Supplementary File 2: Criteria for quality appraisal

	0	**1**	**2**
Validity	Low: not reported	Modest: one type of validity	High: two types of validity
Reliability	Low: one type of reliability <0.5	Modest: one type of reliability > 0.7	High: two or more types of reliability: Cronbach's alpha >0.7 Inter-item correlations >0.5 Test-retest coefficient >0.7
User centeredness	Take no account of users' perceptions	Listen to users' perceptions, but not use their views to define or to modify items	Actively encourage users to specify what items should be stated, or how they should be defined or altered
Sample size	Low: < 5 participants per item	Modest: 5-10 participants per item	High: $\geq$ 300
Feasibility	Not reported	One aspect reported	More than one aspect reported

*Adapted from Hesselink et al. (2013)