

AI Bias, Fairness and Inclusion

Category: Artificial Intelligence in Education

AI bias in education is a systematic distortion that can produce inaccurate or disadvantageous outcomes for particular learners or groups. Fairness concerns how benefits, burdens and errors are distributed; inclusion concerns whether diverse people can participate, be represented and meaningfully influence the system's design and use on equitable terms.

Bias can enter through historical inequality, labels, measurement choices, missing data, model design or institutional practice. Because definitions of fairness can conflict, no single metric can settle every question. Evaluation needs an explicit educational purpose and reasoned account of affected groups ([Baker & Hawn, 2022](#)).

Fairness Review

1. **Define:** Specify the decision, intended benefit, plausible harms and relevant dimensions of difference.
2. **Represent:** Examine whose data, languages, experiences and accessibility needs are present or absent.
3. **Test:** Compare validity, error patterns and outcomes across appropriate groups and intersecting identities.
4. **Participate:** Involve affected learners and staff in design, interpretation and decisions about acceptable use.
5. **Remedy:** Provide human review, correction and redress, then monitor whether changes reduce inequity.

Examples in Practice

Before using automated language scoring, a university tests performance across first-language groups, disability-related access needs and genres. Educators inspect divergent cases and students can challenge results. If the score disadvantages a group, the institution revises or withdraws the use rather than adjusting learners to the model.

Why It Matters

Educational AI can reproduce existing inequality at scale while appearing objective. A fairness-and-inclusion approach makes values and trade-offs visible, protects access to opportunity and treats affected communities as participants rather than data subjects alone ([Huggins-Manley et al., 2022](#)).

Further Reading

6. Baker, R. S., & Hawn, A. (2022). [Algorithmic bias in education](#). International Journal of Artificial Intelligence in Education, 32, 1052–1092.
7. Huggins-Manley, A. C., Booth, B. M., & D'Mello, S. K. (2022). [Toward argument-based fairness with an application to AI-enhanced educational assessments](#). Journal of Educational Measurement, 59, 362–388.
8. UNESCO. (2021). [Recommendation on the ethics of artificial intelligence](#). UNESCO.

Related Terms: Algorithmic Bias | Inclusive Education | Data Justice | Responsible AI in Education