Screening Out Neurodiversity

Jada Wiggleton-Little, Craig Callender

Jada Wiggleton-Little and Craig Callender

Screening Out Neurodiversity

ABSTRACT. Autistic adults suffer from an alarmingly high and increasing un-employment rate. Many companies use pre-employment personality screening tests. These filters likely have disparate impacts on neurodivergent individuals, exacerbating this social problem. This situation gives rise to a bind. On the one hand, the tests disproportionately harm a vulnerable group in society. On the other, employers think that personality test scores are predictors of job performance and have a right to use personality traits in their decisions. It is difficult to say whether these negative disparate impacts are a case of wrongful discrimination. Nevertheless, we will show that pre-employment personality tests prey on several features of autism in an unfair way, and for this reason, we suggest the contours of some regulation that we deem necessary.

1. INTRODUCTION

Lydia Brown remembers that, while in high school, they applied for many jobs that required a personality test as part of the application. They did not get a single job, and to this day, Brown does not know the reason why. In the documentary, *Persona: The Truth Behind Personality Tests*, they voice a lingering suspicion:

> Perhaps one of the reasons that I was not hired was because I failed the personality test; perhaps one of the reasons was that even if people didn't know affirmatively that I'm autistic, I still coded. I was perceived as "you're kind of weird"; we can tell something is going on with you, so we do not want you in this workplace, you are not a good fit.

Is Brown's concern warranted? After showing that it is, this article considers whether the use of pre-employment personality screening tests in the hiring process is just. This question has been a recurrent one for decades. Consideration of neurodivergent people, however, adds a new twist to the ethical and legal landscape.[1]

Investigation of this topic is demanded by the fact that two societal trends are on a collision course: the growing popularity of using personality tests as pre-employment screeners, and the alarmingly high and increasing unemployment rates of neurodivergent people.[2] With companies facing large volumes of job applications due to the application process moving online, many have turned to personality tests to help winnow the pool of applicants. These tests have become a huge $2 billion industry, with an estimated 60%–70% of Americans having taken a personality test as a prospective employee (Hawkins 2021). Moreover, the field seems ready to expand with the Artificial Intelligence (AI) revolution, as companies such as Pymetrics and HireVue offer personality analyses based on video game performances and facial expressions, respectively.

Meanwhile, neurodivergent people face terrible job prospects. A great number desire to work. Research suggests that they typically demonstrate low absenteeism, superior attention to detail, and a high degree of patience toward repetitive, routine-based duties (Hendricks 2010; Solomon 2020). A supervisor at a Goldman Sachs initiative describes autistic people as having a "laser-like focus and great attention to detail, a talent for spotting irregularities, and strong technical skill" (Butcher 2021). Despite these skills, approximately 42% of autistic people have never been employed (Roux et al. 2015; Solomon 2020). This unemployment rate is worse than that reported for ex-convicts and for people with many other disabilities (Roux et al. 2015). Even among autistic college graduates, roughly four out of five are either unemployed or critically underemployed (Barnett 2020).[3] With an estimated 2.21% of US adults having autism, 50,000 autistic teens aging into adulthood every year (Shattuck et al. 2012; Roux et al. 2015), and increasing numbers of autism diagnoses, the number of unemployed autistic adults will continue to grow. Since meaningful employment is a crucial feature of life satisfaction and unemployment demands state support, the unemployment problem for neurodivergent people is a serious challenge for individuals and society at large.[4]

The growing popularity of pre-employment personality tests and the increasing unemployment rates of neurodivergent people seem to be on a collision course because it is likely (as we'll argue) that these personality tests have disparate negative impacts on neurodivergent people, just as Brown speculates. If that is correct, then these tests are exacerbating a major societal problem. Yet the resolution of this conflict is not trivial, as we run headlong into a clash between two time-honored principles. One is society's desire to not allow mechanisms that have disparate negative

impacts on vulnerable classes. Neurodivergent people certainly form a class that has been stigmatized and harmed by society. Those with mental disabilities, including autistic people, are protected, for instance, in the US by the Americans with Disabilities Act (ADA) of 1990. The other principle is that employers should be free to hire based on personality type. Since there is no essential connection between personality type and race, gender, or physical disability, this freedom does not usually have disparate impacts on vulnerable groups in society.[5]

That isn't the case with autistic people. It is important to note that an autism diagnosis captures a cluster of traits, and the manifestation of these traits can range from mild to severe. No personality trait is essential to having autism, but traits can function as statistical proxies, much as zip code can for race. How can we balance the latitude normally given to employers while not aggravating the unemployment problem among neurodivergent people?

In what follows, we begin with a historical overview of the controversial use of personality tests. We then argue that the best evidence we have is that these tests do, in fact, have negative disparate impacts on autistic people. We briefly consider the philosophical question of whether using personality itself leads to wrongful discrimination. Then we return to our main topic, personality tests. We voice some general concerns about these tests and then detail specific ways in which the tests prey on the vulnerabilities of autistic people. Particular features of the tests, we argue, make their use unjust. Finally, we discuss the ethical contours of any possible regulation designed to fix this problem.

2. BACKGROUND: ETHICAL CONCERNS ABOUT PERSONALITY TESTS

2.1 Types of Tests

There are many types of personality tests, and different tests have faced different kinds of ethical questions. This background will set up the context of our worries.[6] Personality tests were developed and employed as the field of psychology emerged in the early 20th century. They quickly moved from the clinic to industrial psychology, where the idea of using personality tests as an employment screening tool can be traced to 1915, when the Bureau of Salesmanship Research at the Carnegie Institute of Technology was tasked with developing scientific methods for selecting salespeople through personality testing (Carrigan 2007). These ideas were perhaps first implemented as a stand-alone test during World War I with the Woodworth Personal Data Sheet. Designed to exclude the "insane, feeble-

minded, psychopathic and neuropathic" from the forces sent to France, it asked 100 yes-or-no questions and filtered out those who answered "Yes" more than 30 times. Not long after, tests such as the Bernreuter Personality Inventory and the Humm–Wadsworth Temperament Scale became prominent in the industry (Gibby and Zickar 2008).

Today, personality screening tools are a big business. Companies such as TalentSorter, Hogan Personality Inventory, and Aspiring Minds each report administering millions of tests. Moreover, 80% of Fortune 500 companies and 89% of Fortune 100 companies are said to use such tests (Winterhalter 2014), in addition to over 200 US federal agencies (Cunningham 2012). Large companies, such as Best Buy, Lowe's, Kroger, McDonald's, CVS, Target, and Walmart, all have used or still use such tests. The rationale for use is ultimately efficiency. An automated test helps a company quickly shrink a large pool of applicants to a smaller one.

Personality tests are not used alone to select a candidate for a position. They are instead a kind of first-pass filter to move to the next stage. The result of a personality test may exclude an applicant from progressing to a second round, such as an in-person interview. Or the result may be part of some algorithm including other factors, such as the strength of their resume. Used either way, a personality test can increase or decrease an applicant's chances of moving forward in a job search.

Personality tests should be distinguished from ability or skill tests, like typing tests, language proficiency tests, and general cognitive tests. Although there can be some overlap, they should also be distinguished from honest/integrity tests, which are also often used as a screening tool. There are hundreds of personality assessments, but most trace their ancestry (and many of their questions) to one of three types of tests. The three tests are the Myers-Briggs Type Indicator (MBTI), the Minnesota Multiphasic Personality Inventory (MMPI), and those based on Big Five personality traits. All three rely on self-report reactions to prompts. Otherwise, they are very different, each with its own fascinating history, and each prompting sometimes different ethical worries (see Emre 2018 and Paul 2005 for popular histories).

The MBTI was the first of the three developed. Based on Jungian personality types, Katharine Cook Briggs and her daughter, Isabel Briggs, created the MBTI in 1917. The MBTI distinguishes people via four opposite pairs of features, such as introversion or extroversion, resulting in 16 possible personality types. According to the Myers-Briggs company website, the MBTI is available in 29 languages and used by more than 88%

of Fortune 500 companies in 115 countries. The MBTI is often claimed to be the most popular personality test in the world.

Designed by Starke Harthaway and J.C. McKinely in 1939 to spot and sort mental illness, the original MMPI asks 537 true–false questions and scores test-takers along ten clinical scales (e.g., schizophrenia, hysteria). It does not sort people into types, like MBTI, but places them on a multidimensional scale based on psychological categories that were common when the test was developed. MMPI and variants like the California Personality Test migrated outside the clinical setting into a wide variety of areas, including courtrooms and pre-employment screening. MMPI claims it is still the most widely used clinical test in the world. However, because MMPI ran afoul of the American Disabilities Act's prohibition on administering medical tests to prospective employees (*Karraker v. Rent-A-Center, Inc.*, 411 F.3d 831, 837), its use in pre-employment screening has dropped and is now confined to certain types of employment, such as law enforcement.[7]

Lastly, a quite different set of tests is based on contemporary psychology. Using factor analysis on language or psychometric data, psychologists found that five personality traits—Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism—emerged as the most parsimonious way to characterize someone's personality. The Big Five personality schema is based on the idea that personality traits should be expressed in natural language (Goldberg 1993), while the Five Factor is based on psychometric data via personality questionnaires (McCrae and John 1992). Although the two differ, both assume personality traits can be discovered empirically instead of extracted from theory.

Big Five and Five Factor are theories and not tests, but many tests are based on them, including the Neuroticism-Extraversion-Openness Personality Inventory-Revised (NEO-PI-R) the Honesty/Humility-Emotionality-Extraversion-Argreeableness-Conscientousness-Opennes to Experience (HEXACO) test, and the Hogan Personality Inventory. NEO-PI-R has 240 items, such as "I get irritated easily," to which one agrees or disagrees on a five-point Likert scale. The others work similarly. For convenience, we'll refer to tests based on Big Five or Five Factor as "Big Five" tests. We will now consider a few ethical concerns raised by these tests.

2.2 Pseudoscience

The first concern is whether it is ethical to use known faulty tools. Like the notorious Rorschach inkblot test, many personality tests originated from now discredited ideas in psychology's early history. It is difficult to ethically justify using a tool known to be problematic in employment decisions. For example, the wildly popular MBTI bears all the hallmarks of pseudoscience. First, it has no internal reliability. When subjects retake the test just five weeks later, half are categorized differently (Pittenger 2005). When reexamined at intervals ranging from five weeks to six years, one study showed that between 39% and 76% of test-takers were assigned a different type (National Research Council 1991). It was the unreliability of polygraph tests that motivated the Employee Polygraph Protection Act in the US, which prohibits private employers (with a few exceptions) from using lie detectors as a pre-screening test for employment. A similar concern motivates demands that personality tests stop being used as pre-employment filters (Nevins 2005).

Second, the MBTI has no validity. Factor analyses showed that 83% of the differences among 1,291 students could not be accounted for by the MBTI. Many aspects of personality are left out of the test. In addition, some factors are not statistically independent of one another (Pittenger 2005).

Third, MBTI fails miserably as a predictor for job performance and job satisfaction. There is no evidence demonstrating that, for example, those identified as personality type ESFP (Extroverted, Observant, Feeling, and Prospecting) make better or worse salespeople than those who were identified as INTJ (Introverted, Intuitive, Thinking, Judging).

Lastly, the test's background theory seems false. The MBTI is based on the idea that the test uncovers one's "true self." But what is a true self, and why do self-reported bimodal preferences reveal it? The assumption that the MBTI reveals one's true self is based on Briggs noticing an association between her friends' and relatives' answers on the questionnaire and what she took their true selves to be.

MBTI knockoffs abound, but these too will likely suffer the same fate. The Enneagram test, which scraps the four Jungian types for nine categories, has been accused of being "so vague and malleable that anything relevant can be shoehorned to fit the theory" (Carroll 2011, 306). While Myers and Briggs's self-taught mastery of psychometrics was impressive for its day, MBTI now seems no more scientific than social media quizzes that determine what Hogwarts house one would belong to in the *Harry Potter* series. It is like a horoscope or fortune telling,

and it seems to "work" the same way (i.e., through the Forer effect, an acceptance phenomenon where people subjectively confirm or find truths in generic descriptions of their personality or lives). Indeed, like astrology, MBTI is often used as a guide to what career you should seek. Showing that MBTI fails to agree with known data and lacks internal consistency and testability, Stein and Swan (2019) claim that MBTI's true value comes from using it as a pedagogical example of "how to distinguish valid science from pseudoscientific 'woo'" (3).

Is it ethical to filter out prospective employees based on such a test? Comparing MBTI to astrology, Paul (2016) describes it as "a Carl Jung-inspired load of nonsense engineered to make everyone who takes it feel good about themselves." Interestingly, there have been cases where employers pre-screened applicants based on their astrological sign. In 2009, an Austrian insurance company ran an ad declaring a preference for Capricorn, Taurus, Aquarius, Aries, and Leo applicants. In defense, the company cited evidence that its best workers came from these signs (Radford 2011). Whatever one thinks of this practice, we suspect few would defend the use of an invalid and unreliable employment filter if (say) Pisces were a vulnerable group in society.

2.3 Privacy

Privacy has always been a major concern about personality tests. The tests ask for disclosure of what many consider to be the private sphere surrounding a person—a person's "backstage" area, or the parts of an individual's life that employers have no right to know (see DeArmond 2012). For example, the original MMPI asked many odd and intrusive true–false questions, such as: "I have never indulged in any unusual sex practices," and "I have difficulty in starting or holding my bowel movements." Alarm over these sorts of intrusive questions resulted in a two-year investigation by a US Senate subcommittee in 1964 and 1965. The subcommittee reviewed personality tests used on government employees and discovered true–false test items that included "Christ performed miracles," "Many of my dreams are about sex matters," and much more along similar lines (Creech 1966). The tests effectively ask applicants to disclose features that employers are not legally (in many countries) or ethically (we would argue) allowed to ask, questions that reveal gender preference, religious identification, and so on. Even apart from such disclosure issues, the tests just seem to pry too deeply into what one normally deems to be part of the private sphere surrounding a

person. The investigation led to 35 members of Congress co-sponsoring a bill in 1965 that would have made the MMPI illegal to use in screening for federal jobs, but the bill failed to pass.

2.4 Bias and Discrimination

All three tests have been plagued by charges of bias and discrimination. These charges are a bit ironic since personality test vendors almost universally advertise that their tests offer a fairer and more objective screening filter than traditional methods of screening. Many vendors claim that their tests will increase workplace diversity. Yet there are serious concerns that they can promote the opposite.

The original MMPI is a prominent example of how a personality test can be racially biased. Scores are generated by comparing a person's answers to those provided by its normative sample. This sample, the so-called "Minnesota Normals," consisted of 724 white, mostly rural Minnesotans (Butcher et al. 1983). If your answers about unusual sex practices or bowel movements differed greatly from those provided by a typical Minnesota farmer in the 1930s, then you would score highly on its clinical scale. Given this original scoring method, it is not surprising that the MMPI discriminated against race. McCreary and Padilla (1977) showed that Black and Hispanic offenders scored higher on negative traits compared to whites (see also Butcher et al. 1983). Within the clinical context, this bias could lead to racial minorities being disproportionately diagnosed with mental illness; within the employment context, it could make members of a racial minority look like trouble. It was a demand to eliminate racial bias in the MMPI that led to the creation of the MMPI-2, which removed many (but not all) of the more controversial items and employed a new normative sample of 2,600 respondents drawn from a representative sample of the whole country.

The threat of bias is almost endemic to tests like this because the scoring involves a tacit normative judgment. To see the problem, remember that MMPI is a long questionnaire designed to reveal mental health problems. Test items include statements such as "I think I would enjoy the work of a librarian." Is agreeing with this a sign of mental health or illness? The answer is not obvious. To tell, one needs to see what antecedently determined healthy and unhealthy people tend to answer. That is where the normative judgment enters.

Even the more academically respected Big Five tests have recently been accused of sexism. For example, a reporter took a Big Five test twice,

once identifying as a woman and the second time identifying as a man, providing identical answers both times (see Goldhill 2019). Interestingly, the results showed two different personality outcomes, revealing gender stereotypes built into the test. Not every test has all three problems, and not every problem is equally severe. However, these sorts of problems do tend to recur with the neurodivergent population. Are test results positively associated with variables that matter, such as job performance? Do they reveal mental disabilities, effectively asking candidates whether they are part of a vulnerable group? And are they biased against autistic people?

3. SCREENING OUT NEURODIVERSITY

In 2014, Kyle Behm applied to a Kroger's supermarket and learned from an insider that he was "red-lighted" by the Unicru personality test, causing his application to be excluded from consideration. The Unicru test was based on the Big Five theory of personality. After applying to six other stores that used Unicru and never getting an interview, Behm filed complaints with the Equal Employment Opportunity Commission (EEOC) against all seven companies under the ADA. He had noticed an overlap between questions on the Unicru test and questions used to clinically diagnose his bipolar disorder (Weber and Dworskin 2014).[8] Unlike Behm, Lydia Brown, with whom we began the article, likely will never know whether their personality test results excluded them from employment considerations. Companies do not usually provide applicants with the reasons they were not hired; and because it is unlawful for companies to ask applicants to disclose if they have a disability, companies also don't know if the individuals that they are screening out have autism. Test vendors likely have not studied the bias against neurodivergent people either. For example, in a lawsuit over an EEOC, the personality test vendor, Kronos, asserted that they do not keep track of potential bias against the mentally disabled (see Timmons 2021). Unfortunately, autistic people are likely to always be kept in the dark about these matters.

Nonetheless, a strong case can be made that personality tests do have a disparate impact on neurodivergent people. In 2011, the Rhode Island Commission for Human Rights claimed there was "probable cause" to think that CVS Pharmacy's use of a pre-employment personality test revealed mental health disabilities, which violated state law. Confined to autism, there now seems to be sufficient evidence for this claim.

Let's restrict ourselves to tests based on Big Five personality traits. The reason for this is that most commentators agree that MBTI should not be

used for employment purposes given the test's poor validity and reliability. Because there are so few researchers who take the MBTI seriously, there are also no good studies on the topic. And as regards MMPI, there isn't great need for discussion because this test isn't used much in employment contexts (since it runs afoul of the ADA) and because the answer is unsurprising and well known. The MMPI-2 is a clinical tool, so there is a strong connection between MMPI-2 results and an autism diagnosis. For example, when the MMPI-2 was administered to both neurodivergent and neurotypical adults—controlling for age, gender, and intelligence—Ozonoff et al. (2005) found that autistic adults had MMPI-2 scores that reflected social isolation, interpersonal difficulties, depressed mood, and coping deficits. According to the researchers, these results were consistent with the DSM-IV clinical description of autism, which suggests that the MMPI-2 may accurately capture the autistic phenotype.

The Big Five tests are "scientific" in a way that the MBTI is not, and unlike the MMPI or MMPI-2, Big Five tests were not designed with clinical psychopathology in mind (although potential clinical applications of Big Five tests is currently a hot topic, e.g., Schwartzman, Wood, and Kapp 2016). However, background theory suggests that links between Big Five traits and autism should exist. There are plenty of links between autism and personality.[9] Part of the autistic pattern of behavior is characterized by problems with social interaction and communication, so it would not be surprising if autistic people scored lower on Extraversion than neurotypical individuals. Problems comprehending personal cues might cause lower Agreeableness scores, problems with self-control may cause lower Conscientiousness scores, and so on.

There are dozens of studies and research is very much ongoing, but fortunately we can draw on a major meta-analysis of such studies by Lodi-Smith et al. (2018). Combining studies to include almost 4,000 participants, they show that although low Big Five traits and autism are not the same, autistic people do score lower on each of the five traits of Openness, Conscientiousness, Extraversion, Agreeableness, and Emotional Stability. The strength of correlation ranged from somewhat weak (–.21 Fischer z for Conscientiousness, –.22 for Openness) to moderate (–.39 for Agreeableness, –.36 for Emotional Stability) to borderline strong (–.50 for Extraversion). This confirms our expectations from background theory and is in line with many previous results. For instance, all previous studies prior to Lodi-Smith et al. (2018) found a strong link between Extraversion and autism, and all but one found one between Emotional Stability and autism.

Still, only one study, Schriber et al. (2014), frames the question almost perfectly for our purposes. They ask: How well do Big Five test results predict belonging to the autistic group or to the group thought to be typically developing (TD)? We ought to be cautious not to draw too large a lesson from one set of studies, of course, but their answer is interesting. In their model, the Big Five scores correctly assigned a person to the autistic group vs. the TD group 70% of the time. For the researchers, Neuroticism (the opposite of Emotional Stability), Agreeableness, and Conscientiousness scores were the best predictors. The best evidence we have at the moment suggests that Lydia Brown's suspicion is correct: they were discriminated against by personality tests. And with 5.5 million neurodivergent individuals in the US, Brown is not alone.

4. IS PERSONALITY-BASED DISCRIMINATION WRONGFUL?

All employment filters discriminate. Whether the discrimination is wrongful is another question. And that is a large and contested topic in philosophy and legal theory (see Hellman and Moreau 2013). Here we will not argue that personality shouldn't be used by employers. That personality is an important ingredient of job success is a deeply ingrained belief. For example, management training specialist Dan Monteiro argues that, in nine out of ten cases of turnover, the turnover is caused not because of a lack of technical skills, but mainly because of a lack of people skills: the individual could not get along with the rest of the team (Yu 2020). We suspect the conventional wisdom is not entirely in conformity with the science, but we will not challenge it here. Our focus will be on personality tests, not personality per se. However, it is an interesting question whether personality should always be considered acceptable in employment screening. So here we simply point out some considerations that would make for a good question for another day.

As mentioned, there is much debate about what makes discrimination wrongful discrimination, and we suspect different theories may diverge on some questions involving personality. But to see what's special about filtering based on personality, let's say an employment practice is wrongful if it: (1) is superficially unbiased in its treatment of different groups but in fact has an adverse impact on a vulnerable group, and (2) the practice cannot be justified to be related to job performance and consistent with business necessity (e.g., Arneson 2006; Lippert-Rasmussen 2012; Scholes 2014). In Section 3 we provided probable cause for thinking that condition (1) holds. Neurodivergent people are certainly a vulnerable group, and

there is plenty of evidence that personality tests, but also personality itself, has negative disparate impacts on this vulnerable group. That may be acceptable if condition (2) does not hold. Typing tests for data entry jobs presumably have disparate negative impacts on the elderly. That may be acceptable because typing speed is relevant to job performance.

The problem is that what constitutes job performance is not so clear. Should employers be allowed to discriminate against the unattractive? Abercrombie & Fitch did just this. The clothing store explicitly sought to hire only attractive salespeople, turning down applicants who were deemed unattractive or "uncool." Their rationale was that part of job performance is to act as brand ambassador, even if one is only hanging clothes or running the cash register. That kind of rationale can only go so far before many theories condemn it. Famously, many theories of wrongful discrimination (and of course, US law) condemn discriminating against Blacks if the justification is not racist but only the desire to appeal to customers in a predominately racist town. The difference here is that the expansive notion of job performance disproportionately hurts a vulnerable group.

Statistical discrimination is also tricky. Should employers be allowed to discriminate against smokers? Smokers can do almost any job non-smokers can do, but statistically, smokers take longer on breaks and are more likely to have health issues that lead to more absences and job turnover (see Scholes 2014). Yet smokers are disproportionately represented in many vulnerable groups in our society, making it likely that this discriminatory practice has an adverse impact on these groups. These conflicting rationales are reflected in practice, as discrimination against smokers is only legal in about half of the US states.

With these examples in mind, we can see that it is hardly cut and dry what facts should count as relevant to job performance.[10] Personality traits are said to be legitimately relevant to the duties of a particular job. For example, suppose it is a fact that salespersons who are extroverted perform better than those who are introverted (this claim is, in fact, unsubstantiated). Condition (2) seems not to apply. If an autistic applicant's autism manifests as having an introverted personality, and the applicant is excluded, the applicant doesn't seem to have been wronged. A case like this seems more like the typing case.

Or does it? In the typing case, the skill is directly tied to the "core" aspects of the data entry job. Slow typists really can't do the job well. Speedy elderly typists aren't harmed by this skill filter. But is extroversion

really a part of being a good salesperson? Is that part of the "core" job description, or is it more like the controversial "lookism" used by Abercrombie & Fitch? And is there really a link between extroversion as manifested in a job interview and extroversion as a salesperson? In the case of typing, typing skills do survive the transition between the test and in the wild.

Typically, what causes membership in a vulnerable group isn't causally responsible for any variable related to the core duties of a job. All races and genders can stack shelves, pilot an airplane, work at a call center, engineer satellites. Except in rare cases, like working in the entertainment industry, race and gender are not causally connected to job performance. A big exception is disability. For example, for some jobs, no reasonable accommodation can be made for those who can't walk. In these cases, what causes membership in a vulnerable group is causally related to the core duties of the job. That is a rare situation, but it helps us see what is special about the case of personality and neurodivergence: personality traits are thought to be a valid predictor of successful job performance no matter the job.

The assumption that the test is a universal job performance predictor is hard to believe. Patience is surely a virtue in wildlife photography and elderly care, but impatience is probably a virtue in managing a high-risk, fast-paced business venture. Tests from some vendors are tailored to particular jobs, but most are not. Personality is not like cognitive skill (which itself is hardly a simple construct). One might be tempted to think that fast, reliable information processing and inference is a good general skill, as useful to a pilot as a baker as a basketball player.

Personality is not like that. There isn't a one-size-fits-all winning personality that is good for all jobs. We take personality to be the individual differences in characteristic patterns of thinking, feeling, and behaving. With neurodivergent people, there is a non-accidental link between membership in a vulnerable group and these characteristic patterns thought to be relevant to most jobs. Lacking "ideal" personality traits can both identify a person as neurodivergent and identify them as not fit for most jobs. What qualifies as an "ideal" personality trait reflects the attitudes of neurotypical individuals and is a byproduct of societal ableist notions of what makes a "good" person or a "good" employee. Ableist norms value and endorse certain abilities, such as productivity and charisma, and assume an ideal worker to be a man, able-bodied, productive, committed to work, and free from family or other responsibilities (Foster and Wass

2013). Like intelligence, which is also a problematic construct, the right personality is thought to help with employment in data entry, policing, nursing, and sales. Personality screening (automatic or in-person) can perpetuate the pejorative societal belief that there is an "ideal" quality or personality trait for job performance. This is different from the cases of discrimination based on race or gender, where the links to job performance are almost all accidental; and it is different from the non-accidental cases of discrimination based on physical disabilities because personality is so widely considered a universal predictor of job success.

To be clear, in this article we are not challenging an employer's right to use personality in deciding among applicants. However, we do think the question warrants a deeper dive elsewhere.

5. AGAINST PERSONALITY SCREENING TESTS

Personality is a problematic construct. It is understood as a constellation of traits that we carry with us as we navigate the world, like a person's physical attributes, such as height, or skills, such as fluency in a foreign language. These traits are supposed to be stable and, importantly, are understood to be causally responsible for much of our language, thought, and behavior. The last is what makes personality relevant to job performance: these traits affect our behavior.

Many decades of research, however, have shown that this picture of personality is not quite right. Personality traits may not be as stable as we think. Beginning with work by Walter Mischel, and then especially by Richard Nisbett in the late 1960s and early '70s, scientists began showing that there is a strong situational element to personality (see Doris 2002). Observing someone to be extroverted in one situation doesn't mean that they are likely to be extroverted in another. Nisbett blamed this on the fundamental attribution error, our tendency to overestimate the power of traits and underestimate the influence of situations. Nisbett's claims are controversial and we're not endorsing situationism; we're simply pointing out that experiments in this area reveal some disparity between our folk concept of personality and the reality as regards stability.

Personality traits are also not viewed by most psychologists as a kind of "metaphysical" feature of a person that is causally responsible for their behavior (Miller 2021). Rather, in the more scientifically respectable literature, it is viewed as an economically powerful way to classify patterns in thought and language. These patterns were discovered by reactions to items drawn from ordinary language or questionnaires and clustered into

"traits" via factor analysis across large groups of neurotypical individuals, not from looking at behavior.

We mention these features of personality not to suggest that personality shouldn't be used in employment screening. We're granting that it can be. Instead, we mention these facts about personality to highlight that it is a hard thing to measure with a test and even harder to say that what the tests measure is a causal difference-maker to actual behavior on a job.

Some complaints about pre-employment screening with personality tests have nothing to do with neurodiversity. Because they provide useful background, we briefly introduce them in Section 5.1. Then we turn to problems that are specific to autistic individuals in 5.2.

5.1 General Concerns

Personality tests are based on self-reports. Self-reports are a notoriously unreliable way of getting at one's true characteristics. Do such self-reports cohere with third-party judgments of personality? Will they cohere with behavior? In an employment context, it may not matter so much whether you think you are extroverted; what may matter more is whether your team finds you congenial and whether you are friendly to customers. A lot of science is done validating the construct formed from self-reports. Here we report two limitations as regards job performance.

First, in media reports, those weeded out by personality tests often complain that the tests at best reveal only a small sliver of what they bring to a job. All filters make mistakes, of course, so the tests may not be unjustly imposed. But to some extent we know that the standard complaint has some merit, that personality tests leave out known significant aspects of personality. Paunonen and Jackson (2000) point out that even the Big Five model leaves out nine important traits. Worse, what's left out can be relevant to job performance. Later, Paunonen et al. (2003) showed that these nine traits gave better predictive validity than the Big Five model across 19 measures of job success in samples from Canada, England, Germany, and Finland. The point is more general than this. Many other factors relevant to success in a job, such as moral character, goals, and so on, may also be good predictors. Unless the "sliver" revealed by the test proves to be a useful or efficient predictor of job success, selecting that aspect seems potentially arbitrary.

Second, and on that very question, how do these tests fare when it comes to predicting job performance? A lot depends on the test. In a recent review, Hughes and Batey (2017) write about tests like the MBTI:

> Thus, due to poor reliability and questionable validity, the current authors recommend that regardless (or perhaps because) of their simplicity, typologies be treated with caution in all organizational contexts, and under no circumstances should be used for selection. That this point still needs to be raised is testament to the gulf between science and practice we raised in the introduction to this chapter. (154)

Although widely used, the MBTI carries very low predictive validity. We believe it is therefore unethical to use. As for MMPI, due to its clinical origins and use, this question is mostly studied in the few places it is used, such as law enforcement. So far as the authors know, there is not a consensus on whether MMPI or its variants have utility in these professions. Of far more interest to our present topic is whether Big Five-based models—which are far more widely used—have predictive validity regarding job performance.

On this question there is a mountain of data, studies, meta-analyses, and even meta- analyses of meta-analyses, plus a range of opinions. We cannot do justice to all of this literature, but we can point the reader to Hughes and Batey's (2017) comprehensive and even-handed review. We believe the following is a fair representation of the situation.

On the negative side, based on decades of studies, Big Five traits account for only about 5%–7% of variance in job performance measures (Hughes and Batey 2017, 159). That places the Big Five near other methods of selection generally deemed unreliable, such as unstructured interviews (Barrick et al. 2001). For this reason, Morgeson et al. (2007) argue that such tests should not be used in job selection.

On the positive side, however, one can dig into the data and find traits and facets that do correlate tolerably well with some particular measures of type of job performance. For example, one's score on Conscientiousness contains information relevant to some measures of job performance, (Barrick, Mount, and Judge (2001 find $r = 0.10$ is an objective rating, and $r = 0.15$ is a supervisor rating). Ones et al. (2007) thus reply to Morgeson et al. (2007) with other ways of finding utility in the Big Five answers.

Hughes and Batey (2017) feel that using the test alone to screen out applicants is "indefensible" and suggest best practices such as tailoring parts of the tests to very specific analyses of job duties. A serviceable summary of the situation might be that Big Five traits are a very poor predictor of job performance as these tests are typically used in the wild; nevertheless, they contain a lot of information that is potentially useful in selection when combined with other tools.

In sum, even if one is not autistic, one may have cause to complain about the use of personality tests. While they may improve in the future, without modifications, off-the-shelf personality tests have very little link to job performance.

5.2 Specific Concerns

We now turn to our main argument. Its conclusion is that personality tests set up autistic individuals to fail in unjust ways. We'll highlight four ways in which this is the case, although there are undoubtedly more.

First, metacognition is a known problem for autistic people. Metacognition is thinking about thinking, an understanding of one's own and other people's cognitive processes. Examples might be knowing that one tends to be sleepy in the afternoon, that one has trouble remembering people's names upon introduction, or that a colleague finds math difficult. Autistic individuals have a hard time with this kind of self-awareness and social recognition (see Williams 2010 for a discussion).

Unfortunately, personality tests based on self-reports are almost premised upon the ability to reliably know and report on one's own mind and that of others. Consider common questions found on personality tests: How do you make important decisions? What do you do to manage stress? If you could change one thing about your personality, what would it be? To answer these kinds of questions, one requires an understanding of one's own cognitive processes. They demand that autistic individuals perform a task with which they're known to struggle.

Now, of course, some jobs may require some element of metacognition, so for these jobs these questions may be appropriate. A position that requires a lot of open-ended planning might be best filled by someone who could consider different cognitive strategies for getting efficiently toward some goal. In these cases, perhaps this ability is fair to assume. However, it's important to recognize that metacognition is not one thing, nor is it always helpful. One might have an awareness of strengths and weaknesses but not be a good planner, or vice versa. There are many different metacognitive skills, and not all will be relevant to every job. Some may even be detrimental. Self-awareness, for instance, can be paralyzing, leading to reluctance to speak up, and so on. Metacognition seems often to be treated as universally beneficial, but that depends on the particular skill and job.

Second, societal stigma against neurodivergent people is in a sense baked right into the scores. Many of the items on personality tests are not about what you think about yourself but rather about how you feel

others perceive you. Consider how "altruism" is determined in McCrae and Costa's Five Factor Theory. This is measured with one's reactions to the following items:

- Some people think I'm selfish and egotistical. (R)
- I try to be courteous to everyone I meet.
- Some people think of me as cold and calculating. (R)
- I generally try to be thoughtful and considerate.
- I'm not known for my generosity. (R)
- Most people I know like me.
- I think of myself as a charitable person.
- I go out of my way to help others if I can.

In these examples, "R" means that the score is reversed. Notice, with Miller (2021), that half of the items ask about how one is perceived by others, not what one thinks of themself. This presents at least two problems.

One, it preys on the very deficits most associated with an autism diagnosis: one's perception of social interactions (and, of course, metacognition). An autistic person may not be able to accurately recognize and report what others think about them.

Two, imagine that you are neurodivergent. Your social world feels strange, confusing, and often hostile. Unfortunately, in a world that is designed for and by neurotypical people, ableist norms dominate, and autistic people are cast as difficult to understand. Studies show that when a non-autistic person finds an autistic person difficult to understand, the autistic person is likely to be liked less, a phenomenon known as the double empathy problem (Alkhaldi, Sheppard, and Mitchell 2019). Thus, the deck is stacked against you if you fill out the questionnaire honestly, for you will not fill out (for instance) "true" to the item "most people I know like me" and so on.[11]

Third, honesty itself is a huge problem with personality tests. If you Google a particular test, your search results will be filled with tutorials on how to "pass" that personality test. To increase their chances of getting the job, neurotypicals provide the answers they believe the employer wants to hear, not what they really think. Hundreds of papers have been written on the problem of people "faking" answers on the test, with test vendors inserting items to try to compensate for faking. Autistic adults tend not to fake answers, however. Studies suggest that autistic adults are less likely to use reputation management compared to neurotypical adults (Cage et al. 2013). This could be due to honesty or limited social cognition or both. Whatever the explanation, autistic adults are more likely to provide

"true" answers than "good" answers on personality tests, at the expense of potential employment opportunities.

Fourth—and perhaps this should be regarded as a general and not specific problem—the test scores aren't nuanced enough to see what personality deficits can be reasonably accommodated in the workplace. That last sentence may sound odd, for on a traditional understanding of personality we don't think of personality deficits as something that can be accommodated, like a physical disability. However, the traits that cause autistic individuals to perform poorly on personality tests can often be accommodated in the workforce.

Training programs and best practices exist and are expanding as we write (for best practices, see (Oesch 2019)). Many behaviors that are viewed as problematic in the workplace can be solved with items as simple as noise-cancelling headphones or a place and time to nap. Assistive technologies are also being developed. A boss may email a document with the instruction, "Take a look at this." The autistic employee might literally just take a look at it. Yet email filters for either the boss or subordinate can be developed that alert one or the other in a way that facilitates communication of the true request. Such an accommodation may be no more intrusive than an email spellchecker. Neurodivergent people cannot get to this stage, however, for personality is not considered something that can be accommodated.

Not everything can be accommodated, of course. What can be is highly context-sensitive and often very tricky to determine. Dimming lights in the workplace may help employees with sensory overload, but for some jobs that may not be practical or even safe. Modifications for social traits will be especially complex. If large meetings are stressful, can autistic employees be allowed to skip them in a way that doesn't harm workplace efficiency? Alleviating social discomfort might suggest a norm of minimizing personal conversations in an office cubicle space; but perhaps those personal conversations are great for team building and employee retention overall. While not everything can be accommodated, the idea that personality is a universal performance enhancer makes it hard to see that so many traits associated with autism can be accommodated.

There are many other worries. Test-taking itself is a problem for many people with autism spectrum disorder, yet that may not be a skill that matters to job performance. Do personality scores mean the same thing when neurodivergent people take a test designed for the neurotypical?[12] Do the links to job performance still hold for a neurodivergent person?

We lack good evidence to answer the last two questions. But we think we've done enough to show that personality tests exploit features of being neurodivergent that will likely cause them to perform worse than neurotypical people, thereby decreasing their employment prospects. Our points show that in screening out neurodiversity we have an important "new" problem for the ethics of personality testing.

6. RECOMMENDATIONS

Recall that even without considering the impact on neurodivergent people, personality tests are on thin ground ethically (see the references in note 6). For privacy reasons, the US Congress almost banned their use in the 1960s. In the 1970s and '80s, issues of racial and gender bias and reliability/validity led to decreased use. Against this background, we believe that what we've shown tips the balance in favor of some kind of regulation or policy governing their use in employment selection. The tests unjustly target a vulnerable population, have poor or unknown connection to job performance, and exacerbate a major societal problem. New policies are needed. Without them we can only expect the employment situation for neurodivergent people to worsen.

This article is not the place for a detailed policy analysis or recommendation. Here we will simply sketch the ethical contours of what might be done. For various legislative or judicial interpretative paths that might be pursued in the US, see DeArmond (2012). To begin, recall that the demand for personality tests in hiring originates from efficiency and cost. With the application process now mostly online, many corporations receive a volume of applications that is unmanageable by traditional screening methods. Personality tests are a way to filter these large pools in a quick and easy manner.

The reader should know, as relevant background, that personality tests aren't the only way to cull applicants in this way. Many companies offer many types of automated screening tools that do not involve personality. Automated resume readers, for instance, can search for educational qualifications or keywords relevant to job performance. More sophisticated AI algorithms are also deployed. Job platforms, such as LinkedIn, ZipRecruiter, Indeed, and Monster, offer companies automated tools to help rank applicants. Changing or eliminating automated personality tests therefore need not leave companies helpless in the face of mountains of applications. If these other tools filter by criteria linked to job performance or qualification, then so much the better.

These tools can go badly wrong. All the fairness and bias issues that arise in thinking about "algorithmic fairness" occur in this space too. For instance, Amazon trialed a ranking tool that was found to discriminate against women. The program was trained on past applicants and hires. Since the tech industry is predominantly male, the program learned that successful applicants tended to be male; hence, it hunted for traits associated with being male. Upon discovery of this bias in 2015, Amazon scrapped the program for one that was less ambitious.

With this in mind, when it comes to regulating pre-employment personality tests, we can either ban them in some way or fix them. Regarding bans, there is precedent for this when the US Congress passed a limited ban on the use of polygraphs in 1988. As we saw, the 1964–65 Congress also considered a ban on personality tests in federal employment.

A ban is certainly worth considering. A ban need not leave large corporations with impossibly large volumes of applications to screen. Not all large companies use personality filters. While we are sympathetic to such a call, we are also mindful of the fact that hiring is a messy business. All the filters in the hiring pipeline have error rates, both false positives and false negatives. We are skeptical that personality tests as currently used are competitive predictors of successful job performance, but we don't want to prejudge the issue of whether they could improve. With advances in data analytics, variables crafted from psychometric data may become comparable or better than what we get from some other initial filters.

Absent a ban, a natural suggestion would be to follow the lead of laws currently being drafted to deal with algorithmic bias in machine learning. One type of proposal is to regulate the use of these tools, insisting on audits that demonstrate the tools are free of racial and sexual bias (Givens, Schellmann, and Stoyanovich 2021). One can imagine something similar being adopted. Because employers cannot ask applicants about their race, gender, or disability, they cannot obtain this evidence. But as Timmons (2021) points out, test vendors can. In development, they test their product on huge samples of people. It's fine to ask for anonymized demographic data in these studies. Some test vendors already have data on racial and sexual groups; voluntarily disclosed disability status could be added. Vendors could then be responsible for showing that their tests do not discriminate against neurodivergent people.

Unfortunately, this strategy won't work. Being neurodivergent is different from race or gender when it comes to personality. Personality traits are only accidentally linked to race or gender. Ultimately, that is

why it seems fair to demand that machine learning tools and personality tests aren't biased against racial or gender groups. But it's not realistic to only allow personality tests whose scores are statistically independent of autism. Part of an autism diagnosis is to have behavior and thought that are statistically connected to certain personality traits. These scores cannot be made statistically independent of an autism diagnosis and still measure what they purport to measure. "De-biasing" tools, a major strategy in AI ethics, is not a realistic path for personality tests and their impact on neurodivergent people.[13]

Personality tests will have disparate impacts on the neurodiverse. That is a hard fact of life, but it suggests a path forward: either design policy proposals that will move us toward the ideal, where disparate impact is ethically unobjectionable, or allow applicants to opt out of these tests. One could contemplate both actions; although, as we'll see, the rationales behind each are in tension.

The way to make the tests less objectionable is to demand tighter connections between personality tests and job performance. In France, for instance, Labor Code art. L 121-6,52 insists that pre-employment screening tools have a "direct and necessary link" to needed skills on the job. With such a regulation, companies or vendors would need to produce evidence that the personality tests they use are predictive of specific performance skills. The NEO-PI-R test, for instance, associates six facets with Extroversion, including one known as Excitement-Seeking. To make the point with a stereotype, perhaps low Excitement-Seeking scores are a predictor for job success as an accountant.

Some personality tests already contain a blend of ability or skill questions. Some are also marketed at specific jobs. We suspect that the upshot of such a regulation would be more tests moving into this mixed space, tests designed specifically for particular types of positions. Customer service jobs might have one test and data entry another. Where this regulation would have teeth is in the additional requirement that there actually be evidence linking the test to the job type. An added bonus of this requirement would be the decreased use of pseudoscientific tests such as MBTI. The neurodiverse would no longer get caught in the trap caused by thinking that personality is a general predictor of job performance across the board.

This remedy will do nothing about the fact that the test preys on some features of autism. Adoption of a new norm by test vendors could be very helpful in this regard: design "autism-safe" tests. It is common in the

testing world to eliminate problematic questions. This is regularly done when items show signs of racial or sexual bias. Worried about divulging mental disabilities, CVS modified its test in response to the Rhode Island Commission for Human Rights finding that the test implicitly disclosed mental disability.[14] And test vendors already strive to design tests that are harder to fake. Vendors can be encouraged to do what they're good at, now keeping neurodivergent people in mind. Autism-safe tests will still have disparate impacts, of course, but items known to be especially problematic for the neurodiverse can be removed.

No changes to the tests will help deal with traits linked to behaviors that can be reasonably accommodated in the workplace. For this reason and others, another path may be preferable: allowing applicants the option of not taking the test. Let us be the first to admit that this fix is not ideal. It invites the well-known "catch-22" of disability disclosure (Stefan 2002). If you opt out of the test, then you have signaled that you have a disability, disclosing what you need not disclose. This effective disclosure opens you up to stigmatization and discrimination. Recall that Ameri et al. (2017) found that applicants who disclosed a disability received 25% fewer call backs than otherwise identical applicants. If you don't opt out of the test, however, then you risk a low score in part because the test preys on your condition—hence, the catch-22.

Note that a tension between our two routes must be navigated. In US employment law an applicant may ask for accommodation for a pre-employment test if they have a relevant disability. The blind may ask for and expect written tests to be delivered in Braille, for instance. Applicants cannot, however, opt out of tests that are directly tied to job performance. There are good reasons for this. Allowing slow typists to opt out of a typing test is not fair to employers or other applicants. So if tests become more "autism-safe" and are better targeted at specific job skills, then the argument for being able to opt out becomes weaker. If the tests are not linked to job performance, the argument for opting out becomes correspondingly stronger. New policy will have to confront this dilemma. Until tests meet the standards roughly characterized here, we believe an "opt out" option is necessary.

7. CONCLUSION

Lydia Brown was right to be suspicious that personality pre-employment screening tests hurt their chances at many jobs. If 2% of the population has autism and most job applicants must take such tests, Brown is hardly

alone. Unbeknownst to most, many autistic individuals have also had their options eliminated by these tests. But disparate impact is not inherently a form of wrongful discrimination. If it happens because the characteristics needed for the job are disproportionately represented in the population, then it may be ethically tolerable. Because personality traits are non-accidentally related to the neurodivergent, who are a vulnerable group in society, and because these same traits are thought to be predictors of job performance, we are in a bind. In this article we characterized this bind and then turned to the screening tests themselves. We argued that they prey on several features of the autism diagnosis and, for this reason, suggested the ethical contours that regulation must navigate.

Regulation by itself will not address the massive neurodivergent unemployment problem facing society. Pre-employment personality screening tests are only one barrier to employment. Standard interviews are also a substantial barrier. The autistic applicant who isn't screened out by a personality screening test must eventually face an in-person interview that the vast majority of applicants find very challenging (Sarrett 2017; Whelpley and May 2022). The question is not whether personality tests impact autistic individuals worse than in-person interviews do. They both do, and they both need to be modified. Maras, Norris, and Crane (2020) show that structured interviews eliminate many unnecessary disadvantages of standard in-person interviews. Instead of asking open-ended questions, such as "tell me a bit about yourself," employers can ask specific questions about skills or educational background. Questions can be shared beforehand or even printed out. Small changes can have large differences. Just as Maras, Norris, and Crane (2020) recommend modifying the in-person interview process to focus on more specific and skill-related questions, we are recommending modifying personality screening tests in the same direction.

Programs to help autistic individuals through the interview process are an important part of addressing the unemployment problem among the neurodivergent population. Fixing the problem, however, is a two-way street. Autistic applicants can find ways to improve their performance in the hiring process. When facing standard personality tests, however, there is little they can do. In addition, employers can and should try to make the recruitment and hiring process better suited to the autistic community. They can take steps in this direction by modifying both personality tests, if used, and interviews. Neither damages their chances of finding the right person for the job; in fact, by not removing qualified applicants from the

pool for no good reason, they can increase the probability of finding the right person.

Employment recruitment and screening need to become more cognizant of screening out neurodiversity for no good reason. Chipping away at barriers will move more neurodivergent people into the workplace. As this becomes more common and employers see them excel in different sorts of jobs, we expect that norms will change and make employment recruitment and screening more neurodivergent-friendly.[15]

NOTES

1. Our focus will be primarily ethical; for a recent excellent discussion of related US legal issues, see Timmons (2021).
2. A related trend we won't discuss is the shifting of so many jobs to remote performance due to the COVID-19 pandemic. This shift has placed new pressure on the meaning of job performance, which often has automatically assumed on-site location. Since some neurodivergent people prefer remote work, an expanded understanding of job performance may open job prospects. It will also influence what personality traits are relevant to job performance.
3. There are some non-discriminatory reasons for the disparity in employment rates, such as different levels of education. But there are also data on wage gaps, attitudes, and more that suggest employer discrimination is also an issue. Ameri et al. (2015) sent out roughly 6,000 applications for positions as accountants that were identical apart from the cover letter. In one-third, the cover letter disclosed having Asperger's syndrome, in one-third a physical disability, and in one-third no disability. The groups disclosing disabilities received roughly 25% fewer expressions of interest—see also Lorenz et al. (2016) and Solomon (2020).
4. Griffiths et al. (2016) estimate that autism services in the US exceed $236 billion annually and will rise to $1 trillion by 2025, with the cost of supporting a single individual exceeding $2 million over their lifetime. These figures exclude the indirect costs associated with a loss of income, career opportunities, and productivity in the workforce due to restricted movement (for both the individual and caregivers). Behind the numbers there are other significant costs. Meaningful employment promotes self-dignity for autistic people and has been shown to improve overall quality of life and cognitive performance (see Hendricks 2010). There is also a loss of human capital due to the underutilization of potential workers.

5. Unfortunately, there is a long and sorry history of attributions of negative personality traits (e.g., miserly, weak, lazy, etc.) targeted at particular groups. But, of course, there is no evidence of genuine associations.

6. For more background, see Cavico et al. (2015), Creech (1996), Emre (2018), Mutjaba (2015), Mulvihill (2006), Paul (2005), Stabile (2002), Timmons (2021), Winterhalter (2014), Gibby and Zickar (2008).

7. *Karraker v. Rent-A-Center, Inc.* tries to carve a narrow path between personality tests that constitute medical tests designed to uncover a mental disability listed in the DSM and those that measure traits relevant to job performance; the court prohibits the first and allows the second. There is a logical distinction between certain personality traits and having autism, so the court's position is not formally inconsistent. But since the former is often a proxy for the latter, and some of the personality tests were originally designed and used as clinical tools for diagnosing the mental, the material distinction almost vanished. For example, Trull et al. (1995) show strong correlations between the ADA non-compliant MMPI and the Big-Five-based ADA-compliant NEO Personality Inventory-Revised (NEO-PI-R). The fundamental problem is that personality psychopathy "can be understood largely in terms of common dimensions of personality" (Trull et al. 1995, 516).

8. After filing the complaints, Behm's father, an attorney, learned that the EEOC was conducting an inquiry of their own into personality tests used by Kroger and PetSmart. So far as we know, the inquiry is still ongoing. Behm settled with two companies. Meanwhile, Lowe's announced that they are partnering with a mental health institute to develop new tests that do not weed out those with mental disabilities. See Timmons (2021, 422) for further details.

9. For a small start on this issue, see Austin (2005), De Pauw and Mervielde (2010), and Schriber et al. (2014).

10. The hard questions have a way of appearing in every theory of wrongful discrimination. In Moreau's (2010) freedom-based theory of wrongful discrimination, one has a right to decide how to live one's life without having to take account of one's "normatively extraneous" traits. Our question in this framework becomes: Is personality sometimes a "normatively extraneous" trait in employment contexts?

11. This problem is similar to the "bias in, bias out" problem for machine learning (e.g., Fazelpour and Danks 2021). Because machine learning algorithms are trained in our biased preferences, they can deliver results that inherit the bias in society. Here the test-taker is inheriting the bias in society against the neurodivergent, harming themselves if they are honest (see also Timmons 2021).

12. Here we have in mind the literature stimulated by Wakabayashi, Baron-Cohen, and Wheelwright (2006).

13. Lydia Brown also argues that AI audits for bias are poorly suited for eliminating bias against disability (see Brown, Shetty, and Richardson 2020).

14. CVS agreed to remove items related to mental health, such as "you change from happy to sad without any reason," "you get angry more often than nervous," and "your moods are steady from day to day."

15. Thanks to two fantastic anonymous referees, Richard Arneson, Jonathan Cohen, Liza Perkins Cohen, UCLA's History and Philosophy of Science Reading Group, and UC San Diego's Neurodiversity in Tech group, especially Leanne Chukoskie and Pamela Cosman. The authors are grateful for financial support from the National Science Foundation's 2019 Future of Work grant FW-HTF-RL: Neurodiversity in Tech: Using Interactive Decision Theory and Augmented Reality to Enable Employment for Adults with Autism Spectrum Disorder.

REFERENCES

Alkhaldi, Rabi Samil, Elizabeth Sheppard, and Peter Mitchel. 2019. "Is There a Link Between Autistic People Being Perceived Unfavorably and Having a Mind That is Difficult to Read?" *Journal of Autism Developmental Disorder* 49 (10): 3973–83. https://doi.org/10.1007/s10803-019-04101-1.

Ameri, Mason, Lisa Schur, Meera Adya, F Scott Bentley, Patrick McKay, and Douglas Kruse. 2018. "The Disability Employment Puzzle: A Field Experiment on Employer Hiring Behavior." *Industrial & Labor Relations Review* 72 (2): 329–64. https://doi.org/10.1177%2F0019793917717474.

Arneson, Richard J. 2006. "What Is Wrongful Discrimination?" *San Diego Law Review* 43 (4): 775–808.

Austin, Elizabeth J. 2005. "Personality Correlates of the Broader Autism Phenotype as Assessed by the Autism Spectrum Quotient (AQ)." *Personality and Individual Differences* 38 (2): 451– 60. https://doi.org/10.1016/j.paid.2004.04.022.

Barnett, Claire. 2020. "Why Autistic Unemployment Is So High." TEDx Talks Vanderbilt University Video. YouTube. https://www.youtube.com/watch?v=FVZu557_k04&t=16s.

Barrick, Murray R., Michael K. Mount, and Timothy A. Judge. 2001. "Personality and Performance at the Beginning of the New Millennium: What Do We Know and Where Do We Go Next?" *International Journal of Selection and Assessment* 9 (1–2): 9–30. https://doi.org/10.1111/1468-2389.00160.

Bates, M. and Hughes, D. J. 2017. "Using Personality Questionnaires for Selection." In *The Wiley Blackwell Handbook of the Psychology of Recruitment, Selection and Retention,* edited by Harold W. Goldstein, Elaine D. Pulakos, Jonathan Passmore, and Carla Semedo, 113–50. Chichester: Wiley Blackwell.

Brown, Lydia X. Z., Ridhi Shetty, and Michelle Richardson. 2020. *Report- Algorithm-driven Hiring Tools: Innovative Recruitment or Expedited Disability Discrimination?* Washington, DC: Center for Democracy & Technology. https://cdt.org/insights/report- algorithm-driven-hiring-tools-innovative-recruitment-or-expedited-disability-discrimination/.

Butcher, J. N., Braswell, L., & Raney, D. 1983. "A Cross-Cultural Comparison of American Indian, Black and White Inpatients on the MMPI and Presenting Symptoms." *Journal of Consulting and Clinical Psychology* 51 (4): 587–94. https://doi.org/10.1037//0022-006x.51.4.587.

Butcher, Sarah. 2021. "Goldman Sacs Explain Why it's Hiring Contractors with Autism." eFinancialCareers. January 7, 2021. https://www.efinancialcareers.com/news/2021/01/goldman-sachs-autism.

Cage, Eilidh, Elizabeth Pellicano, Punit Shah, and Geoffrey Bird. 2013. "Reputation Management: Evidence for Ability but Reduced Propensity in Autism." *Autism Research* 6 (5): 433–42. https://doi.org/10.1002/aur.1313.

Carrigan, Martin. 2007. "Pre-Employment Testing—Prediction of Employee Success and Legal Issues: A Revisitation of Griggs V. Duke Power." *Journal of Business & Economics Research* 5 (8): 35–44. https://doi.org/10.19030/jber.v5i8.2567.

Carroll, Robert Todd. 2011. *The Skeptic's Dictionary: A Collection of Strange Beliefs, Amusing Deceptions, and Dangerous Delusions*. New York: John Wiley & Sons.

Cavico, Frank J., Bahaudin G. Mujtaba, Eleanor Lawrence, and S. C. Muffler. "Personality Tests in Employment: A Continuing Legal, Ethical, and Practical Quandary." *Advances in Social Sciences Research Journal* 2 (3): 60–84. https://nsuworks.nova.edu/hcbe_facarticles/462/.

Creech, William. 1966. Psychological Testing and Constitutional Rights. *Duke Law Journal* 2: 332–71. https://doi.org/10.2307/1371534.

Cunningham, L. 2012. "Myers-Briggs: Does it Pay to Know your Type?" *The Washington Post*. December 14, 2012. https://www.washingtonpost.com/national/on-leadership/myers-briggs-does-it-pay-to-know-your-type/2012/12/14/eaed51ae-3fcc-11e2-bca3-aadc9b7e29c5_story.html.

DeArmond, Elizabeth. 2012. "To Cloak the Within: Protecting Employees from Personality Testing." *DePaul Law Review* 61 (4): 1129–84.

De Pauw, Sarah S. W., and Ivan Mervielde. 2010. "Temperament, Personality and Develop- mental Psychopathology: A Review Based on the Conceptual Dimensions Underlying Childhood Traits." *Child Psychiatry and Human Development* 41 (3): 313–29. https://doi.org/10.1007/s10578-009-0171-8.

Doris, John M. 2002. *Lack of Character: Personality and Moral Behavior*. Cambridge: Cambridge University Press.

Dulaney, Chelsey. 2015. "Target Pays $2.8 Million to Resolve Discrimination: Charge Alleged that Target Discriminated Based on Race and Sex During Hiring Process." *Wall Street Journal*. August 24, 2015. https://on.wsj.com/3mPtwZ5.

EEOC v. Abercrombie & Fitch Stores, Inc., 135 S. Ct. 2028 (U.S. 2015).

Emre, Merve. 2018. *The Personality Brokers: The Strange History of Myers-Briggs and the Birth of Personality Training*. New York: Doubleday.

Fazelpour, Sina, and David Danks. 2021. "Algorithmic Bias: Senses, Sources, Solutions." *Philosophy Compass* 16 (8) 1–16. https://doi.org/10.1111/phc3.12760.

Foster, Deborah, and Victoria Wass. 2013. "Disability in the Labour Market: An Exploration of Concepts of the Ideal Worker and Organisational Fit that Disadvantage Employees with Impairments." *Sociology* 47 (4): 705–21. https://doi.org/10.1177/0038038512454245.

Gibby, Robert E., and Michael J. Zickar. 2008. "A History of the Early Days of Personality Testing in American Industry: An Obsession with Adjustment." *History of Psychology* 11 (3): 164–84. https://doi.org/10.1037/a0013041.

Givens, Alexandra Reeve, Hilke Schellmann, and Julia Stoyanovich. 2021. "We Need Laws to Take on Racism and Sexism in Hiring Technology." *New York Times*. March 17, 2021. https://www.nytimes.com/2021/03/17/opinion/ai-employment-bias-nyc.html.

Goldberg, L. 1993. "The Structure of Phenotypic Personality Traits." *American Psychologist* 48 (1): 26–34. https://doi.org/10.1037/0003-066X.48.1.26.

Goldhill, Olivia. 2019. "We Took the World's Most Scientific Personality Test—and Discovered Unexpected Sexist Results." *Quartz*. February 11, 2018. https://qz.com/1201773/we- took-the-worlds-most-scientific- personality-test-and-discovered-unexpectedly- sexist-results/.

Griffiths, Amy-Jane, Cristina Giannantono, Amy Hurley-Hanson, and Donald Cardinal. 2016. "Autism in the Workplace: Assessing the Transition Needs of Young Adults with Autism Spectrum Disorder." *Journal of Business and Management* 22 (1): 5–22.

Hawkins, Tim, dir. 2021. *Persona: The Dark Truth Behind Personality Test.* Film. HBO Max. https://www.hbomax.com/feature/urn:hbo:feature:GYC1 puQhu1cLCwgEAAAA0.

Hellman, Deborah, and Sophia Reibetanz Moreau, eds. 2013. *Philosophical Foundations of Discrimination Law.* Oxford: Oxford University Press.

Hendricks, Dawn. 2010. "Employment and Adults with Autism Spectrum Disorders: Challenges and Strategies for Success." *Journal of Vocational Rehabilitation* 32 (2): 125–34. doi:10.3233/JVR-2010-0502

Karraker v. Rent-A-Center, Inc., 411 F.3d 831, 837. (U.S. 2005).

Lippert-Rasmussen, Kasper. 2014. "Indirect Discrimination is Not Necessarily Unjust." *Journal of Practical Ethics* 2 (2): 33–57.

Lodi-Smith, Jennifer, Jonathan D. Rodgers, Sara A. Cunningham, Christopher Lopata, and Marcus L. Thomeer. 2018. "Meta-Analysis of Big Five Personality Traits in Autism Spectrum Disorder." *Autism* 23 (3): 556–65. https://doi.org/10.1177/1362361318766571.

Lorenz, Timo, Cora Frischling, Raphael Cuadros and Kathrin Heinitz. 2016. "Autism and Overcoming Job Barriers: Comparing Job-Related Barriers and Possible Solutions In and Outside of Autism-Specific Employment." *PloS ONE* 11 (1): e0147040. doi:10.1371/journal.pone.0147040.

Maras, Katie, Jade Eloise Norris, Jemma Nicholson, Brett Heasman, Anna Remington, and Laura Crane, 2021. "Ameliorating the Disadvantage for Autistic Job Seekers: An Initial Evaluation of Adapted Employment Interview Questions." *Autism* 25 (4), 1060–75. https://doi.org/10.1177/1362361320981319.

McCrae, R. R., and O. P. John. 1992. "An Introduction to the Five-Factor Model and Its Applications." *Journal of Personality* 60: 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x.

McCreary, C., and E. Padilla. 1977. "MMPI Differences Among Black, Mexican-American, and White Male Offenders." *Journal of Clinical Psychology* 33 (1): 171–2.

Miller, Christian B. 2021. "Empirical Approaches to Moral Character" *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Spring 2021. https://plato.stanford.edu/entries/moral-character-empirical/.

Moreau, Sophia. 2010. "What is Discrimination?" *Philosophy and Public Affairs* 38: 143–79. https://doi.org/10.1111/j.1088-4963.2010.01181.x.

Morgeson, Frederick P., Michael A. Campion, Robert L. Dipboye, John R. Hollenbeck, Kevin Murphy, and Neal Schmitt. 2007. "Are We Getting Fooled Again? Coming to Terms with Limitations in the Use of Personality Tests for Personnel Selection." *Personnel Psychology* 60 (4): 1029–49. https://doi.org/10.1111/j.1744-6570.2007.00100.x.

Mulvihill, Maureen. E. 2006. "*Karraker v. Rent-A-Center*: Testing the Limits of the ADA, Personality Tests, and Employer Pre-employment Screening." *Loyola University Chicago Law Journal* 37 (4): 865–910.

Mutjaba, Bahaudin G. 2015. "Personality Tests in Employment: A Continuing Legal, Ethical, and Practical Quandary." *Advances in Social Sciences Research* 2 (4): 60–84. https://doi.org/10.14738/assrj.23.1004.

Myers, Isabel Briggs, Mary H. McCaulley, Naomi L. Quenk, and Allen L. Hammer. 1998. *The MBTI® Manual: A Guide to the Development and Use of the Myers- Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.

National Research Council. 1991. *In the Mind's Eye: Enhancing Human Performance*. Washington, DC: The National Academies Press. https://doi.org/10.17226/1580.

Nevins, Jennifer Leonard. 2005. "Measuring the Mind: A Comparison of Personality Testing to Polygraph Testing in the Hiring Process." *Dickson Law Review* 109 (3): 857–83.

Oesch, Taryn. 2019. "Autism at Work: Hiring and Training Employees on the Spectrum." *Society for Human Resource Management*. August 19, 2019. https://www.shrm.org/resourcesandtools/hr-topics/behavioral-competencies/global-and-cultural-effectiveness/pages/autism-at-work-hiring-and-training-employees-on-the-spectrum.aspx

Ones, Deniz S., Stephan Dilchert, Chockalingam Viswesvaran, Timothy A. Judge. 2007. "In Support of Personality Assessment in Organizational Settings." *Personnel Psychology* 60 (4): 995–1027. https://doi.org/10.1111/j.1744-6570.2007.00099.x.

Ozonoff, Sally, Nicanor Garcia, Elaine Clark, and Janet E. Lainhart. 2005. "MMPI-2 Personality Profiles of High-Functioning Adults with Autism Spectrum Disorders." *Assessment* 12 (1): 86–95. https://doi.org/10.1177/1073191104273132.

Paul, Annie Murphy. 2005. *The Cult of Personality Testing: How Personality Tests Are Leading Us to Miseducate Our Children, Mismanage Our Companies, and Misunderstand Ourselves*. New York: Free Press.

Paul, Annie Murphy. 2016. "Personality Tests are Popular, but Do They Capture the Real You?" *NPR*. June 25, 2016. https://www.npr.org/sections/health-shots/2016/06/25/483108905/personality-tests-are-popular-but-do-they-capture-the-real-you.

Paunonen, Sampo V., Geoffrey Haddock, Friedrich Forsterling, and Mirja Keinonen. 2003. "Broad Versus Narrow Personality Measures and the Prediction of Behaviour Across Cultures." *European Journal of Personality* 17 (6): 413–33. https://doi.org/10.1002/per.496.

Paunonen, Sampo V., and Douglas N. Jackson. 2000. "What Is Beyond the Big Five? Plenty!" *Journal of Personality* 68 (5): 821–35. https://doi.org/10.1111/1467- 6494.00117.

Pittenger, David. 2005. "Cautionary Comments Regarding the Myers-Briggs Type Indicator." *Consulting Psychology Journal: Practice and Research* 57 (3): 210–21. https://doi.org/10.1037/1065-9293.57.3.210.

Radford, Benjamin. 2011. "Scorpios Need Not Apply: Zodiac Signs Inspire Job Bias." *Live Science*. December 2, 2011. https://www.livescience.com/17291-astrology-job-discrimination-china.html.

Robinson, Erin, Laura Hull, and K. V. Petrides. 2020. "Big Five Model and Trait Emotional Intelligence in Camouflaging Behaviours." *Personality and Individual Differences* 152: 109565. https://doi.org/10.1016/j.paid.2019.109565.

Roux, Anne, Jessica Rast, Julianna Rava, Kristy Anderson, and Paul Shattuck. 2015. *National Autism Indicators Report: Transition into Young Adulthood.* Philadelphia, PA: Life Course Outcomes Research Program, A.J. Drexel Autism Institute, Drexel University. https://drexel.edu/~/media/Files/autismoutcomes/publications/National%20Autism%20Indicators%20Report%20-%20July%202015.ashx.

Sarrett, Jennifer. 2017. "Interviews, Disclosures, and Misperceptions: Autistic Adults' Perspectives on Employment Related Challenges." *Disability Studies Quarterly* 37 (2). https://dsq-sds.org/article/view/5524/4652.

Scholes, Vanessa. 2014. "You are Not Worth the Risk: Lawful Discrimination in Hiring." *RMM Journal* 5: 13–29.

Scholz, Frederike. 2020. "Taken for Granted: Ableist Norms Embedded in the De-sign of Online Recruitment Practices." In *The Palgrave Handbook of Disability at Work,* edited by Sandra L. Fielden, Mark E. Moore, and Gemma L. Bend, 451–69. London: Palgrave Macmillian. https://doi.org/10.1007/978-3-030-42966-9_26.

Schriber, Roberta A., Richard W. Robins, and Marjorie Solomon. 2014. "Personality and Self-insight in Individuals with Autism Spectrum Disorder." *Journal of Personality and Social Psychology* 106 (1): 112–30. https://doi.org/10.1037/a0034950.

Schwartzman, Benjamin C., Jeffrey Wood, and Steven Kapp. 2016. "Can the Five Factor Model of Personality Account for the Variability of Autism Symptom Expression? Multivariate Approaches to Behavioral Phenotyping in Adult Autism Spectrum Disorder." *Journal of Autism and Developmental Disorders* 46 (1): 253–72. https://doi.org/10.1007/s10803- 015-2571-x.

Shattuck, Paul T., Sarah Carter Narendorf, Benjamin Cooper, Paul R. Sterzing, Mary Wagner, and Julie Lounds Taylor. 2012. "Postsecondary Education and

Employment Among Youth with an Autism Spectrum Disorder." *Pediatrics* 129 (6): 1042–49. https://doi.org/10.1542/peds.2011-2864.

Solomon, Calvin. 2020. "Autism and Employment: Implications for Employers and Adults with ASD." *Journal of Autism and Developmental Disorders* 50 (11): 4209–17. https://doi.org/10.1007/s10803-020-04537-w.

Stabile, Susan J. 2002. "The Use of Personality Tests as a Hiring Tool: Is the Benefit Worth the Cost?" *University of Pennsylvania Journal of Labor and Employment Law* 4 (2): 279–313.

Stefan, Susan. 2002. *Hollow Promises: Employment Discrimination Against People with Mental Disabilities*. Washington, DC: American Psychological Association.

Stein, Randy, and Alexander B. Swan. 2019. "Evaluating the Validity of Myers-Briggs Type Indicator Theory: A Teaching Tool and Window into Intuitive Psychology." *Social and Personality Psychology Compass* e12434. https://doi.org/10.1111/spc3.12434.

Timmons, Kelly Cahill. 2021. "Pre-Employment Personality Tests, Algorithmic Bias, and the Americans with Disabilities Act." *Penn State Law Review* 125 (2): 389–452.

Trull, Timothy J., David J. Useda, Paul T. Costa, Jr., and Robert R. McCrae. 1995. "Comparison of the MMPI-2 Personality Psychopathology Five (PSY-5), the NEO-PI, and the NEO-PI-R." *Psychological Assessment* 7 (4): 508–16. https://doi.org/10.1037/1040-3590.7.4.508.

Wakabayashi, Akio, Simon Baron-Cohen, and Sally Wheelwright. 2006. "Are Autistic Traits an Independent Personality Dimension? A Study of the Autism-Spectrum Quotient (AQ) and the NEO-PI-R." *Personality and Individual Differences* 41 (5): 873–83. https://doi.org/10.1016/j.paid.2006.04.003.

Whelpley, Christopher E., and Cynthia P. May. 2022. "Seeing is Disliking: Evidence of Bias Against Individuals with Autism Spectrum Disorder in Traditional Job Interviews." *Journal of Autism Development Disorders* (March 2022). https://doi.org/10.1007/s10803-022-05432-2.

Williams, David M. 2010. "Theory of Own Mind in Autism: Evidence of a Specific Deficit in Self-Awareness." *Autism* 14 (5): 474–94. https://doi.org/10.1177/1362361310366314.

Winterhalter, Elizabeth. 2014. "ISTJ? ENFP? Careers Hinge on a Dubious Personality Test." *Boston Globe*. August 31, 2014. https://www.boston-globe.com/opinion/2014/08/30/istj-enfp-careers-hinge-dubious-personality-test/8ptUGXhu6DndFdjCngcxSN/story.html.

Weber, Lauren and Elizabeth Dwoskin. 2014. "Are Workplace Personality Tests Fair?" *Wall Street Journal*. September 29, 2014. https://www.wsj.com/articles/are-workplace-personality-tests-fair-1412044257.

Yu, Andrea. 2020. "Can Personality Testing Improve Retention and Reduce Turnover? These Specialists Think So." *The Globe and Mail.* January 24, 2020. https://www.theglobeandmail.com/business/careers/article-can-personality-testing-improve-retention-and-reduce-turnover-these/.