Although we were not able to compute DIF under polytomous GGUM, we managed to obtain the Differential Test Functioning (DTF) by combining the test characteristic curves (TCC) of the two groups. As shown in Figure 18, under GGUM, the scale exhibited very small DTF, as the two pretty straight TCCs almost completely overlapped. DTF under SGR (Figure 19), on the other hand, was larger and non-uniform. When the well-being level was below 0, the Chinese participants had lower true scores, whereas their true scores were higher than the U.S. participants when the ability level was above 0. The two TCCs were very slightly S-shaped.

TCCs were also computed for the two groups, and were combined to represent the DTF (Figures 20-21). In general, under GGUM, the scale showed smaller DTF than under SGR, and DTF under both models were non-uniform. The two TCCs crossed at almost the same ability level (i.e., approximately -2.25) under the two models. Below this ability point, the Chinese participants had very slightly lower true scores than the American participants, but the trend reversed past this point, and the differences in the true scores became larger.