**To be added to Results:**

Although we were not able to compute DIF for the polytomous GGUM, we examined 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. To be more specific, when the well-being level was below 0, the Chinese participants had lower expected total score, whereas these scores became higher than the U.S. participants when the trait level was above 0. The two TCCs were very slightly S-shaped under SGR.

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

**To be added to Discussion:**

The fact that in our study, both scales were shown to have smaller DTF under GGUM than SGR, points out the importance of considering the use of GGUM in DIF analyses using personality data. Apparently in the current study, according to GGUM, both the Well-being and the Curiosity scales are more equivalent cross cultures than when they are examined with SGR.

**Fig.18.** IRT test characteristic curves (TCCs) of the Well-being scale under polytomous GGUM for the U.S. and the Chinese groups. Note: the red line represents the TCC of the Chinese group, and the green line represent the TCC of the U.S. group. The horizontal axis “Theta” represents the latent continuum from -3.0 to +3.0, and the vertical axis represents the expected item score.

**Fig.19.** IRT test characteristic curves (TCCs) of the Well-being scale under SGR for the U.S. and the Chinese groups. Note: the red line represents the TCC of the Chinese group, and the green line represent the TCC of the U.S. group. The horizontal axis “Theta” represents the latent continuum from -3.0 to +3.0, and the vertical axis represents the expected item score.

**Fig.20.** IRT test characteristic curves (TCCs) of the Well-being scale under polytomous GGUM for the U.S. and the Chinese groups. Note: the red line represents the TCC of the Chinese group, and the green line represent the TCC of the U.S. group. The horizontal axis “Theta” represents the latent continuum from -3.0 to +3.0, and the vertical axis represents the expected item score.

**Fig.21.** IRT test characteristic curves (TCCs) of the Well-being scale under SGR for the U.S. and the Chinese groups. Note: the red line represents the TCC of the Chinese group, and the green line represent the TCC of the U.S. group. The horizontal axis “Theta” represents the latent continuum from -3.0 to +3.0, and the vertical axis represents the expected item score.