

INTRODUCTION

- Use of psychosocial measures with different conceptual meanings across cultural groups may render treatment outcome analyses invalid in social work research.
- Determining measurement invariance allows researchers to assess whether the construct of a measure is similarly comprehended and measured across participant groups (e.g., based on race, ethnicity, gender, age, etc.).
- Nonequivalence is introduced when groups of participants experience or conceptualize a construct differently, or use distinctive criteria to describe the construct.
- Measurement nonequivalence across cultural groups is posited to occur due to **(a)** cultural differences in norms and relevance of the constructs being assessed; **(b)** language of assessment; or **(c)** potential differences in participants' environments and opportunity structures to engage in certain behaviors or develop beliefs due to contextual differences, racism, or other forms of discrimination.
- To illustrate this statistical procedure, measurement invariance properties across racial groups for two commonly used instruments in social work and substance abuse treatment research: *The Revised Helping Alliance Questionnaire (HAQ-II)* and *The Short Inventory of Problems (SIP-R)* are presented.

METHOD

Procedures

Publically available, secondary data were obtained from a randomized clinical trial conducted in the National Institute on Drug Abuse Clinical Trials Network. The selected trial investigated the effect Motivational Enhancement Therapy (MET) on retention and substance use among adults seeking treatment at five community-based outpatient treatment programs (Ball et al., 2007).

Sample

To assess invariance across racial groups, HAQ-II data were examined for 138 African American and 133 non-Latino White participants at their 2nd MET session.

- Mean age of the HAQ-II sample of African American and non-Latino White participants (n = 271) was 36.42 years (SD = 10.33).
- 72% of the HAQ-II sample was male.
- African American participants tended to be older (M = 38.25, SD = 10.06) than non-Latino White participants (M = 34.52, SD = 10.35).

Baseline SIP-R data were examined across 195 African American participants and 194 non-Latino White participants.

- Mean age of the SIP-R sample of African American and non-Latino White participants (n = 389) was 35.93 years (SD = 10.22).
- 71% of the SIP sample was male.
- African American participants tended to be older (M = 37.90, SD = 10.01) than non-Latino White participants (M = 33.93, SD = 10.07).

Measures

The Revised Helping Alliance Questionnaire (HAQ-II; Luborsky et al., 1996).

The HAQ-II is used to assess the therapeutic alliance between therapist and client in substance abuse treatment research and other clinical studies.

- The HAQ-II is a 19-tem instrument that assesses the therapeutic alliance from both the therapist and client perspectives.
- Each item is rated by the respondent on a six-point Likert-type scale (1 = strongly disagree, 6 = strongly agree).
- The HAQ-II has two subscales, positive (13 items) and negative therapeutic alliance (4 items).

The Short Inventory of Problems-Revised (SIP-R; Kiluk et al., 2012). The SIP-R is a 17-item instrument used to assess participants' perceptions of the adverse consequences of their substance use during 3 months prior to assessment.

- Each item is rated by the respondent on a four-point Likert-type scale. Anchors for Items 1-9 are 1 = never to 4 = daily or almost daily; while anchors for 10-17 are 1 = not at all to 4 = very much.
- The SIP-R has five factors: impulse control (3 items), interpersonal (3 items), intrapersonal (3 items), physical (3 items), and social (5 items).

RESULTS – Haq-II

Fit Indices by Measurement Invariance Model Test

	χ^2	df	$\Delta \chi^2$	Δdf	χ^2 difference test p	CFI	ΔCFI	RMSEA (90% CI)	$\Delta RMSEA$
Base Model: African American	176.123	113				0.94		0.06 (0.05, 0.08)	
Base Model: White, Non-Latino	214.66	113				0.91		0.08 (0.06, 0.09)	
Configural Invariance	390.79	226				0.93		0.07 (0.06, 0.09)	
Metric Invariance	408.44	241	17.65	15	.28	0.93	0.0	0.07 (0.06, 0.08)	0.0
Scalar Invariance	424.32	256	15.89	15	.39	0.93	0.0	0.07 (0.06, 0.08)	0.0
Invariance of Residual Variances	475.74	273	51.42	17	<.001	0.91	0.02	0.07 (0.06, 0.09)	0.0

Note. df = degrees of freedom; CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval.

RESULTS – SIP-R

Fit Indices by Measurement Invariance Model Test

	χ^2	df	$\Delta \chi^2$	Δdf	χ^2 difference test p	CFI	ΔCFI	RMSEA (90% CI)	$\Delta RMSEA$
Base Model: African American	186.41	107				0.96		0.06 (0.05, 0.08)	n/a
Base Model: White, Non-Latino	242.14	107				0.94		0.08 (0.07, 0.09)	n/a
Configural Invariance	428.55	214				0.95		0.07 (0.06, 0.08)	n/a
Metric Invariance	448.55	226	19.98	12	.07	0.95	0.00	0.07 (0.06, 0.08)	0.0
Strong/Scalar Invariance	483.56	238	35.03	12	<.001	0.94	-0.01	0.07 (0.06, 0.08)	0.0
Partial strong/scalar invariance model by freeing Items 3, 6, 11, 17	455.23	234	6.69	8	.57	0.95	0.00	0.07 (0.06, 0.08)	0.0
Strict invariance with partial strong/scalar invariance model	493.19	251	37.96	17	<.001	0.94	-0.01	0.07 (0.06, 0.08)	0.0
Partial strict invariance model by freeing Items 1 and 4	476.05	249	20.82	15	.14	0.94	-0.01	0.07 (0.06, 0.08)	0.0

Note. df = degrees of freedom; CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval.

DISCUSSION

- Many psychosocial measures used in social work research are susceptible to measurement nonequivalence because measurement nonequivalence is frequently overlooked by researchers.
- Use of measures with different conceptual meaning across racial and ethnic groups may render invalid analyses comparing such groups. Conclusions drawn from invalid findings can lead to ineffective treatments and policy initiatives.
- Findings supported the comparable understanding of therapeutic alliance and consequences of substance as measured by the HAQ-II and SIP-R in African American and non-Latino White participants.
- Difference in reliability caused by the identified items needs verification in future studies to ensure use of reliable Haq-II and SIP-R latent factors.

ACKNOWLEDGEMENT: The information reported here results from secondary analyses of data from clinical trials conducted as part of the National Drug Abuse Treatment Clinical Trials Network (CTN) sponsored by the National Institute on Drug Abuse (NIDA). Specifically, data from CTN-0004 (MET to Improve Treatment Engagement and Outcome in Subjects Seeking Treatment for Substance Abuse) were included. CTN databases and information are available at www.ctndatashare.org. Presentation was supported by award number P20MD002288 from NIMHD. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIMHD or NIH.