previous studies.

HEXACO. The HEXACO-60 (Ashton & Lee, 2009) includes the traditional "Big Five" personality traits (emotional stability, extraversion, agreeableness, conscientiousness, and openness) plus honesty-humility, which is conceptually and empirically related to psychopathy (Lee & Ashton, 2005). The 60 items measure the six factors listed above (10 items each), and participants respond on a scale of 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). An example item is "I would never accept a bribe, even if it were large" (honesty-humility).

Dark Triad. The Short Dark Triad (SD3; Jones & Paulhus, 2011) is a 27-item measure with three factors: Machiavellianism, narcissism, and psychopathy (9 items each). It was chosen because two widely-used self-report measures of psychopathy were used in Study 3, but it was not yet known how the CRT-WP would relate to other "dark" personality traits. Participants indicate the extent to which they agree with each item on a scale of 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). An example is "Make sure your plans benefit yourself, not others" (Machiavellianism). Two items of the SD3 psychopathy factor were considered irrelevant to workplace psychopathy ("I enjoy having sex with people I hardly know" and "I have never gotten into trouble with the law") and removed to accurately assess convergent validity.

Integrity. The TSD-Integrity (Catano et al., 2018) is a 10-item overt self-report measure of integrous personality. The 10 items are taken from other big five traits but deemed the most critical to the conceptualization of someone who has high integrity. Participants indicate the extent to which each statement or adjective is characteristic of themselves on a 7-point Likert scale from 1 (*extremely uncharacteristic*) to 7 (*extremely characteristic*). An example item is "I always have a place for everything and everything in its place" (conscientiousness).

Impulsivity. The Balloon Analogue Risk Task (BART; Lejuez et al., 2002) is an interactive task that measures impulsivity and risk-taking behaviours in adolescents (Aklin et al., 2005; MacPherson et al., 2010). Participants are presented with an opportunity to earn money by pumping up a series of virtual balloons (i.e., each click on an "inflate balloon" button adds money to a running total, which they can "bank" when they want). Importantly, there is a randomized point of explosion for each balloon: the balloon will explode if the participant chooses to inflate beyond this point, and they lose any unbanked money. It is recommended that participants are presented with anywhere between 10 and 30 balloons, or trials (Wallsten et al., 2005). BART was embedded within a Qualtrics survey using JavaScript (a screenshot can be seen in Figure S5 of the online supplement). A group of graduate students pilot-tested the functionality of this version across different devices and web browsers. Before starting the BART, participants were presented with the same screenshot as Figure S5 and were given detailed instructions about the rules and guidelines for the task. They completed 10 trials (10 balloons), each balloon inflation was set at £0.01 GBP, with the maximum breaking point of each balloon set at 32 inflations, which was the medium breakpoint established in the original BART research (Lejuez et al., 2002). BART outcomes include the total amount of money earned, number of balloons exploded, and the average number of inflations per balloon.

Cheating. The number matrix task presented by Wiltermuth (2011) was used as a measure of cheating. Participants were presented with ten matrices (12 numbers in each) and asked to find the two numbers in each matrix which added up to 10.00 exactly (see examples in Figure S6 of the online supplement). They were told to simply click the "Solved" box below each matrix to indicate that they had solved it (they did not need to specifically identify the two numbers). However, only five matrices actually had a correct solution, but the other five were unsolvable. Thus, participants should be under the impression that the researcher would not be able to verify their answers, signalling an opportunity to cheat. Participants were given 5 minutes to complete as many matrices (presented in a randomized order) as possible, and were paid £0.10 GBP for each matrix they *reported* solving. Cheating behaviours was based on how many unsolvable matrices participants indicated solving (of the ones they attempted).

Selfish Decision-making. A revised version of the traditional dictator game was used to assess selfish decision-making. We adapted the task so that participants are not actually paired with another participant, but are told that they are in order to induce the dilemma of allocating a reward between themselves and another person. First, participants were told that they were randomly paired with another participant in the study using their Prolific IDs, and that their pair had earned extra compensation based on their performance on the previous tasks. When advancing from this page, a fake "loading" icon was displayed for 2.5 seconds to give participants the impression that pairing and calculation were taking place, before automatically advancing to the next page. Thus, every participant was then told that they were randomly assigned the role of "The Decider", which allowed them to decide how much of the compensation they gave themselves and their hypothetical partner. The structure of the response options was taken from Brosig-Koch et al. (2017), where there are 11 payment options (i.e., an 11-point scale) for the participant to select, ranging from a very self-interested choice

(£0.75 GBP for them and £0 for their partner) to least self-interested choice (£0.25/£1.00, respectively). Each consecutive option offers a decrease of £0.05 for the participant but an increase of £0.10 for their partner, so that the total sum of money increases (from £0.75 to £1.25) as the participants themselves receive less (£0.75 to £0.25), giving some incentive to choose the options on the upper end of the scale.

Counterproductive Work Behaviours. CWBs were measured by the CWB-C (Spector et al., 2006) similar to Study 3, with the same instructions.

Cyberloafing. An adapted version of the scale by Akbulut et al. (2016) was used to measure cyberloafing, which is defined as using the internet for non-work activities during work hours. This version used in the current study contained 4 factors (sharing, shopping, accessing online content, and online gaming/gambling) and 25 items overall. We decided to exclude the real-time updating factor because it focuses on Twitter specifically and would thus be irrelevant to participants without Twitter. Participants responded to each item based on how often they engaged in that behaviour during work hours on a 5-point scale from 1 (*Never*) to 5 (*Every day or Almost Everyday*). An example item is "I post status updates on social networks" (sharing).

Results

We used independent samples *t*-tests to test Hypothesis 6, which predicted that faking would be possible for all overt measures of personality but not the CRT-WP, by comparing whether the means were different between the honest and the simulated-hiring conditions. Results are presented in Table 5. As hypothesized, there were significant differences between the two conditions for all overt personality measures (*d*s ranging from |0.30| to |0.73|). Participants were able to "fake good" in the hiring condition and obtained significantly higher mean scores for openness, conscientiousness, agreeableness, extraversion, honesty-humility, integrity, and narcissism, but significantly lower scores for emotionality, Machiavellianism, and psychopathy. However, also as hypothesized, there were no

significant differences between the honest and simulated hiring conditions for the implicit CRT-WP (*d* = 0.07). These results suggest that participants were unable to fake on the CRT-WP, likely because they were unaware of what it was measuring.

Descriptive statistics, reliability coefficients, and intercorrelations for the CRT-WP and all outcome variables are presented in Table 6.[11] The internal consistency of the CRT-WP was consistent with previous studies (*KR-20* = .82). None of the BART outcomes (money won, number of balloons exploded, or the average number of pumps per balloon) were significantly correlated with the CRT-WP, failing to support Hypothesis 7a. Regarding the percentage of unsolvable matrices that participants reported solving, the correlation with CRT-WP scores was not significant, failing to support Hypothesis 7b. However, supporting Hypothesis 7c, the CRT-WP was significantly negatively correlated to the dictator game choice (r = -.24, p < .001), indicating that participants who scored higher on the CRT-WP made more self-interested choices. Moreover, in the *simulated hiring* condition specifically the correlation between selfish behaviour in the dictator game and the CRT-WP (r = -.36) was larger than with any overt personality measure (r's ranging from -.17 to .22).

Finally, CRT-WP scores were significantly predictive of self-reported production deviance (r = .20), theft (r = .20), and the overall CWB composite score (r = .20) measured one year later, supporting Hypothesis 8a. Additionally, CRT-WP scores were significantly positively correlated with three of the four cyberloafing factors (online shopping, r = .24; accessing online content, r = .30; online gaming/ betting, r = .26), and the cyberloafing composite score as well (r = .22), supporting Hypothesis 8b.

[11] Since there were no differences in how participants responded to the CRT-WP between conditions, correlations between the CRT-WP and Wave-2/3 outcomes were computed on the whole sample (see Table 6). However, since participants responded differently to all overt personality measures depending on their condition, we reported these correlations separately per condition in Tables S11 to S16 of the online supplement. For full correlations between the CRT-WP and factor-level scores of the cyberloafing and CWB scales, refer to Table S17 of the online supplement.

Given the number of participants not returning to complete Waves 2/3, we also compared results of models based on available data (using a maximum likelihood estimator) to a full information maximum likelihood model (see Newman, 2014). Results are available in Table S18 and were identical between the two for all outcomes. In addition, in a preliminary attempt to examine the incremental validity of the CRT-WP, we tested a regression model. The self-reported psychopathy score from the "dark triad" was entered first. The CRT-WP score entered subsequently and explained incremental variance in selfish behaviors in the dictator game (ΔR^2 = .04, p = .003). The overt self-reported measure of psychopathy was not a significant predictor of selfish behaviors (β = -.06, p = .39), but the CRT-WP was (β = -.22, p = .003).[12]

Discussion

Overall, the results of Study 5 provide some additional support for the validity of the CRT-WP. Wave-1 results show that participants were unable to engage in faking or impression management to artificially lower their scores on the implicit CRT-WP, while they did for all other (overt) personality measures. The only surprising result was that participants in the hiring condition artificially *raised* their level of narcissism, when this is generally seen as an undesirable trait. Yet, participants in the hiring condition were asked "What job or type of job did you imagine applying for in this scenario?", and many of them responded with jobs commonly associated with leadership and power. Therefore, it is possible that participants believed that some level of narcissism was desirable for such positions, which would be supported by a meta-analysis indicating that narcissism has a curvilinear relationship with leadership success (Grijalva et al., 2015).

[12] Identical findings were observed in an alternative regression model where all three dark triad measures were entered first. CRT-WP scores also accounted for incremental variance beyond the self-reported psychopathy score for the cyberloafing factors of "accessing online content" (ΔR^2 = .05, p = .005; β = .23, p = .005) and "online betting" (ΔR^2 = .02, p = .05; β = .16, p = .047).

Wave-1 findings further supported the construct-related validity of the CRT-WP scores. Indeed, CRT-WP scores were related to scores for other personality variables that they should be theoretically related to. For example, CRT-WP scores were negatively related to honesty-humility and integrity (r = -.25 and -.24), but positively correlated with overt psychopathy (r = .20), as well as Machiavellianism and narcissism (r = .27 and .28) for participants who were responding honestly. The CRT-WP being positively related to all sub-scales of the dark triad is not surprising given the content of the items, considering that aspects of Machiavellianism overlap with some JMs such as social superiority. In addition, there is empirical overlap between measures of the three elements of the dark triad (O'Boyle et al., 2012). Although many of these correlations are relatively small, this is standard for overt and covert/implicit measures of the same construct (Bing et al., 2007).

In terms of criterion-related validity, evidence was mixed. CRT-WP scores were related with more selfish choices in the dictator game, providing promising evidence that it can predict behaviours theoretically related to psychopathy, such as self-interested behaviour. CRT-WP scores also scores predicted CWB and cyberloafing behaviours measured over one year later. Although both measures were self-report and may be influenced by impression management, these results provide additional evidence for the criterion-related validity of the CRT-WP and confirm the findings from Study 3. However, CRT-WP scores were not associated with impulsive behaviour as measured by the BART or cheating as measured by the number matrix task (although relationships were in the expected direction for cheating).

Beyond the CRT-WP, in the honest condition, agreeableness and honesty-humility were related to less, while overt Machiavellianism and psychopathy were related to more, selfish choices made in the dictator game. However, these relationships were not observed in the hiring condition, except for honesty-humility. Interestingly, none of the overt variables were related to the BART outcomes, with

two exceptions: In the hiring condition, agreeableness was negatively correlated with the number of balloons exploded, whereas men had significantly more pumps per balloon than women. This could be due to the randomized explosion points or the number of trials used in the current study, which was at the very minimum recommended (Wallsten et al., 2005), or because the BART has been mainly validated for adolescents while the mean age of this sample was almost 40 years old. Scores on the TSD-integrity were unrelated to any of the performance behaviours in Wave-2, in either condition. Overall, these results further show that overt measures of personality (except maybe honesty-humility) do not function appropriately when used in a hiring situation.

Study 6 – CRT-WP Validity in the Workplace

The objective of this final study was to provide additional evidence for the criterion-related validity of the CRT-WP, as well as evidence for incremental validity over and above self-report measures of personality, given the somewhat mixed findings from Studies 3 and 5, in a workplace setting.

While Studies 3 and 5 already demonstrate that the CRT-WP is associated with CWBs (and other outcomes, such as selfish decision-making and cyberloafing), both these studies relied on online samples (i.e., MTurk and Prolific). It is therefore important to provide evidence of criterion-related validity using an organizational sample. In addition, Studies 3 and 5 were not designed specifically to examine the incremental validity of the CRT-WP beyond other personality measures that organizations might consider as alternative assessments. For example, Catano et al. (2018) proposed a short personality-based measure of integrity derived from Big-Five items that was associated with CWBs. Alternatively, organizations could explore facet-level personality measures that are conceptually related to integrity or elements of psychopathy. For instance, Krueger et al. (2012) developed a short self-report measure of irresponsibility derived from a maladaptive personality traits model, whereas

Goldberg et al. (2006) proposed a short measure of cautiousness as a facet of Conscientiousness. Both these constructs are conceptually associated with the psychopathy elements (i.e., the CRT-WP JMs of carefree impulsivity and fearlessness). While such self-report measures would be more prone to faking or impression management than the implicit CRT-WP (as demonstrated in Study 4), organizations might still consider such tools because of their ease of use. Since psychopathy and conceptually similar constructs are particularly relevant to predict the abuse element of CWB (e.g., Boddy, 2011), we propose the following:

> *Hypothesis 9:* CRT-WP scores will display incremental validity over conceptually related self-report personality measures in the prediction of (a) overall CWBs, and (b) the abuse factor of CWBs.

Methods

Sample and Procedure

Employees (N = 2754) working operations and maintenance jobs (e.g., technicians, engineers) in a branch of the Canadian federal government were invited to participate. Data collection took place in two phases. At Time 1, they were asked to complete an online questionnaire including the CRT-WP, three self-reported personality measures, and demographic information. A total of 820 individuals started the questionnaire (i.e., 30% response rate), but 183 did not provide complete responses, and 22 were excluded for endorsing six or more illogical responses to the CRT-WP. This led to usable data from 615 employees. Most of the sample identified as men (89.8%, with 6.8% as women) and White (91.7%, with 1.8% as visible minority, and 2.8% as Indigenous). Employees were on average 35.92 years old (SD = 7.13) and had been working for this organization for 12.63 years (SD = 6.18). At Time 2 (two weeks later), participants were invited back to complete a CWB measure, and 292 individuals did.

Measures

CRT-WP. We used the same 22-item CRT-WP measure (*KR-20* = .82) as in Studies 3-5.

Self-report personality measures. We included three self-report measures conceptually related to psychopathy (or dark personality traits): First, we used the same 10-item TSD-Integrity scale used in Study 5 (α = .79; Catano et al., 2018) with a 7-point scale (1 = *extremely uncharacteristic* to 7 = *extremely characteristic*). Second, we used a 10-item Cautiousness measure from the IPIP (α = .83; e.g., "I choose my words with care"; Goldberg et al., 2006) with a 5-point scale (1 = very inaccurate to 5 = very accurate). Finally, we used a 7-item Irresponsibility Scale from the PID-5 (α = .72; e.g., "I make promises that I don't intend to keep"; Krueger et al., 2012).

CWB. We used the same 30-item CWB-C measure (α = .84; Spector et al., 2006) as in Studies 3 and 5.

Results

Correlations between our main study variables showed that CRT-WP scores were moderately but significantly correlated with the scores of all self-report personality measures at Time 1 (based on N = 615; integrity, r = -.13, p < .001; cautiousness, r = -.16, p < .001; and irresponsibility, r = .19, p < .001). In addition, CRT-WP scores were significantly associated with both overall CWB scores (r = .20, p < .001), and particularly scores on the abuse factor (r = .22, p < .001), from Time 2 (based on N = 292). A full correlation table is provided in Table S19 of the Online Supplement.

To test Hypothesis 9, we conducted two sets of hierarchical regression analyses aimed at examining the incremental validity of CRT-WP scores in predicting CWB scores and abuse factor scores, above and beyond scores from the three self-report personality measures. Results are presented in Table 7. They also include relative weight analyses (using the RWA Web tool; Tonidandel & LeBreton, 2011) to illustrate the importance of CRT-WP scores in predicting variance for both

outcomes. Supporting Hypothesis 9, CRT-WP scores did provide incremental validity for predicting overall CWB scores ($\beta = .13$, $p = .02$; $\Delta R^2 = .02$; Relative weight = 17.81%) and abuse specifically ($\beta = .17$, $p < .01$; $\Delta R^2 = .03$; Relative weight = 28.12%) over and above the three self-report personality scores.

Discussion

The findings from Study 6 replicate and expand on the results presented in Studies 3 and 5, using a sample of actual employees. First, CRT-WP scores were significantly but only moderately correlated with scores on three self-report measures of personality traits conceptually related to psychopathy (or facets of psychopathy): negatively with integrity and cautiousness, but positively with irresponsibility. This provides additional evidence for the construct-related validity of CRT-WP scores. Second, consistent with Studies 3 and 5, CRT-WP scores measured in a workplace setting were positively associated with CWBs. CRT-WP scores were particularly related to abuse, which is consistent with past psychopathy research (Boddy, 2011; Scherer et al., 2013). These results provide additional evidence of the criterion-related validity of CRT-WP scores. Finally, we found promising evidence that CRT-WP scores accounted for incremental variance above and beyond the three self-reports described above. This is practically meaningful because, although such self-report measures are more prone to faking in selection settings (as shown in Study 5), organizations could consider using them instead of (e.g., because of the shorter time commitment) or alongside an implicit measure.

General Discussion

In summary, the results of our six studies provide preliminary evidence to support our implicit measure of workplace psychopathy (the CRT-WP) in terms of content coverage, theoretical relevance, internal consistency, test-retest reliability, construct-related validity, criterion-related validity (ranging from two-week to one-year time lags), incremental validity, and faking-resistance when compared to

overt self-report measures. Importantly, these findings are based on a variety of samples (including North American MTurk respondents, a large sample of real job seekers, experienced managers from the UK, and an organizational sample), research designs or analytical approaches (e.g., factor analysis, IRT, experimental manipulation), and criteria (e.g., self-reports, behavioural).[13] Although more research is needed to fully validate the CRT-WP, it appears to be a reliable and valid measure of work-related psychopathic tendencies, beliefs, and behaviours, that cannot be faked while the implicit nature of the test is maintained. CRT-WP scores are also largely unrelated to sex, and show small relationships with age and ethnicity (especially in our large and more generalizable sample for Study 4 – but see Table S20 for a summary across studies), suggesting limited potential for adverse impact.

Theoretical Implications

Researchers have repeatedly stated the need for a self-administered measure of psychopathy that is resistant to faking, especially for use in organizational contexts (Lilienfeld et al., 2015; Smith & Lilienfeld, 2013; Wu & LeBreton, 2011), and implicit approaches, for instance using the CRT methodology, have been suggested as a promising way to produce such a measure (O'Boyle et al., 2012; Spain et al., 2014). We created and validated the CRT-WP as an answer to these calls, combining best-practice recommendations in CRT development (e.g., James & LeBreton, 2012; LeBreton et al., 2020; Schoen et al., 2021) and conceptual foundations from established psychopathy theories and models. Theoretically, the present research provides preliminary evidence that the CRT methodology can be used to assess "darker" personality traits, such as psychopathy. Practically, the CRT-WP is still self-administered yet avoids common issues with overt, self-report measures for work-related purposes.

[13] We also provide the item-total correlations for the 22-items of the CRT-WP across all studies in Table S21, confirming that all correlations are positive and significant (except for two items in Study 3 Wave-2). We also present the item-outcome correlations for key variables across Studies 3-6 in Table S22. Both tables can be found in the online supplement.

The development of the CRT-WP, including the content coverage and factor structure, also has implications for future research in the area. The two factors that emerged based on factor analysis in Study 2, labeled as individual-oriented and other-oriented psychopathy, were generally aligned with the structure emerging from self-report psychopathy measures, for instance the "primary" and "secondary" factors in the SRP and LSRP. This is especially relevant, given that JMs and items were created as an amalgamation of key conceptualizations of psychopathy, but using an implicit measurement approach. The interaction analyses performed in Study 3 also supported the channeling hypothesis, confirming that there are benefits to combining overt and implicit measures of psychopathy (Bing et al., 2007).

This research highlights the potential of developing and using implicit measures (e.g., CRTs) to assess dark personality traits in a workplace setting. Although CRTs have not yet become very popular (LeBreton et al., 2020), they are uniquely suited for situations where other measurement approaches face issues. Although any personality construct could benefit from having an implicit measure (given that all overt personality measures were faked in Study 5), they may prove most useful to assess highly desirable (or undesirable) constructs, since these may result in the most faking, as is the case with the CRT for aggression (James et al., 2005). For instance, it could be beneficial to develop implicit measures for other constructs important to organizations (e.g., other "dark" personality traits).

Practical Implications

Although more research is certainly needed before the CRT-WP can be implemented in organizational settings, the results of our six studies provide preliminary evidence that it could represent a valuable tool for organizations, for instance to assess individuals for selection or promotion. Its self-administered nature makes it practical and cost-efficient, as it is not requiring a third-party rater who may inaccurately assess the construct of psychopathy, and which is highly impractical in a hiring context. However, unlike traditional self-report measures, the CRT-WP is not susceptible to impression

management and faking. As such, it could be very useful for jobs or organizations in which deterring abusive and unethical leadership is a priority (e.g., military, law enforcement). The CRT-WP also shows no correlations with participant sex and no or small relationships with ethnicity (see Table S21), while overt measures of psychopathy do correlate with participant sex and could have potential for adverse impact.

One practical weakness of CRTs compared to their overt counterparts is the longer assessment time. The amount of text for each item, along with the fact that participants may spend minutes on each item trying to deduce the "correct" answer, means CRTs may take longer to complete than overt scales. Estimates from the first wave of Study 2 would suggest it takes an average of 18-19 minutes for participants to complete the 22-item CRT-WP. That said, the value of this trade-off likely depends on the job and the organization. For example, the CRT-WP may not be worthwhile to assess or select entry-level employees. Overall, the CRT-WP has potential as a valuable tool to identify potential "red flags" (a) for managers or leaders in government or private organizations, or individuals in positions of power or authority (i.e., that are attractive to individuals higher on psychopathy and where they can do the most "damage"); (b) when used alongside other valid predictors of job performance (e.g., structured interviews, cognitive ability tests); and (c) in the final stages of the selection process to screen a final pool of candidates that has already been narrowed down by other measures.

Consistent with recommendations from Jones and Arnold (2008), we would encourage organizations interested in using the CRT-WP to avoid terms like "psychopathy", both for applicants or employees who complete the assessment or managers potentially making decisions based on CRT-WP scores. For example, referring to the construct of psychopathy directly, when debriefing test-takers about their results or scores, could (a) incorrectly lead applicants to believe that a clinical assessment of mental health disorders was used (when the CRT-WP was designed to be disconnected from such

disorders); and (b) lead to negative reactions from test-takers as non-experts often associate "psychopathy" with serial killers and other extreme cases due to overuse in entertainment and media (Caponecchia et al., 2012). However, providing feedback and transparency in assessment tools is important for organizations, and we recommend using similar but less inciting language such as "dark personality/leadership tendencies". Finally, we recommend that organizations use the number of illogical response options selected by a test-taker in the CRT-WP as an indicator of attentiveness or response quality. In line with past CRT work, scores of individuals selecting over 25% of illogical responses should be considered as invalid, and the person should ideally be asked to retake the CRT-WP.

Limitations and Future Directions

One limitation of the current research is the reliance on self-report measures for some outcome variables, which might create issues with social desirability (see the correlations between CWBs and the BIDR in Study 3). Therefore, future research examining the predictive validity of the CRT-WP should use objective measures of outcome variables, if possible, similar to Wave-2 of Study 5. Studies conducted with employees or job applicants could use performance appraisal data or documented complaints, for example. Second, future research should seek to determine whether the CRT-WP assesses *workplace-specific* psychopathy as it intends to. This aim can be accomplished by including a work-related self-report measure of psychopathy (e.g., B-Scan Self; Mathieu & Babiak, 2016b) and determining whether CRT-WP scores are more strongly related to scores on that work-related measure compared to a self-report measure of general psychopathy. Third, while we include real job seekers in Study 4, they completed the CRT-WP in a low-stakes situation (i.e., for practice/preparation). Thus, before the CRT-WP can be used in hiring processes, studies should be conducted with real job applicants completing the test in a high-stakes context. Future research should also explore the possible

designation of cut-off scores. According to previous CRT research, there is no definitive answer regarding how to consistently apply or justify cut-off scores for use by organizations (James & LeBreton, 2012), especially when considering different jobs (i.e., hospital manager vs. professional football coach).

Future research should also examine test-taker reactions to the CRT-WP (and other implicit CRTs, in general) and the ability of the CRT-WP to prevent faking when participants are aware of what it is actually measuring. Regarding the former, it is important to determine whether the implicit (and somewhat deceptive) nature of the CRT-WP causes test-takers to have less favourable reactions to assessment. This would have significant implications if organizations use the CRT-WP to evaluate employees or applicants, as these test-takers may have negative attitudes toward the organization as a result, which can lead to refusing job offers and more (McCarthy et al., 2017). Second, the faking resistance of CRTs is generally conditional upon their implicit nature remaining intact. Yet, participants *can* fake CRTs when informed about how the test *truly* works beforehand (Bowler et al., 2013; LeBreton et al., 2007; Wiita et al., 2020). Recently, Wiita et al. (2020) and Schoen et al. (2022) added faking detection scales to CRTs, to give them a second layer of protection against faking. Such scales involve embedding additional items using the same format within the CRTs, but alter the response options so that these new items only have *one* correct answer that is high on the trait of interest. Respondents unaware of the trait being measured should (and do) still select that option because it is the *only* "correct" one from a conditional reasoning standpoint. However, those privy to the nature of the CRT and attempting to fake to lower their scores get caught because they consistently avoid that (i.e., more aggressive or low achievement) option on these items. Future research should follow a similar procedure to develop a faking-detection scale for the CRT-WP so that it is both proactively faking-preventative and reactively faking-detecting.

Conclusion

In summary, the current research presents the theoretical foundations and initial psychometric evidence for an implicit measure of workplace psychopathy, and suggests that it has benefits over overt measures of psychopathic personality. The development of the CRT-WP fills multiple gaps in the personality literature as it is the first, to our knowledge, to measure of work-related psychopathy that is both self-reported and faking-resistant. In theory, the CRT-WP could help identify or screen individuals with high levels of psychopathic tendencies who may be causing abusive, toxic, and unethical work environments.

Acknowledgements

We would like to thank Dr. Zvonomir Galić for his insights and comments on an earlier version of this work.

Declarations

Conflicts of interest: The authors have no potential conflicts of interest to report.

Data availability statement: Data and other research materials are available at https://osf.io/3aegn/?view_only=484fec494272431fa8e03298da793153

Ethics approval: This research, and each individual study, was approved by the Research Ethics Board of Saint Mary's University (REB file: 19-045)

Funding: This research was funded by the Social Sciences and Humanities Research Council (SSHRC) via a Canada Graduate Scholarship – Doctoral (767-2020-1448) awarded to the first author.

Consent to participate: Informed consent was obtained from all participants included in this research.

References

ACFE. (2020). *Report to the Nations: Global Study on Occupational Fraud and Abuse*. https://acfepublic.s3-us-west-2.amazonaws.com/2020-Report-to-the-Nations.pdf

Akbulut, Y., Dursun, Ö. Ö., Dönmez, O., & Şahin, Y. L. (2016). In search of a measure to investigate cyberloafing in educational settings. *Computers in Human Behavior*, *55*, 616–625. https://doi.org/10.1016/j.chb.2015.11.002

Aklin, W. M., Lejuez, C. W., Zvolensky, M. J., Kahler, C. W., & Gwadz, M. (2005). Evaluation of behavioral measures of risk taking propensity with inner city adolescents. *Behaviour Research and Therapy*, *43*(2), 215–228. https://doi.org/10.1016/j.brat.2003.12.007

Anderson, J. L., Sellbom, M., Wygant, D. B., & Edens, J. F. (2013). Examining the necessity for and utility of the Psychopathic Personality Inventory—Revised (PPI–R) validity scales. *Law and Human Behavior*, *37*(5), 312–320. https://doi.org/10.1037/lhb0000018

Ashton, M. C., & Lee, K. (2009). The HEXACO-60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, *91*(4), 340–345. https://doi.org/10.1080/00223890902935878

Babiak, P. (1995). When psychopaths go to work: A case study of an industrial psychopath. *Applied Psychology: An International Review*, *44*(2), 171–188. https://doi.org/10.1111/j.1464-0597.1995.tb01073.x

Babiak, P. (2007). From Darkness Into the Light: Psychopathy in Industrial and Organizational Psychology. In H. Hervé & J. C. Yuille (Eds.), *The Psychopath: Theory, Research, and Practice.* (2006-11788-016; pp. 411–428). Lawrence Erlbaum Associates Publishers.

Babiak, P., & Hare, R. D. (2012). *The B-Scan 360 Manual*.

Berry, C. M., Sackett, P. R., & Tobares, V. (2010). A meta-analysis of conditional reasoning tests of aggression. *Personnel Psychology*, *63*(2), 361–384. https://doi.org/10.1111/j.1744-6570.2010.01173.x

Bing, M. N., Stewart, S. M., Davison, H. K., Green, P. D., McIntyre, M. D., & James, L. R. (2007). An integrative typology of personality assessment for aggression: Implications for predicting counterproductive workplace behavior. *Journal of Applied Psychology*, *92*(3), 722–744. https://doi.org/10.1037/0021-9010.92.3.722

Blickle, G., & Schütte, N. (2017). Trait psychopathy, task performance, and counterproductive work behavior directed toward the organization. *Personality and Individual Differences*, *109*, 225–231. https://doi.org/10.1016/j.paid.2017.01.006

Boddy, C., Miles, D., Sanyal, C., & Hartog, M. (2015). Extreme managers, extreme workplaces: Capitalism, organizations and corporate psychopaths. *Organization*, *22*(4), 530–551. https://doi.org/10.1177/1350508415572508

Boddy, C. R. (2011). Corporate psychopaths, bullying and unfair supervision in the workplace. *Journal of Business Ethics*, *100*(3), 367–379. https://doi.org/10.1007/s10551-010-0689-5

Boddy, C. R. P., Ladyshewsky, R., & Galvin, P. (2010). Leaders without ethics in global business: Corporate psychopaths. *Journal of Public Affairs*, *10*(3), 121–138. https://doi.org/10.1002/pa.352

Bowler, J. L., & Bowler, M. C. (2014). Evaluating the fakability of a conditional reasoning test of addiction proneness. *International Journal of Psychology*, *49*(5), 415–419. https://doi.org/10.1002/ijop.12030

Bowler, J. L., Bowler, M. C., & Cope, J. G. (2013). Measurement issues associated with conditional reasoning tests: An examination of faking. *Personality and Individual Differences, 55*(5), 459–464. https://doi.org/10.1016/j.paid.2013.04.011

Brosig-Koch, J., Riechmann, T., & Weimann, J. (2017). The dynamics of behavior in modified dictator games. *PLoS ONE, 12*(4), e0176199. https://doi.org/10.1371/journal.pone.0176199

Burns, G. N., & Christiansen, N. D. (2011). Methods of measuring faking behavior. *Human Performance, 24*(4), 358–372. https://doi.org/10.1080/08959285.2011.597473

Caponecchia, C., Sun, A. Y. Z., & Wyatt, A. (2012). 'Psychopaths' at work? Implications of lay persons' use of labels and behavioural criteria for psychopathy. *Journal of Business Ethics, 107*(4), 399–408. https://doi.org/10.1007/s10551-011-1049-9

Catano, V. M., O'Keefe, D. F., Francis, R. E., & Owens, S. M. (2018). Construct-based approach to developing a short, personality-based measure of integrity. *International Journal of Selection and Assessment, 26*(1), 75–92. https://doi.org/10.1111/ijsa.12197

Catano, V., Wiesner, W., & Hackett, R. (2016). *Recruitment and Selection in Canada* (Vol. 6). Nelson Education.

Coid, J., Yang, M., Ullrich, S., Roberts, A., & Hare, R. D. (2009). Prevalence and correlates of psychopathic traits in the household population of Great Britain. *International Journal of Law and Psychiatry, 32*(2), 65–73. https://doi.org/10.1016/j.ijlp.2009.01.002

Conway, J. M., & Huffcutt, A. I. (2003). A review and evaluation of exploratory factor analysis practices in organizational research. *Organizational Research Methods, 6*(2), 147–168. https://doi.org/10.1177/1094428103251541

DeSimone, J. A., & James, L. R. (2015). An item analysis of the Conditional Reasoning Test of Aggression. *Journal of Applied Psychology, 100*(6), 1872–1886. https://doi.org/10.1037/apl0000026

Edens, J. F., Buffington, J. K., Tomicic, T. L., & Riley, B. D. (2001). Effects of positive impression management on the Psychopathic Personality Inventory. *Law and Human Behavior, 25*(3), 235–256. https://doi.org/10.1023/A:1010793810896

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists.* (pp. xi, 371). Lawrence Erlbaum Associates Publishers.

Fisher, P. A., Robie, C., Christiansen, N. D., & Komar, S. (2018). The impact of psychopathy and warnings on faking behavior: A multisaturation perspective. *Personality and Individual Differences, 127*, 39–43. https://doi.org/10.1016/j.paid.2018.01.033

Forouzan, E., & Cooke, D. J. (2005). Figuring out la femme fatale: Conceptual and assessment issues concerning psychopathy in females. *Behavioral Sciences & the Law, 23*(6), 765–778. https://doi.org/10.1002/bsl.669

Galić, Z. (2016). Conditional Reasoning Test for Aggression: Further evidence about incremental validity. *International Journal of Selection and Assessment, 24*(1), 24–33. https://doi.org/10.1111/ijsa.12126

Galić, Z., Scherer, K. T., & LeBreton, J. M. (2014). Validity evidence for a Croatian version of the Conditional Reasoning Test for Aggression. *International Journal of Selection and Assessment, 22*(4), 343–354. https://doi.org/10.1111/ijsa.12082

Gao, Y., & Raine, A. (2010). Successful and unsuccessful psychopaths: A neurobiological model. *Behavioral Sciences & the Law, 28*(2), 194–210. https://doi.org/10.1002/bsl.924

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe* (Vol. 7, pp. 7–28). Tilburg University Press.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Proceedings of the 2005 Meeting of the Association of Research in Personality*, *40*(1), 84–96. https://doi.org/10.1016/j.jrp.2005.08.007

Grijalva, E., Harms, P. D., Newman, D. A., Gaddis, B. H., & Fraley, R. C. (2015). Narcissism and leadership: A meta-analytic review of linear and nonlinear relationships. *Personnel Psychology*, *68*(1), 1–47. https://doi.org/10.1111/peps.12072

Hare, R. D. (1980). A research scale for the assessment of psychopathy in criminal populations. *Personality and Individual Differences*, *1*(2), 111–119. https://doi.org/10.1016/0191-8869(80)90028-8

Hare, R. D. (1985). Comparison of procedures for the assessment of psychopathy. *Journal of Consulting and Clinical Psychology*, *53*(1), 7–16. https://doi.org/10.1037/0022-006X.53.1.7

Hart, C. M., Ritchie, T. D., Hepper, E. G., & Gebauer, J. E. (2015). The Balanced Inventory of Desirable Responding Short Form (BIDR-16). *SAGE Open*, *5*(4), 1–9. https://doi.org/10.1177/2158244015621113

Hauser, D., Paolacci, G., & Chandler, J. (2019). Common concerns with MTurk as a participant pool: Evidence and solutions. In *Handbook of research methods in consumer psychology*. (pp. 319–337). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9781351137713-17

Henley, A. G. (2002). *Psychopathy and career interest in a noncriminal population* (2002-95012-130; Issues 12-B). The University of Texas at Austin.

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1*(1), 104–121. https://doi.org/10.1177/109442819800100106

James, L. R. (1998). Measurement of personality via conditional reasoning. *Organizational Research Methods, 1*(2), 131–163. https://doi.org/10.1177/109442819812001

James, L. R., & LeBreton, J. M. (2012). *Assessing the implicit personality through conditional reasoning.* (pp. 1, 244). American Psychological Association. https://doi.org/10.1037/13095-000

James, L. R., & McIntyre, M. D. (2000). *Conditional reasoning test of aggression: Test manual*. Innovative Assessment Technology.

James, L. R., McIntyre, M. D., Glisson, C. A., Bowler, J. L., & Mitchell, T. R. (2004). The conditional reasoning measurement system for aggression: An overview. *Human Performance, 17*(3), 271–295. https://doi.org/10.1207/s15327043hup1703_2

James, L. R., McIntyre, M. D., Glisson, C. A., Green, P. D., Patton, T. W., LeBreton, J. M., Frost, B. C., Russell, S. M., Sablynski, C. J., Mitchell, T. R., & Williams, L. J. (2005). A conditional reasoning measure for aggression. *Organizational Research Methods*, *8*(1), 69–99. https://doi.org/10.1177/1094428104272182

Jones, D. N., & Paulhus, D. L. (2011). The role of impulsivity in the Dark Triad of personality. *Personality and Individual Differences*, *51*(5), 679–682. https://doi.org/10.1016/j.paid.2011.04.011

Jones, J. W., & Arnold, D. W. (2008). Protecting the Legal and Appropriate Use of Personality Testing: A Practitioner Perspective. *Industrial and Organizational Psychology, 1*(3), 296–298. https://doi.org/10.1111/j.1754-9434.2008.00050.x

Kelley, S. E., Edens, J. F., Donnellan, M. B., Mowle, E. N., & Sörman, K. (2018). Self- and informant perceptions of psychopathic traits in relation to the triarchic model. *Journal of Personality, 86*(4), 738–751. https://doi.org/10.1111/jopy.12354

Kelsey, K. R. (2016). *The vulnerability of self-report measures of psychopathy to positive impression management: A simulation study with inmates* (2016-26528-172; Issues 1-B(E)). University of North Texas.

Kelsey, K. R., Rogers, R., & Robinson, E. V. (2015). Self-report measures of psychopathy: What is their role in forensic assessments? *Journal of Psychopathology and Behavioral Assessment, 37*(3), 380–391. https://doi.org/10.1007/s10862-014-9475-5

Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine, 42*(9), 1879–1890. https://doi.org/10.1017/S0033291711002674

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika, 2,* 151–160. https://doi.org/10.1007/BF02288391

Landay, K., Harms, P. D., & Credé, M. (2019). Shall we serve the dark lords? A meta-analytic review of psychopathy and leadership. *Journal of Applied Psychology*, *104*(1), 183–196. https://doi.org/10.1037/apl0000357

Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: Arbitrary distinctions between organizational, Mechanical Turk, and other convenience samples. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 8*(2), 142–164. https://doi.org/10.1017/iop.2015.13

LeBreton, J. M., Barksdale, C. D., Robin, J., & James, L. R. (2007). Measurement issues associated with conditional reasoning tests: Indirect measurement and test faking. *Journal of Applied Psychology*, *92*(1), 1–16. https://doi.org/10.1037/0021-9010.92.1.1

LeBreton, J. M., Grimaldi, E. M., & Schoen, J. L. (2020). Conditional reasoning: A review and suggestions for future test development and validation. *Organizational Research Methods*, *23*(1), 65–95. https://doi.org/10.1177/1094428118816366

Lee, K., & Ashton, M. C. (2005). Psychopathy, Machiavellianism, and narcissism in the five-factor model and the HEXACO model of personality structure. *Personality and Individual Differences*, *38*(7), 1571–1582. https://doi.org/10.1016/j.paid.2004.09.016

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, *8*(2), 75–84. https://doi.org/10.1037/1076-898X.8.2.75

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology*, *68*(1), 151–158. https://doi.org/10.1037/0022-3514.68.1.151

Lilienfeld, S. O., & Andrews, B. P. (1996). Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *Journal of Personality Assessment*, *66*(3), 488–524. https://doi.org/10.1207/s15327752jpa6603_3

Lilienfeld, S. O., Waldman, I. D., Landfield, K., Watts, A. L., Rubenzer, S., & Faschingbauer, T. R. (2012). Fearless dominance and the US presidency: Implications of psychopathic personality traits for successful and unsuccessful political leadership. *Journal of Personality and Social Psychology*, *103*(3), 489–505. https://doi.org/10.1037/a0029392

Lilienfeld, S. O., Watts, A. L., & Smith, S. F. (2015). Successful psychopathy: A scientific status report. *Current Directions in Psychological Science, 24*(4), 298–303. https://doi.org/10.1177/0963721415580297

Ljubin-Golub, T., Petričević, E., & Sokić, K. (2020). Predicting academic cheating with triarchic psychopathy and cheating attitudes. *Journal of Academic Ethics, 18*(4), 377–393. https://doi.org/10.1007/s10805-019-09338-0

MacPherson, L., Magidson, J. F., Reynolds, E. K., Kahler, C. W., & Lejuez, C. W. (2010). Changes in sensation seeking and risk-taking propensity predict increases in alcohol use among early adolescents. *Alcoholism: Clinical and Experimental Research, 34*(8), 1400–1408. https://doi.org/10.1111/j.1530-0277.2010.01223.x

Mahmut, M. K., Menictas, C., Stevenson, R. J., & Homewood, J. (2011). Validating the factor structure of the Self-Report Psychopathy Scale in a community sample. *Psychological Assessment, 23*(3), 670–678. https://doi.org/10.1037/a0023090

Marcus, D. K., Church, A. S., O'Connell, D., & Lilienfeld, S. O. (2018). Identifying careless responding with the Psychopathic Personality Inventory–Revised validity scales. *Assessment, 25*(1), 31–39. https://doi.org/10.1177/1073191116641507

Mathieu, C., & Babiak, P. (2016a). Corporate psychopathy and abusive supervision: Their influence on employees' job satisfaction and turnover intentions. *Personality and Individual Differences, 91*, 102–106. https://doi.org/10.1016/j.paid.2015.12.002

Mathieu, C., & Babiak, P. (2016b). Validating the B-Scan Self: A self-report measure of psychopathy in the workplace. *International Journal of Selection and Assessment, 24*(3), 272–284. https://doi.org/10.1111/ijsa.12146

McCarthy, J. M., Bauer, T. N., Truxillo, D. M., Anderson, N. R., Costa, A. C., & Ahmed, S. M. (2017). Applicant perspectives during selection: A review addressing "So what?," "What's new?," and "Where to next?" *Journal of Management*, *43*(6), 1693–1725. https://doi.org/10.1177/0149206316681846

Moor, L., & Anderson, J. R. (2019). A systematic literature review of the relationship between dark personality traits and antisocial online behaviours. *Personality and Individual Differences*, *144*, 40–55. https://doi.org/10.1016/j.paid.2019.02.027

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide: Statistical analysis with latent variables* (Vol. 7). Muthén & Muthén.

Newman, D. A. (2014). Missing Data: Five Practical Guidelines. *Organizational Research Methods*, *17*(4), 372–411. https://doi.org/10.1177/1094428114548590

Nicholls, T. L., & Petrila, J. (2005). Gender and psychopathy: An overview of important issues and introduction to the special issue. *Behavioral Sciences & the Law*, *23*(6), 729–741. https://doi.org/10.1002/bsl.677

O'Boyle, E. H. Jr., Forsyth, D. R., Banks, G. C., & McDaniel, M. A. (2012). A meta-analysis of the Dark Triad and work behavior: A social exchange perspective. *Journal of Applied Psychology*, *97*(3), 557–579. https://doi.org/10.1037/a0025679

Olver, M. E., & Wong, S. C. P. (2015). Short- and long-term recidivism prediction of the PCL-R and the effects of age: A 24-year follow-up. *Personality Disorders: Theory, Research, and Treatment*, *6*(1), 97–105. https://doi.org/10.1037/per0000095

Pardue, A. D., Robinson, M. B., & Arrigo, B. A. (2013). Psychopathy and corporate crime: A preliminary examination, Part 2. *Journal of Forensic Psychology Practice*, *13*(2), 145–169. https://doi.org/10.1080/15228932.2013.765746

Patrick, C. J. (2010). Operationalizing the triarchic conceptualization of psychopathy: Preliminary description of brief scales for assessment of boldness, meanness, and disinhibition. *Unpublished Test Manual, Florida State University, Tallahassee, FL*, 1110–1131.

Paulhus, D. L., Hemphill, J. F., & Hare, R. D. (2012). *Self-report psychopathy scale (SRP-III)*. Toronto, ON, Canada: Multi-Health Systems.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Robinson, E. V., & Rogers, R. (2015). Empathy faking in psychopathic offenders: The vulnerability of empathy measures. *Journal of Psychopathology and Behavioral Assessment*, *37*(4), 545–552. https://doi.org/10.1007/s10862-015-9479-9

Rogers, R., Vitacco, M. J., Jackson, R. L., Martin, M., Collins, M., & Sewell, K. W. (2002). Faking psychopathy? An examination of response styles with antisocial youth. *Journal of Personality Assessment*, *78*(1), 31–46. https://doi.org/10.1207/S15327752JPA7801_03

Roulin, N., & Krings, F. (2016). When winning is everything: The relationship between competitive worldviews and job applicant faking. *Applied Psychology: An International Review*, *65*(4), 643–670. https://doi.org/10.1111/apps.12072

Scherer, K. T., Baysinger, M., Zolynsky, D., & LeBreton, J. M. (2013). Predicting counterproductive work behaviors with sub-clinical psychopathy: Beyond the Five Factor Model of personality. *Personality and Individual Differences*, *55*(3), 300–305. https://doi.org/10.1016/j.paid.2013.03.007

Schoen, J. L., Bowler, J. L., & Schilpzand, M. C. (2018). Conditional reasoning test for creative personality: Rationale, theoretical development, and validation. *Journal of Management*, *44*(4), 1651–1677. https://doi.org/10.1177/0149206315618012

Schoen, J. L., DeSimone, J. A., Meyer, R. D., Schnure, K. A., & LeBreton, J. M. (2021). Identifying, Defining, and Measuring Justification Mechanisms: The Implicit Biases Underlying Individual Differences. *Journal of Management*, *47*(3), 716–744. https://doi.org/10.1177/0149206319889137

Schoen, J. L., Williams, J. L., Reichin, S. L., & Meyer, R. D. (2022). It's a trap! Faking and faking detection on conditional reasoning tests. *Personality and Individual Differences*, *198*, 111803. https://doi.org/10.1016/j.paid.2022.111803

Smith, D. R., Hoffman, M. E., & LeBreton, J. M. (2020). Conditional reasoning: An integrated approach to item analysis. *Organizational Research Methods*, *23*(1), 124–153. https://doi.org/10.1177/1094428119879756

Smith, S. F., & Lilienfeld, S. O. (2013). Psychopathy in the workplace: The knowns and unknowns. *Aggression and Violent Behavior*, *18*(2), 204–218. https://doi.org/10.1016/j.avb.2012.11.007

Smith, S. F., Watts, A. L., & Lilienfeld, S. O. (2014). On the trail of the elusive successful psychopath. *The Psychologist*, *27*(7), 506–510.

Spain, S. M., Harms, P., & LeBreton, J. M. (2014). The dark side of personality at work. *Journal of Organizational Behavior*, *35*(S1), S41–S60. https://doi.org/10.1002/job.1894

Spector, P. E., Fox, S., Penney, L. M., Bruursema, K., Goh, A., & Kessler, S. (2006). The dimensionality of counterproductivity: Are all counterproductive behaviors created equal? *Journal of Vocational Behavior*, *68*(3), 446–460. https://doi.org/10.1016/j.jvb.2005.10.005

Stevens, G. W., Deuling, J. K., & Armenakis, A. A. (2012). Successful psychopaths: Are they unethical decision-makers and why? *Journal of Business Ethics*, *105*(2), 139–149. https://doi.org/10.1007/s10551-011-0963-1

ten Brinke, L., Kish, A., & Keltner, D. (2018). Hedge fund managers with psychopathic tendencies make for worse investors. *Personality and Social Psychology Bulletin*, *44*(2), 214–223. https://doi.org/10.1177/0146167217733080

Tepper, B. J., Duffy, M. K., Henle, C. A., & Lambert, L. S. (2006). Procedural injustice, victim precipitation, and abusive supervision. *Personnel Psychology*, *59*(1), 101–123. https://doi.org/10.1111/j.1744-6570.2006.00725.x

Tonidandel, S., & LeBreton, J. M. (2011). Relative Importance Analysis: A Useful Supplement to Regression Analysis. *Journal of Business and Psychology*, *26*(1), 1–9. https://doi.org/10.1007/s10869-010-9204-3

Vergauwe, J., Hofmans, J., Wille, B., Decuyper, M., & De Fruyt, F. (2021). Psychopathy and leadership effectiveness: Conceptualizing and testing three models of successful psychopathy. *The Leadership Quarterly*, *32*(6), 101536. https://doi.org/10.1016/j.leaqua.2021.101536

Verschuere, B., Uzieblo, K., De Schryver, M., Douma, H., Onraedt, T., & Crombez, G. (2014). The inverse relation between psychopathy and faking good: Not response bias, but true variance in psychopathic personality. *Journal of Forensic Psychiatry & Psychology*, *25*(6), 705–713. https://doi.org/10.1080/14789949.2014.952767

Viswesvaran, C., & Ones, D. S. (1999). Meta-analyses of fakability estimates: Implications for personality measurement. *Educational and Psychological Measurement*, *59*(2), 197–210. https://doi.org/10.1177/00131649921969802

Wallsten, T. S., Pleskac, T. J., & Lejuez, C. W. (2005). Modeling behavior in a clinically diagnostic sequential risk-taking task. *Psychological Review*, *112*(4), 862–880. https://doi.org/10.1037/0033-295X.112.4.862

Wiita, N. E., Meyer, R. D., Kelly, E. D., & Collins, B. J. (2020). Not aggressive or just faking it? Examining faking and faking detection on the conditional reasoning test of aggression. *Organizational Research Methods*, *23*(1), 96–123. https://doi.org/10.1177/1094428117703685

Wiltermuth, S. S. (2011). Cheating more when the spoils are split. *Organizational Behavior and Human Decision Processes*, *115*(2), 157–168. https://doi.org/10.1016/j.obhdp.2010.10.001

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, *34*(6), 806–838. https://doi.org/10.1177/0011000006288127

Wu, J., & Lebreton, J. M. (2011). Reconsidering the dispositional basis of counterproductive work behaviour: The role of aberrant personality. *Personnel Psychology*, *64*(3), 593–626. https://doi.org/10.1111/j.1744-6570.2011.01220.x

Table 1

Existing Measures of Psychopathy and their Issues with Use in Hiring

	Measure	*Target Use*	*Issues for Use in Hiring*
1.	PCL-R	Forensic	- Conceptualized and validated using criminals, which have demonstrated differences from corporate psychopaths - Requires expert rater and lengthy/invasive assessment
2.	SRP	Forensic and Non-Forensic	- Conceptualized and validated using criminals, which have demonstrated differences from corporate psychopaths - Susceptible to faking and social desirability
3.	PPI-R	Non-Forensic and Forensic	- Conceptualized with mainly students and "community" samples, which likely do not capture workplace specifics - Susceptible to faking, although detects social desirability
4.	LSRP	Forensic and Non-Forensic	- Conceptualized and validated using criminals, which have demonstrated differences from corporate psychopaths - Susceptible to faking and social desirability
5.	TriPM	Non-Forensic and Forensic	- Conceptualized with mainly students and "community" samples, which likely do not capture workplace specifics - Susceptible to faking and social desirability
6.	PM-MRV	Corporate	- Uses untrained raters to make assessments of others - The rater must have familiarity with the target, therefore it is unlikely it could be used to rate a new applicant
7.	B-Scan 360	Corporate	- Uses untrained raters to make assessments of others - The rater must have familiarity with the target, therefore it is unlikely it could be used to rate a new applicant
8.	B-Scan Self	Corporate	- Conceptualized using PCL-R framework of psychopathy which was based on criminals - Susceptible to faking and social desirability

Note. The ordering of "forensic and non-forensic" or "non-forensic and forensic" signifies which target population the measure was designed for first, and who it later became used for secondarily. PCL-R = Psychopathy Checklist – Revised; SRP = Self-Report Psychopathy scale; PPI-R = Psychopathic Personality Inventory – Revised; LSRP = Levenson Self-Report Psychopathy Scale; TriPM = Triarchic Psychopathy Measure; PM-MRV = Psychopathy Measure – Management Research Version.

Table 2

CRT-WP Justification Mechanisms and Content Coverage

CRT-WP Justification Mechanisms		*Conceptual Overlap with Models of Psychopathy*			
Name	*Description*	*PCL-R/SRP*	*PPI-R*	*TriPM*	*LSRP*
1. Externalization	A bias to automatically attribute fault for negative occurrences and failures, and an unconscious need to rationalize shifting blame, to other people or external factors. This tendency results in a "global irresponsibility" for undesired actions and outcomes, even when they clearly resulted from choices under their control. Mentally justifying the deflection of blame to others or external causes is a way to absolve themselves of any wrongdoings and protect their ego. Even when consciously analyzing a negative outcome, there is an unconscious need to find an ultimate target of blame that is not the self.	- Irresponsibility - Failure to accept responsibility	- Blame externalization	- Disinhibition	- Primary
2. Carefree Impulsivity	An inclination for actions and decisions guided by impulsivity instead of reasoning, deliberation, or long-term planning. Impulsive and spontaneous decision-making are automatically categorized as exciting and interesting, while more careful thought is categorized as boring and basic. A careful way thinking and reasoning before acting is labelled as a deficiency that most "regular" people value and need, while "above-average" individuals (like themselves) can engage in more exciting and spontaneous ways of thinking.	- Poor behaviour control - Lack of long-term goals - Impulsivity	- Carefree non-planfulness - Impulsive non-conformity	- Disinhibition	- Secondary
3. Social Superiority	A persisting belief that one's social status and social skills are superior to generally everyone around them. Individuals with this bias believe that they can charm and persuade others in any situation and justify doing so simply because they have the skill. They also believe that they are a dominant, alpha social personality that should be considered above others. They also rationalize that these social characteristics (e.g., dominance, influence, persuasiveness) should be valued and rewarded by everyone, and those who do not share them must be inferior.	- Glib/Superficial charm - Grandiose estimation of self - Pathological lying [B]	- Social potency	- Boldness	- Primary & Secondary [A]

4. Fearlessness	A bias in interpreting high-stress situations as stimulating and to frame risky behaviors or decisions as bold and brave. There is a predisposition to interpret the level of fear or anxiety that most people experience in stressful situations as a sign or weakness. This tendency results in an unconscious inclination for high-risk behaviours along with a high tolerance/resilience for uncertainty in potential outcomes. This differs from carefree impulsivity in that individuals rationalize (though in a biased way) their high-risk choices even after deliberation, as fearlessness and boldness are ingrained as core values.	- Need for stimulation [B]	- Fearlessness - Stress immunity	- Boldness	- Secondary
5. Ruthless Self-Interest	A predisposition to see situations as opportunities for self-promotion with complete disregard for anyone or anything other than the self. The belief and value systems are biased toward achieving individual goals and advancement at any cost, thereby justifying exploitation, manipulation, or negative impacts felt by others. It is rationalized that ruthlessness is part of being strong, and that one needs to engage in these behaviours to guard against the threat that someone may surpass or do better than them. There is a survival-of-the-fittest mentality.	- Cunning and manipulativeness - Callousness - Parasitic lifestyle	- Machiavellian egocentricity	- Meanness	- Primary
6. Insensitivity	A disinclination to feel concern, guilt, remorse, or consideration to the feelings of others when interpreting events or relationships. Guided by a deep-rooted belief that the self is more important than others, who are framed as simple background characters, considering the feelings of others is an unconsciously low priority and often automatically ignored. The innate framing of others as low importance results in an indifference toward their emotions, beliefs, or desires, and an overall lack of empathy. This differs from ruthless self-interest in that insensitivity is always present, even in situations where there is nothing to gain.	- Lack of remorse/ guilt [B] - Shallow affect - Lack of empathy	- Coldheartedness	- Meanness	- Primary

Note. It is acceptable for the JMs to have some degree of overlap according to the other implicit CRTs which have already been developed (James et al., 2004). Only the factors of the PCL-R and SRP related to criminal behaviour are not covered, which is by choice. [A] Social Superiority covers both primary and secondary factors of the LSRP. [B] These factors are somewhat covered by other JMs as well.

Table 3

Descriptive Statistics, Reliability Coefficients, and Correlations for the 22-Item CRT-WP in Study 2

	Scale	*M*	*SD*	1.	2.	3.	4.	5.	6.	7.
1. Factor 1	-10 to 10	-5.68	3.21	(.77)						
2. Factor 2	-12 to 12	-2.66	5.00	.22**	(.80)					
3. CRT-WP	-22 to 22	-8.34	6.49	.66**	.88**	(.80)				
4. Sex	1-2	1.40	0.49	-.03	-.10*	-.09	-			
5. Ethnicity	1-2	1.21	0.41	.16**	-.07	.03	.03	-		
6. Education	1-2	1.53	0.50	-.02	.00	-.01	.01	-.04	-	
7. Employment	1-2	1.12	0.33	-.11*	-.03	-.08	.07	-.02	.13**	-
8. Age	-	37.54	10.92	-.11*	-.18**	-.19**	.14**	-.14**	-.01	.07

Note. $N = 384$ (*pairwise* deletion; with 383 for sex and age). Sex (1 = male, 2 = female). Ethnicity (1 = White, 2 = minority). Education (1 = university, 2 = non-university). Employment (1 = employed, 2 = unemployed, student, and retired). *KR-20* reliability coefficients are presented in parentheses. * $p < .05$. ** $p < .01$ (two-tailed).

Table 4

Descriptive Statistics, Reliability Coefficients, and Correlations for Study 3

	M	*SD*	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. W1 CRT-WP	-9.16	6.78	(.83)												
2. W2 CRT-WP	-9.94	7.32	.72**	(.87)											
3. SRP-III Overall	2.01	0.55	.32**	.26**	(.90)										
4. TriPM Overall	2.25	0.41	.19**	.19**	.85**	(.91)									
5. TriPM Bold.	3.02	0.70	.07	.13	.40**	.61**	(.90)								
6. TriPM Mean.	1.83	0.54	.26**	.24**	.83**	.81**	.27**	(.89)							
7. TriPM Disin.	1.91	0.60	.06	.03	.55**	.62**	-.14*	.46**	(.90)						
8. BIDR IM	4.54	1.16	-.10	-.11	-.40**	-.34**	.18**	-.38**	-.54**	-					
9. CWBs	1.30	0.37	.18*	.13	.49**	.52**	.03	.45**	.59**	-.46**	(.95)				
10. Sex	1.40	0.49	-.03	-.03	-.28**	-.23**	-.22**	-.27**	.01	.03	-.14	-			
11. Ethnicity	1.21	0.41	.03	.04	-.00	-.02	-.01	-.03	.01	.02	-.02	.01	-		
12. Education	1.49	0.50	-.04	-.00	.12	.15	.03	.09	.19*	-.02	.08	-.03	-.12	-	
13. Employment	1.10	0.30	-.12	-.06	-.19**	-.16*	-.25**	-.07	.01	-.04	-.07	.05	.05	.06	-
14. Age	37.70	10.79	-.25**	-.22**	-.26**	-.19**	-.08	-.16*	-.15*	.20	-.11	.13	-.13	.02	-.01

Note. N = 198 (*pairwise* deletion; with 197 for BIDR and 196 for CWB and sex). Sex (1 = male, 2 = female). Education (1 = university, 2 = non-university). Ethnicity (1 = White, 2 = non-White). Employment (1 = employed, 2 = unemployed, student, and retired). W1 = Wave one. W2 = Wave two. IM = Impression management. *KR-20* and Cronbach's α reliability coefficients are presented in parentheses. * $p < .05$. ** $p < .01$ (two-tailed).

Table 5

Results of Independent Samples t-Tests for Study 5 Wave-1 Variables – Comparing Honest and Simulated Hiring Conditions

Variable	*Honest Mean*	*Hiring Mean*	*t*	*p*	*d*	95% CI	
						Lower	Upper
CRT-WP	-12.43	-12.83	.51	n.s.	0.07	-1.13	1.93
Openness	3.61	3.80	-2.16	*	-0.30	-.36	-.02
Conscientiousness	3.85	4.18	-4.47	***	-0.62	-.46	-.18
Agreeableness	3.17	3.57	-4.64	***	-0.62	-.57	-.23
Extraversion	3.23	3.73	-5.26	***	-0.70	-.69	-.32
Emotionality	3.22	2.79	4.56	***	0.60	.25	.63
Honesty-Humility	3.46	3.69	-2.66	**	-0.35	-.41	-.06
Machiavellianism	3.06	2.81	2.77	**	0.37	.07	.43
Narcissism	2.45	2.91	-5.46	***	-0.73	-.63	-.30
Psychopathy	2.11	1.89	2.70	**	0.36	.06	.38
TSD-Integrity	4.99	5.53	-4.61	***	-0.63	-.76	-.31

Note. $df = 217$. "Psychopathy" refers to the calculated Psychopathy scores using the Short Dark Triad without the two items which refer to criminal tendencies and sexual deviancy, which are beyond the scope of the CRT-WP. t = observed t-value, two-tailed. d = Cohen's d measure of effect size. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 6

Study 5 Descriptive Statistics, Reliability Coefficients, and Correlations for the CRT-WP and outcomes from Waves-2/3

	Wave	*M*	*SD*	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. CRT-WP	1	-12.64	5.72	(.82)										
2. BART Total Win	2	66.39	24.88	-.09	-									
3. BART Explosions	2	4.28	1.49	-.04	-.26**	-								
4. BART Avg. Pumps	2	10.42	3.35	-.07	.77**	.29**	-							
5. Matrix Cheating %	2	0.18	0.29	.10	-.06	-.06	.01	-						
6. Dictator Choice	2	4.71	2.33	-.24**	-.12	-.11	-.18*	-.10	-					
7. Cyberloafing Overall	3	2.07	0.82	.22*	-.12	-.08	-.11	.09	.07	(.95)				
8. CWB Overall	3	1.22	0.29	.20*	-.00	-.06	-.01	-.08	-.07	.50**	(.87)			
9. Age	1	38.49	10.78	-.04	-.04	.03	-.06	-.01	-.14	-.28**	-.09	-		
10. Sex	1	1.53	0.50	-.01	-.08	-.11	-.13	.22**	-.09	.03	-.03	-.09	-	
11. Ethnicity	1	1.10	0.30	.16*	.07	-.05	.10	-.04	-.08	.04	.10	-.18**	-.05	-
12. Education	1	1.40	0.49	-.01	-.01	-.06	-.08	.09	.03	-.05	-.16*	-.02	.09	-.06

Note. All correlations are based on pairwise deletion due to drastically different sample sizes between waves. $N = 219$ for Wave 1, 186-192 for Wave 2, and 154-155 for Wave 3 (although correlations between wave 2-3 were based on the available 139 responses). Descriptive statistics and reliability coefficients for each variable are given for the wave it was included in which has the largest *N*. BART Explosions refers to the number of balloons where participants reached the explosion point, across 10 trials. BART Avg. Pumps refers to the average number of pumps that participants used across all 10 trials. The average number of pumps for only unexploded balloons was also calculated but resulted in no significant differences. Matrix Cheating % refers to the percentage of unsolvable matrices that participants reported solving of the ones that they attempted during the time limit. Dictator choices ranged from 1 (most selfish) to 11 (least selfish). Sex (1 = male, 2 = female). Ethnicity (1 = White, 2 = minority). Education (1 = university, 2 = non-university). * $p < .05$. ** $p < .01$ (two-tailed).

Table 7

Study 6 Hierarchical Multiple Regression Analysis Predicting CWB (top) and Abuse (bottom)

Predictors	*b*	*SE*	*β*	R^2	ΔR^2	*Rescaled Relative Weight (%)*
				Overall CWB		
Model 1				.12**		
Constant	1.51	.14	-			
TSD-Integrity	-.05**	.02	-.16			
Cautiousness	-.05*	.04	-.16			
Irresponsibility	.07	.04	.12			
Model 2				.14**	.02*	
Constant	1.57	.14	-			-
TSD-Integrity	-.04*	.02	-.16			32.03
Cautiousness	-.05*	.02	-.15			27.74
Irresponsibility	.06	.04	.10			22.42
CRT-WP	.01*	.00	.13			17.81
				Abuse Factor		
Model 1				.10**		
Constant	1.63	.16	-			
TSD-Integrity	-.04	.02	-.13			
Cautiousness	-.08**	.02	-.22			
Irresponsibility	.03	.04	.04			
Model 2				.13**	.03**	
Constant	1.72	.02	-			-
TSD-Integrity	-.04	.02	-.12			19.50
Cautiousness	-.07**	.02	-.21			40.94
Irresponsibility	.01	.04	.02			11.44
CRT-WP	.01**	.00	.17			28.12

Note. $N = 292$. Relative weights bootstrapped 10,000 times. * $p < .05$. ** $p < .01$.