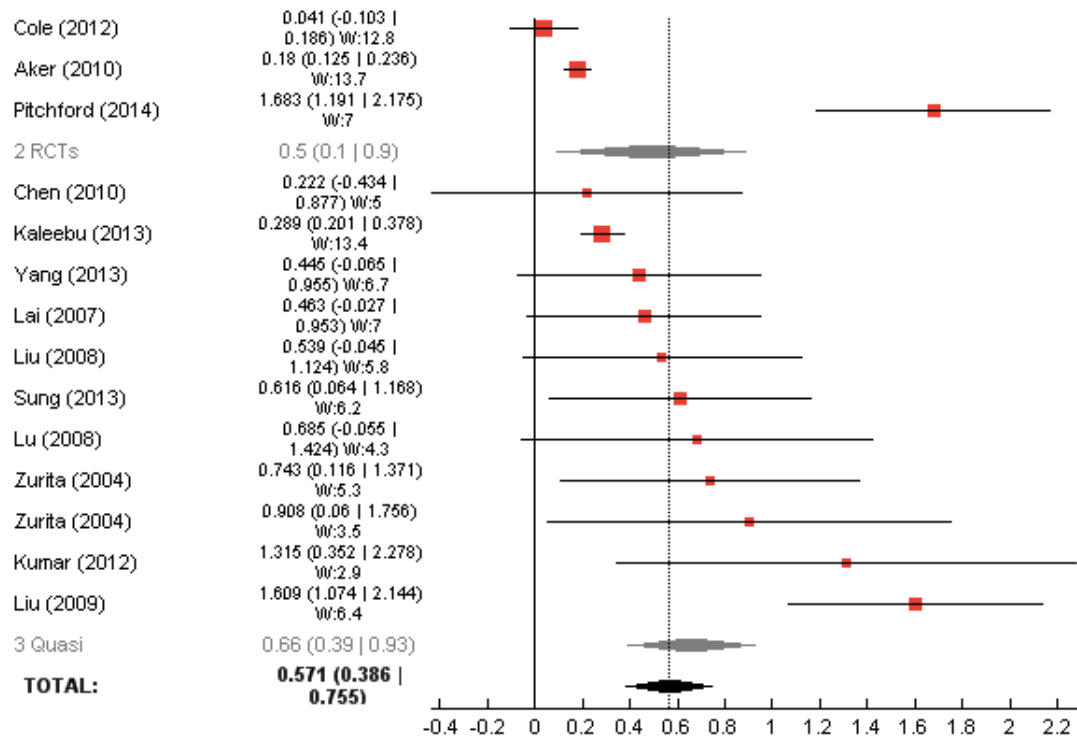


# Online appendix 4: Formal sub-group analyses

## Sensitivity analysis

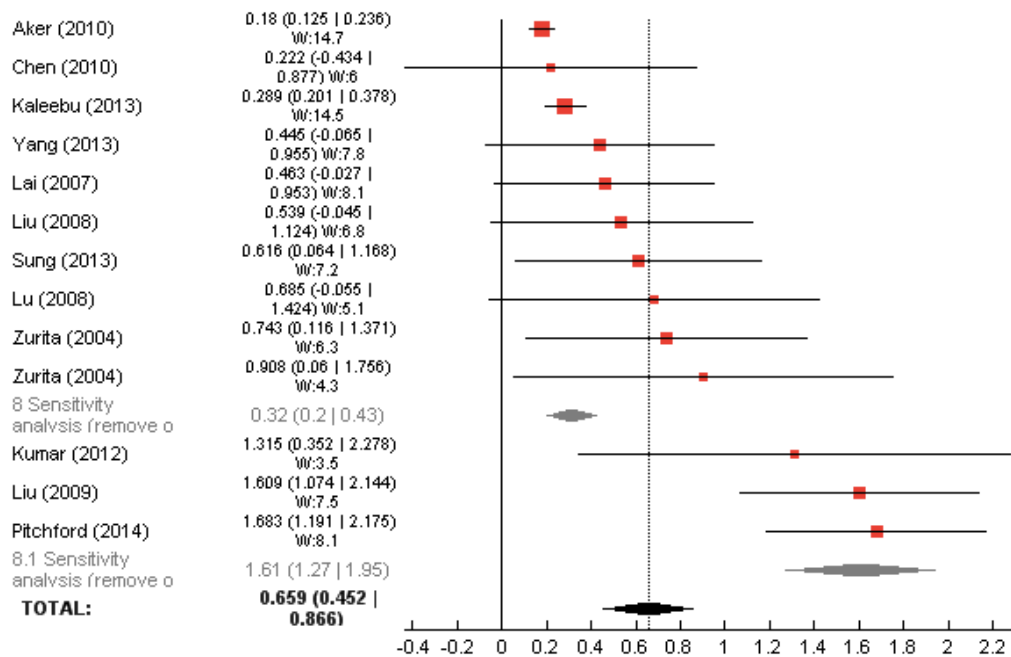
*Design: RCTs vs Quasi-experimental*



Heterogeneity Q (all studies) = 86.2; df = 13; p = 7.45E-13; I-squared = 84.9%.  
 (Group 1 Q = 39.4; df = 2. Group 2 Q = 32.6; df = 10).

Random effects model overall effect: 0.571 (0.386, 0.755) Difference: 0.164; SE difference: 0.247; Z: 0.666; p = 0.506; Q\* within: 19.5; Q\* between: 0.443; (Group 1 Q\*: 11; Group 2 Q\*: 8.5); heterogeneity explained: 0%.

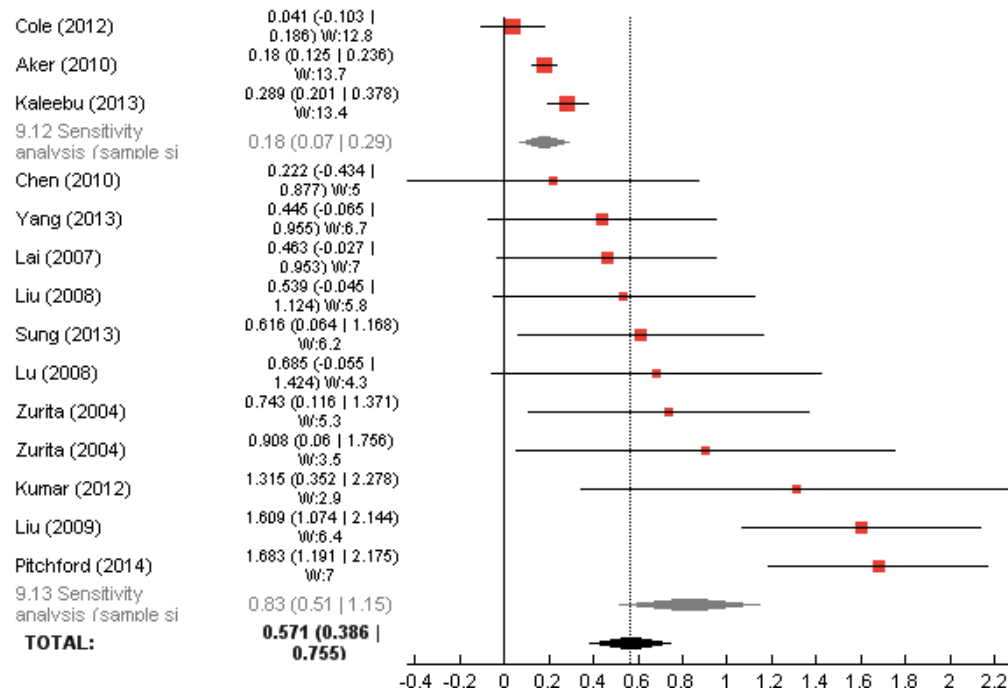
*Design: removal of outliers*



Heterogeneity Q (all studies) = 78.8; df = 12; p = 6.98E-12; I-squared = 84.8%.  
 (Group 1 Q = 15.8; df = 9. Group 2 Q = 0.444; df = 2).

Random effects model overall effect: 0.659 (0.452, 0.866) Difference: 1.29; SE difference: 0.183; Z: 7.07; p = 1.6E-12; Q\* within: 9.08; Q\* between: 49.9; (Group 1 Q\*: 8.64; Group 2 Q\*: 0.444); heterogeneity explained: 92%. \* N.B. the Q\* statistics are calculated using random effects weights and are only used for the analysis of variance.

*Design: sample size*

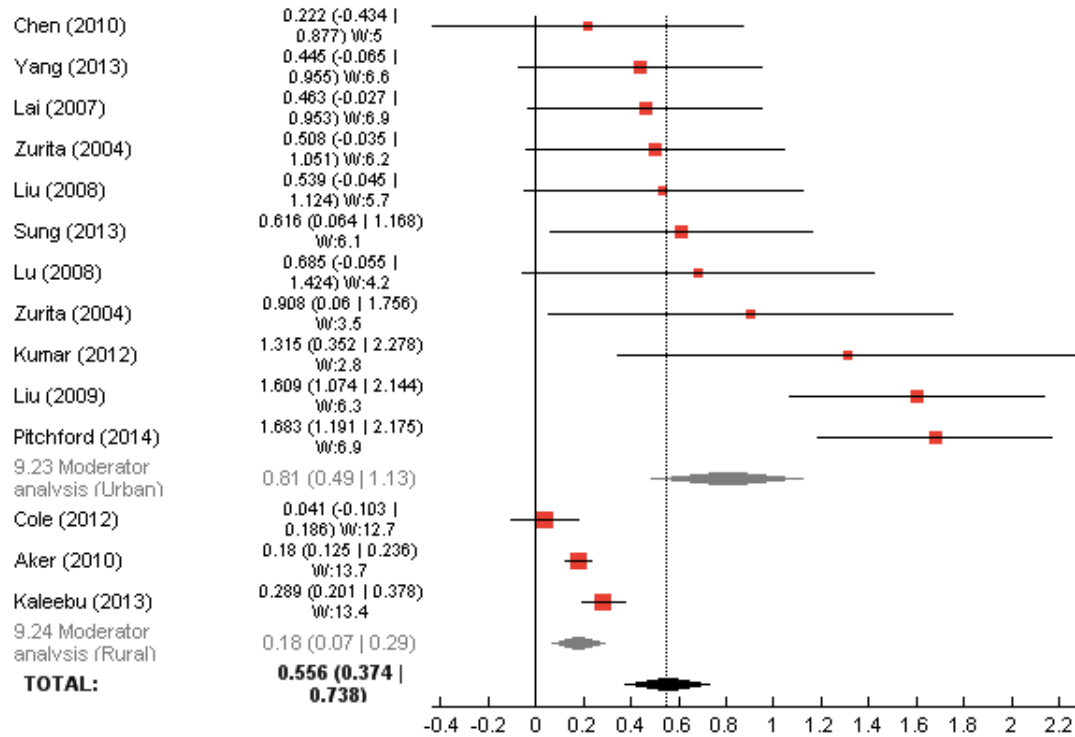


Heterogeneity Q (all studies) = 86.2; df = 13; p = 7.45E-13; I-squared = 84.9%.  
 (Group 1 Q = 8.97; df = 2. Group 2 Q = 30.1; df = 10).

Difference: 0.65; SE difference: 0.172; Z: 3.78; p = 0.000157; Q\* within: 11.5; Q\* between: 14.3; (Group 1 Q\*: 2.81; Group 2 Q\*: 8.7); heterogeneity explained: 61%. \*  
 N.B. the Q\* statistics are calculated using random effects weights and are only used for the analysis of variance.

## Moderator analysis:

### *Rural vs urban*

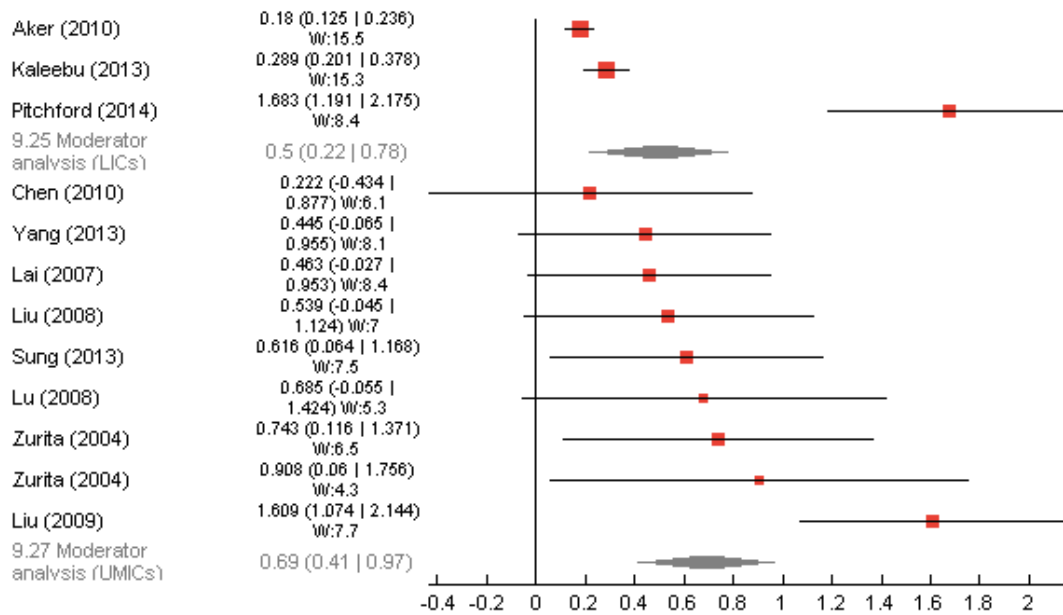


Heterogeneity Q (all studies) = 84.6; df = 13; p = 1.47E-12; I-squared = 84.6%.  
(Group 1 Q = 31.3; df = 10. Group 2 Q = 8.97; df = 2).

Difference: 0.626; SE difference: 0.173; Z: 3.62; p = 0.000297; Q\* within: 11.6; Q\* between: 13.1; (Group 1 Q\*: 8.83; Group 2 Q\*: 2.81); heterogeneity explained: 58%.

\* N.B. the Q\* statistics are calculated using random effects weights and are only used for the analysis of variance.

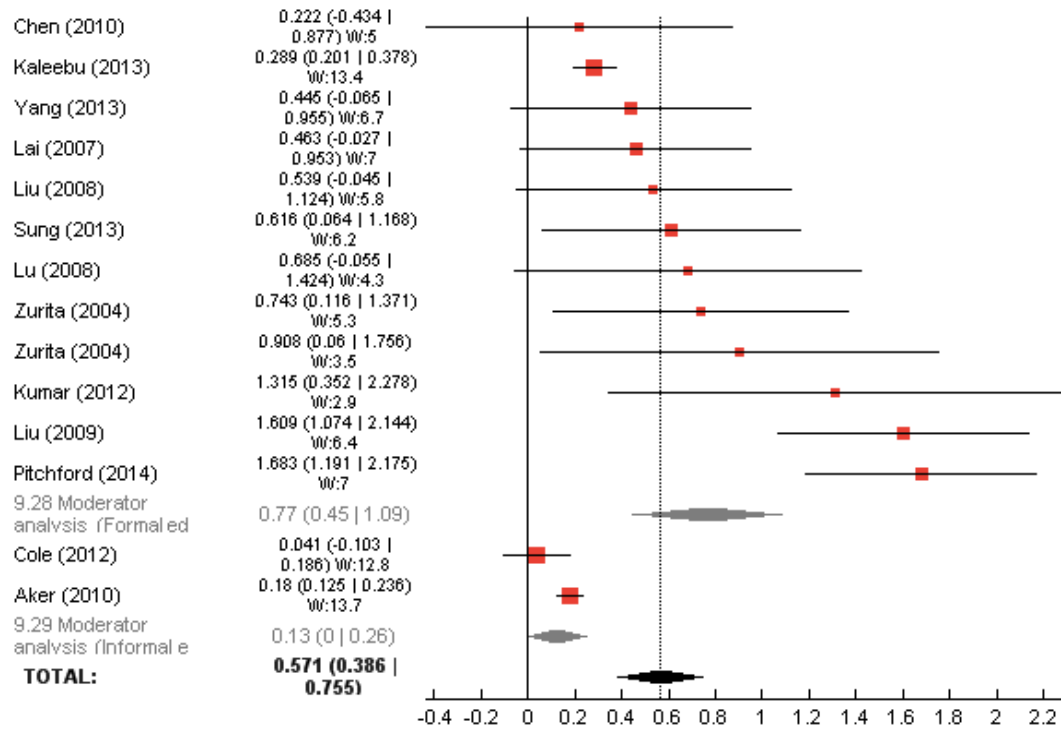
## World Bank Classifications of economies



Heterogeneity Q (all studies) = 74.1; df = 11; p = 2.01E-11; I-squared = 85.2%.  
(Group 1 Q = 38.2; df = 2. Group 2 Q = 15.6; df = 8).

Difference: 0.188; SE difference: 0.203; Z: 0.927; p = 0.354; Q\* within: 22.7; Q\* between: 0.86; (Group 1 Q\*: 15.3; Group 2 Q\*: 7.4); heterogeneity explained: 25%. \* N.B. the Q\* statistics are calculated using random effects weights and are only used for the analysis of variance.

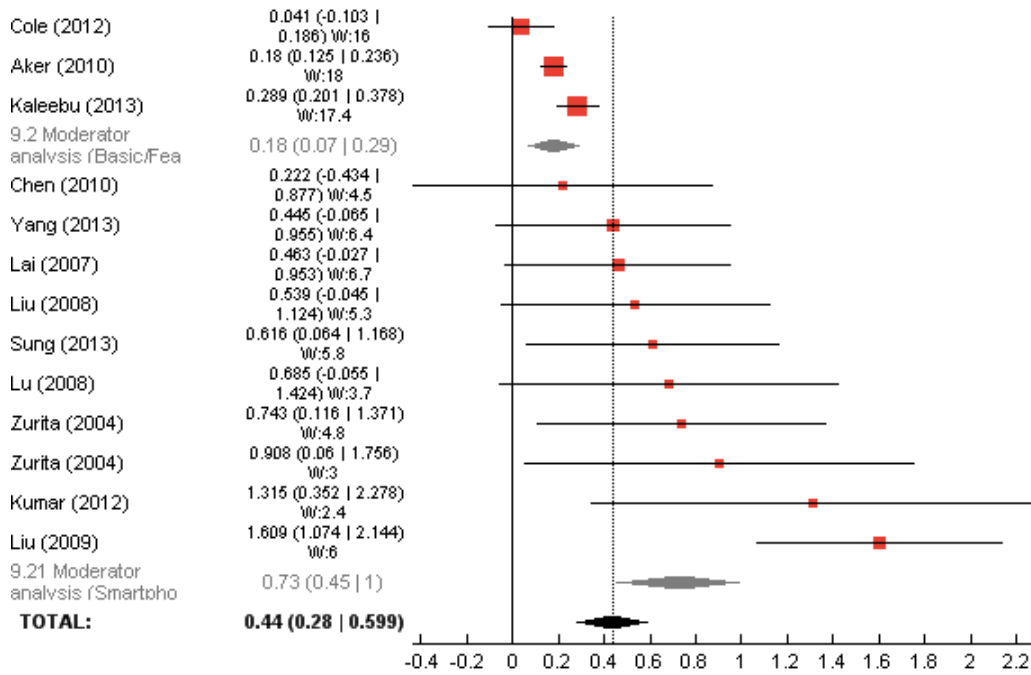
*Formal or informal education*



Heterogeneity Q (all studies) = 86.2; df = 13; p = 7.45E-13; I-squared = 84.9%.  
 (Group 1 Q = 59.5; df = 11. Group 2 Q = 3.08; df = 1).

Difference: 0.643; SE difference: 0.177; Z: 3.64; p = 0.000273; Q\* within: 9.7; Q\* between: 13.2; (Group 1 Q\*: 8.7; Group 2 Q\*: 1); heterogeneity explained: 0%. N.B. the Q\* statistics are calculated using random effects weights and are only used for the analysis of variance.

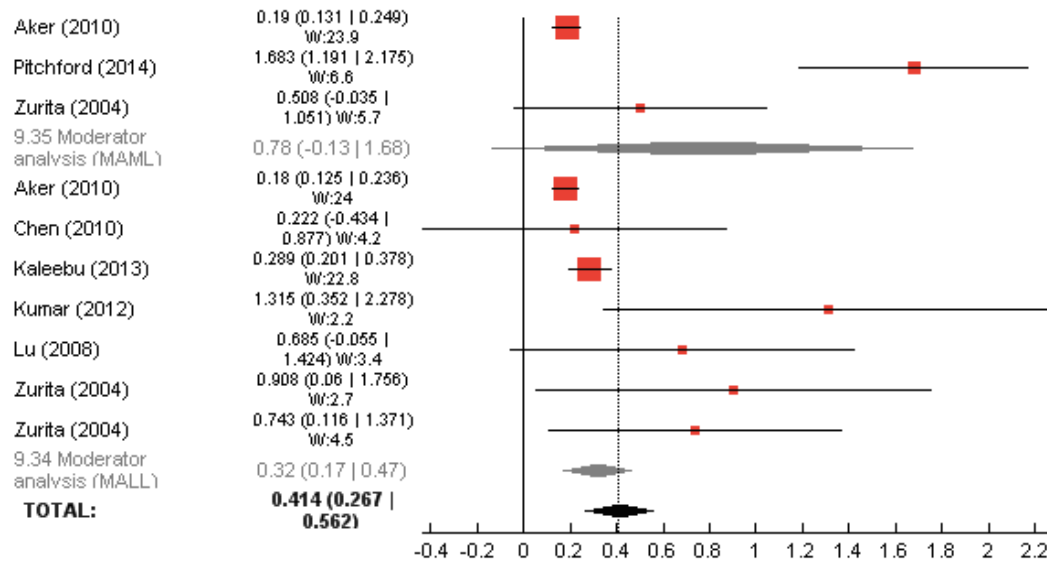
*Applied technology*



Heterogeneity Q (all studies) = 52.6; df = 12; p = 4.92E-7; I-squared = 77.2%. (Group 1 Q = 8.97; df = 2. Group 2 Q = 17.2; df = 9).

Difference: 0.543; SE difference: 0.15; Z: 3.62; p = 0.000292; Q\* within: 11.2; Q\* between: 13.1; (Group 1 Q\*: 2.81; Group 2 Q\*: 8.43); heterogeneity explained: 61%

## MALL vs MAML



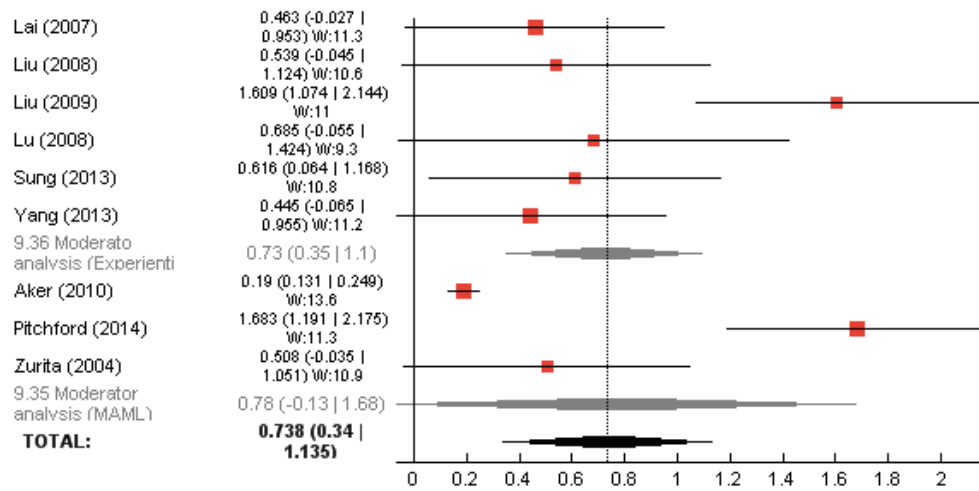
Heterogeneity Q (all studies) = 52; df = 9;  $p = 4.47E-8$ ; I-squared = 82.7%. (Group 1 Q = 16; df = 6. Group 2 Q = 36; df = 2).

Difference: 0.455; SE difference: 0.47; Z: 0.969;  $p = 0.333$ ;  $Q^*$  within: 11.5;  $Q^*$  between: 0.939; (Group 1  $Q^*$ : 1.92; Group 2  $Q^*$ : 9.55); heterogeneity explained: 0%.

\* N.B. the  $Q^*$  statistics are calculated using random effects weights and are only used for the analysis of variance



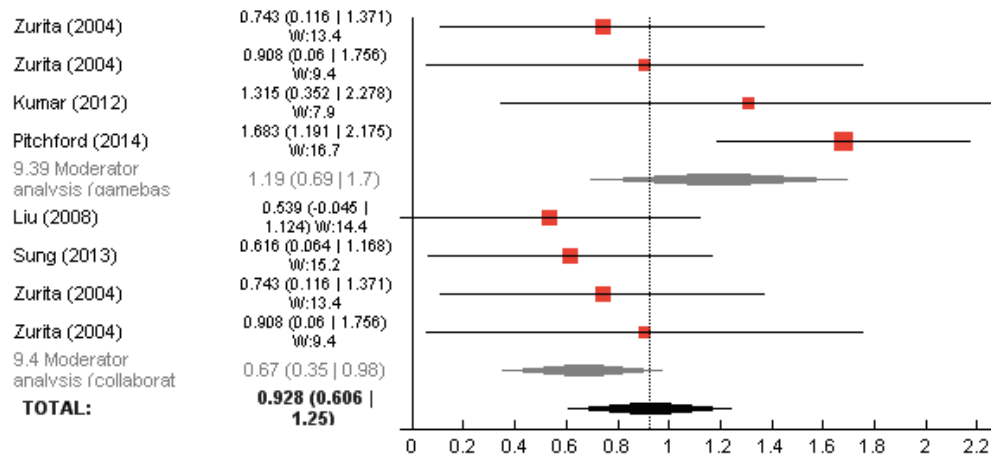
## MALL vs experimental outdoor learning



Heterogeneity Q (all studies) = 67.4; df = 8; p = 1.6E-11; I-squared = 88.1%. (Group 1 Q = 13.3; df = 5. Group 2 Q = 36; df = 2).

Difference: 0.0491; SE difference: 0.501; Z: 0.098; p = 0.922; Q\* within: 6.62; Q\* between: 0.00961; (Group 1 Q\*: 4.7; Group 2 Q\*: 1.92); heterogeneity explained: 0%. \* N.B. the Q\* statistics are calculated using random effects weights and are only used for the analysis of variance

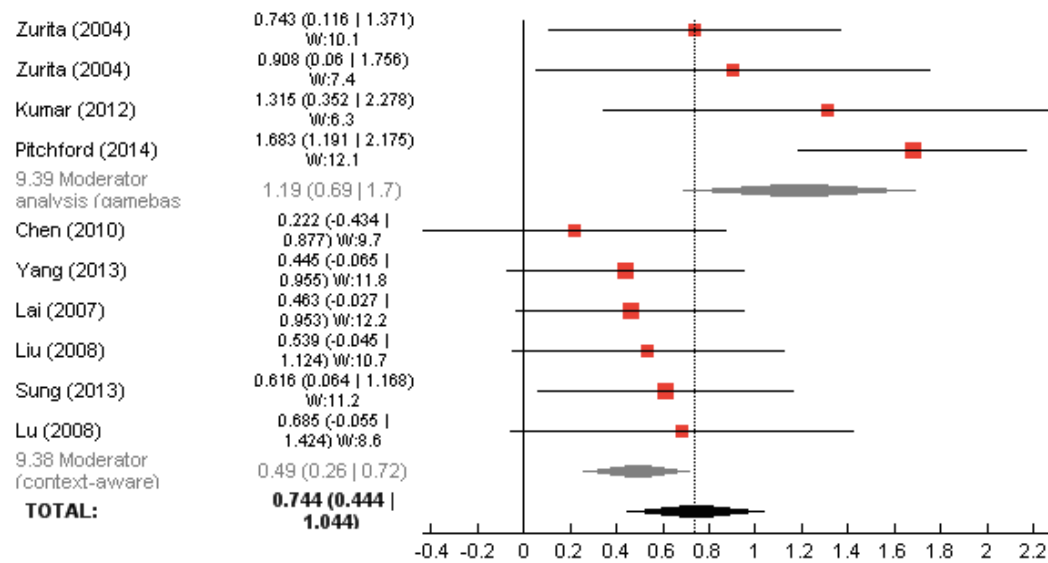
## Game-based learning & collaborative learning



Heterogeneity Q (all studies) = 13.2; df = 7; p = 0.0669; I-squared = 47%. (Group 1 Q = 6.12; df = 3. Group 2 Q = 0.581; df = 3).

Difference: 0.529; SE difference: 0.302; Z: 1.75; p = 0.0133; Q\* within: 4.29; Q\* between: 6.13; (Group 1 Q\*: 2.41; Group 2 Q\*: 0.581); heterogeneity explained: 86%. \* N.B. the Q\* statistics are calculated using random effects weights and are only used for the analysis of variance.

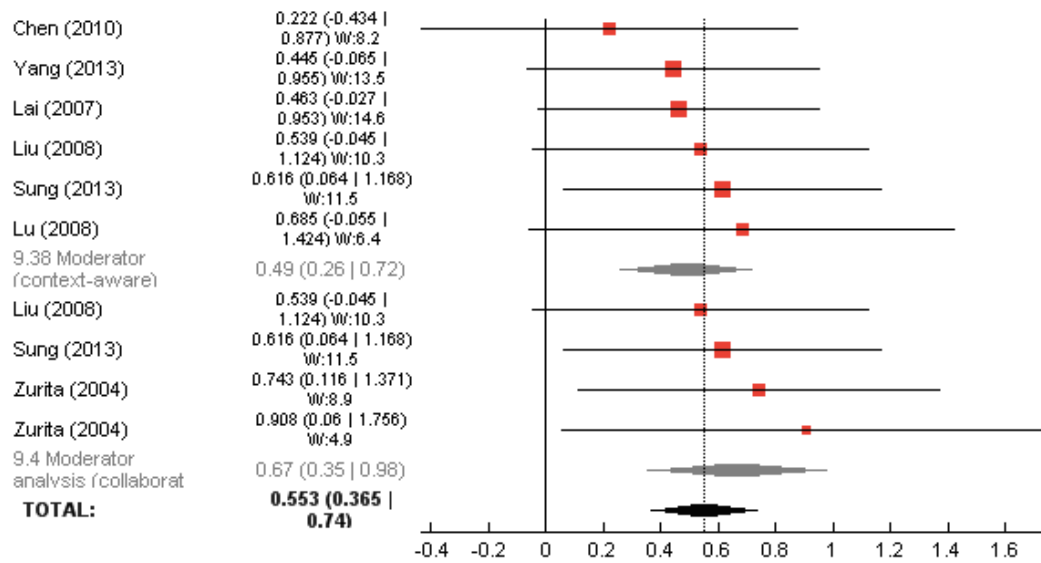
*Game-based mobile learning & context-aware/experiential mobile learning*



Heterogeneity Q (all studies) = 21.2; df = 9; p = 0.0118; I-squared = 57.5%. (Group 1 Q = 6.12; df = 3. Group 2 Q = 1.18; df = 5).

Difference: 0.705; SE difference: 0.282; Z: 2.5; p = 0.0125; Q\* within: 3.58; Q\* between: 6.24; (Group 1 Q\*: 2.41; Group 2 Q\*: 1.18); heterogeneity explained: 100%. \* N.B. the Q\* statistics are calculated using random effects weights and are only used for the analysis of variance.

*Collaborative mobile learning & Context-aware/experiential mobile learning*



Heterogeneity Q (all studies) = 2.53; df = 9; p = 0.98; I-squared = 0%. (Group 1 Q = 1.18; df = 5. Group 2 Q = 0.581; df = 3).

Difference: 0.175; SE difference: 0.2; Z: 0.879; p = 0.379; Q\* within: 1.76; Q\* between: 0.773; (Group 1 Q\*: 1.18; Group 2 Q\*: 0.581); heterogeneity explained: NaN%. \* N.B. the Q\* statistics are calculated using random effects weights and are only used for the analysis of variance.