EAP Speaking teacher reference

Commentary for the matching selected extracts. These are proposed interpretations for discussion and should be opened after teachers complete the exercise.

Fictional AI-generated material for teacher standardisation. Proposed judgements require human review.

SP-X1

Ben's presentation accurately separates six weeks of pilot operation from the response window. It explains a limitation and proposes a safeguard without inventing staffing figures. This supports clear explanation and justification under the first criterion. The full reconstructed performance receives 17/25, within Strong 16–20; this extract alone cannot establish pronunciation, pace or the whole mark.

SP-X2

Dawit communicates the purpose and links his review to the supplied aims. "There is four aims" is an agreement error, but meaning is clear. Consider the frequency and effect of errors across the full performance. The full presentation receives 12/25, within Competent 11–15, using the stated simulated delivery evidence.

SP-X3

Chloe explains why silence does not establish understanding. Dawit checks his interpretation; Chloe answers the specific question. This is two-way interaction with evidence of listening. Dawit's clarification should receive credit rather than being treated automatically as a failure. Full interaction judgements are Chloe 21/25, Exceptional 21–25, and Dawit 12/25, Competent 11–15; these depend on all turns and simulated observations.

SP-X4

Ben asks a relevant follow-up that challenges cumulative workload. Amina responds conditionally and acknowledges missing capacity evidence; Chloe develops the implication. The exchange shows engagement rather than four unrelated speeches. Full interaction marks are Ben 16/25 Strong, Amina 22/25 Exceptional and Chloe 21/25 Exceptional. Use the complete rubric and performance before applying these ranges.

Whole performance reference

Candidate	Presentation	Interaction	Total
Amina	21	22	43
Ben	17	16	33
Chloe	20	21	41
Dawit	12	12	24

All scores are proposed and conditional on simulated delivery evidence. Human markers should record their own reasons before resolving differences. No completion of human standardisation is implied.