







Security Rating Template

Kisida, B., Goodwin, L., & Bowen, D.H. (2020). Teaching History Through Theater: The Effects of Arts Integration on Students' Knowledge and Attitudes. *AERA Open*, 6(1), 1-11, DOI 10.1177/2332858420902712.

Rating	Criteria for Rating		Initial Score	Adjust	Final Score
	Design	Attrition			
5 	Randomised design	0-10%	5		
4 	Design for comparison that considers some type of selection on unobservable characteristics (e.g. RDD, Diff-in-Diffs, Matched Diff-in-Diffs)	11-20%			
3 	Design for comparison that considers selection on all relevant observable confounders (e.g. Matching or Regression Analysis with variables descriptive of the selection mechanism)	21-30%		<div>Adjustment for threats to internal validity [-2]</div>	3
2 	Design for comparison that considers selection only on some relevant confounders	31-40%			
1 	Design for comparison that does not consider selection on any relevant confounders	41-50%			
0 	No comparator	>50%			

Threats to Validity	Threat to Internal Validity?	Comments
Threat 1: Confounding	Low	Stratified randomisation. Descriptive statistics demonstrate equivalence on known/measured characteristics.
Threat 2: Concurrent Interventions	Low	Contextual information and random allocation suggests similar practices across the sample.
Threat 3: Experimental effects	High	No information provided; suggestion that teachers in the intervention group may have modified their other teaching as a result of

		participation; known threats associated with waitlist controls.
Threat 4: Implementation fidelity	Low	One off intervention delivered by experienced professionals with a long history of delivering the programme.
Threat 5: Missing Data	Low	Missing data is low and similar across groups.
Threat 6: Measurement	Moderate	Tests are well aligned to state standards, were concealed from programme operators, and have good face validity and internal validity. No blinding, but safeguards present.
Threat 7: Selective reporting	High	No prospectively registered protocol reported.

- **Initial padlock score:** 5 Padlocks – This used a randomised trial design and attrition was very low.
- **Reason for adjustment for threats to validity:** -2 Padlocks – Two threats to internal validity are classified as high.
- **Final padlock score:** initial score adjusted for threats to validity = 3 Padlocks.