

Minimizing Ambiguities:

A meta-analysis of the reliability and validity of healthcare uncertainty tolerance scales

G.C. Stephens¹, M.N. Karim², M. Sarkar³, A.B. Wilson⁴, M.D. Lazarus^{1,3}

1. Centre for Human Anatomy Education, Department of Anatomy and Developmental Biology, Faculty of Medicine, Nursing and Health Sciences, Monash University, Australia
2. School of Public Health and Preventive Medicine, Faculty of Medicine, Nursing and Health Sciences, Monash University, Australia
3. Monash Centre for Scholarship in Health Education, Faculty of Medicine, Nursing and Health Sciences, Monash University, Australia
4. Department of Anatomy and Cell Biology, Rush University, Chicago, IL

Purpose

Uncertainty tolerance (UT), a construct describing individuals' responses to uncertain stimuli, is considered highly relevant to healthcare practice. Despite a strong desire to evaluate UT, there remain inconsistent results between studies and across UT scales.¹ A first step in resolving these debates is evaluating UT scale efficacy. Accordingly, the present study asked **'How strong is the reliability and validity evidence for healthcare UT scales?'**

Methods

This systematic review evaluated UT scales administered to healthcare professionals and/or trainees with provision of evidence for validity and reliability. A data extraction rubric was developed based on principles outlined in the Standards for Educational and Psychological Testing.² A meta-analysis of studies' Cronbach's alphas evaluated aggregated internal consistency for 1) all included studies, 2) by scale, and 3) by population type.

Results

- **36 studies** met the inclusion criteria, among which **76 Cronbach's alphas** were reported. A meta-analysis of all alphas resulted in an aggregated internal consistency of **0.776** (95% CI = 0.752–0.800), representing respectable internal consistency across included studies.
- 4 scales appeared in ≥3 included (k=) studies: Physicians' Reactions to Uncertainty scale 1990 (**k=7, PRU1990**) and 1995 versions (**k=11 PRU1995**)³, Tolerance for Ambiguity scale (**k=6, TFA**)⁴, and Tolerance of Ambiguity in Medical Students and Doctors scale (**k=4 TAMSAD**)⁵

Table 1: Aggregated internal consistency by scale ranged from very good to only minimally acceptable. n= number of Cronbach alphas. p <0.001.

Scale	n	Aggregated Internal Consistency	95% Confidence Interval	Interpretation of Reliability
1. PRU1990	15	0.833	0.788–0.868	Very good
2. PRU1995	28	0.818	0.782–0.849	Very good
3. TFA	7	0.761	0.740–0.779	Respectable
4. TAMSAD	6	0.697	0.605–0.779	Minimally acceptable

Figure 1: Aggregated internal consistency was significantly (p < 0.001) higher in physicians than medical students. n= number of Cronbach alphas, CI = confidence interval



Physicians

0.797 (CI= 0.766–0.824)

n= 54

Respectable Reliability
(upper CI very good reliability)



Medical Students

0.707 (CI=0.674–0.737)

n= 20

Respectable Reliability (lower CI minimally acceptable reliability)

Table 2: Validity evidence was strong for PRU1990 and PRU1995, and weak for the TFA and TAMSAD. Validity evidence for response processes and testing consequences was critically lacking for all scales. k= number of studies.

Validity Evidence	Scale Each cell represents the sum of validity evidence ratings			
	PRU1990 (k=7)	PRU1995 (k=11)	TFA (k=7)	TAMSAD (k=4)
Test content	2	1	0	2
Response processes	0	0	0	0
Internal structure: Reliability	14	22	12	8
Internal structure: Construct validity	7	4	1	2
Relation to other variables: Convergent/divergent validity	6	4	4	1
Relation to other variables: Predictive/concurrent validity	14	20	12	6
Testing consequences	0	0	0	0
Total rating/maximum possible score	43/98	51/154	29/84	19/56
Overall quality and comprehensiveness	Strong	Strong	Moderate	Weak

Discussion

Of the four UT scales analyzed, none had a comprehensive validity argument, especially within the medical student context. Indeed, the validity evidence of three of the four major scales was initially evaluated in populations of physicians only, which, in combination with the lack of response processes, raises concerns about the use of these scales in medical student populations. It may be the case that the conceptualization of uncertainty differs between learner and expert populations, and thus further research exploring this is warranted.

Significance

Among scales analysed, PRU1990 and PRU1995 offer the strongest evidence for reliability and validity among physician populations. Based on currently available data, there is insufficient evidence for the reliability and validity of UT measures in medical student populations. We would caution against using these measures for high stakes applications such as medical student selection, curriculum assessment and/or program evaluation until further research is undertaken.

1. Strout, T. D., Hillen, M., Guthrie, C., Anderson, E., Hutchinson, R., Ward, H., Kay, H., Mills, G. J., & Han, P. (2018). Tolerance of uncertainty: A systematic review of health and healthcare-related outcomes. *Patient Education and Counseling*, 101(9), 1518–1537.
2. American Educational Research Association, American Psychological Association & National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
3. Gerrity, M. S., White, K. P., DeVellis, R. F., & Dittus, R. S. (1995). Physicians' reactions to uncertainty: Refining the constructs and scales. *Motivation and Emotion*, 19(3), 175–191.
4. Geller, G., Tambor, E. S., Chase, G. A., & Holtzman, N. A. (1993). Measuring physicians' tolerance for ambiguity and its relationship to their reported practices regarding genetic testing. *Medical Care*, 31(11), 989–1001.
5. Hancock, J., Roberts, M., Monrouxe, L., & Matlick, K. (2015). Medical student and junior doctors' tolerance of ambiguity: development of a new scale. *Advances in Health Sciences Education*, 20(1), 113–130.