cogent
psychology

EVOLUTIONARY PSYCHOLOGY | RESEARCH ARTICLE

The psychometric properties of the Arabic version of the mental retardation attitude inventory-revised (MRAI-R) scale

Ghaleb H. Alnahdi[*]

Abstract: Background: This study examined the psychometric properties of the Arabic version of the Mental Retardation Attitude Inventory-Revised (MRAI-R) scale. **Method**: Data were collected from 455 undergraduate college students (214 female, 241 male). A statistical analysis was conducted using confirmatory factor analysis (CFA), exploratory factor analysis (EFA), and Cronbach's alpha. **Results**: The internal consistency of the MRAI-R scale overall was good (0.76). However, it was less than 0.7 for the four subscales. CFA results for the 36-item scale indicated that the observed data did not support the four-factors model. However, separating the scale into two scales based on the phrasing type of items (positively phrased or negatively phrased) resulted in an acceptable fit for the model. **Conclusion**: Replication of this research with a different sample from Saudi Arabia is required to confirm the results of this study. Any future iteration of this study should consider the rephrasing of some items.

Subjects: Test Development, Validity & Scaling Methods; Attitudes & Persuasion; Disability Studies; Disability

Keywords: Arabic; Saudi Arabia; intellectual disability; MRAI-R; validation; psychometric properties; CFA; Negatively phrased items

1. Background

Numerous educational disciplines lack scales whose psychometric characteristics have been studied and tested with different samples. Indeed, one of the challenges facing special education researchers and specialists in the Arab world is the dearth of tested scales with confirmed psychometric properties. To begin to fill that gap, it makes sense to translate into Arabic the

[*]Corresponding author: Ghaleb H. Alnahdi, Department of Special Education, Prince Sattam Bin Abdulaziz University (Prince Salman University), Saudi Arabia.
E-mail: ghalnahdi@gmail.com; g. alnahdi@psau.edu.sa

Additional information is available at the end of the article

ABOUT THE AUTHOR

Ghaleb H. Alnahdi is an associate professor in special education at Prince Sattam bin Abdulaziz University. He earned his Ph.D. in Special Education from Ohio University. He also holds a master's degree in research and evaluation from Ohio University (2012) and another, in special education, from King Saud University (2007). His research focuses on intellectual disability, inclusive education, cross-cultural validation of scales, and teacher preparation. He is involved in several research projects with different research groups at the national and international levels.

PUBLIC INTEREST STATEMENT

The Mental Retardation Attitude Inventory-Revised (MRAI-R) scale was originally developed in English language. Therefore, it is important to ensure that the translated version do maintain the psychometric properties of the original version. This study examines the suitability of the Arabic version of the Mental Retardation Attitude Inventory-Revised (MRAI-R) scale. The information from this study will be important to confirm whether this scale is appropriate for use to measure attitudes towards people with intellectual disability in the Arab world.

scales used in other countries to assess people with intellectual disabilities (PWID) and test them with different samples. Testing the psychometric properties of such scales using local samples begins the process of modification required to adapt these scales to local environments. Toward that end, this study aimed to examine the psychometric properties of the Mental Retardation Attitude Inventory-Revised (MRAI-R).

The number of special education programs for students with intellectual disability (ID) in Saudi Arabia has grown in recent years (Alnahdi, 2014), which has increased the chances of these students to interact with typically developing students. Negative attitudes pose one of the main obstacles to successful inclusion for PWID. Hence, measuring social attitudes can help overcome these obstacles and ensure the successful integration of PWID (Alnahdi, 2019).

The MRAI-R scale is commonly employed to examine people's attitudes towards PWID and has been used over the last 22 years with different samples from various countries. Noted for its reliability and multidimensionality (Rice, 2009), researchers from the United States (Antonak & Harth, 1994; Krajewski & Flaherty, 2000; McManus, Feyes, & Saucier, 2011), China (Hampton & Xiao, 2008, 2009), Japan (Horner-Johnson et al., 2002), Kuwait (Al-Kandari, 2015), and Australia (Yazbeck, McVilly, & Parmenter, 2004) have advocated for its use. Yet, while researchers from many countries have tested the scale, "the majority [of] studies have been conducted in western countries, and the generalisability of the results to non-western countries is relatively unknown" (Hampton & Xiao, 2009, p. 300). Indeed, the present study represents the first use of the MRAI-R with a sample from Saudi Arabia.

Previous studies have examined the psychometric properties of the MRAI-R. The scale's developers hypothesised a four-factor structure (Antonak & Harth, 1994), which successfully fit the data in a study involving 286 Japanese college students (Horner-Johnson et al., 2002). In a study of 534 Chinese college students (Hampton & Xiao, 2008), however, the data did not fit the four-factors model; that study's authors reported that only the social distance (SDIS) subscale of the MRAI-R produced reliable measurements.

Still, though some studies have failed to fit the data to the hypothesised model or generate sufficient reliability statistics, researchers recommend endeavouring to validate and revise existing scales over developing new ones from scratch (Antonak & Livneh, 2000). Thus, this study examined the construct validity of the Arabic version of the MARI-R scale in a sample of Saudi college students.

The study pursued the following research questions: (1) Is the MRAI-R a reliable measure for use with Saudi college students? (2) Does the four-factors model explain this study's data? (This study set out under the assumption that the observed data would indeed fit the hypothesised four-factor model.)

The MRAI-R scale was developed as a revised version of Harth's (1974) original 50-item summated-rating inventory to measure attitudes toward PWID (Antonak & Harth, 1994). The MRAI-R scale has four subscales: the integration segregation subscale (INSE) to measure participants' attitudes about mainstreaming PWID in classrooms and other settings, the SDIS subscale to measure participants' attitudes toward socially interacting with and physical proximity to PWID, the private rights (PRRT) subscale to measure participants' attitudes toward PWID rights in society, and the subtle derogatory beliefs (SUDB) subscale to measure participants' attitudes toward derogatory beliefs about PWID.

2. Methods

2.1. Sample
The sample of this study comprised male and female students at a university in Saudi Arabia. Universities in Saudi Arabia are separated by gender. Questionnaires were distributed to the male and female sections of the university via professors. A total of 455 students participated in the

study, with female participants representing 47 % (214) of the sample and male participants representing 53% (241) of the sample. Ages ranged from 18 to 35 years old, with a mean age of 21.8 years old. Participants were allowed to opt out of completing the scale at any time. The sample met Bryant and Yarnold (1995) minimum sample size recommendation for confirmatory factor analysis with the subject-to-variable ratio equal five (62 parameters *5 = 310), as well as the minimum sample size requirements of Wolf, Harrington, Clark, and Miller (2013).

2.2. Translation

The study adapted the Arabic version of the MRAI-R translated by Al-Kandari and Salih (Al-Kandari & Salih, 2008), retaining only 29 items of the 34 items to match the original scale (the translators had added 5 items). Moreover, the version used in this study dispensed with updates made to 22 of the items for the 1994 revision (Antonak & Harth, 1994). Those modifications had been made in 7 items for political and social reasons, in 6 items for cultural reasons, and 8 items for the sake of simplification. As these reasons do not apply to the Saudi context, these 22 items were translated directly from the original English scale. In sum, due to the modifications that took place in the Arabic version by Al-Kandari and Salih (2008) to the original English scale, all items were comparable to the original version by Antonak and Harth (1994) than to the Al-Kandari and Salih (2008) version.

The Arabic version was checked against the English version of the scale, which were then compared by a bilingual researcher, who made some small modifications for clarity. Essentially, all 29 items retained in this study scale were identical to the original English version. Finally, a pilot study was conducted with 30 college students who provided feedback on the clarity of the items. Twelve items were positively phrased, meaning agreement with them indicates positive attitudes, and seventeen items were negatively phrased, these items were recoded so disagreement with these items indicates positive attitudes.

These items were divided into the four following subscales: First, the social distance (SDIS) subscale which has eight items (e.g. "I would rather not have a person who has ID swim in the same pool that I swim in"); Second, the integration-segregation (INSE) subscale which has seven items (e.g. "school officials should not place children with ID and children without ID in the same classes"); Third, the private rights (PRRT) subscale which has seven items (e.g. "a person should not be permitted to run a day care centre if he or she will not serve children who have ID"); and fourth, the subtle derogatory beliefs (SUDB) subscale which has seven unfavourable statements about PWID (e.g. "children who are ID waste time playing in class instead of trying to do better").

2.3. Analysis procedures

Statistical analysis took place in four steps to assess the structural validity of the Arabic version of the MRAI-R. First, Cronbach's alpha was computed for overall scale and for the four subscales. Second, a confirmatory factor analysis (CFA) determine whether the data fit the four-factors model of the original scale. Third, if the data did not fit the original structure, an exploratory factor analysis (EFA) was conducted to understand how the items clustered together. Fourth, after reviewing the EFA outputs, another CFA was conducted based on EFA results.

3. Results

3.1. Reliability

Table 1 shows results from 9 other studies using the same scale in different countries within the past 22 years to contextualize the results of the internal consistency analysis in this study. The overall review would give better understanding. Cronbach's alpha for the overall scale was > 0.7, but it was < 0.7 for all four subscales. In the other studies reported in Table 1, only the original study by Antonak and Harth (Antonak & Harth, 1994) had a Cronbach's alpha > 0.7 for the overall scale and all four subscales. For the PRRT subscale, Cronbach's alpha ranged from 0.5 to 0.63 for most studies (Hampton & Xiao, 2008; Krajewski & Flaherty, 2000; McManus et al., 2011), with one study (Sam, Li, & Lo, 2016) in addition to the original study (Antonak & Harth, 1994) coming in

Table 1. Reliability statistics of MRAI-R studies

Year	Study	Country	Overall	INSE	SDIS	PRRT	SUDB
1994	(Antonak & Harth, 1994)	USA	0.91	0.81	0.82	0.76	0.73
1999	(MacDonald & MacIntyre, 1999)	Canada	0.9	refer to (Antonak & Harth, 1994) reliability results			
2000	(Krajewski & Flaherty, 2000)	USA	NA	0.71	0.84	0.59*	0.6*
2002	(Horner-Johnson et al., 2002)	Japan		report (Antonak & Harth, 1994) reliability results			
2004	(Yazbeck et al., 2004)	Australia		refer to (Antonak & Harth, 1994) reliability results			
2008	(Hampton & Xiao, 2008)	China	0.78	0.5*	0.78	0.5*	0.21*
2009	(Hampton & Xiao, 2009)	China	0.78	NA	NA	NA	NA
2011	(McManus et al., 2011)	USA	NA	0.82	0.78	0.63*	0.77
2016	Sam et al., 2016	China	NA	.664	.833	.748	0.496*
2018	This study	Saudi Arabia	0.762 (.784[a])	0.49*	0.64*	0.503*	0.59*
			0.793[b] (.726[c])				

*alpha < 0.7, NA = not available, a = 27 items (items 21 and 26 excluded), b = 17 negatively phrased items, c = 12 positively phrased items.

higher than 0.7. Cronbach's alpha for the SDIS subscale was better than for other subscales. Three studies (Hampton & Xiao, 2008; Krajewski & Flaherty, 2000; McManus et al., 2011) in addition to the original scale study (Antonak & Harth, 1994) had a Cronbach's alpha > 0.7, though it was 0.64 for this study; other studies (Horner-Johnson et al., 2002; MacDonald & MacIntyre, 1999; Sam et al., 2016; Yazbeck et al., 2004) did not report Cronbach's alpha for this subscale and simply referred to that of the original study. For the SUDB subscale, Cronbach's alpha ranged from 0.21 to 0.60 (Hampton & Xiao, 2008; Krajewski & Flaherty, 2000; Sam et al., 2016) in most studies, with one study (McManus et al., 2011) in addition to the original study (Antonak & Harth, 1994) with a Cronbach's alpha > 0.7. The INSE subscale was in between the other subscales, and Cronbach's alpha was > 0.7 for two studies (Krajewski & Flaherty, 2000; McManus et al., 2011) in addition to the original study. Two studies (including this study) reported Cronbach's alphas < 0.7, 5 other studies did not report, and 3 of them refer to the results of the original study. In sum, the majority of studies showed acceptable reliability for the overall scale. Still, many of the studies reported Cronbach's alpha < 0.7 in all subscales except the SDIS subscale. In addition, some studies referred to the reliability scores of the original scale, which might indicate that their data could not support internal consistency (Horner-Johnson et al., 2002; MacDonald & MacIntyre, 1999; Yazbeck et al., 2004).

3.2. Correlations
Pearson correlation coefficients were calculated for the overall MRAI-R scale along with the four subscales. All of these correlations were statistically significant (see Table 2). Twenty-seven items from the overall MRAI-R scale correlated significantly at $p = 0.01$, with correlations ranging from 0.210 to 0.535 (see appendix).

Two items (21 and 26) were not significantly correlated with the overall scale. In future iterations of the Arabic MRAI-R, these two items might need to be phrased more simply and directly to facilitate quicker comprehension by college-age participants, who often try to screen questions as fast as possible. For example, item 26 "Even though PWID have some reasons to complain, they would get what they want if they were just more patient" comprises two parts, both necessary to convey the intended meaning; if this were conveyed in a simpler fashion, it might reduce the amount of time required for comprehension and a cogent response. Item 21, "The problem of prejudice against PWID has been exaggerated", was phrased to examine attitudes indirectly. First, there is the problem of prejudice: participants might believe either that the problem does or does not exist. Then, if participants believe it indeed exists, there is another element of the question, namely, whether the problem has been exaggerated or not. The second component of the item does not apply to participants who do not believe that there is a prejudice problem.

3.3. Exploratory factor analysis
After conducting the CFA and finding that the observed data failed to fit the four-factors model as hypothesised, EFA was performed to understand how items were loading and clustering together. Principal components analysis with Varimax rotation (Table 3) was conducted using the statistical package software (SPSS-21) with four fixed factors. In the first factor, 14 items loaded with > 0.30, and all of these items were negatively phrased prompts from all four subscales. Four items loaded on the second factor, though these items were supposed to load on three different subscales according to the original scale. Six items loaded on the third factor, though these items were supposed to load on three different factors. Five items loaded on the fourth factor, though these items were supposed to load on four different subscales according to the original scale.

3.4. Confirmatory factor analysis
CFA is recommended for testing the construct validity of a scale (De Vet, Terwee, Mokkink, & Knol, 2011) and was used in this case to examine whether the four-factors model for the 29 items would explain the observed data in this study, as hypothesised. Chi squares are reported in Table 4 but are further discussed here because of the sensitivity to sample size even with data that fits the model reasonably well (Byrne, 2010). A CFI value ≥ 0.90 (Pugesek, Tomer, & Von Eye, 2003), a root

cogent •• psychology

Table 2. Means, standard deviations, and Pearson correlations

		N	Mean	Std. Deviation	1	2	3	4	5	6	7
1	Integration segregation	453	3.052	0.594	1						
2	Social distance	453	3.606	0.628	0.327**	1					
3	Private rights	453	3.552	0.607	0.198**	0.443**	1				
4	Subtle derogatory beliefs	453	3.166	0.606	0.186**	0.182**	0.112*	1			
5	Overall Mean	453	3.353	0.419	0.628**	0.796**	0.670**	0.549**	1		

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

Table 3. EFA loading with Varimax rotation

	Components			
	1	**2**	**3**	**4**
Item25*	0.639			
Item15*	0.626			
Item18*	0.618			
Item27*	0.598			
Item23*	0.549			
Item13*	0.542	0.348		
Item4*	0.538			
Item12*	0.528			
Item16*	0.513			
Item9*	0.512			
Item10*	0.476			
Item20*	0.47			
Item6*	0.431			
Item1*	0.377	0.352	−.359−	
Item24		0.679		
Item29		0.668		
Item28		0.532		
Item14		0.374		
Item3			0.669	
Item5			0.602	
Item8		0.35	0.571	
Item2			0.484	
Item7*			−.457−	
Item11		0.352	0.353	0.301
Item19				0.631
Item21*				−.624−
Item22				0.602
Item26*				−.471−
Item17		0.375		0.449

Only loadings > .3 are shown in this table, * negatively phrased item.

Table 4. Goodness-of-fit indices for the four-factors model of the MRAI-R scale

Model	SBS-χ^2	p	df	χ^2/df	RMSEA	CFI	SRMR	GFI
M1	1465.88	0.00	371	3.95	0.081	0.498	0.1126	0.778
M2	1338.81	0.00	364	3.67	0.077	0.553	0.1031	0.777
M3[a]	104.26	0.00	47	2.21*	0.052*	0.923*	0.0484	0.962*
M4	197.86	0.00	112	1.76*	0.041*	0.922*	0.0450	0.953*
M5	534.509	0.00	129	4.14	0.083	0.645	0.0900	0.859

*indicator of a good fit, M1 = initial model with 4 factors, M2 = initial model with 4 factors with controlling for items errors, M3 = three factors model with passively phrased items, M4 = four factors model with negatively phrased items, M5 = model with 17 items recommended by (Al-Kandari & Salih, 2008)

mean square error of approximation (RMSEA) of 0.05 or lower (Hu & Bentler, 1998, 1999; Schermelleh-Engel, Moosbrugger, & Müller, 2003), and a Chi square less than 3 (Kline, 1994) were considered indicators of good fit.

The first model (M1) was the four-factors model with data from 432 participants with no covariance between item errors. The fit indices for this model were RMSEA (0.081), CFI (0.498), GFI (0.778), and χ2/df = 3.95, which was a clear indication that four-factors structure did not support the data. The second CFA (M2) added a covariance between some item errors, as recommended by the software (AMOS) to improve the model fit. The fit indices for this model were RMSEA (0.52), CFI (0.553), GFI (0.777), and χ2/df = 3.67, which showed no significant improvement in the fit indices in comparison with M1. Thus, M1 and M2 failed to explain the hypothesised model with the observed data in this study. The third CFA (M3) was done with 12 positively phrased items. M3 only used three factors because the SUDB subscale has no positively phrased items. The fit indices for this model were RMSEA (0.052), CFI (0.923), GFI (0.962), and χ2/df = 2.21. The fourth CFA (M4) was done with only 17 negatively phrased items. The fit indices for this model were RMSEA (0.041), CFI (0.922), GFI (0.953), and χ2/df = 1.76. A fifth run of CFA (M5) was conducted with 17 items that Al-Kandari and Salih (Al-Kandari & Salih, 2008) recommend for use with one factor only. The results showed that this model also failed to explain the data and the fit indices were as follows: RMSEA (0.083), CFI (0.645), GFI (0.859), and χ2/df = 4.14.

In sum, the hypothesised model with 29 items with both positively and negatively phrased items failed to fit the data. However, separating items based on negative or positive phrasing type appeared to improve the situation: M4, which used only 17 negatively phrased items, showed adequate fit indices to the hypothesised four-factors model, and the M3 showed adequate fit indices to a three-factors model (since the fourth subscale has no positively phrased items).

4. Discussion

This study examined the construct validity of the MRAI-R among Saudi college students and failed to support the hypothesised four-factor structure of the MRAI-R with data from participants. This finding is consistent with studies using data from Chinese college students (Hampton & Xiao, 2008). Hampton and Xiao (2008) concluded that the differences in social environment and available services, with China on one side and the United States and Japan on the other, influenced the responses to items which led to findings that data from China did not fit the four-factors model. Horner-Johnson et al. (2002) also pointed out the need to further explore the generalisability of the MRAI-R in different cultures, as the MRAI-R might not elicit nuances of attitudes toward PWID in other cultures.

The results of this study are consistent with findings from other studies that recommend not to use both positively and negatively phrase items in the same scale (Barnette, 2000; Benson & Hocevar, 1985; Stewart & Frye, 2004). Indeed, some studies suggest reversing the answer options for half of the sample (Barnette, 2000). So, half of the sample would have the options start from "strongly agree", and the other half would have the options start from "strongly disagree". The effect of phrasing items on the validity and reliability of scales has been of interest to many studies (Barnette, 2000; Benson & Hocevar, 1985; Pilotte & Gable, 1990; Schriesheim, Eisenbach, & Hill, 1991; Spector, Van Katwyk, Brannick, & Chen, 1997; Stewart & Frye, 2004; Stewart, Roberts, Eleazer, Boland, & Wieland, 2006). For example, in some studies, items loaded on different factors based on type of phrasing Lewis and Sauro (2017), and other researchers had to remove some negatively phrased items to improve the internal consistency (Mogre & Amalba, 2016). Similarly, it was evident in some studies that the negatively phrased items to loaded on a single factor, while the assumption was to load on different factors (Pilotte & Gable, 1990; Stewart et al., 2006). In addition, some researchers believe that having negatively phrased items might result in having "artifactual factor" in the analysis (Spector et al., 1997).

In conclusion, the Arabic version of the MRAI-R scale showed a poor fit with the four-factor structure of the original scale. However, separating the scale into two scales based on the phrasing type of items resulted in an acceptable fit for the model. This study's findings imply some salient points. First, since this is the first MRAI-R study using a Saudi sample, it would be useful to replicate it with a different sample within Saudi Arabia. Second, in a future iteration of the study, it would be optimal to use only one type of phrasing for items. For example, it would be better to have the

**cogent **psychology

entire survey be comprised of positively phrased items, this would make it easier for participants to respond to positively phrased items, since participants would not have to switch between positively and negatively phrased items that might be perceived differently by each participant (Roszkowski, & Soven, 2010).

In addition, it would be advisable to use a scale of "strongly agree" to 'strongly disagree for half of the participants while using "strongly disagree" to "strongly agree" for the other half. Third, the two items, item 21 *"the problem of prejudice towards people with ID has been exaggerated"* and item 26 *"even though people with ID have some cause for complaint, they would get what they want if they were more patient"* showed non-significant correlation with the scale as whole, as it does not seem to be due to translation issues, the items were not simple in phrasing to allow participants to answer them easily in short time. Rephrasing of these items would be recommended, and it would be worthy to examine how it will fit in any future study. Finally, this study had one limitation related to the sample, which was chosen based on convenience; students from other universities across Saudi Arabia might not respond the same to the scale prompts.

Author details
Ghaleb H. Alnahdi
E-mail: ghalnahdi@gmail.com
ORCID ID: http://orcid.org/0000-0003-1692-9765
Special Education Department, College Of Education, Prince Sattam bin Abdulaziz University, Alkahrj, Saudi Arabia.

Conflict of interest statement
The author has no conflicts of interest to declare.

References
Al-Kandari, H. Y. (2015). High school students' contact with and attitudes towards persons with intellectual and developmental disabilities in Kuwait. *Australian Social Work*, *68*, 65–83. doi:10.1080/0312407X.2014.946429

Al-Kandari, H. Y., & Salih, F. A. (2008). Adapting antonak and harth's mental retardation attitude inventory for kuwait's culture. *Digest of Middle East Studies*, *17*(1), 1–31.

Alnahdi, G. H. (2014). Special education programs for students with intellectual disability in Saudi Arabia: Issues and recommendations. *Journal of the International Association of Special Education*, *15*(1), 83–91..

Alnahdi, G. H. (2019). The positive impact of including students with intellectual disabilities in schools: Children's attitudes towards peers with disabilities in Saudi Arabia. *Research in Developmental Disabilities*, *85*, 1–7. doi:10.1016/j.ridd.2018.10.004

Antonak, R. F., & Harth, R. (1994). psychometric analysis and revision of the mental retardation attitude inventory. *Mental Retardation*, *32*, 272–280.

Antonak, R. F., & Livneh, H. (2000). Measurement of attitudes towards persons with disabilities. *Disability and Rehabilitation*, *22*, 211–224. doi:10.1080/096382800296782

Barnette, J. J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, *60*, 361–370. doi:10.1177/00131640021970592

Benson, J., & Hocevar, D. (1985). The impact of item phrasing on the validity of attitude scales for elementary school children. *Journal of Educational Measurement*, *22*, 231–240. doi:10.1111/jedm.1985.22.issue-3

Bryant, F. B., & Yarnold, P. R. (1995). *Principal-components analysis and exploratory and confirmatory factor analysis*. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 99 –136). Washington, DC: American Psychological Association.

Byrne, B. M. (2010). Multivariate applications series. *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York, NY: Routledge/Taylor & Francis Group.

De Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: A practical guide*. Cambridge: Cambridge University Press.

Hampton, N. Z., & Xiao, F. (2008). Psychometric properties of the mental retardation attitude inventory-revised in Chinese college students. *Journal of Intellectual Disability Research*, *52*, 299–308. doi:10.1111/j.1365-2788.2007.01020.x

Hampton, N. Z., & Xiao, F. (2009). Traditional Chinese values and attitudes of Chinese university students toward people with intellectual disabilities. *International Journal of Disability, Development and Education*, *56*, 247–261. doi:10.1080/10349120903102270

Harth, R. (1974). Attitudes towards minority groups as a construct in assessing attitudes towards the mentally retarded. *Education and Training of The Mentally Retarded*, *6*, 142–147.

Horner-Johnson, W., Keys, C., Henry, D., Yamaki, K., Oi, F., Watanabe, K., ... Fugjimura, I. (2002). Attitudes of Japanese students toward people with intellectual disability. *Journal of Intellectual Disability Research*, *46*, 365–378. doi:10.1046/j.1365-2788.2002.00406.x

Horner-Johnson, W., Keys, C., Henry, D., Yamaki, K., Oi, F., Watanabe, K., ... Fugjimura, I. (2002). Attitudes of Japanese students toward people with intellectual

disability. *Journal of Intellectual Disability Research, 46*, 365–378. doi:10.1046/j.1365-2788.2002.00406.x

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424. doi:10.1037/1082-989X.3.4.424

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1–55. doi:10.1080/10705519909540118

Kline, P. (1994). *An easy guide to factor analysis*. New York: Routledge Taylor & Francis Group.

Krajewski, J., & Flaherty, T. (2000). Attitudes of high school students toward individuals with mental retardation. *Mental Retardation, 38*, 154–162. doi:10.1352/0047-6765(2000)038<0154:AOHSST>2.0.CO;2

Lewis, J. J. R., & Sauro, J. (2017). Revisiting the factor structure of the system usability scale. *Journal of Usability Studies, 12*(4), 183–192.

MacDonald, J. D., & MacIntyre, P. D. (1999). A rose is a rose: Effects of label change, education, and sex on attitudes toward mental disabilities. *Journal on Developmental Disabilities, 6*, 15–31.

McManus, J. L., Feyes, K. J., & Saucier, D. A. (2011). Contact and knowledge as predictors of attitudes toward individuals with intellectual disabilities. *Journal of Social and Personal Relationships, 28*, 579–590. doi:10.1177/0265407510385494

Mogre, V., & Amalba, A. (2016). Psychometric properties of the dundee ready educational environment measure in a sample of Ghanaian medical students. *Education for Health, 29*(1), 16. doi:10.4103/1357-6283.178921

Pilotte, W. J., & Gable, R. K. (1990). The impact of positive and negative item stems on the validity of a computer anxiety scale. *Educational and Psychological Measurement, 50*, 603–610. doi:10.1177/0013164490503016

Pugesek, B. H., Tomer, A., & Von Eye, A. (2003). *Structural equation modeling: Applications in ecological and evolutionary biology*. Cambridge: Cambridge University Press.

Rice, C. J. (2009). Attitudes of undergraduate students toward people with intellectual disabilities: Considerations for future policy makers. *College Student Journal, 43*(1), 207–215..

Roszkowski, M. J., & Soven, M. (2010). Shifting gears: Consequences of including two negatively worded items in the middle of a positively worded questionnaire. *Assessment & Evaluation in Higher Education, 35*(1), 113–130.

Sam, K. L., Li, C., & Lo, S.-K. (2016). Validation of the mental retardation attitude inventory-revised (MRAI-R): A multidimensional rasch analysis. *International Journal of Social Science and Humanity, 6*, 519–524. doi:10.7763/IJSSH.2016.V6.703

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*, 23–74.

Schriesheim, C. A., Eisenbach, R. J., & Hill, K. D. (1991). The effect of negation and polar opposite item reversals on questionnaire reliability and validity: An experimental investigation. *Educational and Psychological Measurement, 51*(1), 67–78. doi:10.1177/0013164491511005

Spector, P. E., Van Katwyk, P. T., Brannick, M. T., & Chen, P. Y. (1997). When two factors don't reflect two constructs: How item characteristics can produce artifactual factors. *Journal of Management, 23*(5), 659–677. doi:10.1177/014920639702300503

Stewart, T. J., & Frye, A. W. (2004). Investigating the use of negatively phrased survey items in medical education settings: Common wisdom or common mistake? *Academic Medicine, 79*, S18–S20. doi:10.1097/00001888-200410001-00006

Stewart, T. J., Roberts, E., Eleazer, P., Boland, R., & Wieland, D. (2006). Reliability and validity issues for two common measures of medical students' attitudes toward older adults. *Educational Gerontology, 32*(6), 409–421. doi:10.1080/03601270600685628

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 73*, 913–934. doi:10.1177/0013164413495237

Yazbeck, M., McVilly, K., & Parmenter, T. R. (2004). Attitudes toward people with intellectual disabilities: An Australian perspective. *Journal of Disability Policy Studies, 15*, 97–111. doi:10.1177/10442073040150020401

Cogent Psychology (ISSN: 2331-1908) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com