DATA METHODS PROJECT REPORT

PROJECT

- Generate a survey in Qualtrics for collecting the measures of Universal Diverse Orientation.

- Run a confirmatory factor analysis using Stata on original data collected from the Qualtrics survey.

RESEARCH QUESTION

How well do the observed variables represent the unobserved constructs of Universal Diverse Orientation (UDO): diversity of contact, relativistic appreciation, and comfort with differences? Can using the confirmatory factors analysis (CFA) statistical tool provide useful information about a new/updated Miville-Guzman Universality-Diversity Scale's (M-GUDS) dimensionality and validity to improve the quality of measurement of Universal Diverse Orientation?

FRAMEWORK

This project used the structural equation modeling (SEM) framework to run a confirmatory factor analysis (CFA) of the first implementation of an updated UDO survey measurement (Ho, 2020). SEM allowed testing of unobserved concepts and covariance structure in the new UDO survey measure (Ho, 2020). CFA was used on the UDO survey measure as the constructs were measured with multiple items, the scale items had a linear relationship to the scale total, and the researcher had a priori idea of which items measure which constructs (Levine, 2007). The original theory of Universal Diverse Orientation guided the building of a three-factor structure model (Ho, 2020). The framework prescribed four steps: 1. referencing the theory for the model, 2. using a structural equation model of the theory to build a path diagram, 3. fitting the path diagrams factorial structure to the original sample data using the principal factors method, and 4. checking the goodness-of-fit of the model (Ho, 2020).

1. THEORY

Confirmatory Factor Analysis (CFA) is a theory driven factor analysis that uses sample data to verify the factor structure of the measurement model by analyzing measurement errors (Ho, 2020). Universal Diverse Orientation is measured with the Miville-Guzman Universality-Diversity Scale(M-GUDS) 15-item survey. A Universal Diverse Orientation (UDO) score is comprised of 3 sub-scores: 1. relativistic appreciation, 2. diversity of contact, and 3. comfort with differences (Miville, et al., 1999).

2. STRUCTURAL EQUATION MODEL | PATH DIAGRAM

SEM's features allow researchers to conceptualize variables using path diagrams with boxes, circles, and arrows to visualize the relationships (Ho, 2020).
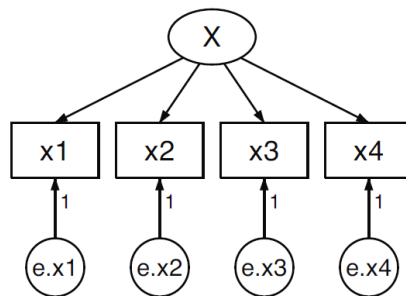
Figure 1: Example Path Diagram (Stata Structural Equation Modeling Reference Manual Release 16, 2019)

The equations examples are: $x1 = a1 + \beta1X + e.x1$, $x2 = a2 + \beta2X + e.x2$, $x3 = a3 + \beta3X + e.x3$, $x4 = $ and $a4 + \beta4X + e.x4$ (Stata Structural Equation Modeling Reference Manual Release 16, 2019).

3. FIT THE FACTORIAL STRUCTURE TO THE DATA | PRINCIPAL FACTORS METHOD

The UDO three-factor structure model is estimated using the principal factors method (Ho, 2020).

4. GOODNESS OF FIT INDEX AND MODIFICATION INDEX

Fit is assessed with the goodness of fit index (GFI), comparative fit index (CFI), and root mean square error of approximation (RMSEA) (Levine, 2007).

DATA COLLECTION REPORT

The sampling frame was 49 masters' students in the fall semester from the computer game development program at Southern Methodist University (SMU) in order to target the population of game developers. The participants were administered the Qualtrics UDO survey via a school address email with a link to the opt-in survey. The students in this non-probabilistic convenience (judgment) sample were from the United States and south-east Asia. This project used this original data set of N=49. The collected Qualtrics data was exported to a CSV file and analyzed using the statistical software package Stata16 (Ho, 2020).

UNIT OF OBSERVATION

Universal Diverse Orientation (UDO) is a social attitude construct used to measure multicultural awareness and acceptance of similarities and differences (Miville, et al., 1999). UDO includes three factors: diversity of contact, relativistic appreciation, and comfort with differences. Relativistic appreciation of oneself and others measures how much an individual understands another based on comparison or contrast (Miville, et al., 1999). Diversity of contact measures the level of diverse individuals the individual interacts or could interact with (Miville, et al., 1999). Comfort with differences measures how comfortable an individual is with cultural differences (Miville, et al., 1999).

Universal Diverse Orientation (UDO) was an ordinal variable treated like a continuous interval variable for statistical modeling. UDO was the dependent variable in other research conducted with additional survey data (demographics) not included in this project analysis.

QUALTRICS SURVEY METHODOLOGY REPORT

"A survey is a systematic method for gathering information from a sample of entities for the purposes of constructing quantitative descriptors of the attributes of the larger population of which the entities are members" (Groves et al., 2009). This project re-created from scratch the quantitative M-GUDS 15-item scale as a Qualtrics survey in order to collect data for ongoing

research. It asked 15 close-ended short questions using a bi-polar 7-point Likert-scale (see figure below) to produce numerical (scored 1-6, 0) descriptions about the Universal Diverse Orientation aspects of the population of the study.

Q18   I would like to go to concerts that feature music from other countries.

Clear	
1	Strongly disagree
2	Disagree
3	Somewhat disagree
4	Somewhat agree
5	Agree
6	Strongly agree
0	N/A

Figure 2: Likert-Scale Score Example and Modified Language in Q18

The survey methodology modified/updated some of the outdated language (see Q18 figure above) in a 15-item M-GUDS questionnaire and created a new web-based survey optimized for phone screen display in Qualtrics. Block 0 gathered browser meta-data and added a force-response question regarding age, 18 years or older, that used skip logic if under age. If appropriate age, then display logic showed consent request to voluntarily participate in study pre-set with the default "yes" choice and skip logic that took the user to end of survey if "no" was selected.

Q1 Browser Meta Info

Browser (1)
Version (2)
Operating System (3)
Screen Resolution (4)
Flash Version (5)
Java Support (6)
User Agent (7)

Q2 I am 18 years or older.

○ Yes (1)

○ No (2)

Skip To: End of Survey If I am 18 years or older. = No

Display This Question:

   If I am 18 years or older. = Yes

◉

Q3 We are conducting a research study to learn more about Universal Diverse Orientation. Your participation in this study is voluntary. If you agree to take part and then change your mind, you can withdraw for any reason. There are no penalties if you withdraw, decline to participate, or skip any parts of the study. If you agree to participate, you will fill out a Qualtrics survey with demographics and Universal Diverse Orientation data. Your participation should take about 5-10 minutes. We will keep your data safe by storing it in a Duo authenticated secure folder. This study could help build more diverse teams. Would you like to participate in this research study?

○ Yes (1)

○ No (2)

End of Block: Block 0

Block 2 contained the randomized five diversity of contact questions, all force response. Block 3

contained the five randomized relativistic appreciation questions, all force response. Block 4

contained the randomized five comfort with differences questions, all force response.

Start of Block: Block 2 UDO Diversity of Contact

Q17 I would like to join an organization that emphasizes getting to know people from different countries.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

Q18 I would like to go to concerts that feature music from other countries.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

Q19 I often listen to the music of other cultures.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

Q20 I am interested in learning about the many cultures that have existed in this world.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

Q21 I attend events where I might get to know people from different cultural backgrounds.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

End of Block: Block 2 UDO Diversity of Contact

Start of Block: Block 3 UDO Relativistic Appreciation

Q22 Persons with disabilities can teach me things I could not learn elsewhere.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

Q23 I can best understand someone after I get to know how they are both similar and different from me.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

Q24 Knowing how a person differs from me greatly enhances our friendship.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

Q25 In getting to know someone, I like knowing both how they differ from me and are similar to me.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

Q26 Knowing about the different experiences of other people helps me understand my own problems better.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

End of Block: Block 3 UDO Relativistic Appreciation

Start of Block: Block 4 UDO Comfort with Differences

Q27 Getting to know someone of another race is generally an uncomfortable experience for me.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

Q28 I am only at ease with people of my own culture.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

Q29 It's really hard for me to feel close to a person from another culture.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

Q30 It is very important that a friend agrees with me on most issues.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

Q31 I often feel irritated by persons of a different culture.

○ Strongly disagree (1)

○ Disagree (2)

○ Somewhat disagree (3)

○ Somewhat agree (4)

○ Agree (5)

○ Strongly agree (6)

○ N/A (7)

End of Block: Block 4 UDO Comfort with Differences

STATA DATA FORMATTING REPORT

I exported the Qualtrics survey data to a .csv file and opened it in MS Excel to format the data to import into Stata. I removed the 2nd and 3rd rows of header information and renamed each of the columns with the sub-scale question type in a sequential order: DiversityOfContact1, DiversityOfContact2, DiversityOfContact3, DiversityOfContact4, DiversityOfContact5, RelativisticAppreciation1, RelativisticAppreciation2, RelativisticAppreciation3, RelativisticAppreciation4, RelativisticAppreciation5, ComfortWithDifference1, ComfortWithDifferences2, ComfortWithDifferences3, ComfortWithDifferences4, ComfortWithDifferences5. I anonymized the participant names by assigning a number from 1-49 and re-labeled the column to "ID".

In MS Excel I looked for missing data and found the blank cells where the "not applicable" response had been selected and manually entered a "0" using the "general" number data type. However, when I imported this file into Stata, any column that had this "0" data entry was incorrectly determined to be string data rather than numeric data. I had to select each of these six columns and convert them from string data to numeric in Stata and saved the file as a native Stata data file (.dta).

SEM REPORT

I created a path diagram depicting the three factorial structures that underlie the 15 measurements of Universal Diverse Orientation. Once the data was properly formatted, I used the Stata SEM Builder GUI, enlarging the canvas size to 10x6 inches in order to fit 15 items vertically on the graph. I assigned an observed variable from the data to each of the 15 variable boxes, creating the endogenous variables in the path diagram. I named the three latent variables for the three UDO factors: "DiversityOfContact", "RelativisticAppreciation", and "ComfortWithDifferences", creating the exogenous variables in the path diagram. I mapped the single direction path of the corresponding five items from each one of the three factors:

DiversityOfContact1-5 to "DiversityOfContact", RealivisticAppreciation1-5 to

"RelativisticAppreciation" and ComfortWithDifferences1-5 to "ComfortWithDifferences". I then

mapped each of the covariances of the three factors to one another: "DiversityOfContact"

to/from "RelativisticAppreciation", "DiversityOfContact" to/from "ComfortWithDifferences", and

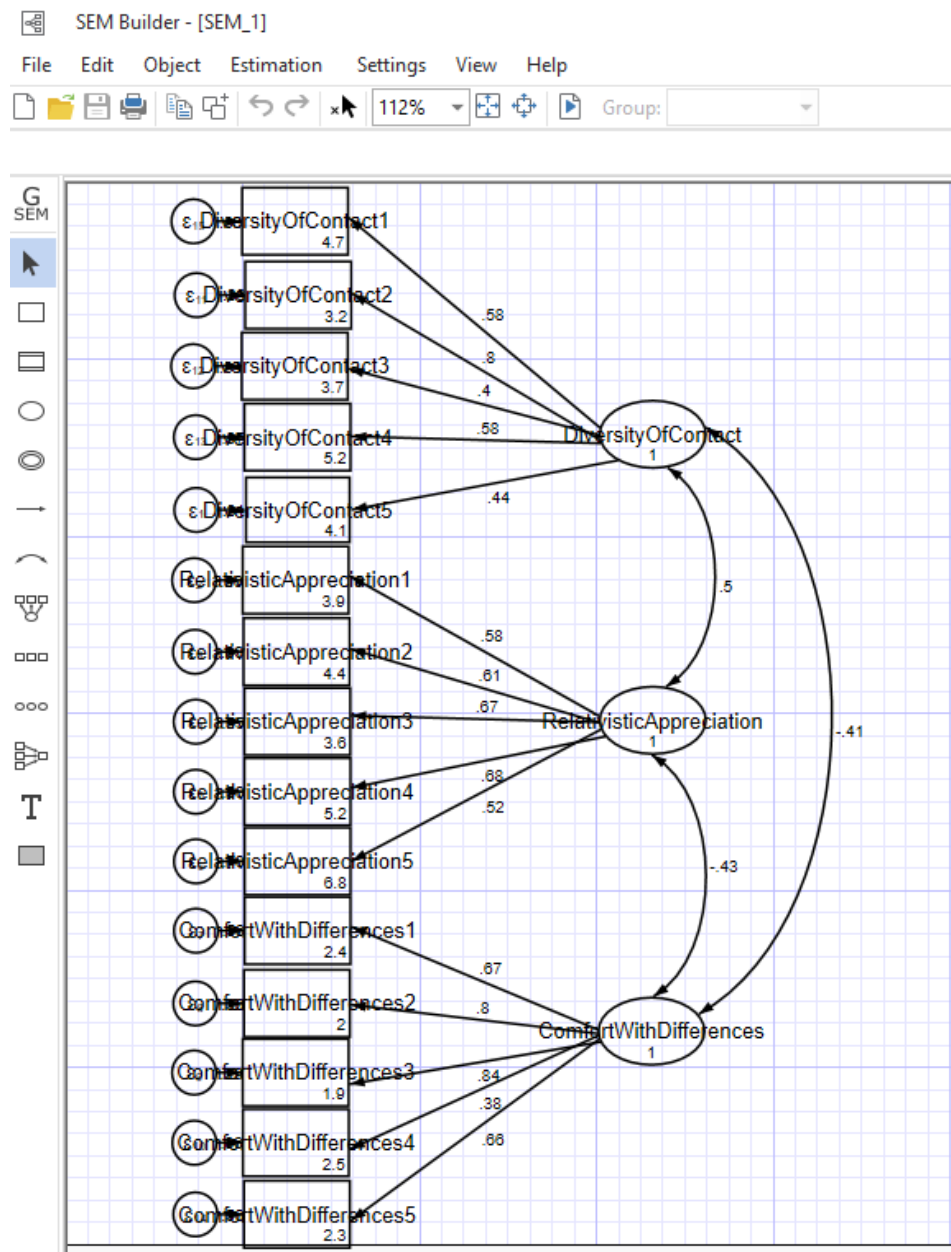"RelativisticAppreciation" to/from "ComfortWithDifferences".

Figure 3: UDO Path Diagram in Stata SEM Builder

ESTIMATION REPORT

After the path diagram was complete, I used the Stata SEM Builder GUI window to run the

estimation. In the SEM estimation options model tab, I selected "maximum likelihood" and in the

reporting tab checked "display standardized coefficients and values" with the 95% confidence

level selected in the drop down and the "report p-values", "test statistics", and "confidence

intervals" radio dials selected. In the .do file, the command was:

"sem (DiversityOfContact -> DiversityOfContact5,) (DiversityOfContact -> DiversityOfContact2,)
(DiversityOfContact -> DiversityOfContact3,) (DiversityOfContact -> DiversityOfContact4,)
(DiversityOfContact -> DiversityOfContact1,) (RelativisticAppreciation ->
RelativisticAppreciation1,) (RelativisticAppreciation -> RelativisticAppreciation2,)
(RelativisticAppreciation -> RelativisticAppreciation3,) (RelativisticAppreciation ->
RelativisticAppreciation4,) (RelativisticAppreciation -> RelativisticAppreciation5,)
(ComfortWithDifferences -> ComfortWithDifferences1,) (ComfortWithDifferences ->
ComfortWithDifferences2,) (ComfortWithDifferences -> ComfortWithDifferences3,)
(ComfortWithDifferences -> ComfortWithDifferences4,) (ComfortWithDifferences ->
ComfortWithDifferences5,), covstruct(_lexogenous, diagonal) standardized
latent(DiversityOfContact RelativisticAppreciation ComfortWithDifferences) cov(
DiversityOfContact*RelativisticAppreciation DiversityOfContact*ComfortWithDifferences
RelativisticAppreciation*ComfortWithDifferences) nocapslatent".

The estimation generated the results:

```
Fitting target model:

Iteration 0:   log likelihood =  -1025.151  (not concave)
Iteration 1:   log likelihood = -1013.5282
Iteration 2:   log likelihood = -1008.8685  (backed up)
Iteration 3:   log likelihood = -1005.0575
Iteration 4:   log likelihood = -1003.7495
Iteration 5:   log likelihood = -1003.6626
Iteration 6:   log likelihood = -1003.6584
Iteration 7:   log likelihood = -1003.6583

Structural equation model                      Number of obs    =         49
Estimation method  = ml
Log likelihood     = -1003.6583

 ( 1)   [DiversityOfContact5]DiversityOfContact = 1
 ( 2)   [RelativisticAppreciation1]RelativisticAppreciation = 1
 ( 3)   [ComfortWithDifferences1]ComfortWithDifferences = 1
```

var(e.DiversityOfContact5)	.8095701	.1239077			.5997564	1.092783
var(e.DiversityOfContact2)	.3601769	.1833609			.1327949	.9769006
var(e.DiversityOfContact3)	.8403964	.1144701			.6434908	1.097554
var(e.DiversityOfContact4)	.6687172	.1595177			.418981	1.06731
var(e.DiversityOfContact1)	.6644814	.140839			.4385993	1.006695
var(e.RelativisticAppreciation1)	.6643987	.1409998			.438314	1.007099
var(e.RelativisticAppreciation2)	.6327874	.1419503			.4076727	.9822092
var(e.RelativisticAppreciation3)	.5465803	.1513168			.317691	.9403791
var(e.RelativisticAppreciation4)	.5330794	.1495487			.3076083	.9238166
var(e.RelativisticAppreciation5)	.7268668	.1327322			.5081806	1.039661
var(e.ComfortWithDifferences1)	.5560694	.1288768			.3530622	.8758036
var(e.ComfortWithDifferences2)	.35759	.1203959			.1848404	.6917894
var(e.ComfortWithDifferences3)	.292212	.1154193			.1347369	.6337381
var(e.ComfortWithDifferences4)	.8578142	.1026213			.6785206	1.084485
var(e.ComfortWithDifferences5)	.5633475	.12673			.3624858	.8755112
var(DiversityOfContact)	1	.			.	.
var(RelativisticAppreciation)	1	.			.	.
var(ComfortWithDifferences)	1	.			.	.
cov(DiversityOfContact,RelativisticAppreciation)	.5029863	.1697357	2.96	0.003	.1703103	.8356622
cov(DiversityOfContact,ComfortWithDifferences)	-.4116922	.1688523	-2.44	0.015	-.7426367	-.0807478
cov(RelativisticAppreciation,ComfortWithDifferences)	-.4298937	.1556948	-2.76	0.006	-.7350499	-.1247376

LR test of model vs. saturated: chi2(87) =   138.41, Prob > chi2 = 0.0004

Standardized	OIM Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
Measurement						
DiversityOfContact5						
DiversityOfContact	.4363828	.1419714	3.07	0.002	.158124	.7146415
_cons	4.12946	.4409215	9.37	0.000	3.265269	4.99365
DiversityOfContact2						
DiversityOfContact	.7998895	.1146164	6.98	0.000	.5752454	1.024533
_cons	3.185226	.3520439	9.05	0.000	2.495233	3.87522
DiversityOfContact3						
DiversityOfContact	.3995042	.1432652	2.79	0.005	.1187095	.6802988
_cons	3.692916	.3994588	9.24	0.000	2.909991	4.47584
DiversityOfContact4						
DiversityOfContact	.5755717	.1385733	4.15	0.000	.3039731	.8471704
_cons	5.179477	.5423576	9.55	0.000	4.116475	6.242478
DiversityOfContact1						
DiversityOfContact	.5792397	.1215723	4.76	0.000	.3409624	.8175169
_cons	4.677983	.4936685	9.48	0.000	3.71041	5.645555
RelativisticAppreciation1						
RelativisticAppreciation	.5793111	.1216961	4.76	0.000	.3407911	.817831
_cons	3.903677	.4194104	9.31	0.000	3.081648	4.725706
RelativisticAppreciation2						
RelativisticAppreciation	.6059807	.1171245	5.17	0.000	.3764209	.8355404
_cons	4.362069	.4632147	9.42	0.000	3.454185	5.269953
RelativisticAppreciation3						
RelativisticAppreciation	.6733645	.1123588	5.99	0.000	.4531453	.8935836
_cons	3.569456	.3878383	9.20	0.000	2.809307	4.329605
RelativisticAppreciation4						
RelativisticAppreciation	.6833159	.1094287	6.24	0.000	.4688396	.8977922
_cons	5.238237	.5480869	9.56	0.000	4.164006	6.312467
RelativisticAppreciation5						
RelativisticAppreciation	.5226215	.1269869	4.12	0.000	.2737318	.7715112
_cons	6.841542	.7057107	9.69	0.000	5.458374	8.224709
ComfortWithDifferences1						
ComfortWithDifferences	.6662812	.0967135	6.89	0.000	.4767262	.8558362
_cons	2.435867	.2845234	8.56	0.000	1.878211	2.993522
ComfortWithDifferences2						
ComfortWithDifferences	.8015048	.0751062	10.67	0.000	.6542994	.9487102
_cons	2.005543	.2478932	8.09	0.000	1.519681	2.491404
ComfortWithDifferences3						
ComfortWithDifferences	.8413013	.0685957	12.26	0.000	.7068562	.9757465
_cons	1.880988	.2377213	7.91	0.000	1.415063	2.346914
ComfortWithDifferences4						
ComfortWithDifferences	.3770753	.1360753	2.77	0.006	.1103725	.643778
_cons	2.499718	.2901194	8.62	0.000	1.931094	3.068341
ComfortWithDifferences5						
ComfortWithDifferences	.6607969	.0958918	6.89	0.000	.4728524	.8487414
_cons	2.320698	.2745244	8.45	0.000	1.78264	2.858756

GOODNESS OF FIT REPORT

After Stata fit the factorial structure to my data, I checked for goodness-of-fit with the "estat gof,

stats(all)" command and generated the results:

```
─────────────────────────────────────────────────────────────
Fit statistic          │     Value   Description
─────────────────────────────────────────────────────────────
Likelihood ratio       │
        chi2_ms(87)    │   138.409   model vs. saturated
           p > chi2    │     0.000
        chi2_bs(105)   │   322.037   baseline vs. saturated
           p > chi2    │     0.000
─────────────────────────────────────────────────────────────
Population error       │
             RMSEA     │     0.111   Root mean squared error of approximation
 90% CI, lower bound   │     0.075
        upper bound    │     0.145
           pclose      │     0.006   Probability RMSEA <= 0.05
─────────────────────────────────────────────────────────────
Information criteria   │
               AIC     │  2103.317   Akaike's information criterion
               BIC     │  2194.124   Bayesian information criterion
─────────────────────────────────────────────────────────────
Baseline comparison    │
               CFI     │     0.763   Comparative fit index
               TLI     │     0.714   Tucker-Lewis index
─────────────────────────────────────────────────────────────
Size of residuals      │
              SRMR     │     0.108   Standardized root mean squared residual
                CD     │     0.988   Coefficient of determination
─────────────────────────────────────────────────────────────
```

DISCUSSION

The goodness-of-fit p-value is significant. The RSMA at 0.111 was too far above the acceptable

threshold of .08. The CFI is acceptable at 0.763, but low for this set of sample data. Iteration was

part of the framework's process and thus a version two of the survey was created. In the

subsequent survey design, I removed the "not applicable" option, thus eliminating the 0 score

for an item. I administered the survey to a new sample of masters' students from the computer

game development program at Southern Methodist University (SMU) in the fall of 2020. Once the

results are in (May 2021), I will run the CFA again to compare against the first sample. I can then

determine if the second iteration of the survey is a better overall fit to the model than the first

iteration.

BIBLIOGRAPHY

R. M. Groves, et al, *Survey methodology*, 2nd Edition. Hoboken, New Jersey, USA: Wiley-Interscience, 2009.

K. Ho, "Confirmatory factor analysis in stata," lecture presentation notes for University of Texas at Dallas, Richardson, TX, USA, September 2, 2020.

T. Levine, "Confirmatory factor analysis and scale validation in communication research", *Communication Research Reports*, vol. 22, no. 4, pp. 335–338, 2005, 2007, https://doi.org/10.1080/00036810500317730.

M. L. Miville, C. J. Gelso, P. Touradji, R. Pannu, W. Liu, P. Holloway and J. Fuertes, "Appreciating Similarities and Valuing Differences: The Miville-Guzman Universality-Diversity Scale," *Journal of Counseling Psychology,* vol. 46, no. 3, pp. 291-307, 1999.

Stata structural equation modeling reference manual release 16. Stata Press, College Station, TX, USA, 2019.