RICE

Department of Psychological Sciences
Fred Oswald
Professor and Herbert S. Autrey Chair in Social Science

January 23, 2022

Department of Consumer and Worker Protection
Commissioner Vilda Vera Mayuga
42 Broadway, 9th Floor
New York NY, 10004

RE: Proposed Rules to implement Local Law 144, Automated Employment Decision Tools

Dear Commissioner of the New York City Department of Consumer and Worker Protection (DCWP),

Thank you for the opportunity to share a set of interrelated concerns with the DCWP's Proposed Rules to implement Local Law 144, Automated Employment Decision (updated).

Definitional concerns

1. "Machine learning, statistical modeling, data analytics, or artificial intelligence" - Selection tools would be exempted from this definition if they did not* use cross-validation (required in part iii of the definition), and yet cross-validation is a recommended practice to help ensure, all other things being equal, that selection models are not overfitting data and overstating their benefits.

2. "Simplified output"

(2a) The definition depends on the definition of “machine learning, statistical modeling, data analytics, or artificial intelligence" which itself is problematic (see above);

(2b) The end of the definition states that "[i]t does not refer to the output from analytical tools that translate or transcribe existing text, e.g., convert a resume from a PDF or transcribe a video or audio interview," and yet this output very often produces or contributes meaningfully to a "score...tag or categorization...or ranking." Therefore, it is very unclear, and concerning, why these forms of processed applicant data -- involving some of the most machine-learning-heavy processes -- would be excluded from simplified output.

3. "Scoring rate" - Although the scoring rate is used in machine learning, reporting subgroup scores above the overall median in the selection context is an unorthodox and (at best) indirect way to indicate the performance of an AEDT. For instance, the scoring rate ignores underlying subgroup score distributions and proportions, and the scoring rate does not necessarily reflect the actual effect of selection on the overall distribution of scores.

(continued)

Email foswald@rice.edu • Office 713-412-2147 • Rice University Department of Psychological Sciences–MS 25
6100 Main St. • Houston, TX 77005 • workforce.rice.edu

RICE

Department of Psychological Sciences
Fred Oswald
Professor and Herbert S. Autrey Chair in Social Science

Additional concerns

4. Sampling distortions - The independent bias audit is based on "historical data regarding applicant selection that the vendor has collected from multiple employers." Thus, information from the audit report is a gross aggregate: e.g., across wide-ranging industries, jobs, and specific applicant populations. The selection rates and impact ratios in employer settings may also vary greatly, and yet they will be averaged away in the vendor audit report

5. Statistical inaccuracies - Setting aside issue #4 above, small sample sizes are associated with greater inaccuracies in the selection rates and impact ratios reported, yet the audit report does not come with any indication of these inaccuracies (e.g., 95% confidence intervals). Note that because intersectional samples refine the overall sample, they can be more important yet will be even less accurate than higher-level aggregates (a potential tradeoff).

Biographical information

My professional background involves 23+ years of psychometrically developing and evaluating selection measures (e.g., job knowledge and skills, personality, interests) used within organizational, military, and educational settings. In the last several years, I have engaged in a range of activities relevant to the use of AI assessments in employment settings: e.g., teaching machine learning workshops and seminars, publishing peer-reviewed research articles, and serving on technical advisories (e.g., with the Society for Industrial-Organizational Psychology, and the Institute for Workplace Equality). Current roles of relevance include serving as Chair of the Board on Human-Systems Integration (BOHSI) at the National Academies (which recently produced a report on human-AI teams) and member of the National AI Advisory Committee (NAIAC), which advises the Secretary of Commerce and the President. Comments expressed herein are solely my own, and do not represent the views of these aforementioned groups.

With appreciation for these continued important deliberations, please reach out any time to discuss further.

Sincerely,

Frederick L. Oswald
Professor and Herbert S. Autrey Chair in Social Sciences
Department of Psychological Sciences
Rice University

Email foswald@rice.edu • Office 713-412-2147 • Rice University Department of Psychological Sciences–MS 25
6100 Main St. • Houston, TX 77005 • workforce.rice.edu