1. Most importantly, the proposed rule amendments could benefit from further clarifications to state that human guided processes (e.g., traditional psychometric measurement techniques) are not intended to fall under the definition. Otherwise, employers may fall back into relying upon inherent human bias (who you know, similar-to-me) at the expense of job-related assessment tools (such as job knowledge exams or self-report personality inventories). To that end, I would recommend the following (blue highlighted) language be added to clarify the definition of “machine learning, statistical modeling, data analytics, or artificial intelligence.”

Machine learning, statistical modelling, data analytics, or artificial intelligence. “Machine learning, statistical modelling, data analytics, or artificial intelligence” means a group of mathematical, computer- based techniques:

1. that generate a prediction, meaning an expected outcome for an observation, such as an assessment of a candidate’s fit or likelihood of success, or that generate a classification, meaning an assignment of an observation to a group, such as categorizations based on skill sets or aptitude; and
2. for which a computer at least in part identifies the inputs, the relative importance placed on those inputs, and other parameters for the models in order to improve the accuracy of the prediction or classification; and
3. for which the inputs and parameters are refined through cross-validation or by using training and testing data.

Techniques wherein human judgment determines the inputs, the relative importance placed on those inputs, and other parameters, do NOT constitute “machine learning, statistical modelling, data analytics, or artificial intelligence,” even if statistical information is used to help inform human judgment.

1. There remains confusion regarding numerous scenarios posed during the previous comment period. I would recommend that the agency issue a questions-and-answers document to clarify outstanding questions posed during both comment periods. An example follows which would benefit from such a questions-and-answers document to guide employers.

*“The proposed definition for “automated employment decision tool” could benefit from further explanation and illustrations, particularly with regard to the phrase “or to use a simplified output as one of a set of criteria where the output is weighted more than any other criteria in the set.” What about a situation where a candidate score is generated that is based 90% on structured panel interview ratings (numerical ratings averaged) and 10% on a tool that uses machine learning? Or a score that is based 90% on a job knowledge exam (multiple choice test score) and 10% on a tool that uses machine learning?”*

1. The bias audit section lacks clarity, particularly around what statistics must be reported in varying scenarios. For example, the distinction between the two statements below is not clear.

“Where an AEDT selects candidates for employment or employees being considered for promotion to move forward in the hiring process or classifies them into groups, a bias audit must, at a minimum:”

*Versus*

“Where an AEDT scores candidates for employment or employees being considered for promotion, a bias audit must, at a minimum:”

Is the distinction here “selects” and “classifies” versus “scores?” If so, that distinction is unclear and would benefit from elaboration and/or further examples.

1. The use of intersectional category reporting for impact ratios is not called for by the federal Uniform Guidelines on Employee Selection Procedures and therefore adds administrative burden to employers. Further, sample sizes can dwindle with intersectional analysis and result in unstable results. Similarly, the published, peer-reviewed literature on assessments rarely reports intersectional group difference results, leading to difficulty with interpreting findings in the context of existing scientific knowledge. I would strongly advise against requiring intersectional reporting, which adds burden with little associated benefit.
2. “Scoring rate” is not a term commonly used in assessment science, nor is the interpretation of “scoring rate” clear. There also remains a serious concern over equating “bias” with differential “scoring rates” or selection rates, without consideration of the validity of the underlying assessment scores.
3. A bias audit using selection rates from multiple employers is unlikely to be practical or meaningful in many, if not most situations. Selection rates are a product of both (a) the selection tool itself and (b) how employers use the tool, which can vary substantially from employer to employer. For example, employers vary widely in the extent to which they place weight on selection tools and the cutoff scores they use, if any, based on selection tool scores.