

Manual Verification of AiModerators Fact Checking.

Two researchers conducted a manual verification of the AiModerator’s in house design Fact Checker. Our process involved selecting fact checks from four reputable sources (FactCheck.org, Politifact.com, washingtonpost.com, apnews.com) from corresponded across validation sources regarding race and immigration, national security, and climate change. Each researcher independently compared AiModerator's response to the validation sources without knowledge of the other's initial decision. AiModerator’s factual output was cross-validated, ensuring the facts aligned and included relevant contextual information from the validation source. Afterwards, for each reputable source they assigned one of the following labels:

- Factual:** AiModerators output aligns correctly with the validation source.
- Partially Factual:** AiModerators output partially aligns with the validation source. The facts are correct, but there is missing information and context or slight misrepresentation.
- Partially Unfactual:** AiModerators output partially aligns with the validation source; however, the statements are incorrect, unclear, or lack sufficient context.
- Unfactual:** AiModerators output does not align with the validation source, meaning the statements are false.

The two researchers then discussed each comparison openly, agreeing upon a label before writing a brief note to summarize the logic. The label most consistent with all validation sources was considered the final result.

AiModerators responses were consistently Partially Factual. AiModerator correctly identified the facts; however, additional context was missing, which could have provided a whole-rounded explanation. AiModerator fact checked many other debate discussion points missing from these validation sources. For example, Joe Bidens claims against Donald Trump regarding the Central Park Five case and Donald Trump’s claims of success with Opportunity Zones and support of Historically Black Colleges and Universities (HBCUs). The output of these fact checks is available in an example interaction log output for the race and immigration clip (“AiModerator-InteractionLog-Race-Immigration.txt”).

Claim	AiModerator’s Fact Checking Response	Validation Source: Factcheck [1]	Validation Source: Politifact [2]	Validation Source: Associated Press [3]	Validation Source: Washington Post [4]	Median Score
“Joe got \$3.5 million from Russia and it came through Putin because he was very friendly with the former mayor of Moscow. And it was the mayor Moscow’s wife. And you got \$3.5 million. Your family got \$3.5 million. And, you know, some day you’re going to have to explain	1: Hunter Biden's alleged \$3.5 million payment from a Russian mayor's wife lacks verified evidence. 2: The claim originates from a Republican Senate report but has been labeled as "unproven" and lacks context. Hunter Biden's lawyer asserts he was not involved with the firm receiving the funds. 3: The allegation that Hunter Biden received \$3.5 million from Elena Baturina, the wife of Moscow's former mayor, stems from a report by Republican-led Senate committees. However, this claim is based on a	Partially Factual Notes: AiModerator correctly states lack of evidence and the lawyer’s denial but fails to emphasize that the report does not directly implicate Hunter Biden personally.	Partially Factual Note: AiModerator correctly identifies lack of evidence. However, omits the democratic staff dismissing the accusation.	Partially Factual Note: AiModerator captures the core elements of the claim's lack of verification and the lawyer’s denial. However, it omits key clarifications about Joe Biden’s lack of involvement, the absence of evidence for Putin’s connection, and the mischaracterization of the Republican Senate report.	Partially Factual Note: AiModerator correctly identifies lack of verification, skepticism, and the lawyers denial. However, omits additional context provided.	Partially Factual

<i>why you got \$3.5 [million]."</i>	confidential document that has not been publicly verified, leading to significant skepticism. Hunter Biden's legal representative has stated that he was not a co-founder of the firm involved, Rosemont Seneca Thornton, and thus was not entitled to any payment.					
<i>"They keep talking about \$750, which I think is a filing fee. ... Tens of millions of dollars (in income taxes) I prepaid."</i>	1: Trump inaccurately refers to his \$750 tax payment as a "filing fee," which is misleading. 2: Tax experts clarify that there is no such thing as a filing fee for personal income taxes, and the \$750 payment relates to tax owed, not a fee. 3: President Donald Trump's claim that his \$750 tax payment is merely a "filing fee" has been widely criticized as misleading. Tax professionals emphasize that there is no filing fee for personal income tax returns; instead, the \$750 represents the amount owed after accounting for various tax credits and deductions. The New York Times reported that this payment was part of a broader analysis of Trump's tax returns, which revealed he paid minimal federal income taxes in certain years.	Partially Factual Note: AiModerator's output does a good job of addressing the core inaccuracy of the "filing fee" claim and contextualizing the \$750 payment as part of minimal federal taxes paid. However, it fails to address Trump's claim of "prepaid taxes" and the broader context of his tax filings, such as the rolling forward of payments and refunds, which are critical to fully understanding the issue.	<i>Source did not fact check. Not possible to compare.</i>	Partially Factual Note: AiModerator effectively identifies the inaccuracy of the filing fee claim and provides some context about minimal tax payments, but it misses important details about Trump's claims of prepaid taxes, the \$72.9 million refund, and the IRS audit.	<i>Source did not fact check. Not possible to compare.</i>	Partially Factual
<i>"He called [Black people] super predators. And he said that, he said it, super predators, and they have never lived that down."</i>	1: Trump's claim that Biden labeled young black Americans as "super predators" is false. 2: Biden actually rejected the term "super predators" and emphasized that most youth in the justice system are not violent. His comments were made in the context of discussing the 1994 crime bill,	Partially Factual Note: AiModerator's output effectively refutes the central claim and provides relevant context about Biden's comments and the crime bill. However, it omits critical historical context, including Hillary Clinton's use of "super	Partially Factual Note: AiModerator accurately refutes Trump's claim and provides relevant context about Biden's remarks on predators and the 1994 crime bill. However, it lacks specific details about the broader context of Biden's	<i>Source did not fact check. Not possible to compare.</i>	Partially Factual Note: AiModerator's output correctly identifies the core falsehood of Trump's claim and provides relevant context about Biden's comments on predators and the 1994 crime bill. However, it omits critical details about the origin of	Partially Factual

	<p>which has faced criticism for its impact on mass incarceration.</p> <p>3: During the debate, President Donald Trump accused Vice President Joe Biden of referring to young black Americans as "super predators," a claim that has been fact-checked and deemed false. In reality, Biden warned against "predators on our streets" in support of the 1994 crime bill but later clarified that most youth in the justice system were not violent and rejected the notion of them being "super predators."</p>	<p>predators," Biden's nuanced comments about society's role in creating predators.</p>	<p>speech, the historical origin of the term "super predators" (Hillary Clinton, not Biden), and the fact that Biden did not single out Black Americans.</p>		<p>the term "super predators" (Hillary Clinton, not Biden) and the broader context of Clinton's remarks.</p>	
<p>"We changed the policy." <i>The Obama administration</i> "built the cages."</p>	<p>1: The Trump administration's family separation policy was a deliberate choice, not a Democratic mandate.</p> <p>2: Contrary to President Donald Trump's claims, there was no law requiring family separation; it was a policy decision made by his administration. This approach has led to widespread criticism and outrage across the political spectrum.</p> <p>3: The family separation policy implemented by President Donald Trump's administration was part of a "zero-tolerance" approach to illegal immigration, which aimed to deter asylum seekers by separating children from their parents. This policy sparked significant backlash, as it was not mandated by any existing law but rather a strategic decision made by the administration. Reports indicated that thousands of children were held in facilities described as cages, with many</p>	<p>Partially Factual</p> <p>Note: AiModerator accurately identifies that family separation was a deliberate policy choice by the Trump administration under its zero-tolerance approach. However, it omits key historical context about limited separations under previous administrations, the origin of "cages."</p>	<p>Partially Factual</p> <p>Note: AiModerator accurately captures the deliberate nature of the Trump administration's family separation policy. However, does not give additional context regarding cages were instead chain-linked enclosures and the term "cages" was coined by critics.</p>	<p><i>Source did not fact check. Not possible to compare..</i></p>	<p><i>Source did not fact check. Not possible to compare.</i></p>	<p>Partially Factual</p>

	parents deported without their children.					
<i>"Look, their real [climate] plan costs \$100 trillion."</i>	<p>1: President Donald Trump claims that Joe Biden's environmental plan would cost \$100 trillion, labeling it unrealistic.</p> <p>2: Trump's assertion about the \$100 trillion cost of Biden's environmental plan has been met with skepticism, as it lacks detailed context. Critics argue that such figures can be misleading without considering the potential economic benefits of green initiatives.</p> <p>3: President Donald Trump's claim that Joe Biden's environmental plan would cost \$100 trillion has sparked debate regarding its accuracy and feasibility. Critics point out that such a high figure may not fully account for the long-term economic benefits associated with transitioning to a greener economy, such as job creation in renewable energy sectors and reduced healthcare costs from pollution-related illnesses.</p>	<p>Partially Factual</p> <p>Notes: AiModerator effectively conveys skepticism about Trump's \$100 trillion claim and raises valid points about the broader economic debate around green initiatives. However, it fails to clarify that the \$100 trillion figure is associated with the Green New Deal, not Biden's plan.</p>	<p>Partially Factual</p> <p>Notes: AiModerator's output effectively captures skepticism about Trump's \$100 trillion claim and highlights the broader economic debate about environmental policies. However, it fails to clarify the distinction between Biden's plan and the Green New Deal,</p>	<i>Source did not fact check. Not possible to compare.</i>	<p>Partially Correct</p> <p>Notes: AiModerator's output effectively identifies skepticism about Trump's \$100 trillion claim and provides valid context regarding green initiatives' potential benefits. However, it omits critical distinctions between Biden's plan and the Green New Deal, details about the Green New Deal's scope.</p>	Partially Factual

- [1] <https://www.factcheck.org/2020/10/factchecking-the-final-2020-presidential-debate/>
- [2] <https://www.politifact.com/article/2020/oct/23/fact-checking-donald-trump-joe-biden-final-preside/#Climate%20change>
- [3] <https://apnews.com/article/ap-fact-check-final-debate-trump-biden-4d304cf7ce7dee9c228f48bd9b76e8f7>
- [4] <https://www.washingtonpost.com/politics/2020/10/23/fact-checking-second-trump-biden-debate/>