Combining adult and children's invitalscores

Bruce Mallory

3/15/2021

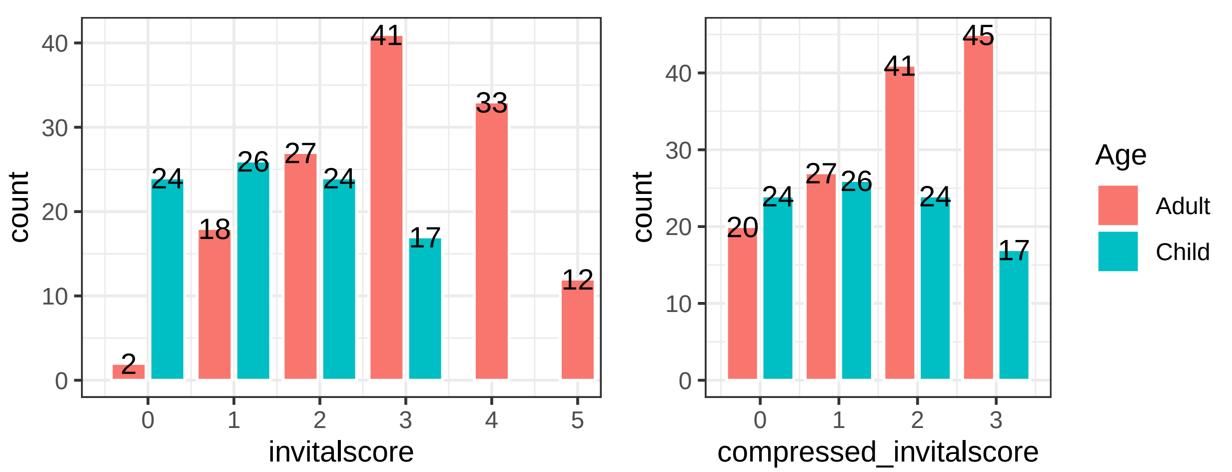
Below are three general ways to normalize the adult and children's `invitalscores`. The comparison below show the original `invitalscore` distribution on the left, and the normalized method on the right. The second method (using standardized scores) seems the least preferable, due to the clumped nature of the resulting adult distribution - either clumped at the low end for the straight standardization (#2a below), or clumped in the middle for the binned standardization (#2b below).

If we transform the `invitalscores` by dividing # true by # of questions asked, we get a percent-true measure (#3a). Binned percent-true (#3c), where we round to the nearest children's percent-true, also results in a a clumped adult distribution, with almost no adults in the bottom bin and almost half the adults in the third bin. And binned percent-true (#3b), where we round up to the nearest 25%, results in an identical outcome to the compressing method (#1).

Which then leaves the decision between the compressing method (#1) with four levels, or the "continuous" percent-true method (#3a). The "continuous" percent-true method results in four levels for the children and six levels for the adults. We're guessing that the normalizing choice (#1 vs. #3a) will not result in a difference in the model fit, nor in the assessment of whether or not `invitalscore` and `inpsychscore` have an effect on any of the dependent variables. Which would argue for leaving the normalizing choice as a contextual question. If you have a judgment about which normalizing method (#1 or #3a) is truer to the context of the questions that you asked of the adults and children, then it makes sense to go with that judgement.

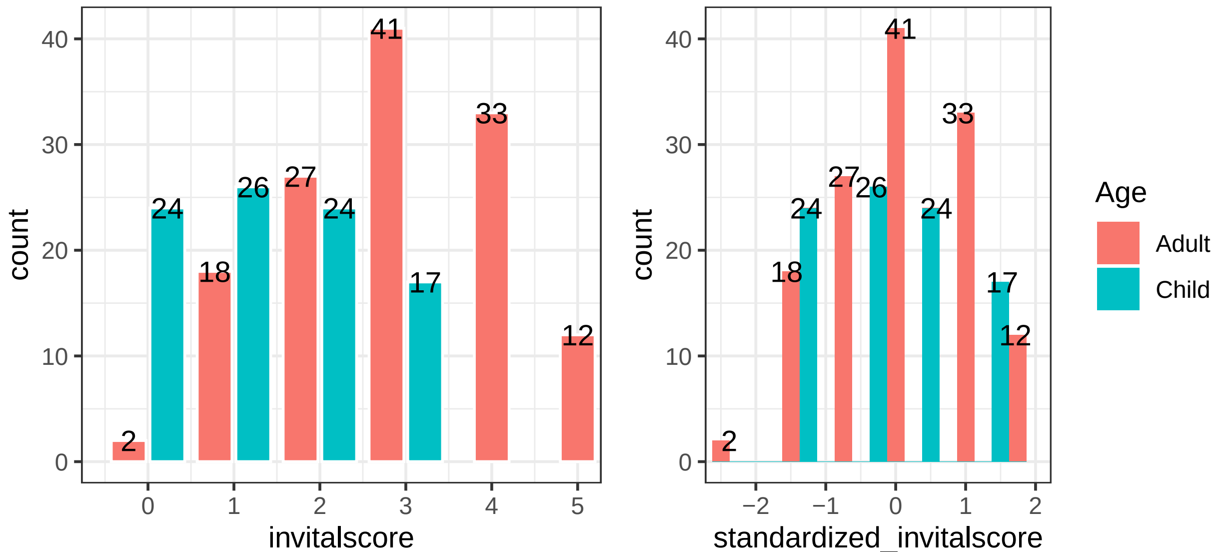
Finally, this analysis has only been done with `invitalscore`. And whatever the decision that is made, we will do the same thing for `inpsychscore`, and then use those when we analyze whether `invitalscore` and/or `inpsychscore` have an effect on the four dependent variables.

1. Compress adult's `invitalscore` to fit children's `invitalscore`



$[0, 5]$ -> $[0, 3]$ by compressing adult 0 & 1 into 0, and adult 4 & 5 into 3, with adult 2 -> 1 and adult 3 -> 2

2a. Standardize both adult & children's `invitalscore`



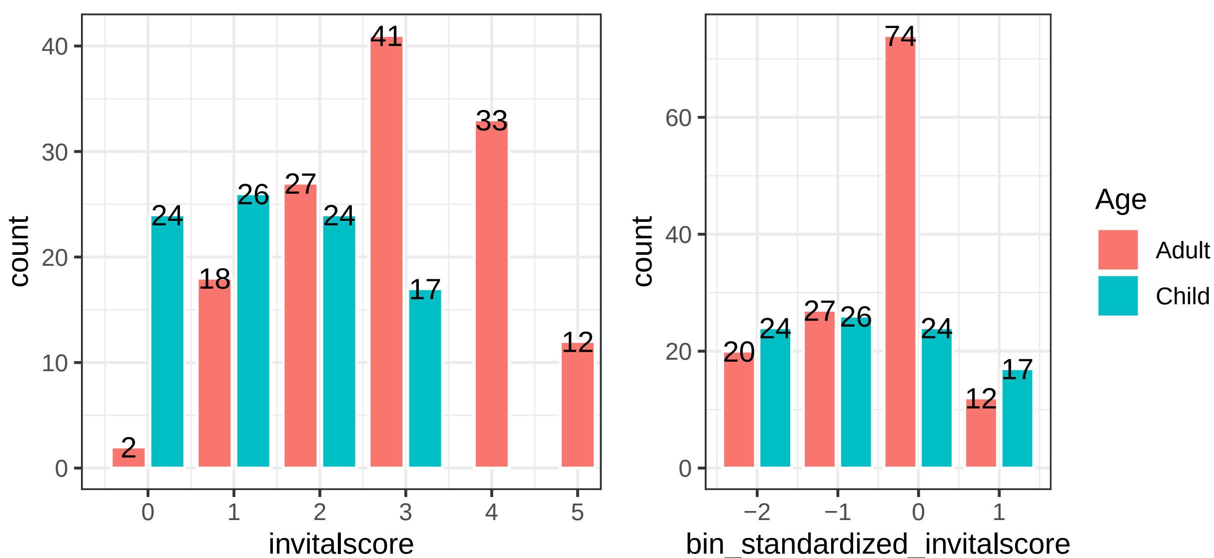
$standardized score = z = \frac{x - \bar{x}}{\sigma}$

The range of standardized scores for adults $= [-2.38, 1.71]$.

The range of standardized scores for children $= [-1.28, 1.52]$.

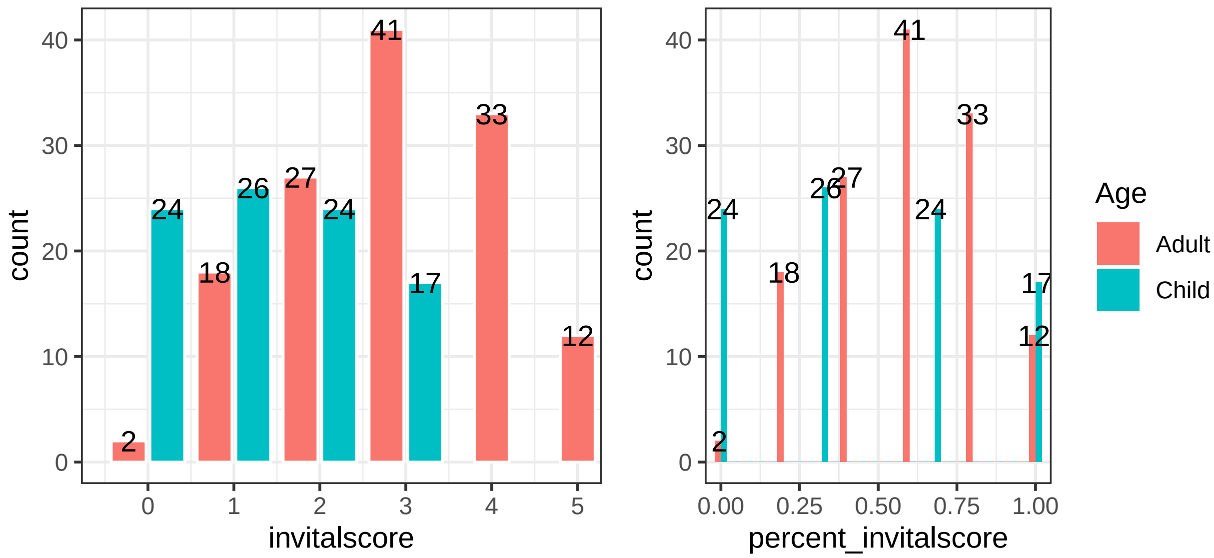
Note: In context, it doesn't make sense for an adult score of 0 or 1 to be lower than the child score of 0.

2b. Standardize both adult & children's `invitalscore`, then round down to the nearest integer to put scores into four bins



$z = \frac{x - \bar{x}}{\sigma}$, four bins: $[z < -2, z < -1, z < 0, z < 1]$ NOTE: Over half the adults end up in the third bin.

3a. Calculate "percent-true" for adult & children's `invitalscore`

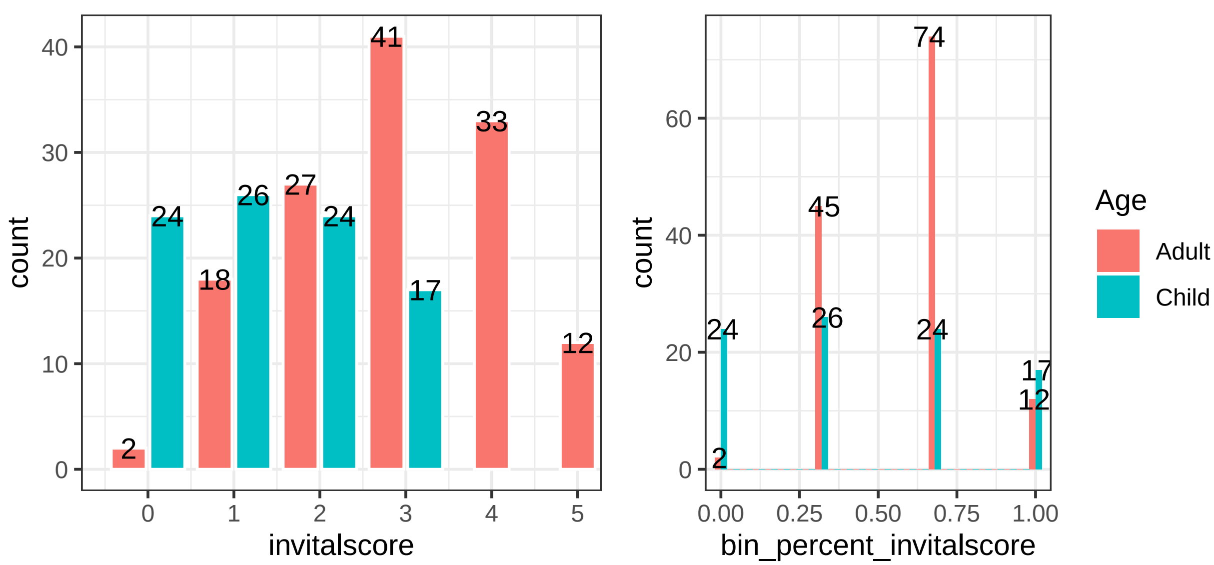


For adults the six possible scores are $[0.0, 0.2, 0.4, 0.6, 0.8, 1.0]$.
For children the four possible scores are $[0.0, 0.33, 0.67, 1.0]$.

3b. Calculate "percent-true" for adult & children's `invitalscore`, then round up to the nearest 25% to put scores into four bins

Four bins: $[\% < 0.25, \% < 0.50, \% < 0.75, \% < 1.0]$. Possible scores are $[0.25, 0.5, 0.75, 1.0]$.
This is identical to normalizing by compressing (#1 above)

3c. Calculate "percent-true" for adult & children's `invitalscore`, then round to the nearest third to put scores into four bins



Four bins: $[\% > 0.0, \% > 0.33, \% > 0.67, \% = 1.0]$. Possible scores are $[0.0, 0.33, 0.67, 1.0]$.