Overall log for all tests

Test 1

The analysis was underpowered because 86 respondents were excluded because of extreme values, leaving 125 responses to be analysed. Mean person location was -1.02 (SD 0.65). Mean item fit residuals and person fit residuals for all items indicated problems with the model: -0.79 (SD 0.85) and 0.88 (SD 0,99) respectively. Although the test was satisfactory in terms of unidimensionalty, local dependency was found in 12 correlations. Two MCQs gave warrant for improvement, under discriminating and with one ability class deviating (q11 and q19.1).

Test 2

Test 2 had a good fit to the Rasch model in all tests. Mean person location was 1.448 (SD 1,026), Mean item fit residuals and person fit residuals were satisfactory: -0.32 (SD 1.56) and -0.22 (SD 0,69) respectively. The Cronbach’s alpha score was 0.7. Two items showed evidence of under discriminating and one ability class deviating (q3 and q18.3).

Test 3

Overall the test 3 showed satisfactory fit to the Rasch model. Mean item fit residuals and person fit residuals for all items were satisfactory: -0.090 (SD 1.29) and -0.30 (SD 1,08) respectively. All items fit well to the ICC-curve and no DIF on gender was identified. The Cronbach’s alpha score was 0.6.

Separate analyses of the two item-sets (MCQs and intended behaviour and self-efficacy items) suggest that the MCQ sub-test work very well, with no apparent validity issues.

Overall, the items evaluating self-efficacy also show satisfactory fit, but two of the intended behaviour items does not follow the logic of the Rasch model with disordered thresholds. Furthermore, the test for unidimensionality resulted in 17 significant T-tests (out of 160). Suggesting that this sub-test may not be unidimensional. This needs to be explored in further analysis, however one interpretation of this finding is that self-efficacy items and intended behaviour items can be considered as to measure two different dimensions. For example, intended behaviour may be more complex dependent on self-efficacy, but also values, knowledge and other factors. We identified no important local dependency in either item sets.

Test 4

Test 4 had a good fit to the Rasch model in all tests but was underpowered. Mean person location was 0,451 (SD 0,451). Mean item fit residuals and person fit residuals were satisfactory: 0.32 (SD 0,924) and -0.22 (SD 0,82) respectively. One item showed evidence of under discriminating and one ability class deviating (q5).