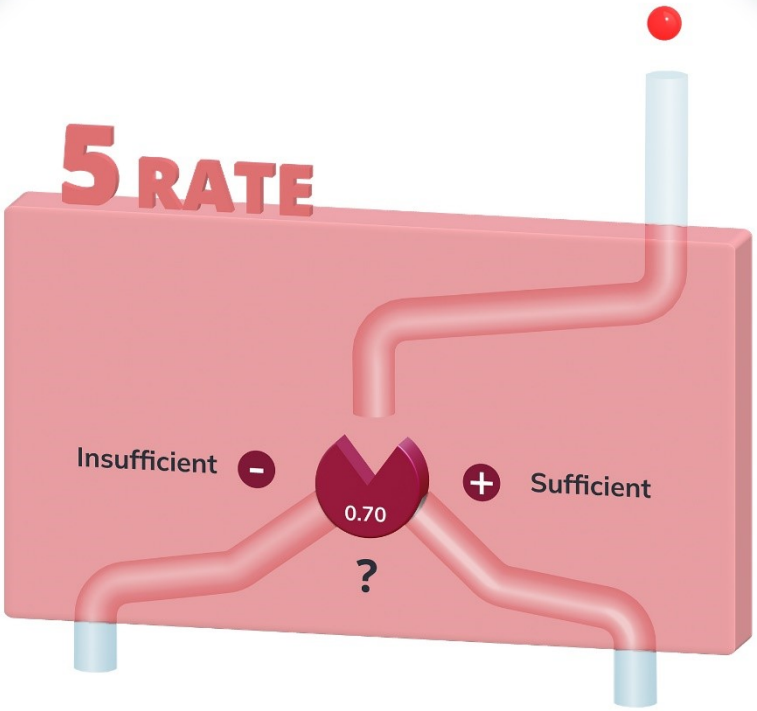
COSMIN

COSMIN Tool VERSION 2.0

Criteria for good measurement properties



Measurement property	**Rating**	**Criteria**
Content validity	+	Included items are relevant for the construct, target population, and context of use, and response options and recall period are appropriate *AND* No key concepts are missing *AND* PROM items and response options are appropriately worded and PROM instructions, items and response options understood by the population of interest as intended
	?	Not enough information reported
	-	Included items are not relevant for the construct or target population *OR* Key concepts are missing *OR* PROM items and response options are not appropriately worded or not understood by the population of interest as intended
Structural validity	+	**CTT:** EFA/PCA: factor loadings of each item on its factor ≥0.30 *AND* Maximum 10% of the items have factor loadings ≥0.30 on multiple factors *AND* Explained variance ≥50% and structure is in line with the theory about the construct to be measured *OR* results on scree plot or Kaiser criterion (Eigenvalues >1) are in line with the theory about the construct to be measured CFA: CFI or TLI or comparable measure >0.95 *OR* RMSEA <0.06 *OR* SRMR <0.08 **IRT/Rasch:** No violation of <u>unidimensionality:</u> CFI or TLI or comparable measure >0.95 *OR* RMSEA <0.06 *OR* SRMR <0.08 *AND* No violation of <u>local independence:</u> residual correlations among the items after controlling for dominant factor <0.20 *OR* Q3s <0.37 *AND* No violation of <u>monotonicity:</u> adequate looking graphs *OR* item scalability >0.30 *AND* Adequate <u>model fit:</u> IRT: χ^2 >0.01 Rasch: infit and outfit mean squares ≥0.5 and ≤1.5 *OR* Z-standardized values >-2 and <2
	?	Not enough information reported
	-	Criteria for '+' not met
Internal consistency	+	At least low evidence for sufficient unidimensionality *AND* Cronbach's alpha ≥0.70

	?	Criteria for "at least low evidence for sufficient unidimensionality" not met *OR* Evidence for insufficient unidimensionality *OR* Not enough information reported
	-	At least low quality evidence for sufficient unidimensionality *AND* Cronbach's alpha <0.70
Cross-cultural validity\ measurement invariance	+	No important differences found between group factors (such as age, gender, language) in multiple group factor analysis *OR* no important DIF for group factors (McFadden's R^2 <0.02)
	?	Not enough information reported
	-	Important differences between group factors *OR* DIF was found
Reliability	+	ICC or (weighted) kappa or Pearson/Spearman correlation ≥0.70
	?	Not enough information reported
	-	ICC or (weighted) kappa or Pearson/Spearman correlation <0.70
Measurement error	+	SDC or LoA <MIC
	?	MIC not defined *OR* not enough information reported
	-	SDC or LoA > MIC
Criterion validity	+	Correlation with gold standard ≥0.70 *OR* AUC ≥0.70
	?	Not enough information reported
	-	Correlation with gold standard <0.70 *OR* AUC <0.70
Hypotheses testing for construct validity	+	≥75% of the results is in accordance with predefined hypotheses
	?	No relevant results were found
	-	≥75% of the results deviates from predefined hypotheses
Responsiveness	+	≥75% of the results is in accordance with predefined hypotheses *OR* AUC ≥0.70
	?	No relevant results were found
	-	≥75% of the results deviates from predefined hypotheses *OR* AUC <0.70

AUC = area under the curve, CFA = confirmatory factor analysis, CFI = comparative fit index, CTT = classical test theory, DIF = differential item functioning, EFA = exploratory factor analysis, ICC = intraclass correlation coefficient, IRT = item response theory, LoA = limits of agreement, MIC = minimal important change, PCA = principal component analyses, RMSEA: Root Mean Square Error of Approximation, SEM = Standard Error of Measurement, SDC = smallest detectable change, SRMR: Standardized Root Mean Residuals, TLI = Tucker-Lewis index