Screening Out The Neurodiverse



Jada Wiggleton-Little and Craig Callender
Philosophy, Institute for Practical Ethics; UC San Diego

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...



Lydia Brown In **Persona: The Truth About Personality Tests**

Trends On Collision Courses

- 1. Terrible unemployment and under-employment among those with ASD
- 2. More and more people are diagnosed with ASD each year. 2.2% of population and growing. 50,000 teens aging into adulthood each yr.
- 3. Personality pre-employment screening tests are a huge business. \$2 billion industry. 80% of Fortune 500 companies use them.
- [4. Personality tests are moving into the AI revolution.]
- [5. The pandemic has moved more and more jobs remote, putting pressure on what qualities counts as necessary for job performance.]

Plan

- 1. Personality Tests
- 2. Screening Out the Neurodiverse
- 3. Is Discrimination Based on Personality Wrongful?
- 4. Problems with Personality Tests
- 5. Recommendations

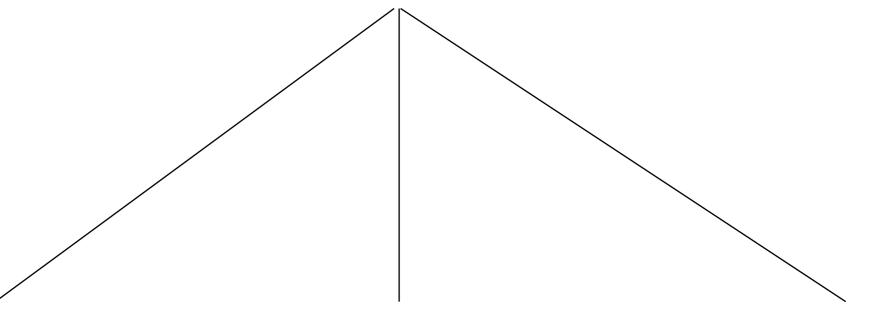


1. Personality Tests

Woodworth Personality Data Sheet: "The most important recommendation to be made is that of rigidly excluding insane, feebleminded, psychopathic and neuropathic individuals from the forces which are to be sent to France and exposed to the terrific stress of modern war."

Moved to industry. Always ethical questions. According to Zickar, employers believed "that people who advocated for labor unions were people who were unsettled and neurotic themselves."

Personality Tests



Myers-Briggs M.B.T.I. Minnesota Multiphasic Personality Inventory

Big Five



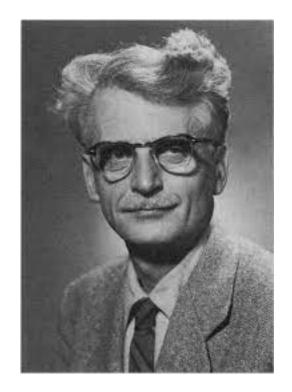
Myers-Briggs

Developed in 1917, published in handbook in 1944. Based on Jungian theory of types. It sorts people into one of 16 personality types, e.g., INFP. The company claims that it is "used by more than 88 percent of Fortune 500 companies in 115 countries, and available in 29 languages" (Myers-Briggs, 2020)

Minnesota Multiphasic Personality Inventory



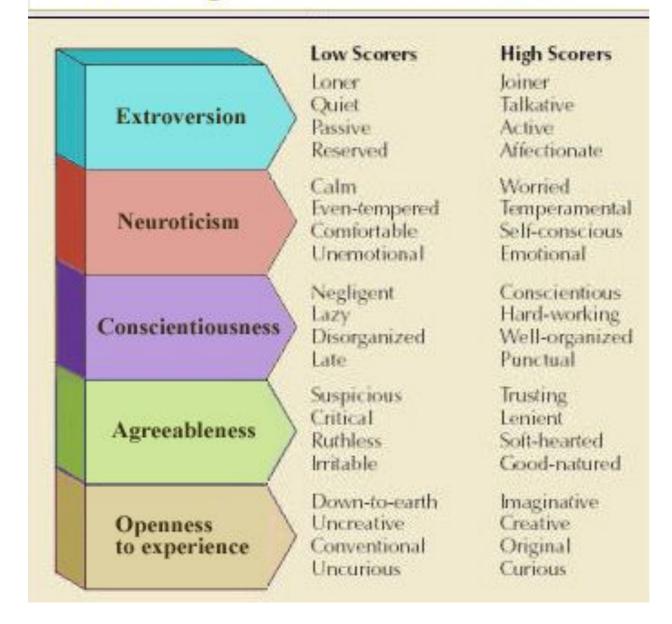
J.C. Mckinley



Stark Hathaway

Developed in 1939 to sort mental illnesses in a hospital, it asks 537 true-false questions and scores test-takers along ten different clinical scales, e.g., schizophrenia, hysteria. The company 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, confined now to certain types of employment like law enforcement.

The Big Five dimension traits



Based on factor analysis of either natural language or psychometric data

Big Five

or

Five Factor

Controversies

Pseudoscience

MBTI fails every way a test can fail.



- Background theory is false
- No internal reliability (e.g., 5 wks later, half are given different type)
- Not valid (83% of differences on the questionnaire for 1291 students could not be accounted for by MBTI); many aspects of personality left out; many not statistically ind of one another
- No connection to job performance or job satisfaction

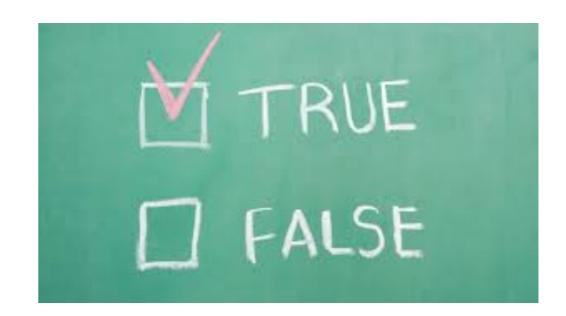
Comparing MBTI to astrology, Paul 2016 describes it as "a Carl Jung-inspired load of nonsense engineered to make everyone who takes it feel good about themselves."

(Interestingly, there have been cases where employers pre-screened applicants based on astrological sign. In 2009 an Austrian insurance company ran an ad declaring a preference for Capricorn, Taurus, Aquarius, Aries and Leo applicants.)

Controversies

Privacy

MMPI used to ask:



"I have never indulged in any unusual sex practices"

"I have difficulty in starting or holding my bowel movements."

Federal employees were asked:

"Christ performed miracles"

"Many of my dreams are about sex matters"

The tests ask for disclosure of what many consider the private sphere. This concern led to a Congressional investigation in 1954/65 that almost made MMPI illegal to use in screening for federal jobs.

YES, I BELIEVE I AM BEING FOLLOWED

AT the Departments of State, Labor, and Health, Education & Welfare, the Export-Import Bank, the Peace Corps and, of all things, the Bonneville Power Administration, job applicants are tested for their personality as a prerequisite of employment. The tests differ, but the most widely used and the one upon

which the others are based is the Minnesota Multiphasic Personality Inventory (MMPI), a 566-question, true-false quiz developed at the University of Minnesota. The test has recently come under strong congressional criticism on the ground that it requires answers that are nobody else's business. Sample questions:

T F		TF		T	F	
	Evil spirits possess me at times.		I have been disappointed in love.			I am a special agent of God.
	I have had very peculiar and		I dream frequently about things			When a man is with a woman he
	strange experiences.		that are best kept to myself.			is usually thinking about things
	I have never been in trouble be-		I believe there is a Devil and a			related to her sex.
	cause of my sex behavior.		Hell in afterlife.	-		I have never seen things dou-
	My soul sometimes leaves my body.	шш	Someone has control over my mind.			bled (that is, an object never looks like two objects to me
	I have not lived the right kind		Once in a while I laugh at a dirty			without my being able to make
	of life.		joke.			it look like one object).
	Sometimes I am strongly attract-		Many of my dreams are about			Dirt frightens or disgusts me.
	ed by the personal articles of	-	sex matters.	님	Н	I dislike to take a bath.
	others, such as shoes, gloves, etc.,		I have strange and peculiar	H	H	I like mannish women.
	so that I want to handle or steal		thoughts.	-		There is something wrong with
	them though I have no use for		Bad words, often terrible words,			my sex organs. I practically never blush.
	them. At times I have a strong urge		come into my mind and I cannot get rid of them.	ŏ	H	I am not afraid of picking up
	to do something harmful or		I feel uneasy indoors.		_	a disease or germs from door
	shocking.		A windstorm terrifies me.			knobs.
	I believe women ought to have		I deserve severe punishment for			I would like to hunt lions in
	as much sexual freedom as men.		my sins.			Africa.
	I believe I am being followed.		At times I think I am no good			Several times a week I feel as if
	I have never indulged in any		at all.			something dreadful is about to
	unusual sex practices.		I have had some very unusual re-			happen.
	Sometimes I feel as if I must		ligious experiences.		П	I never attend a sexy show if I
	injure myself or someone else.		Usually I would prefer to work			can avoid it. I like repairing a door latch.
	The top of my head sometimes		with women. I am bothered by people out-	Н	H	I sometimes feel that I am about
	feels tender. I am worried about sex matters.		side, on streetcars, in stores, etc.,		Н,	to go to pieces.
HH	I believe my sins are unpardon-		watching me.			A large number of people are
	able.	ПП	I have never seen a vision.			guilty of bad sexual conduct.
	I brood a great deal.		I am fascinated by fire.			I like movie love scenes.

Controversies

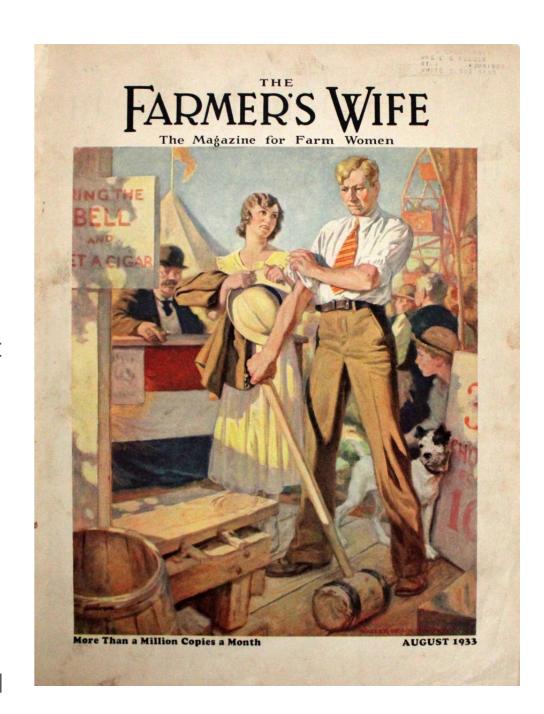
Bias and Discrimination

MMPI and the "Minnesota Normals":

On the original, the normative sample consisted of 724 entirely white, mostly rural Minnesotans (Butcher et al. 1983). As early as 1977 McCreary and Padilla showed that Black and Hispanic offenders scored higher on negative traits than whites (see also Butcher et al 1983).

MMPI-2 uses 2600 more rep sample from 1989 — but demographic shift, of course!

The choice of normative sample will always cause this kind of worry. Even Big Five has been accused of sexism (Goldhill 2018)



2. Screening Out The Neurodiverse

Unlike Kyle Behm, who was told he was "red-lighted", Lydia Brown and others in their situation will likely never know in any specific case whether personality tests screened them out.

Bracket MBTI (terrible) and MMPI-2 (clinical)

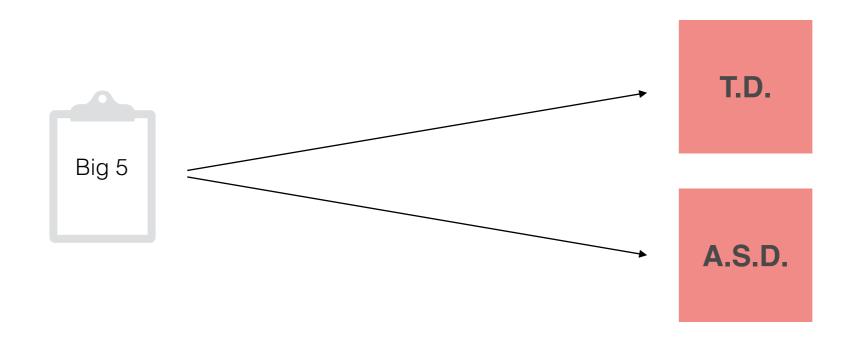
Does Big Five have disparate impact on the neurodiverse?

Background theory: YES

Meta-analysis: YES. Lodi-Smith et al 2018 combine studies to include almost four thousand participants and show that those with ASD score lower on the five traits of Openness, Conscientiousness, Extraversion, Agreeableness, and Emotional Stability. This is in line with past studies, all but one of which found a link between Emotional Stability and ASD.

Screening Out The Neurodiverse

Given your Big Five test results, how well does that predict that you are in the group diagnosed with ASD or in the group thought to be typically developing (TD)?



Schriber et al 2014: scores correctly assigned a person to the ASD vs TD group 70% of the time. Big Five cannot (yet) diagnose ASD, but the traits redundantly code for it. It creates disparate impact upon the neurodiverse like (say) zip code would with racial groups.

The best evidence currently supports Lydia Brown's suspicion. And with 5.5 million like her in the US, she is not alone.

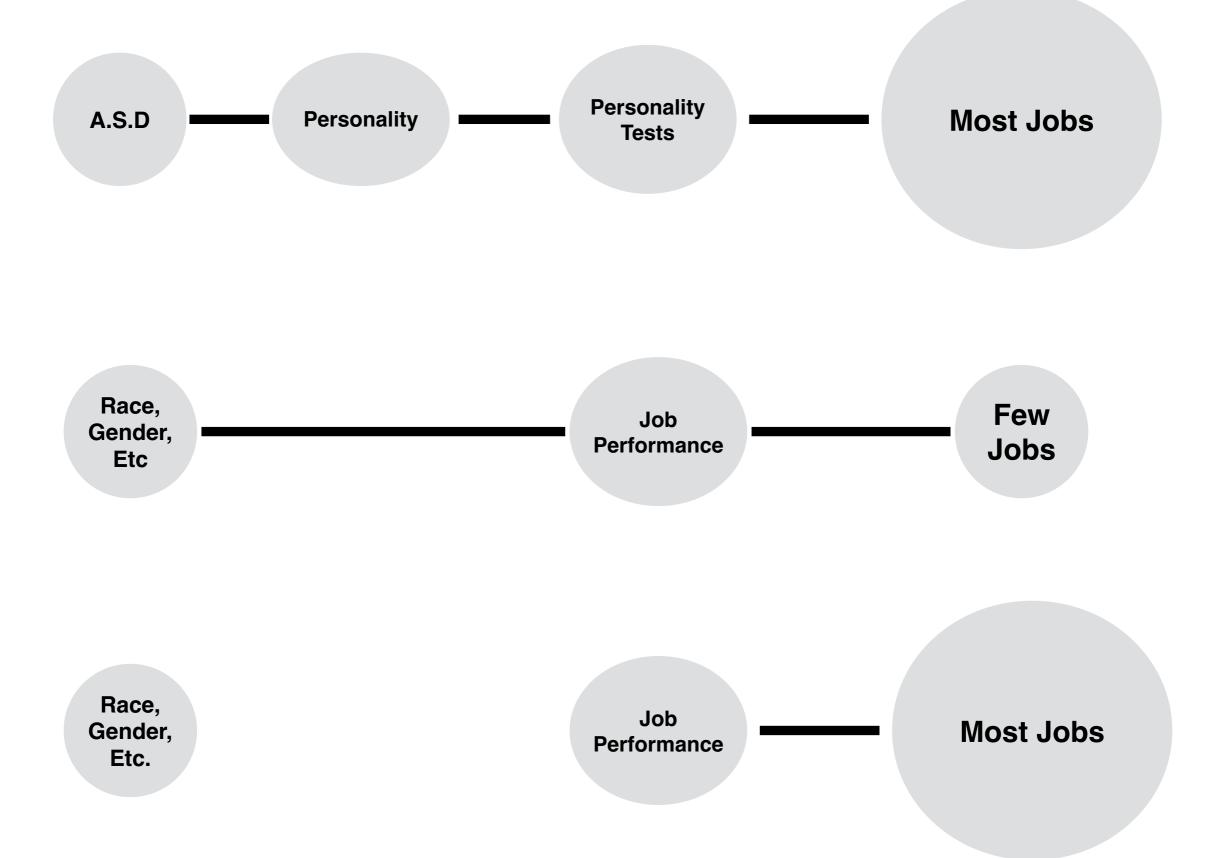
3. Is Discrimination Based On Personality *Wrongful*?

What makes discrimination <u>wrongful</u> is a huge tricky issue. See Hellman and Moreau 2013. There are equality, freedom and consequentialism based theories.

- Is it okay to discriminate based on <u>looks</u>?
- Is it okay to discriminate against smokers?
- Is it okay to discriminate against <u>racial groups</u> for non-racist reasons?

What's easy is if the discrimination is not based on anything tied to job performance, thick or thin. Then FEO is violated.

What's tricky about this case is that personality is thought to be a general enhancer of job performance. Here we have a <u>non-accidental link</u> between group membership in a protected class and a feature thought to be relevant to <u>most</u> jobs.



4. 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 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.

Personality traits are not viewed by most psychologists as a kind of "metaphysical" feature of a person 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 together into "traits" via factor analysis across large groups of neurotypical individuals, not from looking at behavior.

Hard thing to measure with a test and even harder to say that what the tests measure is causal difference-maker to actual behavior on a job...



Donut Shop #1. In this shop the owner does a structured interview with everyone on the short list. Sally strikes the owner as shy and possibly unreliable. The owner wants a cashier who is extraverted and conscientious. The owner does not invite this person to the next stage of the process.

Donut Shop #2. Due to the volume of applicants, the owner asks everyone to take a personality test. Sally scores poorly along the dimensions of Extraversion and Conscientiousness. For the same reasons as above, the owner does not invite this person to the next stage of the process.

General Concerns

Generalizability. Do self-reports cohere with third-party judgements? Connect to actual behavior? Are they reliable? And so on.

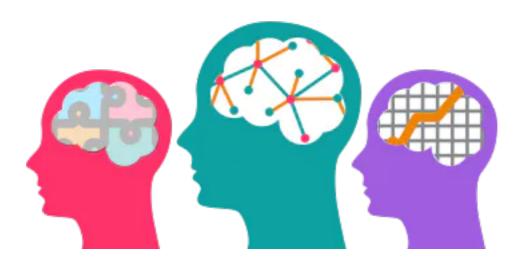
Left Out. Many qualities are not considered, such as moral character, goals, and so on. Even many personality traits are left out, including some that can give better predictive validity than the Big Five model.

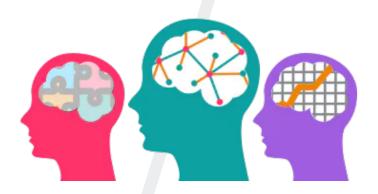
General predictor across jobs? Patience seems good for some jobs, bad for others. Not like cognitive skill (which itself is problematic).

Job performance. Based on decades of studies, Big Five traits account for only about 5-7% of variance in job performance measures (Hughes and Matey 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, Morgenson et al. (2007) argue that such tests should not be used in job selection. However, one can dig into the data and find 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 objective rating, r=0.15 supervisor rating).

The two donut shop cases aren't as symmetrical as they first appear.

The tests set up individuals with ASD to fail in unfair ways. They prey on features of the condition.





- 1. Societal stigma baked into the test. 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.
- I try to be courteous to everyone I meet.
- Some people think of me as cold and calculating.
- I generally try to be thoughtful and considerate.
- I'm not known for my generosity.
- 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.

Notice, with Miller (2021), that <u>half</u> of the items ask about how one is <u>perceived by others</u>, not what one thinks of themself.

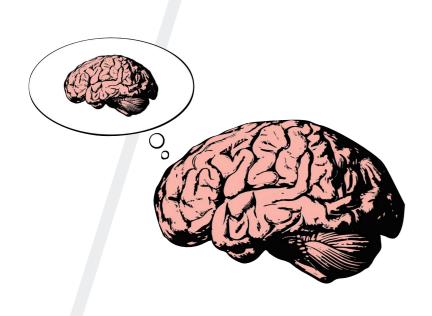


- 1. Societal stigma baked into the test. Consider how "altruism" is determined in McCrae and Costa's Five Factor Theory. This is measured with one's reactions to the following items:
- Some people think I'm selfish and egotistical.
- I try to be courteous to everyone I meet.
- Some people think of me as cold and calculating.
- I generally try to be thoughtful and considerate.
- I'm not known for my generosity.
- 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.

Notice, with Miller (2021), that <u>half</u> of the items ask about how one is <u>perceived by others</u>, not what one thinks of themself.

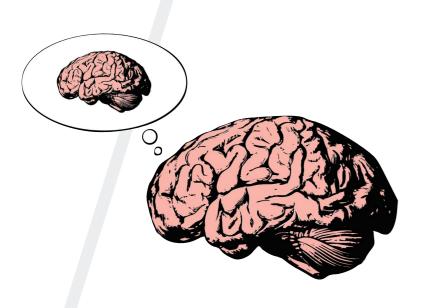


2. Honesty. Put bluntly, everyone fakes their answers, except those with ASD.adults with ASD are less likely to use reputation management compared to neurotypical adults (Cage et al. 2013). Whether this behavior is due to honesty or limited social cognition or both (or other explanations) doesn't matter. Adults with ASD are more likely to provide "true" answers than "good" answers on personality tests, at the expense of potential employment opportunities.

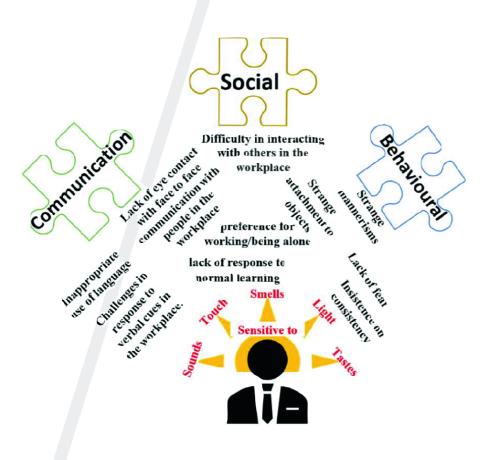


3. Metacognition.

Williams, David M. (2010) Theory of own mind in autism: Evidence of a specific deficit in self-awareness



4. Test taking. Many individuals with ASD struggle to take written tests that take a lot of time. If not relevant to the job, test-taking skills may interfere with result even though it doesn't measure anything apart from the skill to take the test.



5. Accommodation. Personality is not considered something that can be accommodated. Yet the traits that cause the neurodiverse to score badly on personality tests are sometimes features that can be accommodated in the workplace—yet individuals aren't given the chance because they are screened out.



Other worries where we lack evidence

Do results mean the same thing when neurodiverse take test designed for neurotypical?

Do links to job performance generalize from neurotypical to neurodiverse?

Even without considering the neurodiverse, personality tests are on thin ground ethically. For privacy reasons, the US Congress almost banned their use federally in the 1960s. In the 1970s and 80s issues of racial and gender bias and reliability/validity — the latter took out many uses of the polygraph in 1988.

All of these issues and more apply to the neurodiverse. Their privacy is threatened, the tests have poor or unknown connection to job performance for them, and they form a vulnerable group in society. And as we've just shown, the tests unfairly stack the deck against the neurodiverse.

Upshot

Regulation?

Ban?

With advances in data analytics, variables crafted from psychometric data may become comparable or better than what we get from background checks, structured interviews, and so on.

Follow the lead of policies being drafted for ethical Al use? E.g., make vendors provide evidence of "neurodiversity safe"?

Won't work. 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 ASD.

5. Recommendation

Personality tests will have disparate impact upon 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.

Regulation

Demand tighter evidence-based connection to job performance from companies or vendors. For instance, in France, 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

- Some personality tests already contain a blend of ability or skill questions. Others are already targeted at specific jobs.
- Make tests more "ASD safe". Worried about divulging mental disabilities, CVS modified its test in response to the Rhode Island Commission for Human Rights finding mentioned in the introduction. 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."

Vendors may welcome this regulation, as it asks them to do what they're good at.

Regulation

Opt out option

- "Catch 22" of disability disclosure (Stefan 2002). Ameri et al 2017 found that applicants who disclosed a disability received 25% fewer call backs than otherwise identical applicants. But if you don't opt out of the test, then you risk a low score in part because the test preys on your condition.
- Tension between our two suggestions must be navigated. The blind may ask for and expect written tests to be delivered in Braille, but allowing slow typers to opt out of a typing test is not fair to employers or other applicants. The more job related and "ASD safe" the test become, the less reason to allow opting out.

Conclusion

Pre-employment personality screening tests are only one (small) barrier to employment for the neurodiverse. Society needs to confront the problem head-on, with training programs, hiring initiatives, and more. But chipping away at barriers will get more neurodiverse in the workplace. As this becomes more common and employers see them excel in different sorts of jobs, we expect norms will change that make employment recruitment and screening more neurodiverse-friendly.

