The Accuracy of Instant Talent Analytics: A Personal Network Validation

Dr. Tom Janz, I/O Psychologist, Founder: The Talent Analytics Group

Instant Talent Analytics is a new technique that can provide an assessment of an individual without requiring the individual to take a test. It requires no time for the talent—the people from the current or future workforce-- to answer self-report test questions, solve puzzles, play games, react to scenarios, answer interview questions, or "sell me this pen". What gets analyzed has to exist already. Public data mounted by people online includes: LinkedIn profiles, Twitter feeds, authored blogs, stored gameplay and code samples. Accurate instant talent analytics has many use cases ranging from: (1) pre-deal human capital due diligence for corporate acquisitions, (2) fast mapping of pivotal teams into 'high value' vs. 'development-worthy' vs. 'better elsewhere' talent types, (3) locating the best talent for executive search firms to tactfully pry from their client's competitors, (4) fast, cost-effective sorting of all new job applicants in talent acquisition flows, and (5) sales acceleration via optimizing first impression impact with target sales prospects. This paper investigates the construct validity of one of the most reported-on instant talent analytics services, DeepSense. It addresses Instant Talent Analytics validity by correlating my professional judgement on 14 personal characteristics with DeepSense machine-generated scores for a sample of 120 professionals in my personal network.

An active and growing body of field research documents the increasingly credible power of machine learning algorithms to score this type of accessible data on personal characteristics related to personality and ultimately job performance. In **Predicting Personality with Social Media***,* the authors (Golbeck, Robles, and Turner) extracted 17 linguistic usage features from Facebook, used M5'Rules machine learning to produce correlations (from .48 to .65) between predicted and measured scores on the big five personality factors. In **Twenty-five tweets to know you** the authors (Arnoux, Xu, Boyette, Mahmud, Akkiraju, and Sinha, all from IBM) used Gaussian processes with Word Embedding to achieve an average big five correlation of .3, with the best performance for Neuroticism at .45. These results are consistent with the Golbeck, Robles, and Turner study when they deployed Gaussian processes instead of the M5'Rule machine learning method.

This paper is focused on the machine learning-generated algorithms active on the DeepSense persona analytics service, focusing on intentionally public data harvested primarily from LinkedIn and Twitter.

My previous writing on this topic has appeared in the HR Examiner—**1 Short, Shorter, Shortest: Online Tests vs. Social Media Analytics (10/2018)**, and **2 Deep Thinking about Deep Learning (2/2019)**. Those posts provided first looks at external vs. machine scores with small datasets. They also surveyed other published findings and thinking. This post extends that early work by expanding the sample size from 56 to 120 and by adding the big five and DeepSense performance factors. It reveals the correlations between machine-produced and personal expert judgement for: 1] Four DISC factors, 2] Five 'Big five' personality factors, and 3] Seven behavioral factors generated by the data scientists at DeepSense. This time, these correlations rested on a data set of 120 professionals, most from the talent management and recruiting professions.

The Full Norm Sample

The full sample, consisting of all persons I have assessed using DeepSense, includes both persons well known to me and people whom others asked me to analyze.

The total N of persons machine scored was 432, 121 females and 310 males. There are 108 PhDs. This it is not a random sample, but rather people I know well enough to rate on 14 of the 16 characteristics, plus people I don't know, but who were included in projects or promotions.

The Sixteen Personal Characteristics

The Seven DeepSense Performance Characteristics (Means and SDs on the machine scores, N=432)

Factor Name	Description	Mean	SD
1 Attitude and Outlook	Keep a positive attitude even when facing problems and setbacks; Recover quickly from setbacks; Optimistic vs. pessimistic	7.3	.76
2 Need for Autonomy	The need to be free of the command and control of others; Low scores like having others make the tough decisions, know what is expected o them	5.6	.52
3 Team Skills	Work well within teams; Put the best interests of the team over personal interests; Notice and assist struggling team members when possible	6.3	.82
4 General Regard	Earn respect and is generally liked by peers, bosses & customers; Acts with consideration for others, controls impulsive instincts.	6.1	.74
5 Bias for Action	Quickly leap into action when faced with opportunities/problems; Act first and seek forgiveness later instead of gaining team support/approval	7.4	.62
6 Role Stability	Adjust well to the needs of the role ; Remain in the chosen role over time, moving up the ladder in that role vs. changing jobs and roles frequently.	5.1	.79
7 Learning Ability	Quickly master new material, pick up subtle clues and patterns; curious about why things work; always learning new things relevant	5.8	.66

The Four DISC Characteristics

Factor Name	Description	Mean	SD
1 Dominance	Taking control, makes decisions, buys off or coerces others to get their way	5.6	1.3
2 Influence	Gains agreement through relationships, clarifying common interests.	6.7	1.1
3 Steadiness	Maintain predictability—the status quo, shows patience and sympathy	6.1	1.2
4 Compliance	Follow the rules and do what you say you are going to do, be punctual	7.0	1.5

The Big Five Personality Characteristics

Factor Name	Description	Mean	SD
1 Open to Experience	Curious, creative, artistic, risk tolerant, sometimes emotional, adventurous	5.7	1.9
2 Extroversion	Assertive, energetic, gains energy from relationships, values others approval	5.6	1.8
3 Emotional Stability	Initially Neuroticism (now reversed)- Control impulses, remain calm and cool	5.9	2.2
4 Agreeableness	Compassionate and cooperative, trusting and helpful, values collaboration	6.3	1.5
5 Conscientiousness	Organized, dependable, trustworthy, diligent, and honest	6.2	1.8

Intercorrelations among the sixteen machine scored talent factors (N=432)

	Need (Autonomy)	Team over Self	General Regard	Action Orientation	Stability	Writing Ability	D: Dominance	D: Influence	D: Steadiness	D: Calculative	O:OpenExper	O:Extraversion	O: EmotStability	O: Agreeableness	O:Conscientiousness
Positive Resilience	0.009	0.849	0.847	-0.049	0.414	-0.348	-0.343	0.550	0.514	0.388	-0.497	-0.287	-0.113	0.616	0.628
Need (Autonomy)		-0.019	0.035	0.590	-0.448	-0.128	0.331	-0.256	-0.119	0.223	-0.105	0.400	0.039	0.099	-0.111
Team over Self			0.902	-0.184	0.338	-0.431	-0.551	0.709	0.695	0.269	-0.396	-0.310	-0.034	0.527	0.595
General Regard				-0.149	0.290	-0.455	-0.526	0.638	0.713	0.442	-0.425	-0.266	-0.078	0.582	0.616
Action Orientation					-0.509	-0.055	0.706	-0.651	-0.480	0.515	0.015	0.309	0.058	-0.066	-0.181
Stability						0.099	-0.371	0.463	0.260	-0.139	-0.322	-0.289	-0.155	0.340	0.427
Writing Ability							0.271	-0.230	-0.470	-0.322	0.035	0.140	-0.105	-0.229	-0.255
D: Dominance								-0.844	-0.872	0.162	0.113	0.296	0.026	-0.254	-0.369
D: Influence									0.818	-0.123	-0.283	-0.431	-0.083	0.378	0.529
D: Steadiness										0.132	-0.192	-0.282	-0.005	0.368	0.459
D: Calculative											-0.185	0.009	-0.006	0.232	0.202
O:OpenExper												0.053	0.538	-0.567	-0.608
O:Extraversion													-0.041	0.019	-0.390
O: EmotStability														-0.217	-0.357
O: Agreeableness															0.753
O:Conscientiousness															

Of the three sources of measurement factors, the correlations among the DISC factors are extraordinarily high, with Dominance correlating in the -.80s with Influence and Steadiness. Only the Compliance or Calculatedness DISC scale shows independent variance. The average intercorrelation for DISC factors is .61, for Big Five factors is .41, and for the DeepSense performance factors is .145.

The Known Personal Network Sub-sample

From the full dataset of 432 people, 120 were selected based on my history of working or spending non-work time with the person. There were 98 males and 22 females in this sub-set. The next table holds the simple correlations between my personal factor ratings (on a simple 1-10 scale) and the

DeepSense machine-scores for that scale. The DISC scale 'Compliance', the Big Five scale 'Emotional Stability' (Neuroticism reversed), were not rated, since in the larger norm sample they correlated poorly with a simple linear unit-weighted composite of the 16 factor scores. Learning Ability was rated in view of the historically strong role of mental ability in predicting future job performance.

Correlations between personal ratings and DeepSense factors (n=120)**.**

Talent Factor	**r (My Rating, Machine Score) N=120**	**r (DS Factor, Overall Index) N=432**
Attitude and Outlook	.75	.74
Need for Autonomy	.51	-.29
Team Skills	.46	.75
General Regard	.46	.74
Bias for Action	.47	-.50
Role Stability	.47	.62
Learning Ability	.65	-.20
D: Dominance	.45	-.68
D: Influence	.47	.82
D: Steadiness	.49	.68
D: Compliance		.09
B5: Openness to Experience	.53	-.53
B5: Extroversion	.39	-.45
B5: Emotional Stability		-.17
B5: Agreeableness	.43	.61
B5: Conscientiousness	.37	.75

Since the overall career or performance effectiveness of the sample participants was not estimated, the correlation between the individual factor score and a simple, unit-weighted composite of all the factors was used to guide forming an overall Instant Talent Analytics score.

The DISC Compliance (some call Calculatedness) score showed almost no relationship to the simple sum composite. The Big Five Neuroticism factor (reversed so it could be less pejoratively named), produced a low negative correlation, meaning that there is a slight positive relationship between neuroticism and overall score. The Learning Ability performance factor also correlated low and negatively with the simple sum composite. That is most unfortunate considering that general mental ability has performed so strongly and reliability over the past 100 years of published research (See Schmidt, Oh, and Shaffer, 2017). Perhaps larger and higher quality training data sets will boost the power of the deep learning algorithms to relate to the cognitive component of work and life success.

Combining the Factors into a Business Leadership Index

The Disc, Big Five, and performance factor scores available from DeepSense provide a comprehensive profile of personal characteristics that power making stronger first impressions. For talent professionals, a clear picture of top candidates, new coaching clients, and assigned business leaders before meeting them for the first time drives better outcomes. Interpersonally savvy talent professionals have the time to think through **how** to best deliver their value proposition to new contacts, adjusting their style to fit the target person. Without it, they have to figure out style preferences at the same time as making their opening pitch. For sales professionals meeting new client prospects for the first time, they can focus on **what** to sell, having an edge over competitors that have to focus both on **what to sell** and **how to sell it** at the same time.

Sometimes, it's all about making quick, accurate talent decisions vs. getting off to a great first impression. For the current workforce, it can be about a quick mapping to reveal functions, roles, or business units most in need of talent adjustment. Then mapping the current talent into those people best suited for: [1] Self-guided development vs. [2] Coach-guided development, vs. [3] Career counseling. For the future workforce, it's about quickly sorting candidates into the following action categories: [1] Schedule for decision interviews ASAP, [2] Gather and review further validated self-report assessments, and [3], Collect performance or risk confirmation assessments.

Carrying out any of these decision-based tasks requires combining the DeepSense performance factors into a single number. Focusing on the personal factors that correlated with the unit weighted composite of the 16 factors, I created a weighted composite to reflect those factors that correlated significantly (both positively and negatively) with the simple unit-weighted, overall score.

The composite was named —the Business Leadership Index. The 10 biggest drivers were, in order:

1. DISC: Influence
2. Big Five: Conscientiousness
3. DS: Team before Self
4. DS: Positive Resilience
5. DS: General Regard
6. DISC: Dominance (Reversed)
7. DISC: Steadiness
8. DS: Emotional Control
9. Big Five: Agreeableness
10. Big Five: Openness to Experience (Reversed)

This composite, while driven entirely by empirical findings, takes on a distinctly human tone. The reversal on Dominance makes sense. People who insist on winning every battle, forcing those who disagree into submission face increasingly costly resistance over time. The reversal on Openness to Experience was a bit more puzzling. Openness to experience would be a good thing for someone working on creating innovative value, but not so positive when applied to those whose role centers around plan execution, or replication. People high on Openness to Experience can lack a strong focus on execution and results, being too distracted by bright shiny objects that come along.

Extroversion, a positive correlate of sales effectiveness, turns up negative here. Extroverts can spend too much time talking and debating, enjoying the interaction a bit too much. They can end up trying to please everyone, a sure path to reduced effectiveness. A bias for action contributes positively to sales and innovation value, but can leave team members resentful that they were not consulted. It can also lead to unanticipated consequences and errors of omission, and we see a negative weight in the Business Leadership Index. A high need or autonomy can lead to having a difficult time fitting into and harnessing corporate teams and resources, making it harder to scale what might be productive ideas.

When the weighted composite of the DeepSense machine scores was correlated with a similarly weighted composite of the personal network member ratings, the correlation was a substantial .71.

In Conclusion

This study reported the relationship between the author's professional evaluation on 16 personal characteristics of 120 members of his professional network and machine-derived scores on the same factors. The characteristics included scales from two widely used personality frameworks—the DISC and Big Five. An additional seven performance factors created directly by DeepSense, the Instant Talent Analytics service investigated in this study. Correlations between the expert judgments and the machine scores ranged from .35-.75, demonstrating strong construct validity. A linear weighted composite of 14 of the characteristics, judged to index Business Leadership, produced a .71 correlation between the expert scored and machine scored summary scores. The DeepSense results rise substantially above the IBM reported results, and rise slightly above the Golbeck, Robles, and Turner results as well.

Further research is needed into the inter-correlations among the machine-derived scores that show much higher within instrument correlations than is typical for the same constructs measured via multi-item, self-report assessments (The SHL OPQ, Saville Wave, AAI-WBI, Hogan HPI, for example). But that does not make one covariance structure "true" and the other "false". Self-reports on agree-disagree or response frequency items are not verifiable in any meaningful way. They reflect how people subjectively react, but don't tell us how often person A performs action Y. The covariance of self-report scales may not be any more real than the covariance among machine scores.

Limitations to generalizing from this research include the personal and non-random nature of the dataset. Further my ratings were exposed to possible contamination. While I did not memorize, recall, or refer to individual target person DeepSense scores while making my ratings, the validation project was begun after about 25% of the sample had been machine scored. I no doubt recalled the overall standing of some of the target persons while making my professional ratings. This contamination would have the effect of raising the level of correlation above what it would be had all my ratings been made before any of the machine scores were seen. Future research should be conducted on larger, less selective samples. It should also include participant scores on validated psychometric instruments that tap the same constructs as measured by DeepSense. Most importantly, it should include participant performance scores, either on objective measures or supervisory ratings of output value,

References

An Overview of Automated Scoring of Essays. Semire Dikli. (2006) The Journal of Technology, Learning, and Assessment. Volume 5, Number 1 · August 2006.

Computer-based personality judgments are more accurate than those made by humans. Wu Youyou, Michal Kosinski, and David Stillwell (2015) Proceedings of the National Academy of Sciences. January 12, 2015

Deep Thinking about Deep Learning. Tom Janz (2019) HRExaminer, February 2019.

Perceptions of personality in text-based media and OSN: A meta-analysis. Konstantin O. Tskhay, Nicholas O. Rule. (2014) Journal of Research in Personality, 49, 25–30.

Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers. Tal Yarkoni. (2010) Journal of Research in Personality. Jun 1; 44(3): 363–373.

Predicting Personality with Social Media. Golbeck, Robles, and Turner (2011) Alt.chi: Playing Well With OthersMay 7–12, 2011 • Vancouver, BC, Canada

Private traits and attributes are predictable from digital records of human behavior. Michal Kosinski, David Stillwell, and Thore Graepel. (2013) Proceedings of the National Academy of Sciences, April 9, 110 (15) 5802-5805

Short, Shorter, Shortest – Online Tests vs. Candidate Social Media Analytics. Tom Janz (2018) HRExaminer, September 2018.

25 Tweets to Know You: A New Model to Predict Personality with Social Media. Pierre-HadrienArnoux, Anbang Xu, Neil Boyette, Jalal Mahmud, Rama Akkiraju, Vibha Sinha. (2017) IBM Research -Almaden, San Jose, CA, USA

Understanding Personality through Social Media. Yilun Wang. (2015) Department of Computer Science, Stanford University.