Towards Measuring Relational Embeddedness: 2 Factor Analyses of TRENDS Pilot Survey Data

Factor Analysis Conducted by Tim Walker, PhD student in Educational Inquiry, Measurement & Evaluation
In partial fulfillment of the course requirements of Sociology 706r, taught by Dr. Joseph Olson
Social networks studies data collected by Dr. Julie Hite, Department of Educational Leadership & Foundations, McKay School of Education, BYU

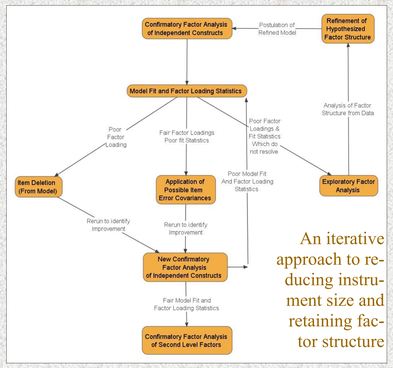
Study 1: Background

The two previous TRENDS pilot survey validation studies had utilized the BYU teaching faculty as the survey population. Each professor who agreed to participate filled out the survey form regarding one of their work-related relationships.

The data utilized in this study was similarly generated by surveying a sample of the BYU faculty. Each professor who agreed to participate was asked to select a single individual and answer 45 items dealing with their work-related relationship with that individual. The participants were asked to choose a person with whom they have interacted, but who is not a member of their own college.

Among the 45 items in the TRENDS III were several items which had not been piloted as part of the TRENDS I and II pilots. This was due to the poor performance of some of the existing items in the TRENDS II analyses, which included CFA of TRENDS items for the first time.

The data were collected on paper copies of the survey. Undergraduate research assistants (URA's) contacted the sampled faculty and made arrangements to invite them to participate. The URA's then arranged for the retrieval of the surveys.



An iterative approach to reducing instrument size and retaining factor structure

What is …?

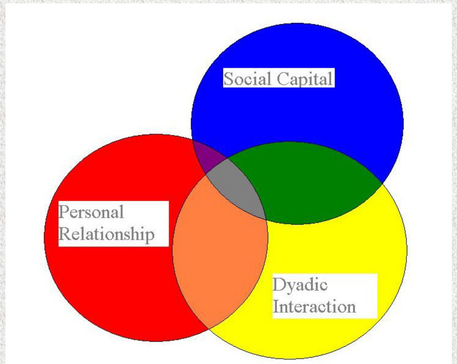
Relational Embeddedness. A theoretical construct that attempts to describe reasons why persons maintain certain interpersonal relationships.

The specific theory of interest was formulated by Hite (2001). Simply stated relational embeddedness is a function of the level to which an individual's relationship involves more or less of three components:

Dyadic Interaction: The extent and quality of interpersonal interaction.

Personal Relationship: Amounts of the emotional connections in the relationship.

Social Capital: The level of mutual and communal reciprocity affecting the relationship



What is…?

Factor Analysis. A statistical method which uses analysis of the ways in which survey or test items tend to be answered in the same ways to empirically estimate the degree to which items are related to one another and to latent constructs.

TRENDS. A survey designed to measure the levels of three theoretical constructs present in relationships. Ultimately the survey is designed to be used in network studies.

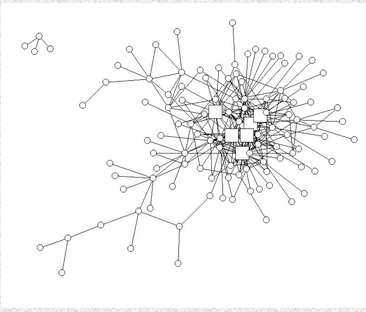
Network Studies: A genre of research which analyzes relationship patterns. To use a survey such as TRENDS in a network setting will require methods of assessing and controlling for non-independence as participants will be asked to complete surveys for as many relationships as appropriate given their network standing and the purpose of the instrument.

Study 2: Background

The data analyzed for this study was generated as part of a larger survey conducted with school head teachers in Uganda. The head teachers were asked to answer questions about relationships with other head teachers which provided them with resources beneficial to the accomplishment of their work. This network of school administrators was defined geographically by district (a Ugandan political division, not equivalent to a US school district). However, the respondents were not limited in choosing the relationships they rated to only their relationships with other head teachers in the same district.

This type of study design may be helpfully pictured with a network diagram or map like the one below in which individuals are displayed as circles and the relationships between them are line segments.

As part of this study a number of items which had been included in the TRENDS II piloting were asked regarding each relationship. Many of these items were eliminated from the final TRENDS II factor models due to factor loadings which did not correspond to the theory-based latent constructs. The first step in this analysis was to conduct exploratory factor analyses of these items to determine an appropriate factor model which could be tested in CFA using the M+ program.



Study 1: Hypotheses

Two essential hypotheses were tested in this study. First, the hypothesized factor structure of the 45 items identified in the TRENDS II piloting were analyzed to determine if they still represented a model for the survey which was statistically significant and returned good model fit statistics.

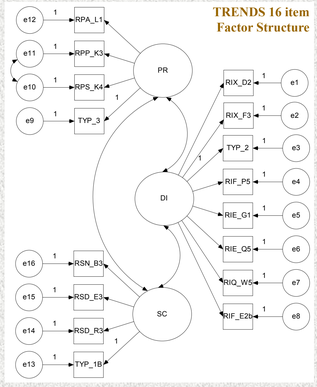
The second hypothesis was that through an iterative process like that outlined above a significant, well-fitting model could be identified with a reduced number of items.

The second hypothesis was crucial as it represents a vital step in the evolution of the TRENDS from the piloted single-response studies to the desired network study.

Study 1: Discussion

The hypothesized TRENDS III factor structure, based upon TRENDS II analysis and Hite's relational embeddedness theory did not exhibit good model fit. This was largely due to the performance of the newly piloted items.

By eliminating these items from models used within the iterative process illustrated above to identify a shortened survey instrument, and by simplifying the factor structure to three factors, good model fit was obtained for the shortened model. In short, while only batting .500 we still managed to hit "at least a double, and perhaps a homerun." (Olsen 2010)



Study 1: Results

Model Fit Statistics

Model	CFI	TLI	RMSEA
Full TRENDS III	0.87	0.93	0.149
16-Item Short	0.96	0.98	0.083

Standardized Factor Loadings

DIEASE by		PERREL by	
RIEG1	0.78	RPAC1	0.656
RIEQ5	0.493	RPAL1	0.833
RIET5	0.715	RPAR4	0.883
DIQUAL by		RPPI5	0.786
RIQF4	0.629	RPPK3	0.399
RIQL2	0.713	RPSK4	0.766
RIQS2	0.933	RPSS2	0.669
RIQV1	0.737	TYP3	0.916
RIQV2	0.853	SOCCAP by	
RIQV3	0.827	RSDE3	0.55
RIQW5	0.927	RSDR3	0.623
DIEFFORT by		RSNB3	0.739
RIFE2	0.482	RSNS4	0.628
RIFE2B	0.596	TYP1B	0.192
RIFE4A2	0.872	DIFACT by	
RIFR3	0.462	DIEASE	1.081
RIFW2	0.518	DIQUAL	0.997
RIFP5	0.639	DIEFFORT	0.881
DIEXTENT by		DIEXTENT	0.762
RIXA5	1.006	TRENDS by	
RIXD2	0.84	DIFACT	0.912
RIXF1	0.714	PERREL	0.96
RIXF2	0.721	SOCCAP	0.914
RIXF3	0.843		
RIXM4	0.889		

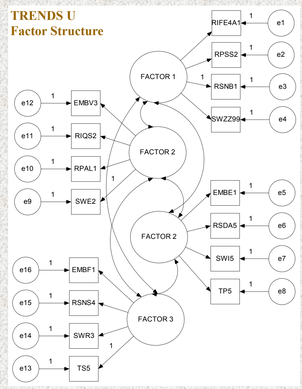
Standardized Factor Loadings & Covariances

DIFACT by	
RIEQ5	0.817
RIEG1	0.48
RIQW5	0.802
RIFE2B	0.715
RIFP5	0.627
RIXD2	0.675
RIXF3	0.641
TYP2	0.861
PERREL by	
RPAC1	0.946
RPPK3	0.837
RPSK4	0.758
TYP3	0.887
SOCCAP by	
RSDE3	0.69
RSDR3	0.705
RSNB3	0.622
TYP1B	0.726
DIFACT with	
PERREL	0.835
SOCCAP	0.836
SOCCAP with	
PERREL	0.722

Theoretical pedigree and validation process of TRENDS

Future Studies

1. The sixteen item survey should be utilized in a network setting. This would enable analysis of the appropriateness of these specific items and factors in this type of study setting. Additionally, it might allow work on the following questions:
   A. When respondents answer multiple surveys what are appropriate ways of measuring respondent influence on the subsequent relationship scores.
   B. When individuals are the subject of multiple surveys what are appropriate ways of measuring subject influence on the subsequent relationship scores.
   C. At what level of clustering in respondent and subject is statistical control necessary?



Study 2: Results & Discussion

Values closer to 1.0 are desired for CFI and TLI measures of model fit. RMSEA ideal values are closer to zero.

Model Fit Statistics

Model	CFI	TLI	RMSEA
Non-Clustered	0.90	0.94	0.12
Subject Clustered	0.96	0.97	0.06
Object Clustered	0.92	0.94	0.12

The respondent-clustered CFA returns the best model fit statistics. This is in keeping with the hypotheses.

As TRENDS moves towards use in the intended network settings clustered CFA analyses will be crucial.

Standardized Factor Loadings

FAC1 by	Model 1	Model 2	Model 3
RIFE4A1	0.647	0.649	0.642
RPSS2	0.727	0.72	0.722
RSNB1	0.815	0.817	0.82
SWZZ99	0.808	0.81	0.83
FAC2 by			
EMBV3	0.679	0.68	0.676
RIQS2	0.851	0.835	0.857
RPAL1	0.73	0.739	0.729
SWE2	0.652	0.657	0.655
FAC3 by			
EMBE1	0.54	0.538	0.55
RSDA5	0.627	0.626	0.631
SWI5	0.818	0.816	0.822
TP5	0.859	0.856	0.864
FAC4 by			
EMBF1	0.806	0.818	0.801
RSNS4	0.564	0.558	0.562
SWR3	0.614	0.61	0.606
TS5	0.363	0.359	0.377

Standardized Factor Covariances

FAC1 with			
FAC2	0.672	0.691	0.67
FAC3	0.796	0.793	0.775
FAC4	0.813	0.796	0.83
FAC2 with			
FAC3	0.655	0.659	0.651
FAC4	0.793	0.794	0.788
FAC3 with			
FAC4	0.87	0.867	0.886

Study 2: Hypotheses

Having established an empirical factor model, the next step was to test the model in a confirmatory analysis to determine its model fit and factor loading characteristics.

Steps 2a and b were to conduct two additional CFA's in which the identity of the survey respondent and the identity of the survey's "target" (subject) was used to cluster the data in order to control for the effects of the same person filling out multiple questionnaires or being the "target" of multiple filled out questionnaires.

Specifically, the hypotheses were that:
The identified factor structure would be statistically significant and exhibit fair model fit, factor loading and covariance statistics.

The effect of survey respondent would be significant, leading to improvements in model fit over the 1st model due to the clustering of the survey respondents.

The effect of survey subject would not make significant improvement in model fit due to the diffuse nature of the network subjects as illustrated in the above network diagram.

Works Cited

Granovetter M. 1983. The strength of weak ties: A network theory revisited. Sociological Theory.1: 201-233.
Hite J. 2003. Patterns of multidimensionality among embedded network ties: A typology of relational embeddedness in emerging entrepreneurial firms. Strategic Organization. 1:9-49.
Muthen & Muthen. 2008. M+ 5.2 (Statistical Analysis Software)
Olsen, J. 2010. Personal Communication.