

Solutions

2.1 Four appointment preparation steps; immediately after the workshop, from memory.

2.2 Morning: ten of twelve; afternoon: seven of twelve. Both numerator and denominator are needed for each mark.

2.3 For example, people chose their sessions, and different tutors taught the sessions. Either could affect the groups or instruction, so the checklist effect was not isolated. Accept different times or missing initial recall scores as justified alternatives.

2.4 Appointment attendance or health improvement was not measured. The card contains no later attendance record, so it cannot show whether participants subsequently missed appointments. Equivalent correction about no health measurement qualifies.

3.1 Six points ($6 \times B1$): Chen's fictional card attributed; ten of twelve; seven of twelve; two accurate limits (2 points); coherent paraphrase preserving recall as the outcome. Sample: "Chen's fictional practice card (2026) reports complete recall of four preparation steps by ten of twelve morning volunteers and seven of twelve afternoon volunteers. Only the morning session used a checklist, but session choice and tutors also differed. The comparison therefore does not isolate a checklist effect. It provides no earlier recall scores or later appointment records and cannot establish an improvement in participants' health."

3.2 Eight points ($8 \times B1$): accurately bounded recall comparison; no clinical claim; self-selection limit; tutor or time limit; no prior scores acknowledged; feasible clearer comparison; same measured outcome proposed; coherent explanation. Sample: "The card supports a limited statement about immediate recall: complete recall was more common in the morning session than in the afternoon session. It does not establish that the checklist caused the difference or improved health. Volunteers selected their sessions, tutors differed, and no earlier memory scores were recorded. A further classroom activity could use the same tutor, instructions and recall task for both groups, with the checklist as the main planned difference. Assigning willing participants to groups rather than letting them choose would reduce one source of imbalance. Recording an initial recall task could also help interpret the later results. These changes would strengthen the comparison, but its conclusion should still concern recall, not an unmeasured health outcome."