Introduction to Factor Analysis

Hsueh-Sheng Wu
CFDR Workshop Series
June 18, 2018

BGSU
Center for Family and Demographic Research

Outline

- Why do sociologists need factor analysis?
- What is factor analysis?
- Sternberg's triangular love theory
- Some data questions
- Types of factor analysis
- Steps of conducting exploratory factor analysis
- Steps of conducting confirmatory factor analysis
- Extension of confirmatory factor analysis
- Conclusion

Why Do Sociologists Need Factor Analysis?

- Most social phenomena of interest are multi-dimensional constructs and cannot be measured by a single question, for example:
  - Well-being
  - Violence
- When a single question is used, the information may not be reliable because people may have different responses to a particular word or idea in the question.
- The variation of one question may not be enough to differentiate individuals.
- Factor analysis is a data reduction tool that helps decide whether and how the information of these questions should be combined to measure a construct.

BGSU Center for Family and Demographic Research

What Is Factor Analysis?

- Factor analysis is a statistical method that identifies a latent factor or factors that underlie observed variables.
- Specifically, factor analysis addresses the following questions:
  - How many latent factors underlie observed variables?
  - How are these latent factors related to observed variables?.
  - What do these factors mean?
- Factor analysis helps answer the question of how accurate the sum of variables measure the latent factor or factors.

Sternberg's Triangular Love Theory

- Sternberg's triangular love scale (45 items) measures three components of the love: *passion, intimacy*, *commitment.*

Table 1. Sternberg's Triangular Love Theory

Type of Love	Passion	Intimacy	Commitment
Infatuation	x		
Liking		x	
Empty Love			x
Romantic Love	x	x	
Companionate Love		x	x
Fatuous Love	x		x
Consummate Love	x	x	x

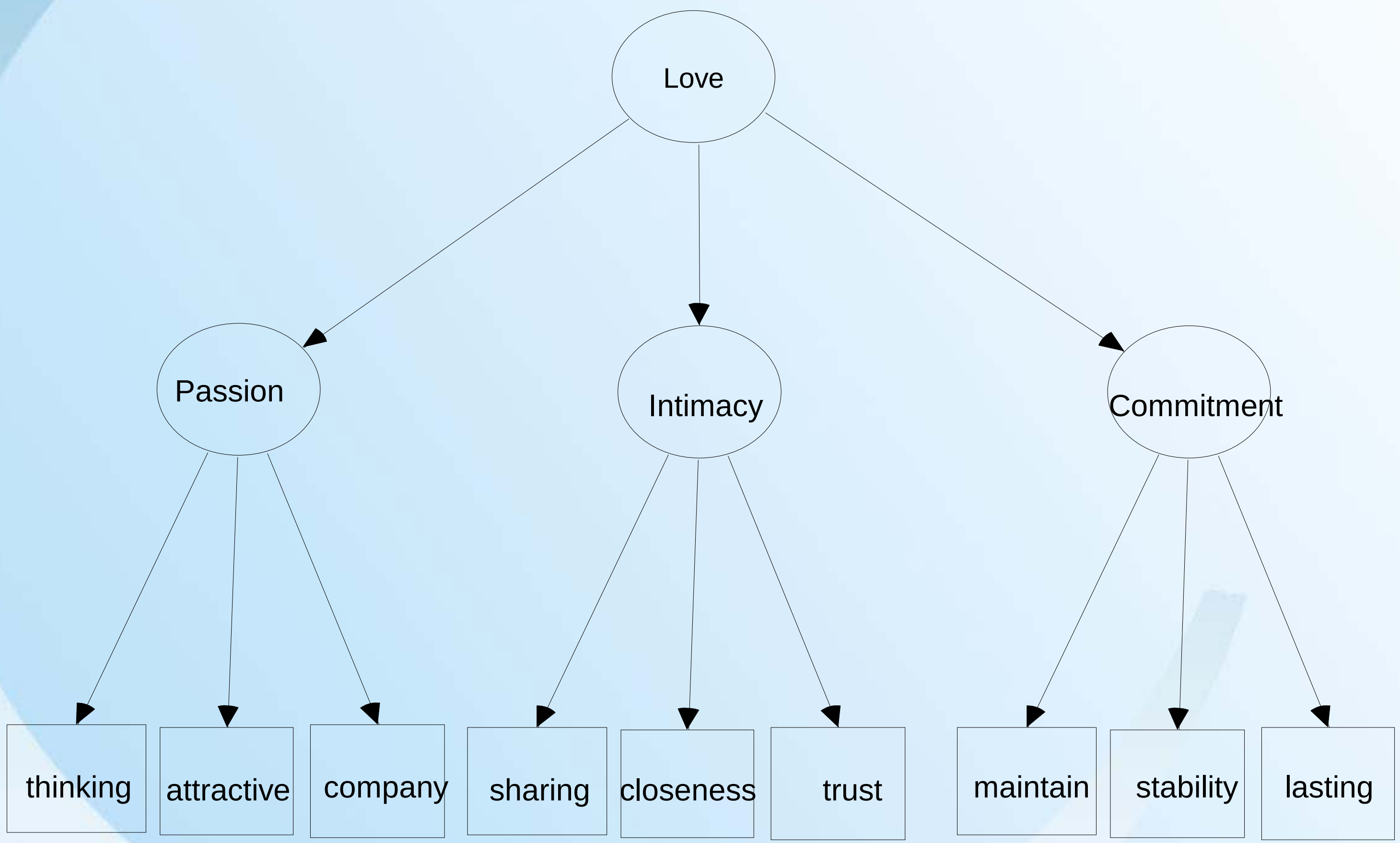
BGSU Center for Family and Demographic Research

Sternberg’s Triangular Love Theory

Table 2. Description of 9 Variables Measuring Love

Dimension	#	Variable	Description
Passion	1	thinking	I find myself thinking about "X" frequently during the day.
	2	attractive	I find "X" to be very personally attractive.
	3	company	I would rather be with "X" than with anyone else.
Intimacy	1	sharing	I am willing to share myself and my possessions with "X."
	2	closeness	I feel close to "X."
	3	trust	I feel that I can really trust "X."
Commitment	1	maintain	I am committed to maintaining my relationship with "X."
	2	stability	I have confidence in the stability of my relationship with "X."
	3	lasting	I expect my love for "X" to last for the rest of my life.

Conceptualization and Operationalization of Sternberg's Triangular Love Theory



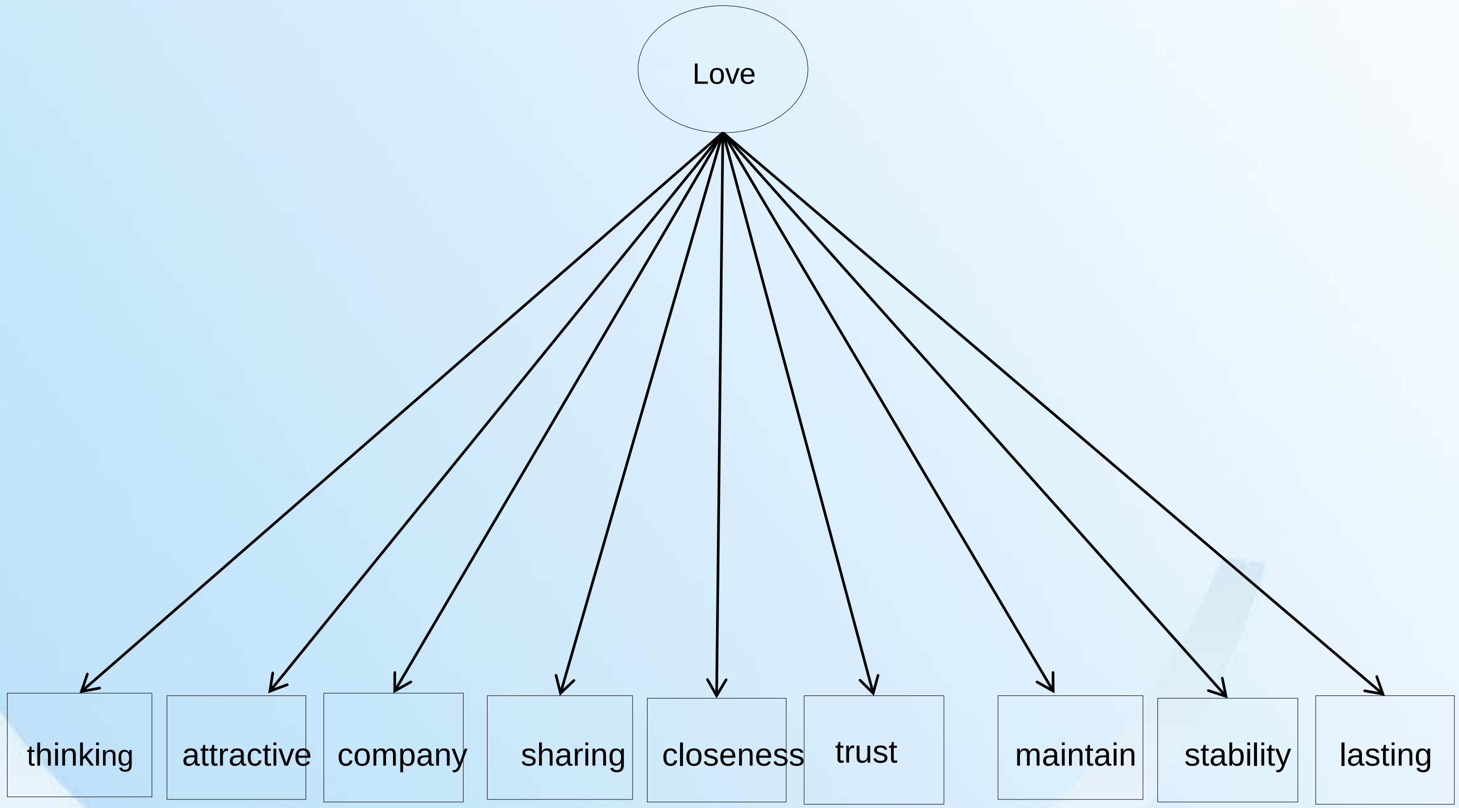
Some Data Questions

- We know that these 9 items should be used to measure love. However, we still need to answer several data construction questions:
  (1) Do these 9 items also measure some constructs other than love?
  (2) Should each of these 9 items be weighted equally in measuring love?

- Factor analysis helps address these two questions.

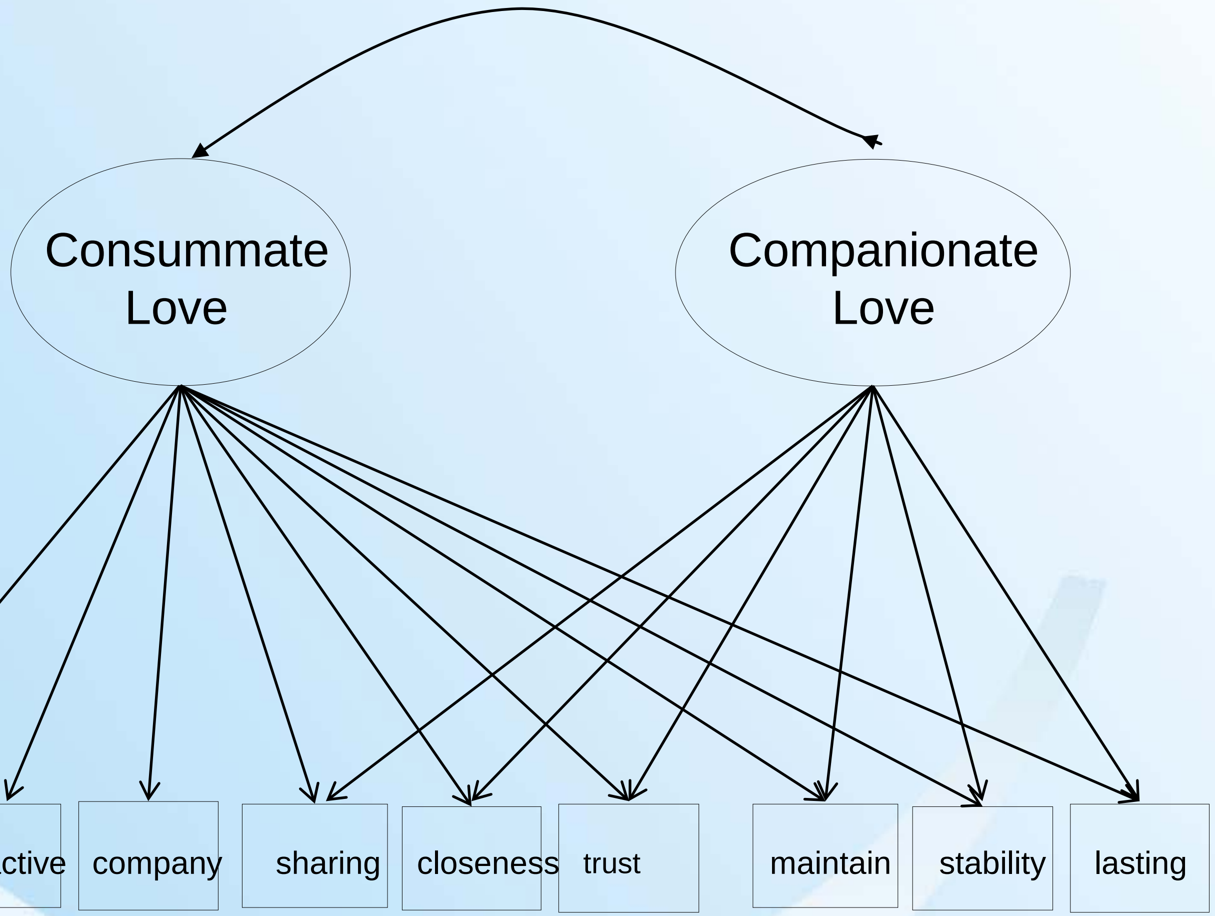
Two Hypothesized Factorial Models of the Love Measures

Model 1:



Two Hypothesized Factorial Models of the Love Measures

Model 2:



BGSU

Center for Family and Demographic Research

Types of Factor Analysis

- Exploratory Factor Analysis (EFA)
  - EFA examines (1) how many factors a measure estimates and (2) what these factors are.
  - EFA is used when an old phenomenon is re-conceptualized or a new phenomenon emerges .
  - SAS, SPSS, Stata, AMOS, LISREL, and Mplus all can conduct EFA.
- Confirmatory Factor Analysis (CFA)
  - CFA examines whether the number of latent factors, factor loadings, factor correlations, and factor means are the same for different populations or for the same people at different time points.
  - CFA is used when the factorial structure of the measures has been established.
  - SAS, Stata, AMOS, LISREL, and Mplus all can conduct CFA .
  - CFA is a more powerful technique than EFA because CFA allows researchers to test their hypotheses about the factorial structure of observed variables.

Steps of Conducting Exploratory Factor Analysis

Step 1. Examine the data

sum

Variable	Obs	Mean	Std. Dev.	Min	Max
female	596	1.538591	.4989273	1	2
thinking	597	4.433836	.7311806	1	5
attractive	598	4.513378	.7198932	1	5
company	598	4.413043	.7732746	1	5
sharing	598	3.884615	1.036227	1	5
closeness	596	4.055369	.9589672	1	5
trust	592	3.592905	1.134664	1	5
maintain	597	3.730318	.9827209	1	5
stability	599	3.766277	.9653519	1	5
lasting	596	3.619128	.9024169	1	5

Steps of Conducting Exploratory Factor Analysis

Step 2. Identify the number of factors underlying these items

factor thinking-lasting, pf

```
(obs=583)

Factor analysis/correlation                        Number of obs    =      583
    Method: principal factors                      Retained factors =        3
    Rotation: (unrotated)                          Number of params =       24
```

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	4.44340	3.83145	0.9479	0.9479
Factor2	0.61195	0.38943	0.1306	1.0785
Factor3	0.22251	0.23062	0.0475	1.1260
Factor4	-0.00811	0.03362	-0.0017	1.1242
Factor5	-0.04173	0.05915	-0.0089	1.1153
Factor6	-0.10088	0.03147	-0.0215	1.0938
Factor7	-0.13235	0.01202	-0.0282	1.0656
Factor8	-0.14437	0.01865	-0.0308	1.0348
Factor9	-0.16302	.	-0.0348	1.0000

LR test: independent vs. saturated: chi2(36) = 2677.62 Prob>chi2 = 0.0000

BGSU Center for Family and Demographic Research

Factor loadings (pattern matrix) and unique variances

Variable	Factor1	Factor2	Factor3	Uniqueness
thinking	0.7018	-0.3697	0.0181	0.3704
attractive	0.6904	-0.3979	0.0709	0.3600
company	0.7119	-0.2409	0.1349	0.4169
sharing	0.7233	0.2408	0.1596	0.3934
closeness	0.6155	0.3281	0.1769	0.4822
trust	0.6147	0.2685	-0.0147	0.5498
maintain	0.7285	0.1202	0.0329	0.4538
stability	0.8175	0.0284	-0.2358	0.2753
lasting	0.6990	0.0767	-0.2921	0.4202

BGSU
Center for Family and Demographic Research

Steps of Conducting Exploratory Factor Analysis

Step 3. Understand what these factor means

rotate, varimax horst blanks(.3)

```
. rotate, varimax horst blanks(.3)

Factor analysis/correlation                        Number of obs    =      583
    Method: principal factors                      Retained factors =        3
    Rotation: orthogonal varimax (Kaiser on)       Number of params =       24
```

Factor	Variance	Difference	Proportion	Cumulative
Factor1	2.09940	0.02794	0.4479	0.4479
Factor2	2.07146	0.96446	0.4419	0.8898
Factor3	1.10700	.	0.2362	1.1260

```
LR test: independent vs. saturated:  chi2(36) = 2677.62 Prob>chi2 = 0.0000
```

BGSU

Center for Family and Demographic Research

Rotated factor loadings (pattern matrix) and unique variances

Variable	Factor1	Factor2	Factor3	Uniqueness
thinking	0.7201			0.3704
attractive	0.7457			0.3600
company	0.6563	0.3479		0.4169
sharing	0.3093	0.6807		0.3934
closeness		0.6769		0.4822
trust		0.5577	0.3258	0.5498
maintain	0.3744	0.5522	0.3179	0.4538
stability	0.4389	0.4365	0.5843	0.2753
lasting	0.3161	0.3699	0.5857	0.4202

(blanks represent abs(loading)<.3)

Factor rotation matrix

	Factor1	Factor2	Factor3
Factor1	0.6271	0.6313	0.4563
Factor2	-0.7465	0.6544	0.1205
Factor3	0.2225	0.4162	-0.8816

estat common

Correlation matrix of the varimax rotated common factors

Factors	Factor1	Factor2	Factor3
Factor1	1		
Factor2	0	1	
Factor3	0	0	1

Steps of Conducting Exploratory Factor Analysis

rotate, promax horst blanks(.3)

```
Factor analysis/correlation                        Number of obs    =      583
    Method: principal factors                      Retained factors =        3
    Rotation: oblique promax (Kaiser on)           Number of params =       24

    --------------------------------------------------------------------------
         Factor  |     Variance   Proportion    Rotated factors are correlated
    -------------+------------------------------------------------------------
        Factor1  |      3.52815       0.7527
        Factor2  |      3.51566       0.7500
        Factor3  |      3.38960       0.7231
    --------------------------------------------------------------------------
    LR test: independent vs. saturated:  chi2(36) = 2677.62 Prob>chi2 = 0.0000
```

estat common

BGSU

Center for Family and Demographic Research

Steps of Conducting Exploratory Factor Analysis

Rotated factor loadings (pattern matrix) and unique variances

Variable	Factor1	Factor2	Factor3	Uniqueness
thinking		0.7482		0.3704
attractive		0.8058		0.3600
company		0.6657		0.4169
sharing	0.6909			0.3934
closeness	0.7465			0.4822
trust	0.5124			0.5498
maintain	0.4585			0.4538
stability			0.5932	0.2753
lasting			0.6515	0.4202

(blanks represent abs(loading)<.3)

Factor rotation matrix

	Factor1	Factor2	Factor3
Factor1	0.4268	-0.4787	0.0839
Factor2	0.8754	0.8712	0.8658
Factor3	0.2269	0.1086	-0.4932

BGSU Center for Family and Demographic Research

Steps of Conducting Exploratory Factor Analysis

estat common

Correlation matrix of the promax(3) rotated common factors

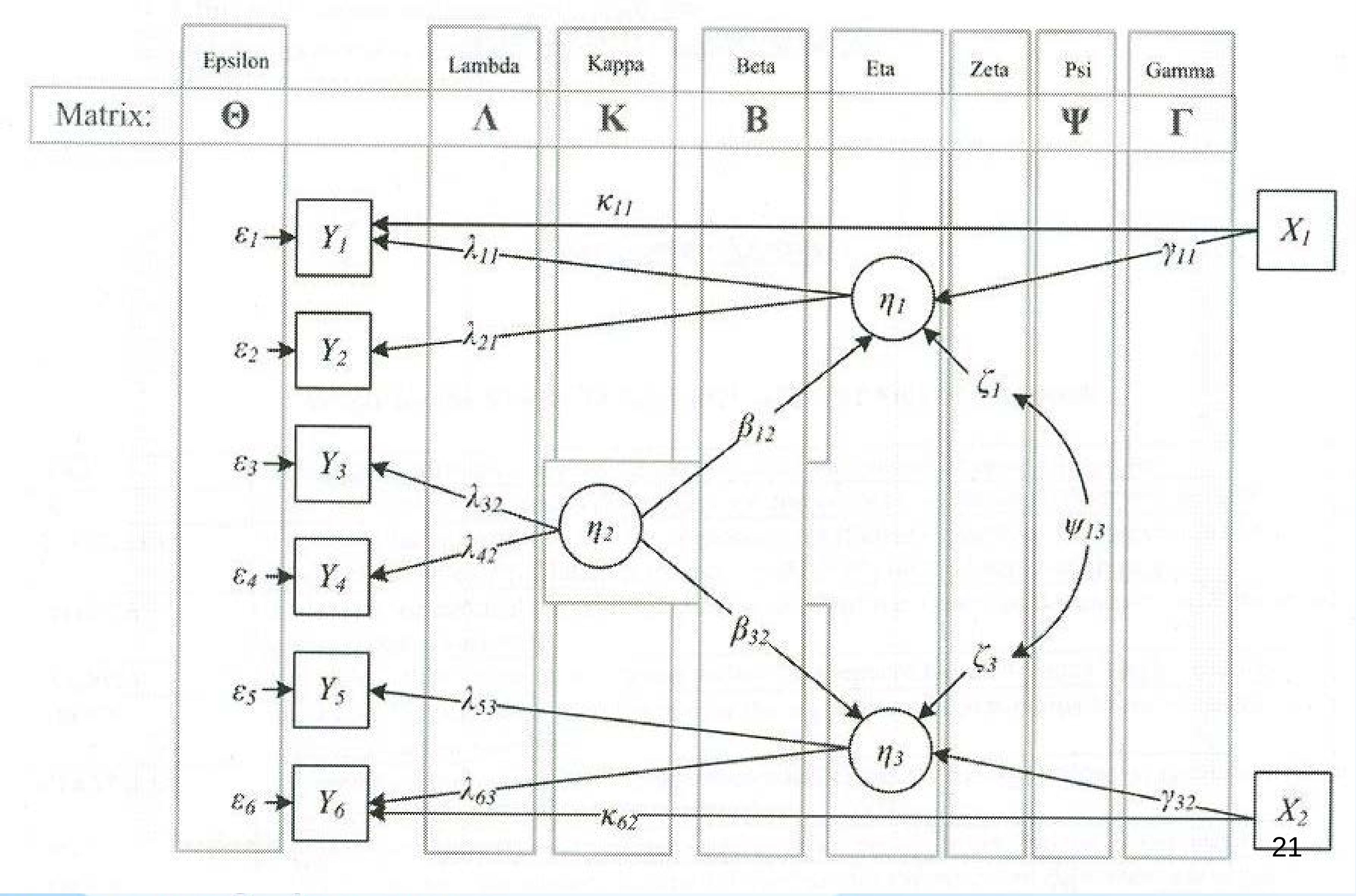
Factors	Factor1	Factor2	Factor3
Factor1	1		
Factor2	.583	1	
Factor3	.6819	.6606	1

Confirmatory Factor Analysis

- Confirmatory Factor Analysis (CFA) is more powerful than Exploratory Factor Analysis (EFA).
- CFA can check the validity and reliabilty of the measures.
- CFA examines whether the underlying factorial structures are the same across different populations or across different time points.
- Any SEM software can estimate CFA.

BGSU
Center for Family and Demographic Research

Graphs and Parameters in SEM



Jargon of SEM

- Variables in SEM
  - Measured variable
  - Latent variable
  - Exogenous variable
  - Error
  - Disturbance
- 8 sets of parameters

BGSU
Center for Family and Demographic Research

Steps of Conducting of Confirmatory Factor Analysis

Step 1: Create a path diagram depicting the factorial structure underlie the measures

Step 2: Fit the factorial structure to the data

Step 3: Examine the goodness of fit index and modification index

Step 4: Consider what types of changes can be made to fit the data better

Repeat the steps 2 through 4.

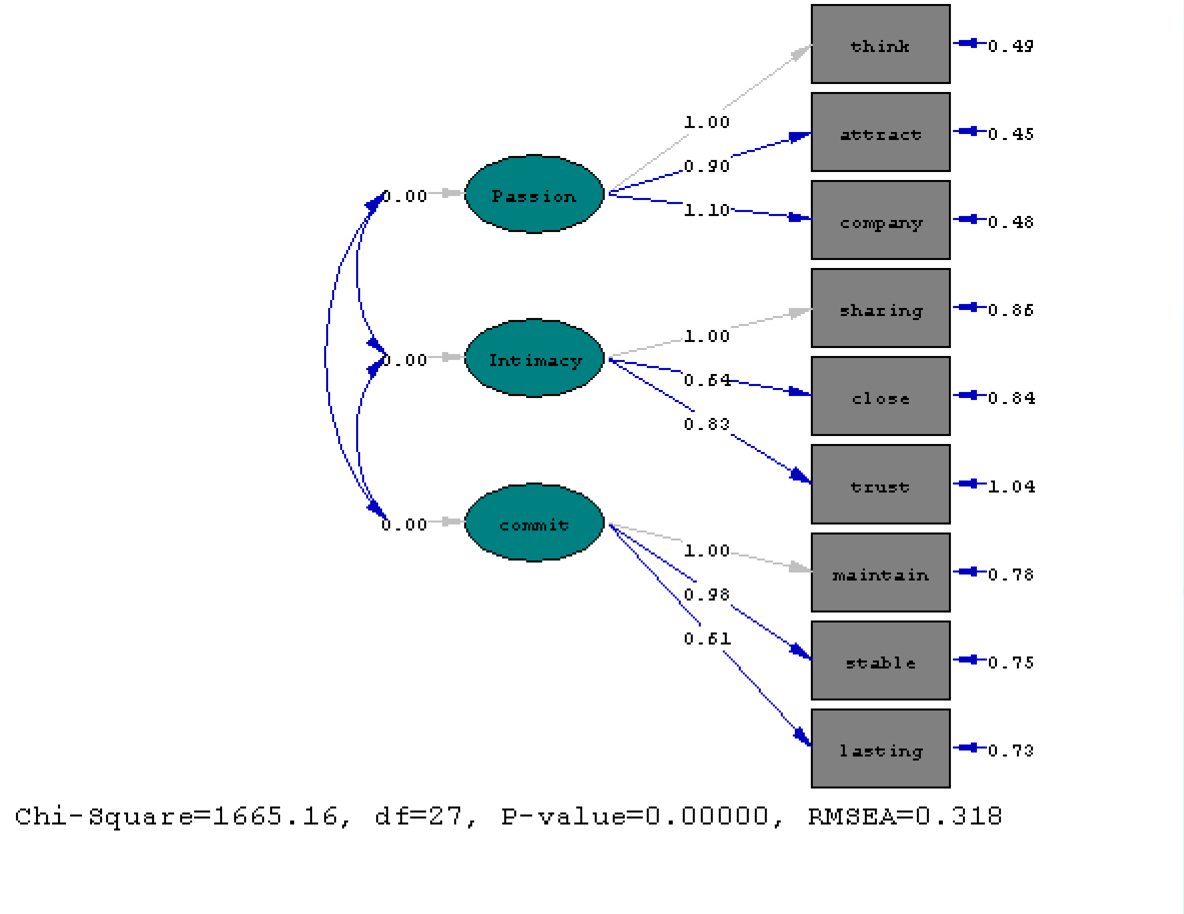
Center for Family and Demographic Research

Confirmatory Factor Analysis

Model 1: A Factorial Model Based on Sternberg's Theory

```
DA NI=10 NO=600 MI=9
RA FI='D:\factor\factor.txt'
LA
think attract company sharing close trust maintain stable lasting
SE 1 2 3 4 5 6 7 8 9/
MO NY=9 NE=3 LY=FU,FI PS=SY,FR TE=DI,FR
FR LY 1 1 LY 2 1 LY 3 1
FR LY 4 2 LY 5 2 LY 6 2
FR LY 7 3 LY 8 3 LY 9 3
FI PS 1 1 PS 2 2 PS 3 3
LE
Passion Intimacy commit
PD
OU SE TV RS ND=3 MI
```

Model 1: A Factorial Model Based on Sternberg's Theory



BGSU Center for Family and Demographic Research

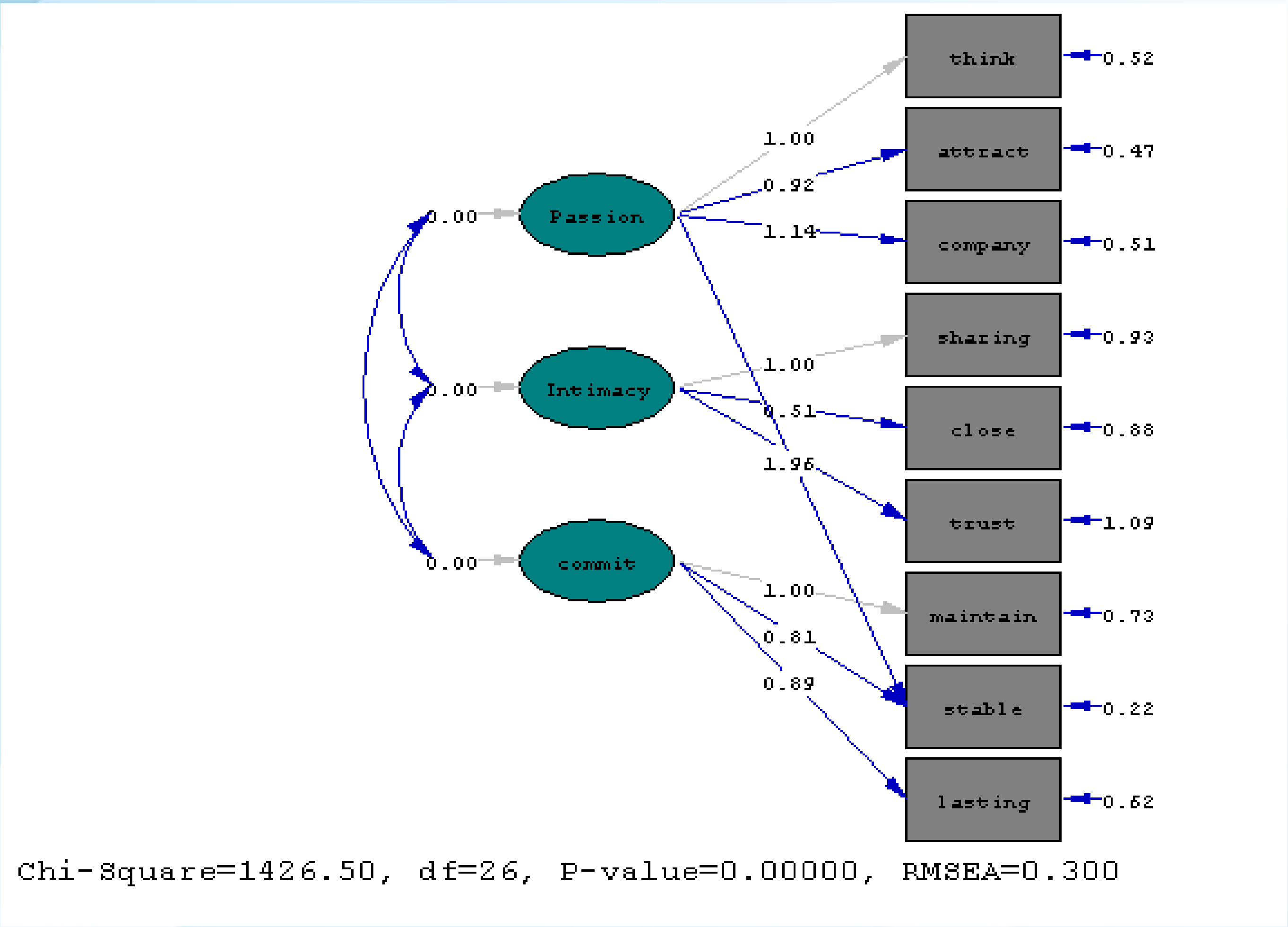
Confirmatory Factor Analysis

Model 2: Allow Stability to Be Loaded onto the Passion Factor

```
DA NI=10 NO=600 MI=9
RA FI='D:\factor\factor.txt'
LA
think attract company sharing close trust maintain stable lasting
SE 1 2 3 4 5 6 7 8 9/
MO NY=9 NE=3 LY=FU,FI PS=SY,FR TE=DI,FI
FR LY 1 1 LY 2 1 LY 3 1 LY 8 1
FR LY 4 2 LY 5 2 LY 6 2
FR LY 7 3 LY 8 3 LY 9 3
FI PS 1 1 PS 2 2 PS 3 3
LE
Passion Intimacy commit
PD
OU SE TV RS ND=3 MI
```

BGSU Center for Family and Demographic Research

Model 2: Allow Stability to Be Loaded onto the Passion Factor



BGSU Center for Family and Demographic Research

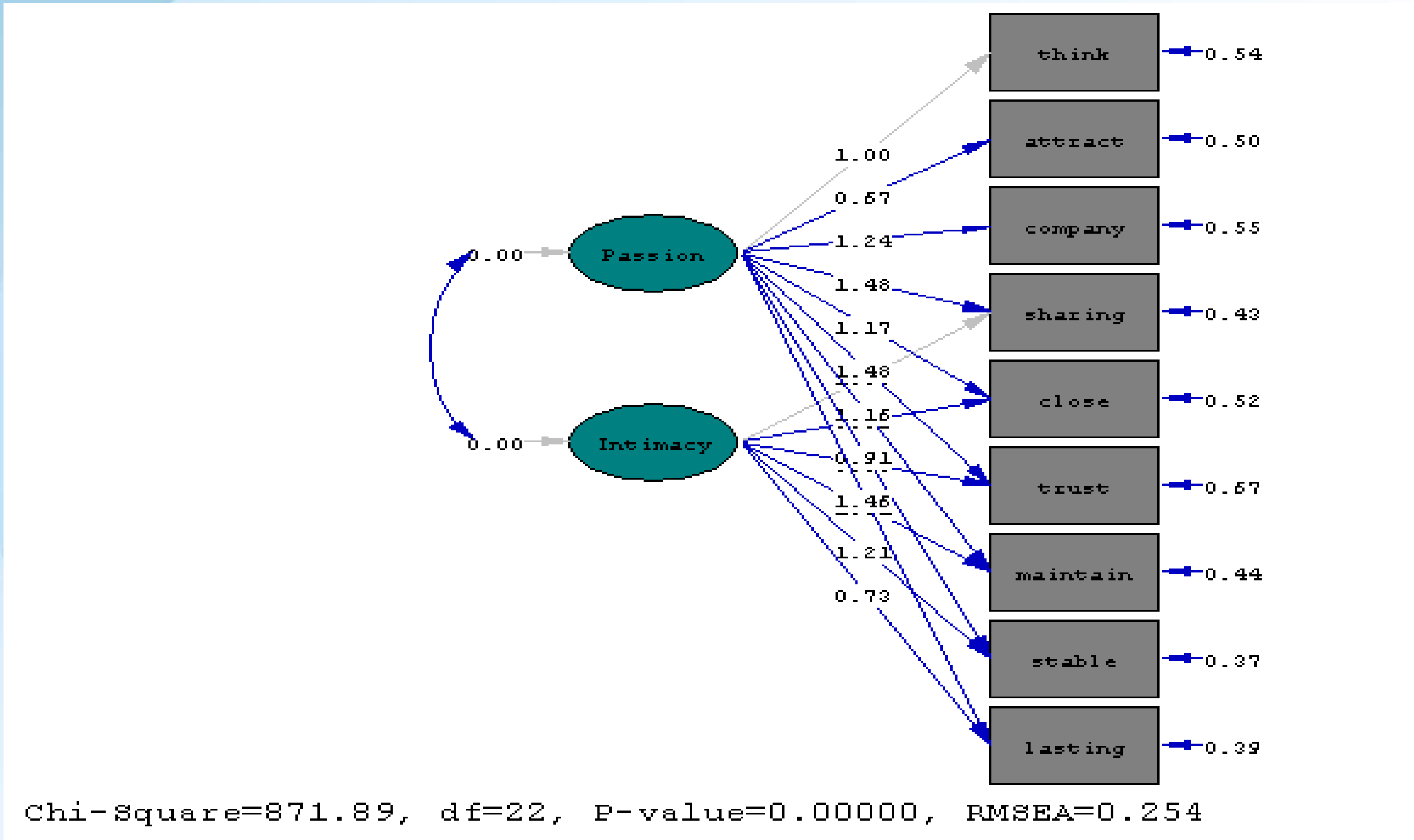
Confirmatory Factor Analysis

Model 3: A Two-factor Model

```
DA NI=10 NO=600 MI=9
RA FI='D:\factor\factor.txt'
LA
think attract company sharing close trust maintain stable lasting
SE 1 2 3 4 5 6 7 8 9/
MO NY=9 NE=2 LY=FU,FI PS=SY,FR TE=DI,FR
FR LY 1 1 LY 2 1 LY 3 1 LY 4 1 LY 5 1 LY 6 1 LY 7 1 LY 8 1 LY 9 1
FR LY 4 2 LY 5 2 LY 6 2 LY 7 2 LY 8 2 LY 9 2
FI PS 1 1 PS 2 2
LE
Passion Intimacy commit
PD
OU SE TV RS ND=2 MI
```

BGSU
Center for Family and Demographic Research

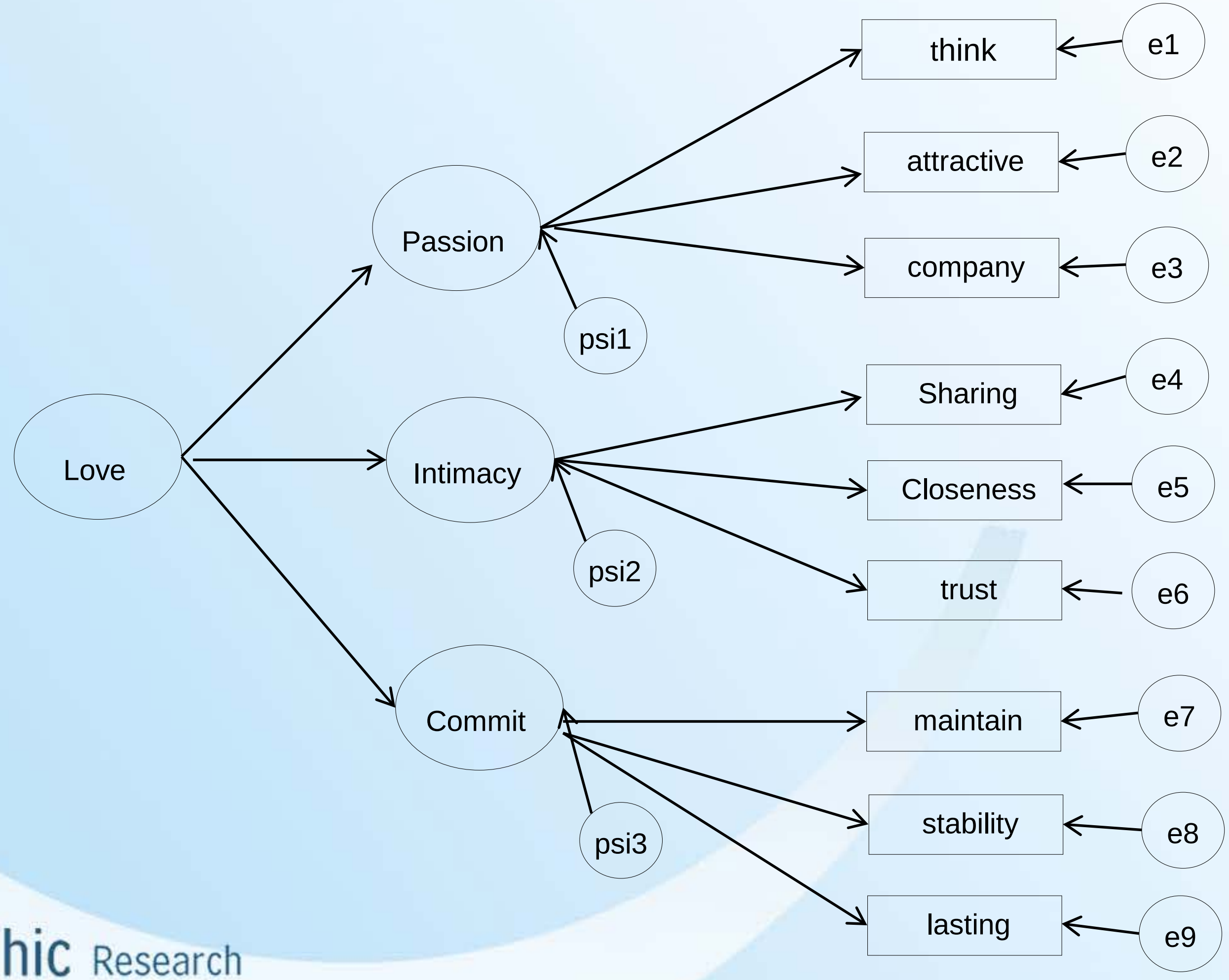
Model 3: A Two-factor Model



Center for Family and Demographic Research

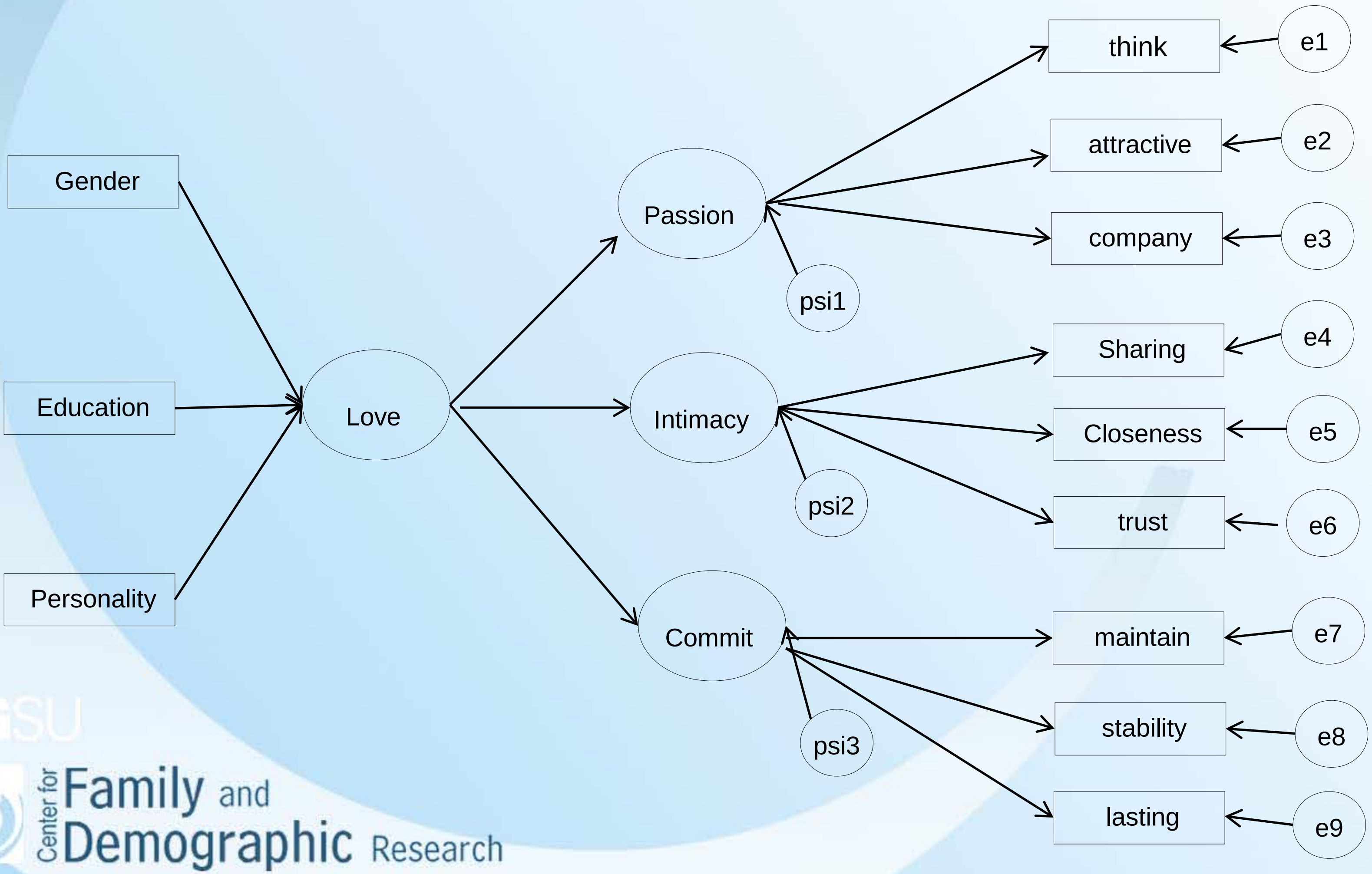
Two Hypothesized Factorial Models of the Love Measures

Model 1: Second-order Factor Analysis



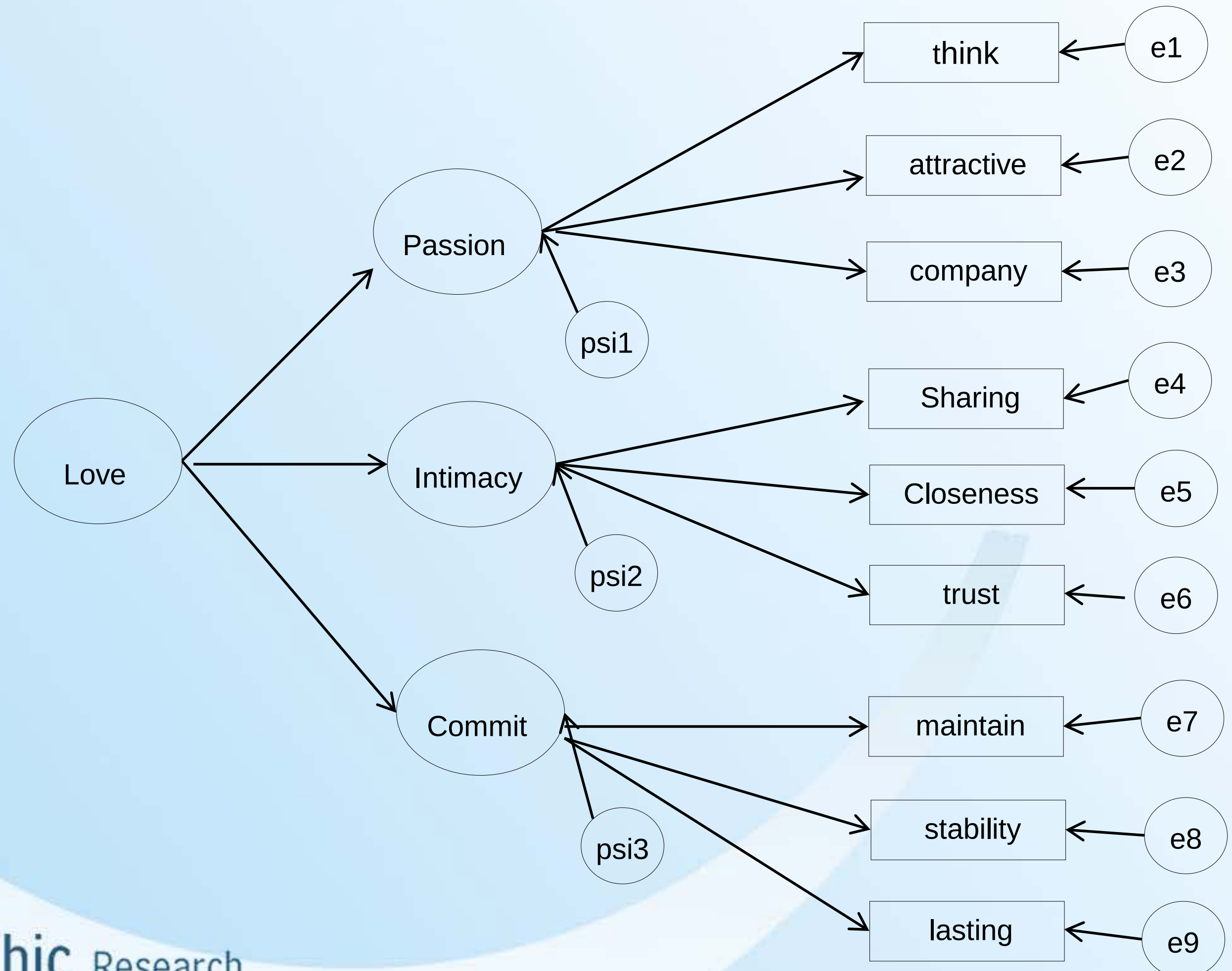
BGSU
Center for Family and Demographic Research

A Multiple indicators and Multiple Causes (MIMIC) model



Multiple Group Comparison

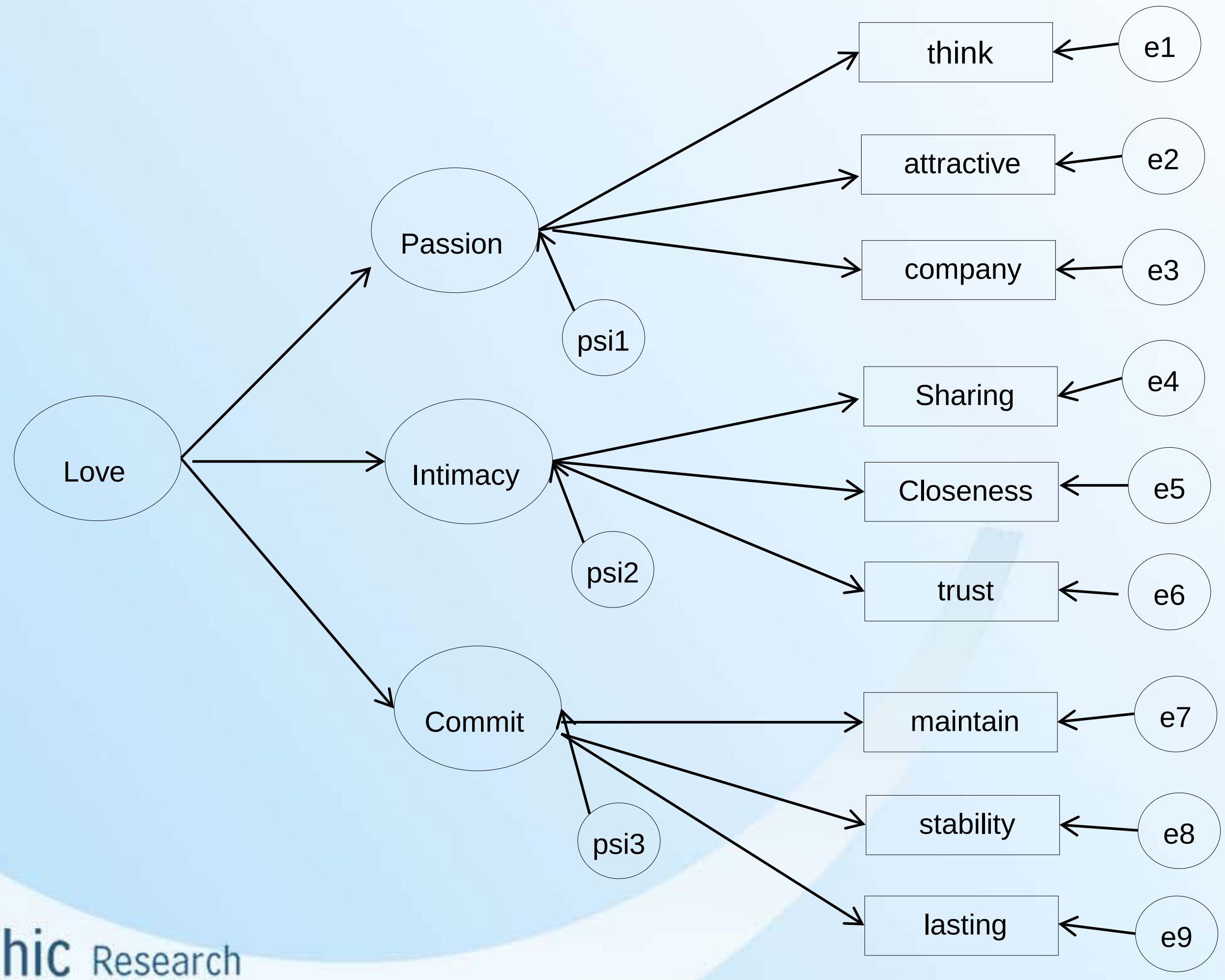
- The factorial Model for Men



BGSU
Center for Family and Demographic Research

Multiple Group Comparison

- The factorial Model for Women



BGSU
Center for Family and Demographic Research

Conclusion

- Factor analysis is a statistical technique to understand what factors underlie observed variables
- Factor analysis is closely related with the measures of validity and reliability.
- CFA is more powerful than EFA. With EFA and CFA, researchers can explore what has been measured, but only with CFA, researchers can fully test their hypotheses about the factorial structure of the measure
- Although modification index can hint on how to modify your models. You should remember that model modification should also be guided by the theory.
- If you encounter problems running factor analysis, please contact me at wuh@bgsu.edu