APPLIED PSYCHOLOGY

IAAP
THE INTERNATIONAL ASSOCIATION OF APPLIED PSYCHOLOGY

APPLIED PSYCHOLOGY, 2018, *0* (0), 1–30
doi: 10.1111/apps.12163

Dark Motives and Elective Use of Brainteaser Interview Questions

Scott Highhouse*

Bowling Green State University, USA

Christopher D. Nye

Michigan State University, USA

Don C. Zhang

Louisiana State University, USA

Brainteaser interview questions such as "Estimate how many windows are in New York" are just one example of aggressive interviewer behaviour that lacks evidence for validity and is unsettling to job applicants. This research attempts to shed light on the motives behind such behaviour by examining the relation between dark-side traits and the perceived appropriateness of brainteaser interview questions. A representative sample of working adults (n = 736) was presented with a list of interview questions that were either traditional (e.g., "Are you a good listener?"), behavioural (e.g., "Tell me about a time when you failed"), or brainteaser in nature. Results of a multiple regression, controlling for interviewing experience and sex, showed that narcissism and sadism explained the likelihood of using brainteasers in an interview. A subsequent bifactor analysis showed that these dark traits shared a callousness general factor. A second longitudinal study of employed adults with hiring experience demonstrated that perspective-taking partially mediated the relationship between this general factor and the perceived helpfulness and abusiveness of brainteaser interview questions. These results suggest that a callous indifference and a lack of perspective-taking may underlie abusive behaviour in the employment interview.

> On the hiring side, we found that brainteasers are a complete waste of time… They don't predict anything. They serve primarily to make the interviewer feel smart. (Laszlo Bock, senior vice president of people operations at Google, quoted in Bryant, 2013)

In 1921, applicants who answered an anonymous job advertisement posted by famed American inventor Thomas A. Edison were surprised to find that they needed to answer a series of brainteasers such as "Is Australia larger

* Address for correspondence: Scott Highhouse, Department of Psychology, Bowling Green State University, Bowling Green, Ohio 43403, USA. Email: shighho@bgsu.edu

than Greenland in area?" "If you were to inherit $1,000,000 within the next year, what would you do with it?" and "How is leather made?" The public reaction to Edison's questions was almost uniformly negative (Dennis, 1984). The *New York Times* published 23 articles about the Edison questions in the month of May alone. Prominent psychologist E.L. Thorndike wrote: "The demonstration of the possession of such diffuse information as Mr. Edison by his questionnaire requires of all candidates for positions means almost nothing at all" (*New York Times*, 1921).

History has repeated itself in the form of brainteaser interview questions (Wright, Sablynski, Manson, & Oshiro, 2012), colloquially known as *oddball* interview questions (www.glassdoor.com/Top-25-Oddball-Interview-Questions-LST_KQ0,34.htm). Companies such as Xerox, Microsoft, and Zappos are purported to ask applicants such questions as "Why is a tennis ball fuzzy?" "Why are manhole covers round?" and "How many cows are in Canada?" These oddball questions are not limited to employers in the United States, as several European employers have adopted the practice as well (https://www.glassdoor.com/blog/top-10-oddball-interview-questions-2016/). The difference between now and the time of Edison, however, is that the general public seems unfazed by such practices. Executive recruiters extol the virtues of brainteasers, noting that they enable interviewers to elicit atypical responses, and demonstrate creativity. One book, entitled *Are you smart enough to work at Google?* (Poundstone, 2012) presents a number of commonly used brainteasers in order to forearm people interviewing for technology jobs. As the opening quote from the Google executive shows, however, experience suggests that such questions are (1) of little utility, and (2) a way for interviewers to feel good about themselves. In describing the use of brainteasers by Microsoft interviewers, for instance, Poundstone (2003) observed:

> Microsoft is famous for its brutal refinement of the form, subjecting candidates to a withering barrage of brainteasers and "unanswerable" questions to separate out the logical, motivated, unflappable, innovative thinkers. (p. 18)

Although there is little empirical research on the topic, this approach to interviewing is likely to lack reliability and may be prone to interviewer biases. At the very least, existing studies suggest that brainteasers distress job applicants (Honer, Wright, & Sablynski, 2006; Wright et al., 2012).

A major objective of our research was to give insight into the motives behind the elective use of insensitive and potentially offensive hiring procedures. To do so, we focused on individual differences related to the perceived appropriateness of brainteaser interview questions—questions that are likely to be perceived by the applicant as insensitive or even as inappropriate. We

know of no way to estimate how many interviewers are using brainteaser questions. We suspect that these kinds of questions are used on an individual-by-individual basis, rather than on a company-wide basis. One can imagine a manager reading about brainteasers in the popular press and deciding to throw in a question or two in his or her own interview. The costs of using assessment tools that are aversive to applicants, however, include reduced job-pursuit intentions, negative word-of-mouth about the organisation, and possibly misleading performance on the assessment itself (see McCarthy, Bauer, Truxillo, Anderson, Costa, & Ahmed, 2017, for a review).

The purpose of the present study, therefore, was to examine whether individual differences reflecting dark motives are associated with the desire to use brainteasers in a hiring context. As we note below, this would allow us to better understand the degree to which use of these types of questions is motivated by general insensitivity, self-aggrandisement, or neither. A finding that elective use of such questions is at least partially associated with dark traits should result in efforts to identify appropriate interviewers, train interviewers in appropriate behaviour, or develop a culture of concern for the applicant (vs. the self) in the hiring process. More broadly, we expected to shed light on the motives behind the elective use of insensitive and potentially offensive hiring procedures.

BACKGROUND

The employment interview has been a staple of applicant assessment for over a century (Buckley, Norris, & Wiese, 2000). This is despite considerable evidence to suggest that traditional (i.e., unstandardised) interviews account for, at best, approximately 4 per cent of the variance in the explanation of job performance (Huffcutt & Arthur, 1994). This is especially troubling because the interview is often the only method of assessment in use, and it is expected to assess both ability and motivation to do the job (Highhouse, 2008). In the first comprehensive review of the employment interview, Wagner (1949) mused: "Yet one wonders why it remains so popular when, if there is any preponderance of evidence it seems to be against the interview as a valid method of selection and placement" (p. 33). More recent reviews have been more charitable to the interview as a selection method (e.g., Levashina, Hartwell, Morgeson, & Campion, 2014). Indeed, the interview can exhibit considerable utility over the use of tests of cognitive ability and conscientiousness if one standardises (a) the questions asked of applicants, and (b) the scoring of the applicants' answers (Cortina, Goldstein, Payne, Davison, & Gilliland, 2000). There is also considerable evidence to indicate that the use of behavioural and situational questions impacts validity (see Macan, 2009). Many employers, however, resist such restrictions on their freedom to size up candidates in their own way (Chapman & Zweig,

2005; Nowicki & Rosse, 2002). Structured interviews get scant attention in the popular management literature and employers often feel that structured interviews restrict their own ability to have a high-quality interaction with applicants (Roulin & Bangerter, 2012).

Fletcher (1992) asserted that the employment interview represents a power imbalance, in which the interviewer is typically in the dominant position. Where there is power, according to Fletcher, there is the potential for its abuse. Dipboye (1997) similarly suggested that there was a potential "dark side" to employment interviewing, wherein the decision-maker is provided with considerable power over the questions asked and qualities considered. According to Dipboye, the inherent ambiguity in the typical interview prevents close scrutiny and monitoring by outside parties. Ferris and King (1991) added that interviewers commonly hire on the basis of "fit". Such a vague criterion, according to the authors, invites manipulation from both sides of the relationship. Given the power imbalance created by the interview situation, it seems only natural that some interviewers will abuse their power.

One such example is the historical variant of the traditional interview, known as the "stress interview" (Freeman, Manson, Katzoff, & Pathman, 1942, p. 428). The stress interview involved supplementing the "characteristically bland type of interviewing" with stressful situations. In their example interview used for police officers, Freeman and colleagues followed initial small talk with the job candidate with:

> …astonishment that the candidate has done so poorly. The atmosphere of the interview quickly snaps from one of friendly interest to one of cold disdain. Interviewers now try to find ground for a rejection. Doubt is cast on the candidate's experience and character. Questions come so rapidly that the candidate cannot adjust to them; any break in the candidate's defense is used to his further disadvantage. (Freeman et al., 1942, pp. 432–433)

Freeman and colleagues acknowledged that such interview tactics would not work unless the candidate is highly motivated to receive a job offer. To the extent that tolerance for stress is an important characteristic of the job, stress interviews may often be job related. In such cases, the stress interviewer may not be indulging his or her own private need for power. Nevertheless, the interview serves as an inviting setting for abusive behaviour—at least from those inclined toward insensitivity, callousness, or cruelty.

DARK INTERVIEWER MOTIVES

Brainteaser questions are merely one variant on the "stress" interview (Corcodilos, 2014; Goldman, 2013). But they have the same goal of keeping the

candidate off-balance and examining his or her ability to tolerate high-pressure interview tactics. Wright et al. (2012) for example, showed people interviews that were either brainteasers or behavioural in nature, and asked them to rate the interview on, among other things, perceived validity and procedural justice. Although the two interviews did not differ in perceived validity, they differed by nearly three quarters of a standard deviation on procedural justice. The brainteasers were seen as less fair than the behavioural questions. Subjecting candidates to subjectively unfair and anxiety-producing questions, therefore, would represent a lack of empathy and concern for fair play. As such, *callous indifference* may underlie the use of brainteasers.

In addition, as suggested by the introductory quote by Google's senior vice president, brainteasers may be a way for certain people to maintain their inflated sense of self. Consider the famous Microsoft brainteaser question: "Why are manhole covers round?" The simple answer is that a round manhole cover cannot fall through, and circular covers do not need to be precisely aligned when placing them on the circular opening.[1] Many people, upon hearing this question the first time, would be baffled—and may even draw a blank in a stressful job interview. When provided the answer in advance, however, the question seems easy (see Wood, 1978). Interviewers using brainteasers are in the position of a game show host, in that they often know, in advance, the answer to the questions they are asking. People in such a position have been repeatedly shown to overestimate the likelihood that they would have known the answer had they not seen it previously (see Roese & Vohs, 2012, for a review). Roese and Vohs (2012) posited that, when people are mere bystanders to events such as elections or sports defeats, they claim greater foreseeability in order to protect and enhance their self-esteem. According to the authors, "…people enhance their self-esteem by taking credit for their apparent knowledgeability" (p. 416). Interviewers using brainteasers, therefore, may be motivated by the *desire to protect and enhance their self-esteem*. They may simply want to show others how smart they are.

DARK INTERVIEWER TRAITS

Dark Triad

There has been considerable interest in recent years in what Paulhus and Williams (2002) termed the "dark triad" of personality—especially as it applies to behaviour in the workplace (see O'Boyle, Forsyth, Banks,

[1] A number of other plausible answers can be found on the Wikipedia entry "manhole covers".

& McDaniel, 2012; Spain, Harms, & LeBreton, 2014). The dark triad includes Machiavellianism, narcissism, and psychopathy—all traits that have been linked with callousness, low humility, and interpersonal antagonism (Furnham, Richards, & Paulhus, 2013; Jones & Figueredo, 2013; Wai & Tiliopoulos, 2012). These dark traits are aversive, but remain within the normal range of functioning (i.e., they are considered sub-clinical). The three traits are theorised to develop from the desire to self-enhance and harm others (Spain et al., 2014), and thus seem especially relevant to the prediction of a preference for brainteaser interview questions. Given that the characteristics of narcissists (i.e., self-importance, need for admiration, lack of empathy) seem most associated with the motives thought to underlie abusive behaviour in the interview (i.e., callous indifference, desire to protect and enhance self-esteem), we expected that narcissism would be more related with desire to use brainteaser interview questions, compared with the other two dark-triad traits.

Hypothesis 1: Relative to the other dark-triad traits, self-reported narcissism will relate most strongly with perceived appropriateness of brainteaser interview questions.

Sadism

There has also been a recent call to extend the dark triad to include sadism (Buckels, Jones, & Paulhus, 2013; Chabrol, Van Leeuwen, Rogers, & Séjourné, 2009; Furnham et al., 2013). Sadism, characterised by cruelty and demeaning behaviour, correlates with dark triad traits, but exhibits unique relations with some outcomes (Buckels et al., 2013; Chabrol et al., 2009). Given their enjoyment of hurting others, we expected that people scoring high on a measure of sadism would be more disposed toward using abusive interview questions, and that this relation would hold after controlling for the other traits in the dark triad.

Hypothesis 2: Self-reported sadism will be positively related with perceived appropriateness of brainteaser interview questions, after controlling for dark-triad traits.

Other Individual Differences

Although we have argued that the use of brainteaser interview questions is likely to be associated with dark motives, we recognise that there exist more innocuous reasons for using the questions. Some people, for example,

may believe, despite the lack of evidence, that brainteasers effectively shed light on interviewee temperament, ability, and/or creativity. Others may not recognise that brainteaser questions could cause applicants distress. Accordingly, we examine individual differences in preference for intuition in hiring and in the ability to recognise and manage others' emotions. These were included for exploratory purposes.

OVERVIEW AND PILOT

The paradigm we used involved presenting lay people with a long list of interview questions of three types: traditional (e.g., "Why do you want to work here?"), behavioural (e.g., "Tell me about a time when you misjudged a person"), and brainteaser (e.g., "How would you explain what a chair was to an alien?"). Although we focused on a representative lay population, we also collected information on interview experience in order to determine whether this would be a significant covariate in the analyses. As the research was focused on psychological factors relating to attitudes toward brain teaser questions (regardless of interviewing experience), the research was largely exploratory, and our analyses required a very large sample of respondents, we believe that the lay sample used in this research is appropriate.

We collected the different types of questions from websites that present interview questions used by companies and collected by applicants (e.g., glassdoor.com). In order to establish that the brainteaser questions were indeed disturbing in nature for the typical applicant, we piloted the 49 items presented in Table 1 with a sample of workers (n = 100) on Amazon's Mechanical Turk (MTurk) crowdsourcing internet marketplace. Specifically, participants were instructed:

> We expect that most every adult has been interviewed for a job at some time in their lives. Below are a number of interview questions that have been used by various companies. Some interview questions are more DISCONCERTING than others. For each one, we ask that you indicate how THROWN OFF BALANCE you would be if you were asked each interview question.

Pilot participants responded how "thrown off balance" they would be on a scale ranging from 1 (not at all) to 5 (extremely).

Cronbach's alpha (.91) for the brainteaser items was quite high. Results showed that the disturbing nature of the interview questions was substantially higher for the brainteaser questions (M = 3.76; SD = .79) than for the behavioural questions (M = 1.82; SD = .79), differing by nearly two and one-half standard deviations (d = 2.48). The brainteaser mean differed from the mean for the traditional questions (M = 1.91; SD = .65) by more than two

TABLE 1
Interview Questions

Brainteaser*	Behavioural	Traditional
On a scale from one to ten, rate me as an interviewer (Kraft Heinz Co.)	Describe a decision you made that was a failure. What happened and why?	Why should we hire you?
What songs best describe your work ethic? (Dell)	If you could start your career over again, what would you do differently?	Do you consider yourself a leader?
What do you think about when you are alone in your car? (Gallup)	Tell me about a time that you participated in a team. What was your role?	What do you look for in a job?
Name three previous Nobel Prize winners. (BenefitsCONNECT)	Tell me about a time when you failed.	Can you work under pressure and deal with deadlines?
How many quarters do you need to reach the height of the Empire State Building? (Vistaprint)	Give an example of a goal you reached and tell me how you achieved it.	What do you know about our organisation?
Have you ever stolen a pen from work? (Jiffy Software)	Have you handled a difficult situation? How?	Are you a good listener?
Pick two celebrities to be your parents. (Urban Outfitters)	Give an example of an occasion when you used logic to solve a problem.	What led you to this point in your life?
Estimate how many windows are in New York. (Bain & Company)	Did you ever postpone making a decision? Why?	How long would it take you to make a meaningful contribution to our firm?
What is your favourite song? Perform it for us now. (Living Social)	Tell me about a time when you were faced with conflicting priorities. How did you determine the top priority?	If you could be anyone else, who would it be?
Why are manhole covers round? (Microsoft)	How do you balance life and work?	What do you love?
How would you determine the weight of a commercial airplane without a scale? (McKinsey & Company)	Tell me about a time when you misjudged a person.	How do you get along with older coworkers?
If you could be any animal on a carousel, what would you pick and why? (Journeys)	Tell me about a time that you helped someone.	Do you check voicemail and email while on vacation?

How would you explain what a chair is to an alien? (anonymous)	Have you handled a difficult situation with another department? How?	How successful do you think you've been so far?
Calculate the angle of two clock pointers when time is 11:50. (Bank of America)	What would you do if you disagree with a coworker?	Tell us why you want to work for us.
If Hollywood made a movie about your life, who would you like to see play the lead role as you? (First Student)	What do you do when your schedule is interrupted? Give an example of how you handle it.	What salary do you think you deserve?
If you were to get rid of one state in the country, what would it be and why? (Forrester)	Describe a situation when you took a risk professionally. What was the outcome?	
	Tell me about a creative solution you developed for a challenging situation or problem.	
	What is the most stressful situation you have handled and what was the outcome?	

*In parentheses are the companies associated with the interview question.

and one-half standard deviations (d = 2.58). The brainteaser questions were clearly seen as more disturbing.

STUDY 1

Participants

MTurk workers (n = 760) from the United States were paid 50 cents to participate. Twenty-four people were eliminated from our sample for failing items checking whether they were reading the instructions carefully (i.e., "In order to show that you are carefully reading the interview questions, please leave this item blank"). This resulted in the final sample of (n = 736) participants. Fifty-six per cent of the participants were male, and 77 per cent were Caucasian. Mean age was 34 years of age. We asked them to indicate how many interviews they have conducted in their career. The response options were 0 (52%), 1–10 (31%), 11–30 (9%), 31–50 (3%), and 51 or more (5%). Seventy-nine per cent worked outside of the home.

Predictor Constructs

All responses to the predictors were made on a five-point scale (*disagree strongly* to *agree strongly*). Items were intermixed.

Dark Triad. The dark-triad traits were assessed using the Short Dark Triad (SD3; Jones & Paulhus, 2014). The instrument consists of 27 items and three subscales: Machiavellianism (nine items), Psychopathy (nine items), and Narcissism (nine items). Example items from the Machiavellianism subscale include "It's not wise to tell your secrets", and "Avoid direct conflict with others because they may be useful in the future". Example items from the psychopathy subscale include "People who mess with me always regret it", and "People often say I'm out of control". Example items from the narcissism subscale include "I have been compared to famous people", and "I insist on getting the respect I deserve".

Sadism. Everyday sadism was assessed using the Short Sadistic Impulse Scale (SSIS; O'Meara, Davies, & Hammond, 2011). This is a 10-item, unidimensional, self-report inventory assessing sadistic impulse. Items include "I have hurt people for my own enjoyment", and "People would enjoy hurting others if they gave it a try".

Intuition in Hiring. Preference for using intuition-based hiring was assessed using six items from Lodato, Highhouse, and Brooks (2011). Items were designed to measure beliefs about relying on feelings when making hiring

decisions (e.g., "I believe it is important to rely on your intuition when hiring employees"), as well as beliefs about using evidence-based practices (e.g., "Hiring an employee is more of an art than a science").

Social Competence. Self-reported ability to recognise and manage others' emotions was assessed using seven items from the Values in Action (VIA) Social subscale (Peterson & Seligman, 2004). Items include "I am able to fit into any situation", and "I have the ability to make others feel interesting".

Procedure

Participants in this study were told the following:

> We are interested in your opinion of the appropriateness of interview questions that have been used in other organisations. By appropriate, we mean that YOU WOULD CONSIDER USING THE INTERVIEW QUESTION when hiring someone.

The traditional, behavioural, and brainteaser interview questions in Table 1 were presented in randomised order, and participants indicated the appropriateness of each on a seven-point scale ranging from "absolutely inappropriate" (1) to "absolutely appropriate" (7).

After completing the appropriateness ratings, participants completed demographic items, followed by the predictor scales.

RESULTS AND DISCUSSION

Table 2 shows the means and standard deviations for the primary study variables. The perceived appropriateness of brainteaser questions was significantly lower than the perceived appropriateness of behavioural questions, $t(660) = -76.35$; $p = .000$; $d = -3.71$, and lower than the perceived appropriateness of traditional questions, $t(660) = -71.13$; $p = .000$; $d = -3.10$. Behavioural questions were viewed as more appropriate than traditional questions, $t(660) = 22.46$; $p = .000$; $d = 0.75$. We found significant sex differences on appropriateness of the brainteaser questions. Males ($M = 2.44$; $SD = 1.05$) saw them as more appropriate than did females ($M = 1.99$; $SD = .82$), $t(660) = 6.17$; $p = .000$; $d = 0.48$). There were no significant sex differences on the behavioural and traditional questions.

Table 2 also shows that the internal consistency for the brainteasers was quite high, and the correlations with other types of questions were small to moderate. As expected, the dark traits showed a positive manifold, with the highest correlations between Machiavellianism and psychopathy, and between sadism and psychopathy.

TABLE 2
Relations among Appropriateness Ratings for Interview Question Types and with Individual Difference Variables (n = 736)

	M	SD	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
1. Brainteasers	2.24	.98	.91	.22**	.38**	.18**	.27**	.18**	.24**	.07†	−.06	−.04	.23**
2. Behavioural	5.50	.77		.89	.63**	.02	−.12**	.11**	−.18**	.11**	.21**	.13**	.03
3. Traditional	4.94	.74			.80	.03	−.01	.17**	−.02	.18**	.16**	−.01	.04
4. Machiavellian	2.85	.67				.71	.62**	.31**	.48**	.10**	−.01	−.15**	.27**
5. Psychopathic	1.97	.65					.79	.31**	.66**	.07†	−.03	−.13**	.39**
6. Narcissistic	2.70	.62						.76	.18**	.19**	.48**	.11**	.17**
7. Sadistic	1.51	.61							.85	.02	−.20**	−.16**	.22**
8. Intuitive	3.37	.61								.82	.10**	.09*	−.05
9. SocialComp.	3.63	.61									.74	.19**	−.04
10. Experience	1.79	1.1											−.05
11. Sex	1.56	.50											

Note:. Cronbach's alpha in the main diagonal. † $p < .10$; *$p < .05$; **$p < .01$ (two-tailed); Sex was coded Female = 1; Male = 2.

As expected, the dark triad traits all correlated positively with perceived appropriateness of brainteaser interview questions. In order to tease out the relative contribution of the dark traits to prediction, we ran a multiple regression, controlling for interviewing experience and participant sex. The results of this analysis are presented in Table 3. Consistent with Hypothesis 1, narcissism emerged as the primary predictor among the dark-triad traits. Consistent with Hypothesis 2, sadism was a significant predictor when all of the variables were in the model.

It is notable that sex remained a strong predictor of perceived appropriateness of brainteasers, even when the other traits were in the model. It appears that some other factors (e.g., relational values) may help explain sex differences on this variable.

Table 3 also shows that, when controlling for the effects of the other variables, preference for intuition-based hiring was positively related with perceived appropriateness of brainteaser questions. Social competence was negatively associated with perceived appropriateness of brainteaser questions. These findings seem logical and provide construct validity for our inferences from the brainteaser questions (Benson, 1998). It should be noted, however, that only 13 per cent of the variance in perceived appropriateness of brainteasers was explained by all of the individual-difference variables in

TABLE 3
Results of Multiple Regression Controlling for Sex and Interview Experience

	Behavioural Qs		*Brainteaser Qs*	
Predictor	β_1	β_2	β_1	β_2
Step 1:				
Interview Experience	.13**	.08†	.00	.02
Sex	.05	.10*	.24**	.17**
Step 2:				
Machiavellian		.13*		−.05
Narcissistic		.06		.14**
Psychopathic		−.14*		.09
Sadistic		−.14*		.15**
Socially Competent		.12*		−.09*
Intuitive		.08†		.09*
R^2	.02**	.10**	.06**	.13**
ΔR^2		.08**		.08**

$^{\dagger}p < .10$; $^{*}p < .05$; $^{**}p < .01$ (two-tailed)

the model. There is clearly room to explore other situational and attitudinal factors.

Although we did not hypothesise relations with perceived appropriateness of behavioural interview questions, we have included these analyses in Table 3 as well. Interestingly, interview experience was a significant covariate for this, but *not* for the perceived appropriateness of brainteaser questions. Machiavellianism and social competence were positively related with the perceived appropriateness of behavioural questions, and psychopathy and sadism were negatively related with this outcome.

Bifactor Analysis

Given that narcissism and sadism were the dominant dark traits in the regression analysis, we suspected that a general (i.e., callousness) factor might underlie the two measures (Hepper, Hart, & Sedikides, 2014; Jones & Figueredo, 2013; Ronningstam, 2014; Wai & Tiliopoulos, 2012). To examine this, we used a form of confirmatory factor analysis (CFA) known as bifactor analysis (Thissen, Wainer, & Wang, 1994). A bifactor solution constrains all of the items to load on one general factor while simultaneously loading on an orthogonal specific factor. In the current study, the general factor reflects the callousness factor that may underlie narcissism and sadism, whereas the specific factors reflect narcissism and sadism independently of the shared factor.

Responses to the narcissism and sadism items were non-normally distributed. As a result, we dichotomised responses such that those who indicated they disagreed with an item were coded 0 and all other responses were coded 1. Previous research has indicated that this is necessary to address problems with non-normally distributed variables that violate the assumptions of both Pearson correlations and the CFA estimation process (Nye & Drasgow, 2011). Table 4 shows that the bifactor model fit the dichotomous data quite well, and better than a single-factor model. Thus, these results suggest that a single general factor is not sufficient to model the data. However, the bifactor

TABLE 4
Fit Indices for the CFA Models with the General Factor

Model	*RMSEA*	*TLI*	*CFI*	*WRMR*
One-Factor	.10	.92	.93	2.33
Two-Factor	.04	.99	.99	1.23
Bifactor	.04	.99	.99	0.82

Note:. RMSEA = Root Mean Square Error of Approximation; TLI = Tucker-Lewis Index; CFI = Comparative Fit Index; WRMR = Weighted Root Mean Square Residual.

model fit only slightly better than the two-factor model, as evidenced by the lower weighted root mean square residual (WRMR). Thus, there is some evidence that modelling the two dark constructs (i.e., narcissism and sadism) alone is insufficient to explain the data in this sample and that both general and domain-specific factors may exist. In order to justify the existence of both general and specific dimensions, however, it is important to demonstrate that the general factor predicts the outcomes of this study above and beyond the residual factors corresponding to narcissism and sadism. We do this by estimating a path model next. Table 5 shows the factor loadings for the bifactor model.

TABLE 5
Factor Loadings in the Bifactor Model of Narcissism and Sadism

Item	*General Factor*	*Narcissism*	*Sadism*
People see me as a natural leader	.37	.56*	
I hate being the centre of attention^{r}	.46*	.28	
Many group activities tend to be dull without me	.64*	.21	
I know that I am special because everyone keeps telling me so	.64*	.29	
I like to get acquainted with important people	.41*	.37	
I feel embarrassed if someone compliments me^{r}	.31	.44*	
I have been compared to famous people	.46*	.23	
I am an average person^{r}	.28	.29	
I insist on getting the respect I deserve	.56*	.03	
I enjoy seeing people hurt	.44*		.80*
Hurting people would be exciting	.49*		.80*
I have hurt people for my own enjoyment	.52*		.81*
People would enjoy hurting others if they gave it a try	.45*		.80*
I have fantasies which involve hurting people	.42*		.83*
I have hurt people because I could	.46*		.79*
I wouldn't intentionally hurt anyone^{r}	.33*		.66*
I have humiliated others to keep them in line	.50*		.67*
Sometimes I get so angry I want to hurt people	.33		.81*

^{r}Reverse scored; *$p < .05$ (two-tailed)

We conducted a path model in which the latent factors from the bifactor analysis are used to predict the perceived appropriateness of brainteaser interview questions. Again, because the specific factors are orthogonal to the general factor and represent the shared variance that is left over after accounting for the general factor, correlations for the domain-specific dimensions illustrate the incremental validity of these narrow factors for predicting perceived appropriateness. We found that the general factor was a significant predictor of perceived appropriateness of brainteaser interview questions ($\gamma = .34$; $p = .00$). After parcelling out the variance related to the general factor in these items, narcissism was not significantly related to perceived appropriateness, $\gamma = .09$; $p = .108$, nor was sadism, $\gamma = .11$; $p = .074$. This suggests that narcissism and sadism do not provide significant incremental validity beyond the general factor for predicting the perceived appropriateness of brainteaser interview questions.

The subsequent bifactor analyses seem to support the callous-indifference explanation for why some interviewers prefer to use brainteaser questions. Study 2 directly examines the mechanisms thought to mediate the relation between dark traits and choice of using brainteaser questions in a sample of experienced interviewers.

STUDY 2

Study 2 was designed to further confirm the existence of a general factor, as well as to explore the possible mediating role of traits related to callousness (i.e., empathy) and self-enhancement (i.e., self-deception), in the relation between dark motives and the disposition toward brainteaser interview questions. The study employs a three-wave data collection wherein MTurk workers completed narcissism and sadism items as part of a mass data collection at Time 1. We identified those workers who had previous hiring experience and followed up with them over 12 months later (Time 2), with measures of empathy (i.e., empathic concern and perspective taking) and self-enhancement. We also administered, at the end of the survey, the interview items as presented in Study 1. This was done to replicate relations of perceived appropriateness of brainteasers with narcissism and sadism. At Time 3, we presented participants with questions directly assessing whether they believed that brainteaser questions are abusive or useful in the hiring process. Figure 1 shows the theoretical model that we tested in this study.

Participants and Procedure

Wave 1 participants were 1,428 MTurk workers who earned $2.50 for completing a long screening survey containing over 25 brief measures of

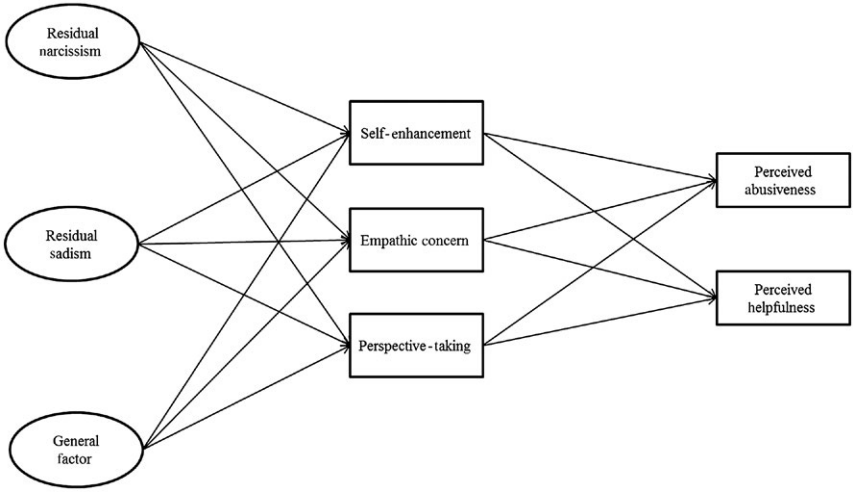

FIGURE 1. Structure of the partial mediation model examined in Study 2. *Note:* Direct paths from the latent factors in the bifactor model to Perceived Abusiveness and Perceived Helpfulness were also estimated but are not illustrated here to increase the clarity of the figure.

organisational constructs administered in randomised order. The narcissism and sadism constructs assessed for this study represented only a small portion of the items in the overall survey, which included enquiries about various job experiences and working conditions—including an item about their experience hiring job candidates.

Wave 2 participants were sampled over 12 months after the initial screening survey. We surveyed only those (n = 496) who provided an affirmative response to the item: "Have you ever made the final decision about whether or not a job applicant should be hired for a job?" Given the large number of items in the initial screening survey, the participants would have no reason to anticipate that this particular question would be used as a screening item in subsequent data collections. Thus, we do not believe that participants would be motivated to falsify reports of their hiring experience. Participants were paid $1.00 for completing the items, along with the opportunity to earn an additional $1.00 for participating in the survey at Time 3. We received responses from 213 people. The sample was 51 per cent female, and 78 per cent had conducted a job interview within the last year.

Wave 3 participants (n = 167) were presented two weeks later with items concerning their views about the potential abusiveness and usefulness of brainteaser interview questions. Specifically, participants were told:

> A currently popular practice in employment interviewing is the use of brainteaser questions. Companies such as Xerox, Microsoft, and Zappos are said to ask applicants such questions as: "Why is a tennis ball fuzzy?" "Why are manhole covers round?" "How many cows are in Canada?" We are interested in your thoughts about the use of brainteaser questions in an employment interview.

This was followed by a series of items inquiring about the perceived abusiveness and usefulness of such items for hiring.

Measures

Dark Traits. Narcissism and Sadism were measured at Wave 1 using the same instruments from Study 1.

Mediators. In order to more directly assess the proposed mechanisms behind the dark trait relation with attitudes about brainteaser interview questions, we used measures of callous indifference (i.e., low empathy) and self-enhancement. Callousness was assessed with the *perspective-taking* and *empathic concern* scales developed by Davis (1983). Perspective taking was assessed with seven items about one's ability to see things from another's point of view (e.g., "I sometimes try to understand my friends better by imagining how things look from their perspective"). Empathic concern was assessed with seven items about one's compassion for the plight of others (e.g., "When I see someone being taken advantage of, I feel kind of protective toward them"). Self-enhancement was assessed with 20 items from the Paulhus (1991) self-deceptive enhancement scale. The scale assesses unrealistically positive self-evaluations (e.g., "I am a completely rational person").

Brainteaser Attitudes. Wave 3 participants were presented with items assessing the *perceived abusiveness* and *perceived helpfulness* of brainteaser interview questions. The items followed the instructions presented in the above procedure and are presented in the Appendix. Responses were made on a five-point scale (*disagree strongly* to *agree strongly*).

RESULTS AND DISCUSSION

Table 6 presents the means, standard deviations, and intercorrelations for variables measured in Study 2. As in Study 1, narcissism and sadism were significant predictors of the perceived appropriateness of brainteaser questions. Table 6 also shows that these dark traits were not, however, correlated with perceived abusiveness and helpfulness as measured in the third wave. Note that the perceived abusiveness of the brainteaser questions was, however, significantly correlated with perspective-taking tendency. There was no relation with empathic concern or self-enhancement.

TABLE 6
Intercorrelations among Variables in Study 2 by Wave of Data Collection

n		M	SD	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
	1. Narcissism	2.75	.63	.78							
1,428	2. Sadism	1.50	.64	.19*	.92						
	3. Self-Enhance	3.16	.50	.12	−.07	.73					
	4. Empathic Concern	3.71	.90	.01	−.36*	−.03	.93				
496	5. Perspective Taking	3.69	.79	.08	−.28*	.12†	.57*	.89			
	6. Perceived Appropriate	2.29	1.08	.19*	.21*	−.06	−.10	−.11	.93		
	7. Perceived Abusive	3.17	.56	−.08	−.08	−.01	.09	.23*	−.19*	.85	
167	8. Perceived Helpful	3.07	.92	.15	−.03	.03	−.03	−.14	.23**	−.68*	.89

Note: Cronbach's alpha in the main diagonal. † $p < .10$; *$p < .05$; **$p < .01$ (two-tailed).

Consistent with the analyses presented in Study 1 and with previous research (Nye & Drasgow, 2011), we again recoded responses to the narcissism and sadism items into dichotomous agree/disagree variables to address problems with non-normality and its effects on the CFA estimation process. As in Study 1, the bifactor model fit these data quite well (RMSEA = .03; TLI = 1.00, CFI = 1.00, WRMR = 1.19), and better than a single-factor model or the two-factor model. Recall that these results suggest that neither a single general factor nor the two dark constructs (i.e., narcissism and sadism) alone are sufficient to explain the data in the sample. Instead, both general and domain-specific factors exist.

We next tested the path model shown in Figure 1, in which the latent factors from the bifactor analysis are included along with the mediators and the outcome variables measured at wave 3. As noted above, some attrition occurred between waves 1 and 3 due to the longitudinal nature of the study. Consequently, fewer participants responded to the survey in wave 3 than in wave 1. Therefore, to address this issue, we estimated the model in Figure 1 in MPLUS version 7.31 using full information maximum likelihood (FIML) estimation. FIML estimation uses all of the available data in a study (rather than eliminating cases if they do not respond to some of the variables) and has been shown to be one of the best approaches to handling missing data in survey research (Newman, 2003, 2009).

The path coefficients for the model in Figure 1 are shown in Table 7. This table shows that the general factor was a significant predictor of the perceived helpfulness of brainteaser interview questions for hiring quality job candidates. Perspective taking is positively associated with the perceived abusiveness of brainteaser questioning, and is negatively associated with the perceived helpfulness of brainteaser questions. In other words, interviewers with a higher ability to see a situation from another person's point of view were more likely to believe that using brainteaser questions in the interview process is abusive and inappropriate and less likely to view these questions as helpful for making employment decisions. Figure 2 shows only the significant paths from the model. This provides a more transparent method for interpreting the mediation results. First, this confirms our findings from Study 1—narcissism cancels out after accounting for the general factor and has no significant paths to any of the variables. In addition, residual sadism clearly shows a lack of empathy but not necessarily a desire for self-enhancement.

Figure 2 shows that perspective taking is a significant mediator for residual sadism and a partial mediator of the general factor. As described above, interviewers who can see another's point of view are more likely to see brainteasers as abusive and unhelpful. Note that the general factor is significantly related to perceived helpfulness, after controlling for this relation. In other words, people who are callous are more likely to view brainteaser

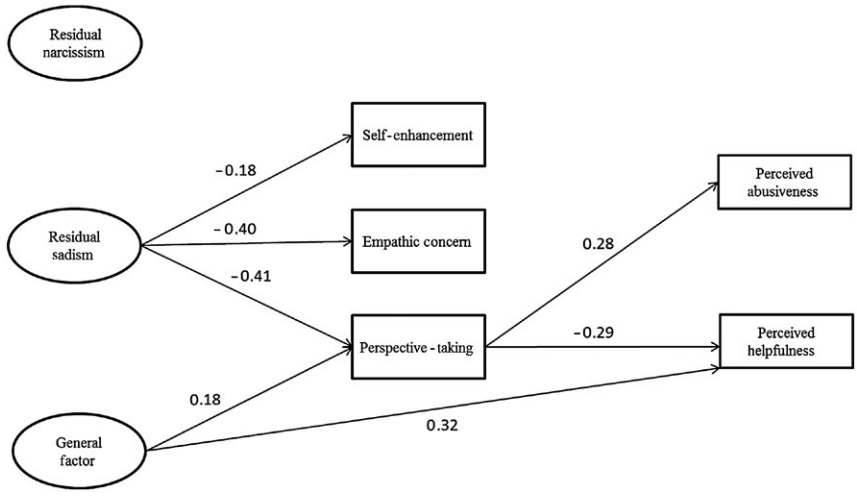

FIGURE 2. Significant paths in the partial mediation model examined in Study 2.

questions as helpful after taking into account perspective taking. We believe this makes sense, as callous people are more likely to view the ends as justifying the means. That is, callous people are likely to focus on the benefits, rather than the costs, of taking advantage of others.

GENERAL DISCUSSION

We began this article with a quotation by a Google human resources executive whose internal analyses showed that brainteaser questions had no validity—suggesting that there may be insincere motives behind the use of such questions in the hiring process. It seemed reasonable, therefore, to examine the role of dark personality traits in the elective use of such questions. The potential validity of brainteasers remains to be seen, the focus of our research was on elective use of questions that have been shown to be distressing to interviewees. Although the research was largely exploratory, we expected that perceived appropriateness of brainteasers among interviewers would be associated with traits reflecting callousness and self-enhancement.

Study 1 showed that people who would consider using brainteaser interview questions when hiring someone are more narcissistic, more sadistic, less socially competent, and believe more strongly in the power of intuition in the hiring process. Bifactor analyses showed that the hypothesised callousness general factor explained preferences for brainteasers, and the specific factors for narcissism and sadism provided no incremental prediction. Study

2 confirmed these findings and showed that the relation between the general factor and perceived helpfulness was partially mediated by perspective taking. The indirect effect of the general factor on the perceived helpfulness of brainteaser questions, when mediated by perspective-taking, was negative. However, there was also a direct relation between the general factor and perceived helpfulness, suggesting that interviewers who score high on the general factor are more likely to view brainteaser questions as helpful. This direct relation was one of the strongest in the final model. Thus, it appears that the general factor (i.e., callousness) is related to beliefs about the perceived helpfulness of these questions, and presumably the use of these questions as well, but that these effects can be mitigated by perspective taking.

It is interesting that the relation between the general factor and perspective-taking was significant and in the positive direction. Although the direction of this finding is counterintuitive, interpreting this relation is complicated by the use of the bifactor model because the general factor represents the shared variance among all of the items, and the specific factors represent the unique variance that is left over after accounting for the shared variance. To explore this relation further, we also estimated a model in which only the general factor was related to the three mediators (i.e., self-enhancement, empathic concern, and perspective-taking). In this model, the general callousness factor was negatively related to perspective-taking (−.33). This suggests that the positive relation between the general factor and perspective taking shown in Figure 2 is only observed after controlling for the negative relation between residual sadism and perspective-taking (the relation between residual narcissism and perspective-taking was not significant). In other words, the variance attributed to the sadism items seems to be negatively related to perspective taking, even though overall callousness may not be. This conclusion is supported by the zero-order correlations reported in Table 6, which show a significant negative relation between sadism and perspective taking and a positive (though non-significant) relation between narcissism and perspective taking. This finding is also consistent with previous research indicating that general callousness does not tend to have a substantial effect on one's ability to see a situation from another person's point of view but does negatively influence the experience of affective empathy (Oliver, Neufeld, Dziobek, & Mitchell, 2016). In the present study, this phenomenon was manifest in the positive relation between the general callousness factor and perspective taking but the negative relation with the perceived helpfulness of the abusive questions. In other words, after accounting for sadistic impulses, those scoring high on the general factor may have the ability to understand how another person feels about receiving brainteaser questions but may not care about his or her negative reactions if these questions are viewed as useful in some way.

It is also notable that the relations between residual factors and the endogenous variables in the path model are stronger than the zero-order bivariate correlations. These findings suggest a possible cooperative suppressor effect of the general factor (Paulhus, Robins, Trzeisniewsky, & Tracy, 2004), where the association between the lower order facets of the bifactor model and the criterion is strengthened when their covariance (general factor) is removed. The opposing patterns of associations also suggest that the psychological measure may be tapping distinct latent constructs. Similar suppressor effects have been observed in measures of psychopathy (Patrick, Hicks, Nichol, & Krueger, 2008). Certainly, these findings need to be replicated in subsequent studies and more research is needed to examine the underlying structure of the SD3.

We believe that more attention needs to be paid to interviewer individual differences as they relate to outcomes other than validity. That is, studies with enough interviewers to interviewees suggest that differences in validity from interviewer to interviewer are merely due to sampling error (Pulakos, Schmitt, Whitney, & Smith, 1996). This does not mean, however, that interviewers do not differ in behaviour during the interview, as insensitive interviewer behaviour may result in withdrawal from the selection process as well as damaged interviewee self-efficacy and subsequent job search behaviours (Ali, Ryan, Lyons, Ehrhart, & Wessel, 2016).

One of the few studies to examine the behavioural patterns that are characteristic of successful interviews (i.e., positive evaluations from the interviewer) showed that interviewers react differently to different interviewee behaviour (Tullar, 1989). Tullar videotaped and coded 28 campus interviews by seven corporate recruiters. Interviewer behaviours were coded as either dominant or structuring. Relevant to our study were dominant behaviours, which included things such as non-supportive responses, demanding answers to questions, changing conversations topics, and talking over the interviewee. Tullar found that candidates scored better when they behaved submissively than when the interviewer behaved dominantly. We might expect that narcissists would be especially influenced by interviewer submissiveness. Spain et al. (2014) also make a number of predictions about the influence of dark traits on *interviewee* behaviours. With this in mind, examining the interaction of dark traits in both interviewees and interviewers would be particularly interesting.

Another line of research, aside from replicating our findings in the model, would be to further pursue an understanding of sex differences in perceived appropriateness of brainteasers. It appears that men are more likely to use brainteaser questions because they are also more callous, on average. This seems to suggest that rogue use of such questions may occur more often in male-dominated professions (e.g., those in Silicon Valley). Given that we

explained only 13 per cent of the variance in perceived appropriateness, future research should examine other dispositional and attitudinal correlates.

It is also important to examine the aspects of brainteasers that make them more or less unfair in the eyes of interviewers and applicants. As one reviewer observed, certain brainteasers (e.g., manhole covers) may seem more fair than other, more peculiar, ones. It is notable, however, that we observed a high degree of internal consistency for the brainteaser questions as a group (.91). Moreover, the brainteaser item-total correlation for the round manhole covers (.69) is comparable to that for choosing to be an animal on the carousel (.73) and for picking two celebrities to be one's parents (.66). Although we agree that different questions may differ in the degree to which they might be disturbing to the interviewee, we would argue that this is similar to different levels of item difficulty found in any set of items.

Limitations and Advantages

Although the limitations of MTurk as a sample of convenience are often overstated (cf., Landers & Behrend, 2015), a field study with hiring managers across several companies, or a longitudinal one looking at interviewers given wide latitude in decision-making, would certainly supplement the findings observed here. A sample of professional HR managers or interviewers would ideally include people who conduct only a few interviews a year as well as those who conduct daily interviews.

We are not really sure who fits the category of "interviewers" as many people are involved in interviewing throughout their careers. We could envision a grocery store manager becoming enamoured with the idea of throwing brainteaser questions into an interview for baggers or cashiers. Indeed, we commonly read articles about interview questions such as "Is your mother a prostitute?" used to evaluate the suitability of prospective NFL players (James, 2010). Users may include sophisticated assessors, human-resource generalists, or managers provided with little or no training in assessment (König et al., 2011; MacIver et al., 2014). Our results in Study 1 showed that the number of interviews conducted was not a significant covariate in our analysis of appropriateness of brainteasers. In other words, experienced interviewers were no more likely to view these interview questions as inappropriate than less experienced lay people. It was, however, a significant covariate in our analysis of appropriateness of behavioural questions. This suggests that experienced interviewers can recognise good questions when they see them, but they do not differ from inexperienced interviewers in their views about brainteaser questions. Study 2 replicated the findings from Study 1, using only people with hiring experience. The primary goal of our study was to identify the types of people who might be inclined to use insensitive

interviewing tactics. Brainteasers appear to be one way for people to exploit the power imbalance in the interview.

Use of the bifactor model is somewhat new to organisational research. There are, however, several advantages of the bifactor solution over a higher-order model (Chen, West, & Sousa, 2006). One of these advantages is that the bifactor model can be used to investigate the role of specific factors that are independent of the general construct. In this regard, the specific factors are comprised of the residual variance that is not accounted for by the general dimension. In our case, the general factor constituted a callousness dimension and most would agree that narcissism and sadism are defined by more than just a lack of empathy (Spain et al., 2014). Therefore, the bifactor model is most appropriate for modelling constructs like these. Although similar information is provided in higher-order models, it can only be found in non-significant disturbance terms that are easily overlooked. Another advantage of the bifactor model is that the relations between the indicators and the specific factors are explicitly tested and one can determine the amount of variance attributable to both general and specific dimensions. A similar test is not available in higher-order models where specific factors are modelled solely as the disturbance terms in the first-order latent constructs. A final advantage is that the bifactor model can be used to assess the effects of the specific factors over and above the general factor. To do this in a second-order model may require non-standard structural equations models that are not easily implemented in commonly used statistical software.

CONCLUSIONS

The present study provides evidence of individual differences in the perceived appropriateness of brainteaser questions. These results are important because these types of questions can potentially create negative perceptions of fairness (Wright et al., 2012) and generally negative applicant outcomes (Ali et al., 2016). Understanding the impetus behind using these potentially detrimental questions, therefore, can help organisations to curb their use during the hiring process and identify those who may be more effective recruiters and interviewers. Employers might also consider limiting individual latitude in interview questioning, training interviewers on the potential impact of brainteaser questions, and establishing an organisational culture that discourages insensitive behaviour toward applicants. Based on the results presented here, it appears that callous interviewers who lack perspective taking ability will be more likely to use inappropriate or offensive hiring tactics.

REFERENCES

Ali, A. A., Ryan, A. M., Lyons, B. J., Ehrhart, M. G., & Wessel, J. L. (2016). The long road to employment: Incivility experienced by job seekers. *Journal of Applied Psychology*, *101*(3), 333.

Benson, J. (1998). Developing a strong program of construct validation: A test anxiety example. *Educational Measurement: Issues and Practice*, *17*, 10–17.

Bryant, A. (2013). In head-hunting, big data may not be such a big deal. *New York Times*, 19 June.

Buckels, E. E., Jones, D. N., & Paulhus, D. L. (2013). Behavioral confirmation of everyday sadism. *Psychological Science*, *21*, 790–792. https://doi.org/10.1177/0956797610369494.

Buckley, M. R., Norris, C. A., & Wiese, D. S. (2000). A brief history of the selection interview: May the next 100 years be more fruitful. *Journal of Management History*, *6*(3), 113–126.

Chabrol, H., Van Leeuwen, N., Rodgers, R., & Séjourné, N. (2009). Contributions of psychopathic, narcissistic, Machiavellian, and sadistic personality traits to juvenile delinquency. *Personality and Individual Differences*, *47*(7), 734–739.

Chapman, D. S., & Zweig, D. I. (2005). Developing a nomological network for interview structure: Antecedents and consequences of the structured selection interview. *Personnel Psychology*, *58*(3), 673–702.

Chen, F.F., West, S.G. & Sousa, K.H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate Behavioral Research*, *41*(2), 189–225.

Corcodilos, N. (2014). How to walk out of an abusive job interview. *PBS News Hour*, January. Retrieved 20 October 2014 from: https://www.pbs.org/newshour/making-sense/ask-the-headhunter-how-to-take-an-abusive-job-interview-and-shove-it/.

Cortina, J. M., Goldstein, N. B., Payne, S. C., Davison, H. K., & Gilliland, S. W. (2000). The incremental validity of interview scores over and above cognitive ability and conscientiousness scores. *Personnel Psychology*, *53*(2), 325–351.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113.

Dennis, P. M. (1984). The Edison questionnaire. *Journal of the History of the Behavioral Sciences*, *20*, 23–37.

Dipboye, R. L. (1997). Structured selection interviews: Why do they work? Why are they underutilized? (pp. 455–473). In N. Anderson, & P. Herriot (Eds.), International handbook of selection and assessment. Chichester: John Wiley and Sons.

Ferris, G. R., & King, T. R. (1991). Politics in human resources decisions: A walk on the dark side. *Organizational Dynamics*, *20*, 59–71.

Fletcher, C. (1992). Ethical issues in the selection interview. *Journal of Business Ethics*, *11*(5–6), 361–367.

Freeman, G. L., Manson, G. E., Katzoff, E. T., & Pathman, J. H. (1942). The stress interview. *Journal of Abnormal and Social Psychology*, *37*(4), 427.

Furnham, A., Richards, S. C., & Paulhus, D. L. (2013). The Dark Triad of personality: A 10 year review. *Social and Personality Psychology Compass*, *7*(3), 199–216.

Goldman, A. (2013). 8 tips for acing a tough job interview. *Forbes*, April. Retrieved 20 October 2014 from: 4https://www.forbes.com/sites/learnvest/2013/04/24/8-tips-for-acing-a-tough-job-interview/.

Hepper, E. G., Hart, C. M., & Sedikides, C. (2014). Moving Narcissus: Can narcissists be empathic? *Personality and Social Psychology Bulletin*, *40*(9), 1079–1091. https://doi.org/10.1177/0146167214535812.

Highhouse, S. (2008). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology*, *1*(3), 333–342.

Honer, J., Wright, C. W., & Sablynski, C. J. (2006). Puzzle interviews: What are they and what do they measure. *Applied HRM Research*, *11*, 79–96.

Huffcutt, A. I., & Arthur, W. (1994). Hunter and Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology*, *79*(2), 184–190.

James, F. (2010). NFL execs: "Is your mother a prostitute?" to player draws union ire. *NPR*, 28 April. Retrieved 18 March 2016 from: https://www.npr.org/sections/thetwo-way/2010/04/nfl_execs_is_your_mother_a_pro.html.

Jones, D. N., & Figueredo, A. J. (2013). The core of darkness: Uncovering the heart of the Dark Triad. *European Journal of Personality*, *27*(6), 521–531.

Jones, D. N., & Paulhus, D. L. (2014). Introducing the Short Dark Triad (SD3): A brief measure of dark personality traits. *Assessment*, *21*(1), 28–41.

König, C.J., Jöri, E. & Knüsel, P. (2011). The amazing diversity of thought: A qualitative study on how human resource practitioners perceive selection procedures. *Journal of Business and Psychology*, *26*(4), 437–452.

Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: Arbitrary distinctions between organizational, Mechanical Turk, and other convenience samples. *Industrial and Organizational Psychology*, *8*(02), 142–164.

Levashina, J., Hartwell, C. J., Morgeson, F. P., & Campion, M. A. (2014). The structured employment interview: Narrative and quantitative review of the research literature. *Personnel Psychology*, *67*(1), 241–293.

Lodato, M. A., Highhouse, S., & Brooks, M. E. (2011). Predicting professional preferences for intuition-based hiring. *Journal of Managerial Psychology*, *26*(5), 352–365.

Macan, T. (2009). The employment interview: A review of current studies and directions for future research. *Human Resource Management Review*, *19*(3), 203–218.

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.

New York Times (1921). More slams at Edison: Experts pronounce his questions only one-tenth effective in gaining their purpose. *New York Times*, 22 May, 83.

MacIver, R., Anderson, N., Costa, A.-C., & Evers, A. (2014). Validity of Interpretation: A user validity perspective beyond the test score. *International Journal of Selection and Assessment*, *22*(2), 149–164.

Newman, D. A. (2003). Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods*, *6*, 328–362.

Newman, D. A. (2009). Missing data techniques and low response rates: The role of systematic nonresponse parameters. In C. E. Lance, & R. J. Vandenberg (Eds.), Statistical and methodological myths and urban legends (pp. 7–36). New York: Routledge.

Nowicki, M. D., & Rosse, J. G. (2002). Managers' views of how to hire: Building bridges between science and practice. *Journal of Business and Psychology*, *17*(2), 157–170.

Nye, C. D., & Drasgow, F. (2011). Assessing goodness of fit: Simple rules of thumb simply don't work. *Organizational Research Methods*, *14*, 548–570.

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*, 557.

O'Meara, A., Davies, J., & Hammond, S. (2011). The psychometric properties and utility of the Short Sadistic Impulse Scale (SSIS). *Psychological Assessment*, *23*(2), 523.

Oliver, L. D., Neufeld, R. W., Dziobek, I., & Mitchell, D. G. (2016). Distinguishing the relationship between different aspects of empathic responding as a function of psychopathic, autistic, and anxious traits. *Personality and Individual Differences*, *99*, 81–88.

Patrick, C. J., Hicks, B. M., Nichol, P. E., & Krueger, R. F. (2007). A bifactor approach to modeling the structure of the Psychopathy Checklist-Revised. *Journal of Personality Disorders*, *21*(2), 118–141.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), Measures of personality and social psychological attitudes (pp. 17–59). San Diego, CA: Academic Press.

Paulhus, D. L., Robins, R. W., Trzesniewski, K., & Tracy, J. L. (2004). Two replicable suppressor situations in personality research. *Multivariate Behavioral Research*, *39*(2), 303–328.

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality*, *36*(6), 556–563.

Peterson, C., & Seligman, M. E. P. (2004). Character strengths and virtues. Oxford: Oxford University Press.

Poundstone, W. (2003). Beware the interview inquisition. *Harvard Business Review*, *81*(5), 18–19.

Poundstone, W. (2012). Are you smart enough to work at Google?. New York, NY: Little, Brown and Company.

Pulakos, E. D., Schmitt, N., Whitney, D., & Smith, M. (1996). Individual differences in interviewer ratings: The impact of standardization, consensus discussion, and sampling error on the validity of a structured interview. *Personnel Psychology*, *49*(1), 85–102.

Roese, N. J., & Vohs, K. D. (2012). Hindsight bias. *Perspectives on Psychological Science*, *7*(5), 411–426.

Ronningstam, E. (2014). Beyond the diagnostic traits: A collaborative exploratory diagnostic process for dimensions and underpinnings of narcissistic personality disorder. *Personality Disorders: Theory, Research, and Treatment*, *5*(4), 434.

Roulin, N., & Bangerter, A. (2012). Understanding the academic–practitioner gap for structured interviews: "Behavioral" interviews diffuse, "structured" interviews do not. *International Journal of Selection and Assessment*, *20*(2), 149–158.

Spain, S. M., Harms, P., & LeBreton, J. M. (2014). The dark side of personality at work. *Journal of Organizational Behavior*, *35*, S41–S60.

Thissen, D., Wainer, H., & Wang, X. B. (1994). Are tests comprising both multiple-choice and free-response items necessarily less unidimensional than multiple-choice tests? An analysis of two tests. *Journal of Educational Measurement*, *31*(2), 113–123.

Tullar, W.L. (1989). Relational control in the employment interview. *Journal of Applied Psychology*, *74*(6), 971.

Wagner, R. (1949). The employment interview: A critical summary. *Personnel Psychology*, *2*(1), 17–46.

Wai, M., & Tiliopoulos, N. (2012). The affective and cognitive empathic nature of the dark triad of personality. *Personality and Individual Differences*, *52*(7), 794–799.

Wood, G. (1978). The knew-it-all-along effect. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(2), 345.

Wright, C. W., Sablynski, C. J., Manson, T. M., & Oshiro, S. (2012). Why are manhole covers round? A laboratory study of reactions to puzzle interviews. *Journal of Applied Social Psychology*, *42*, 2834–2857.

APPENDIX

A currently popular practice in employment interviewing is the use of brainteaser questions. Companies such as Xerox, Microsoft, and Zappos are said to ask applicants such questions as:

"Why is a tennis ball fuzzy?"
"Why are manhole covers round?"
"How many cows are in Canada?"

We are interested in your thoughts about the use of brainteaser questions in an employment interview.

1. These types of interview questions are useful for assessing applicant creativity. [PH]
2. These questions are unnecessarily anxiety-provoking for the applicant. [PA]

3. These questions are useful for assessing an applicant's tolerance for stress. [PH]
4. It is unfair to ask an applicant one of these types of interview questions. [PA]
5. These interview questions assess applicants' ability to "think on their feet". [PH]
6. This type of interview question is useful for assessing one's adaptability. [PH]
7. These types of questions are abusive to the applicant. [PA]