**Appendix D.**

**Quality Rating Tool Study ID Number**

Study and outcome ratings are based on data extracted using McMaster quality appraisal tool template (Law et al. 1998) combined with rating scale from EPHPP (1998), modified for current review question and study type.

**Full citation**

|  |
| --- |
|  |

**Sample / selection**

1) Sample n = Was the sample size justified?Yes No

2) Was there evidence that informed consent was obtained? Yes No

3 Was there evidence of sample / selection bias? Yes No

3) Was the sample taken likely to be representative of the

target population? Yes No

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| **≥ 80% participation** | **60-79% participation** | **< 60% participation** |
|  |  |  |
| **Strong** | **Moderate** | **Weak** |

**Study design**

1) Analyse the quality of the study design:

Did the method of randomisation used allow each study participant to have the same chance of receiving the intervention? Consider:

1. Was study described as randomised? Yes No
2. Was method of randomisation described? Yes No N/A
3. Was method used appropriate? Yes No

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| **Randomisation process well described and appropriate method for randomisation used.** | **Randomisation method stated but methodologically weak, lacking in detail / transparency.** | **Randomisation did not occur****(Quasi-experiential design) or allocation of participant to study group was not described.** |
|  |  |  |
| **Strong** | **Moderate** | **Weak** |

**Confounders**

Were potential confounders accounted for in the study? Examples of how this could be addressed:

a) Stratification

b) Matching

c) Accounted for within the analysis.

Were the groups balanced at baseline with respect to confounders?

Examples of potential confounders most relevant to outcomes:

1. Sex
2. Age
3. Level of study, education, previous experience
4. Professional group (if diverse)
5. Pre-test score on outcome

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| **≥ 80% of relevant confounders controlled for** | **60-79% of relevant confounders are controlled for** | **< 60% of relevant confounders are controlled for. Or control of confounders was not described** |
|  |  |  |
| **Strong** | **Moderate** | **Weak** |

**Blinding**

1) Were the outcome assessors aware of the intervention or exposure status of participants?

***Knowledge outcome***

|  |  |  |
| --- | --- | --- |
| **Measures taken to blind assessor to participant intervention** | **Blinding measures attempted but deduction possible, lacking in consistency or rigour.** | **No attempt made to blind assessor to participant intervention, blinding not possible, or was not described.** |
|  |  |  |
| **Strong** | **Moderate** | **Weak** |

***Psychomotor skills outcome***

|  |  |  |
| --- | --- | --- |
| **Measures taken to blind assessor to participant intervention** | **Blinding measures attempted but deduction possible, lacking in consistency or rigour.** | **No attempt made to blind assessor to participant intervention, blinding not possible, or was not described.** |
|  |  |  |
| **Strong** | **Moderate** | **Weak** |

***Affective or non-technical skills outcomes***

|  |  |  |
| --- | --- | --- |
| **Measures taken to blind assessor to participant intervention** | **Blinding measures attempted but deduction possible, lacking in consistency or rigour.** | **No attempt made to blind assessor to participant intervention, blinding not possible, or was not described.** |
|  |  |  |
| **Strong** | **Moderate** | **Weak** |

**Intervention integrity**

What are the risks that performance bias occurred? Consider:

1. The consistency of interventions.
2. Is it likely that subjects received contamination from an unintended intervention or received a co-intervention that may have influenced the outcomes?

***Knowledge outcome***

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| **Study free of intervention contamination or of any co-intervention that may have favoured one group over another. Measures taken to ensure interventions were consistently applied.** | **Some evidence that either minor inconsistencies of intervention, contamination, or a co-intervention may have influenced results.** | **Evidence of considerable lack of consistency in interventions and / or contamination or co-interventions favouring one group over another.** |
|  |  |  |
| **Strong** | **Moderate** | **Weak** |

***Psychomotor skills outcome***

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| **Study free of intervention contamination or of any co-intervention that may have favoured one group over another. Measures taken to ensure interventions were consistently applied.** | **Some evidence that either minor inconsistencies of intervention, contamination, or a co-intervention may have influenced results.** | **Evidence of considerable lack of consistency in interventions and / or contamination or co-interventions favouring one group over another.** |
|  |  |  |
| **Strong** | **Moderate** | **Weak** |

***Affective or non-technical skills outcomes***

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| **Study free of intervention contamination or of any co-intervention that may have favoured one group over another. Measures taken to ensure interventions were consistently applied.** | **Some evidence that either minor inconsistencies of intervention, contamination, or a co-intervention may have influenced results.** | **Evidence of considerable lack of consistency in interventions and / or contamination or co-interventions favouring one group over another.** |
|  |  |  |
| **Strong** | **Moderate** | **Weak** |

**Data Collection Methods**

Tools for primary outcome measures must be reliable and valid. Demonstration of face or content validity is acceptable. Reliability and validity can be reported either within the study or using standard assessment tools with known reliability & validity.

***Knowledge outcome***

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| **Data collection tools have been shown to be reliable and valid** | **Data collection tools have been shown to be valid but not reliable, or reliability is not described** | **Data collection tools have not been shown to be valid, or both reliability are not described** |
|  |  |  |
| **Strong** | **Moderate** | **Weak** |

***Psychomotor skills outcome***

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| **Data collection tools have been shown to be reliable and valid** | **Data collection tools have been shown to be valid but not reliable, or reliability is not described** | **Data collection tools have not been shown to be valid or both reliability are not described** |
|  |  |  |
| **Strong** | **Moderate** | **Weak** |

***Affective or Non-technical skills outcomes***

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| **Data collection tools have been shown to be reliable and valid** | **Data collection tools have been shown to be valid but not reliable, or reliability is not described** | **Data collection tools have not been shown to be valid or both reliability are not described** |
|  |  |  |
| **Strong** | **Moderate** | **Weak** |

**Withdrawals and Dropouts**

Have the authors described both the numbers and reasons for dropouts?

The % of participants completing the study refers to the % of subjects remaining in the study at the final data collection period in all groups (i.e. control and intervention groups).

|  |
| --- |
|  |

|  |  |  |
| --- | --- | --- |
| **Follow up rate ≥ 80%** | **Follow up rate 60-79%** | **Follow up rate < 60% or was not described** |
|  |  |  |
| **Strong** | **Moderate** | **Weak** |

**Results**

Factors to consider:

Were results reported in terms of significance? Yes No

Was the sample big enough to show significance? Yes No

 No evidence of sample size calculation

Were the statistical analysis methods used appropriate (and if multiple outcomes used was this taken into account)?

Is there evidence of selective reporting? Yes No

|  |
| --- |
|  |

**RATING SUMMARY - Component Ratings for Risk of Bias**

|  |  |  |  |
| --- | --- | --- | --- |
| **Component** | **Strong** | **Moderate** | **Weak** |
| **Sample / selection** |  |  |  |
| **Study design** |  |  |  |
| **Confounders** |  |  |  |
| **Blinding**KnowledgePsychomotorAffective / NT |  |  |  |
| **Intervention integrity**KnowledgePsychomotorAffective/NT |  |  |  |
| **Data Collection**KnowledgePsychomotorAffective / NT |  |  |  |
| **Withdrawals and Dropouts** |  |  |  |

**Global rating for paper**

**EPHPP Recommended Protocol**

Strong = no weak ratings

Moderate = one weak rating

Weak = two or more weak ratings

**Individual learning outcome ratings**

Knowledge outcome rating =

Psychomotor outcome rating =

Affective / Non-technical skills outcome rating =

|  |  |  |
| --- | --- | --- |
| **Strong** | **Moderate** | **Weak** |
|  |  |  |

|  |  |  |
| --- | --- | --- |
| **Yes** | **No** | **N/A** |
|  |  |  |

Was decision agreed by second reviewer?

Final decision on global rating (after discussion by reviewers)

|  |  |  |
| --- | --- | --- |
| **Strong** | **Moderate** | **Weak** |
|  |  |  |