Appendix

These are the similar results that were not reported in the article. You may find the relevant information on the study design in the notes.

Table 1

Recovery Rates and Exposure Balance Indices for the DINA Model (Fixed-length)

| Item selection | Attribute | | | | | | Pattern | Exposure balance $\chi^2$ | Number of overused (>0.2) | Number of underused (<0.02) |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | | | | |
| PWKL | 0.999 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 | 310.470 | 398 | 33 |
| RT | 0.987 | 0.991 | 0.985 | 0.987 | 0.990 | 0.986 | 0.930 | 1.745 | 0 | 0 |
| RP | 0.986 | 0.990 | 0.991 | 0.988 | 0.985 | 0.994 | 0.935 | 0.263 | 0 | 0 |
| SDBS | 0.992 | 0.987 | 0.988 | 0.992 | 0.986 | 0.9863 | 0.931 | 4.542 | 0 | 0 |
| random | 0.933 | 0.931 | 0.941 | 0.949 | 0.951 | 0.939 | 0.714 | 0.477 | 0 | 0 |

Note1: PWKL = posterior weighed Kullback-Leibler method; RT= restrictive threshold method; RP= restrictive progressive method; SDBS = stratified dynamic binary searching.
Note2: This is a 24-item fixed-length CD-CAT for the DINA model. Item bank and examinees were simulated in the same way as Study II. The tuning parameter was set to be 10 for RT and RP; the number of strata was 6 for SDBS.

Table 2

The Measurement Accuracy and Test Length for the NC_RUM Model (Variable-length)

| | Item Selection | Attribute | | | | Pattern | Test Length | |
|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | | Mean | SD |
| 0.7 | PWKL | 0.943 | 0.948 | 0.949 | 0.900 | 0.806 | 7.20 | 1.845 |
| | SHTVOR | 0.903 | 0.926 | 0.934 | 0.920 | 0.736 | 20.28 | 10.083 |
| | DBS | 0.924 | 0.924 | 0.911 | 0.931 | 0.730 | 21.57 | 7.490 |
| 0.8 | PWKL | 0.938 | 0.951 | 0.960 | 0.972 | 0.833 | 8.59 | 2.369 |
| | SHTVOR | 0.946 | 0.933 | 0.949 | 0.950 | 0.801 | 23.87 | 10.341 |
| | DBS | 0.946 | 0.933 | 0.945 | 0.950 | 0.801 | 24.46 | 7.468 |
| 0.9 | PWKL | 0.984 | 0.978 | 0.987 | 0.978 | 0.929 | 11.86 | 3.232 |
| | SHTVOR | 0.971 | 0.971 | 0.983 | 0.979 | 0.907 | 35.73 | 13.258 |
| | DBS | 0.977 | 0.971 | 0.975 | 0.980 | 0.910 | 35.65 | 11.585 |

Note1: PWKL= posterior-weighted Kullback-Leibler method; SHTVOR = Sympson–Hetter method, which comprises test overlap control, variable length, online update, and restricted maximum information; DBS= dynamic binary searching.
Note2: This is a variable-length CD-CAT for the DINA model. Item bank and examinees were simulated in the same way as Study I and other were the same as in Study II except that the target test overlap rate $\overline{T}_{\max}$ for SHTVOR was set to be 0.06.

Table 3

Item Exposure and Item Bank Use for the NC_RUM Model (Variable-length)

| Stopping criteria | Item selection | Test overlap | Overused (>0.2) | Underused (<0.02) |
|---|---|---|---|---|
| 0.7 | PWKL | 0.757 | 8 | 434 |
| | SHTVOR | 0.061 | 0 | 34 |
| | DBS | 0.045 | 0 | 0 |
| 0.8 | PWKL | 0.754 | 10 | 425 |
| | SHTVOR | 0.075 | 0 | 25 |
| | DBS | 0.055 | 0 | 0 |
| 0.9 | PWKL | 0.740 | 14 | 418 |
| | SHTVOR | 0.081 | 0 | 18 |
| | DBS | 0.075 | 0 | 0 |

Note1: PWKL=posterior-weighted Kullback-Leibler index; SHTVOR = the Sympson–Hetter method, which comprises test overlap control, variable length, online update, and restricted maximum information; DBS= dynamic binary searching strategy.