

Covariance Structure Modeling (CSM)

Allyson L. Holbrook

Attributions/thanks: Michael Browne
and Robert MacCallum

What is CSM?

- A statistical technique that allows a researcher to
 - Specify a model of relationships among variables
 - Estimate best-fitting parameters for that model
 - Evaluate the extent to which the proposed model explains the data

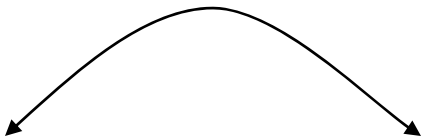
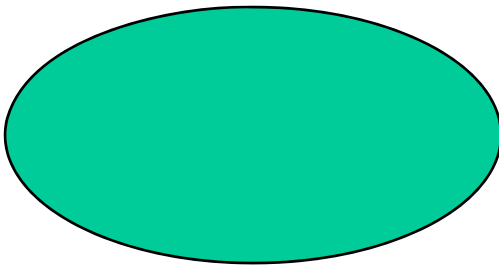
What is CSM used for?

- It is confirmatory not exploratory.
- It is typically used for observational data.
- It uses as it's input a covariance matrix (listwise deletion) and often variable means and standard deviations rather than the typical data structure (observations in rows and variables in columns)
- Tests a model that can be disconfirmed
- Particularly useful when you have multiple indicators of constructs being measured and/or complex causal pathways are hypothesized
 - Versus indices
 - Versus regressions

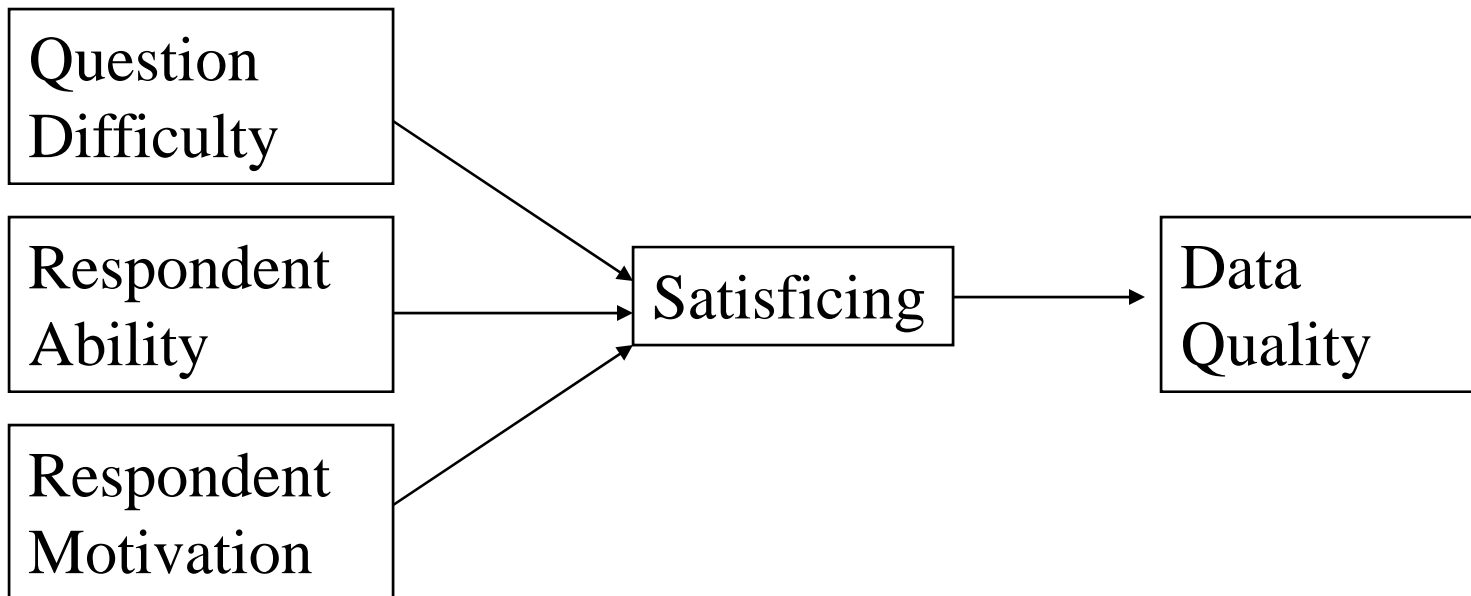
- CSM is an overarching system that encompasses
 - Path Analysis – relations among single variables
 - Confirmatory Factor Analysis – multiple indicator measurement models
 - Models that combine elements of all these
 - Aka structural equation modeling

Key concepts

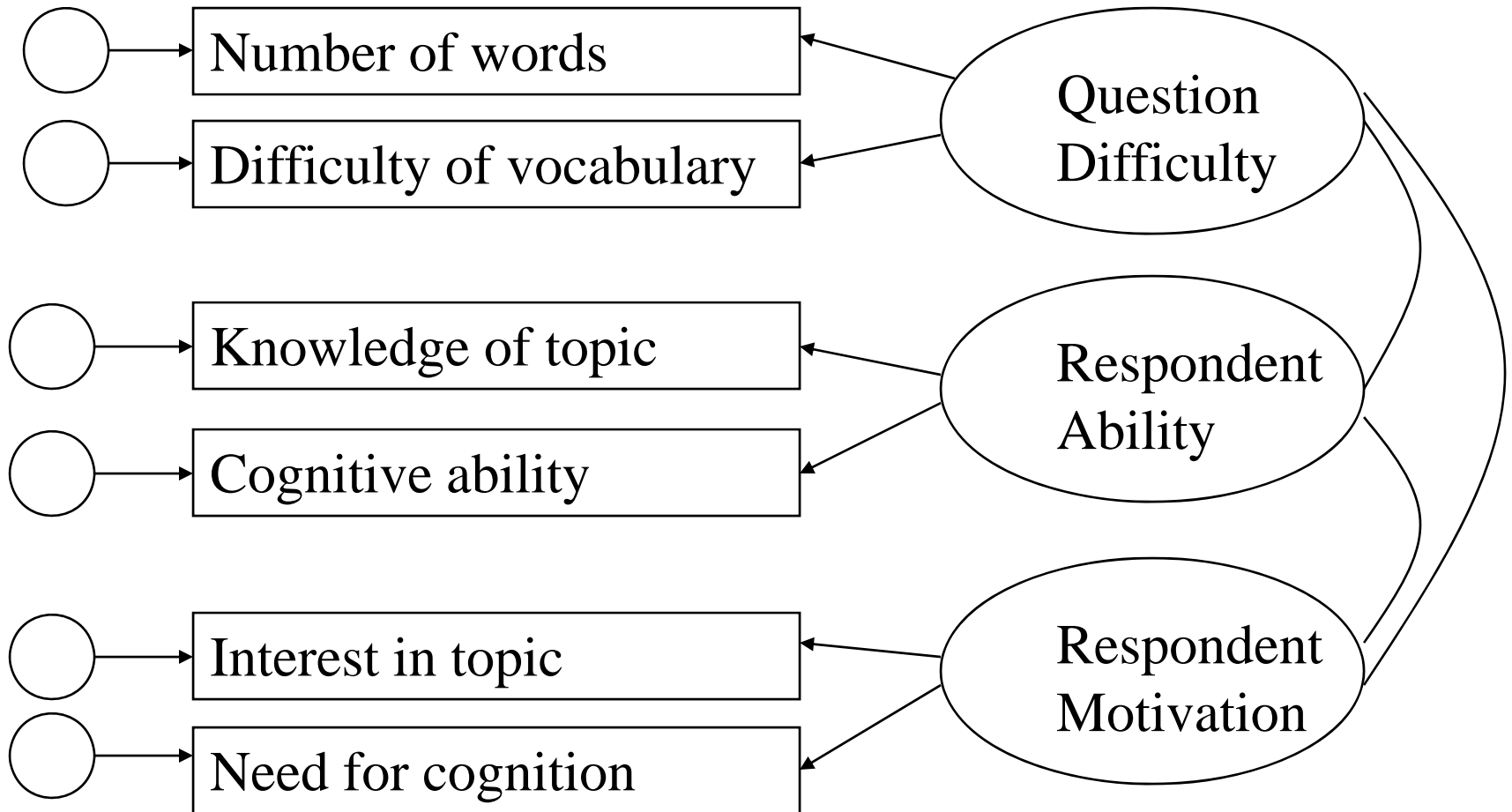
- Manifest variables
- Latent variables
 - Endogenous – receives at least one one-headed arrow
 - Exogenous – does not receive any one-headed



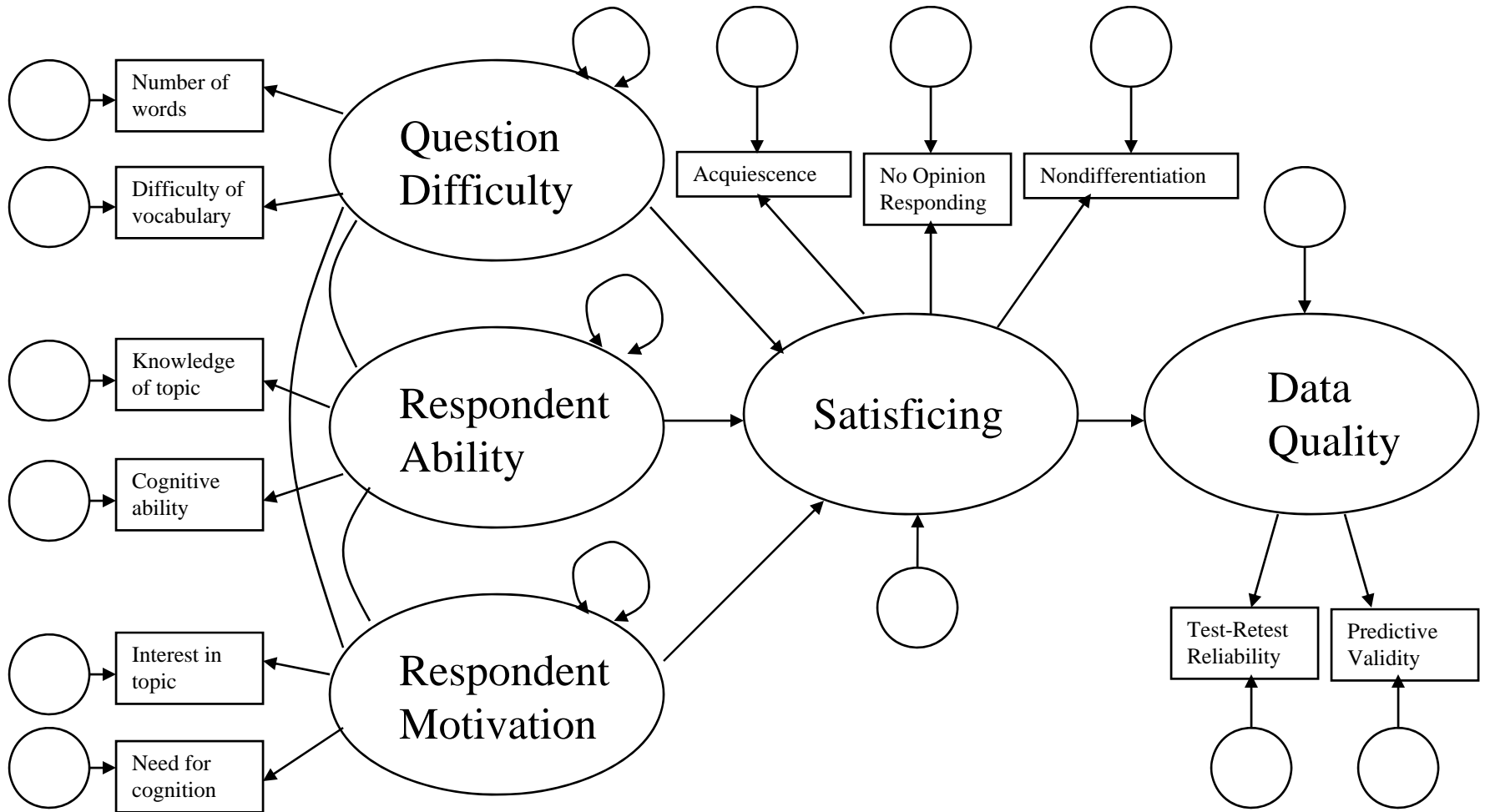
Example path analysis (all manifest variables)



Example factor analysis ("measurement model")



Example CSM:



Nuts and Bolts – How do I do CSM?

- Lots of different programs that can do CSM including...
 - LISREL*
 - RAMONA
 - EQS
 - AMOS
 - Calis
 - SEPATH
 - MECOSA
 - LISCOMP
 - Mx

CSM Steps

- Determine model (drawing)
- Develop input – covariance/correlation matrix, variable means and standard deviations, overall N (listwise deletion) - PRELIS
- Specify model
 - Variables in matrix (manifest variables)
 - Identify endogenous and exogenous latent variables
 - Number of groups (typically 1 – more on this later)
 - Type of input (covariance or correlation matrix)
 - Type of matrix to be analyzed (covariance)
 - matrices specify relationships as fixed (no relationship or 1.0 relationship), free (estimated by the model), or constrained

LISREL Matrices

- Different matrices specify different parts of the model (move toward SIMPLIS)
 - Lambda-X – measurement model for exogenous variables
 - Theta-Epsilon – error variances and covariances for indicators of exogenous variables
 - Phi – covariance and variances for exogenous latent variables

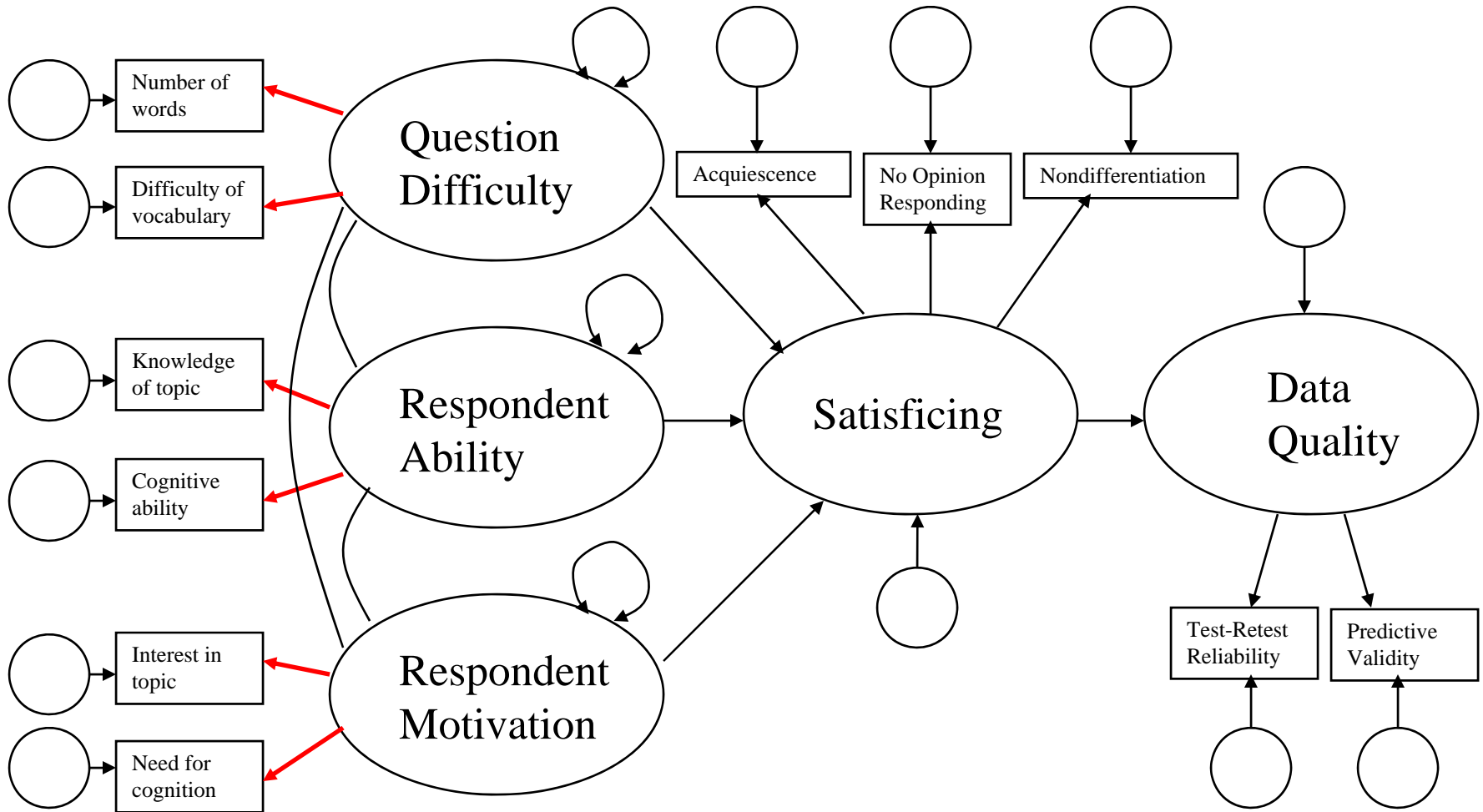
 - Lambda-Y – measurement model for endogenous variables
 - Theta-Delta – error variances and covariances for indicators of endogenous latent variables
 - Gamma – effect of exogenous variables on endogenous variables
 - Psi – Residual terms for endogenous latent variables
 - Beta – Causal relationships among endogenous latent variables
 - Can specify each element or form

LX - Example

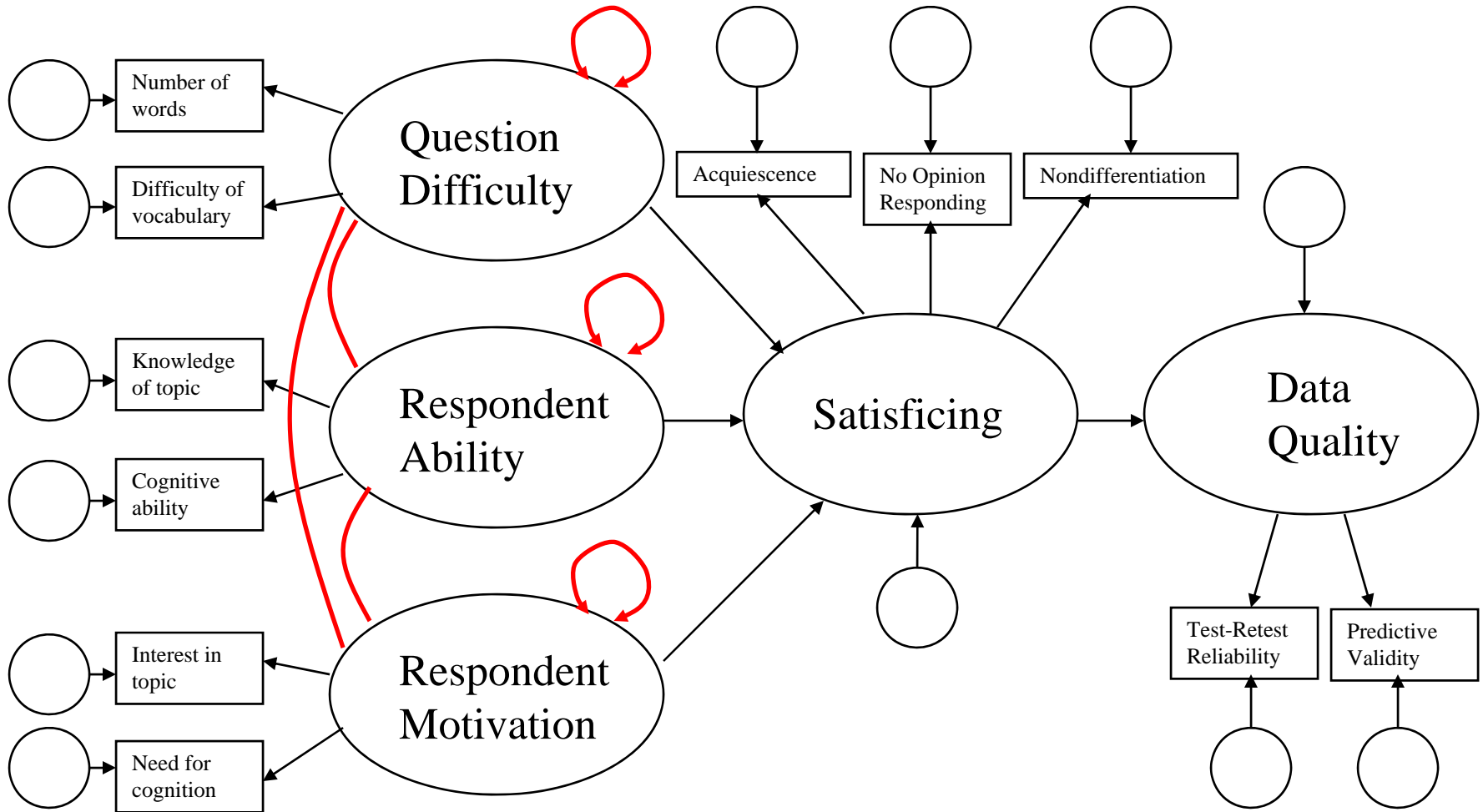
	Question Difficulty	Respondent Ability	Respondent Motivation
Number of words	X	0	0
Difficulty of vocab	X	0	0
Knowledge of topic	0	X	0
Cognitive ability	0	X	0
Interest in topic	0	0	X
Need for cognition	0	0	X

Setting the scale of latent variables – usually set one loading to 1.0. Necessary for the model to converge.

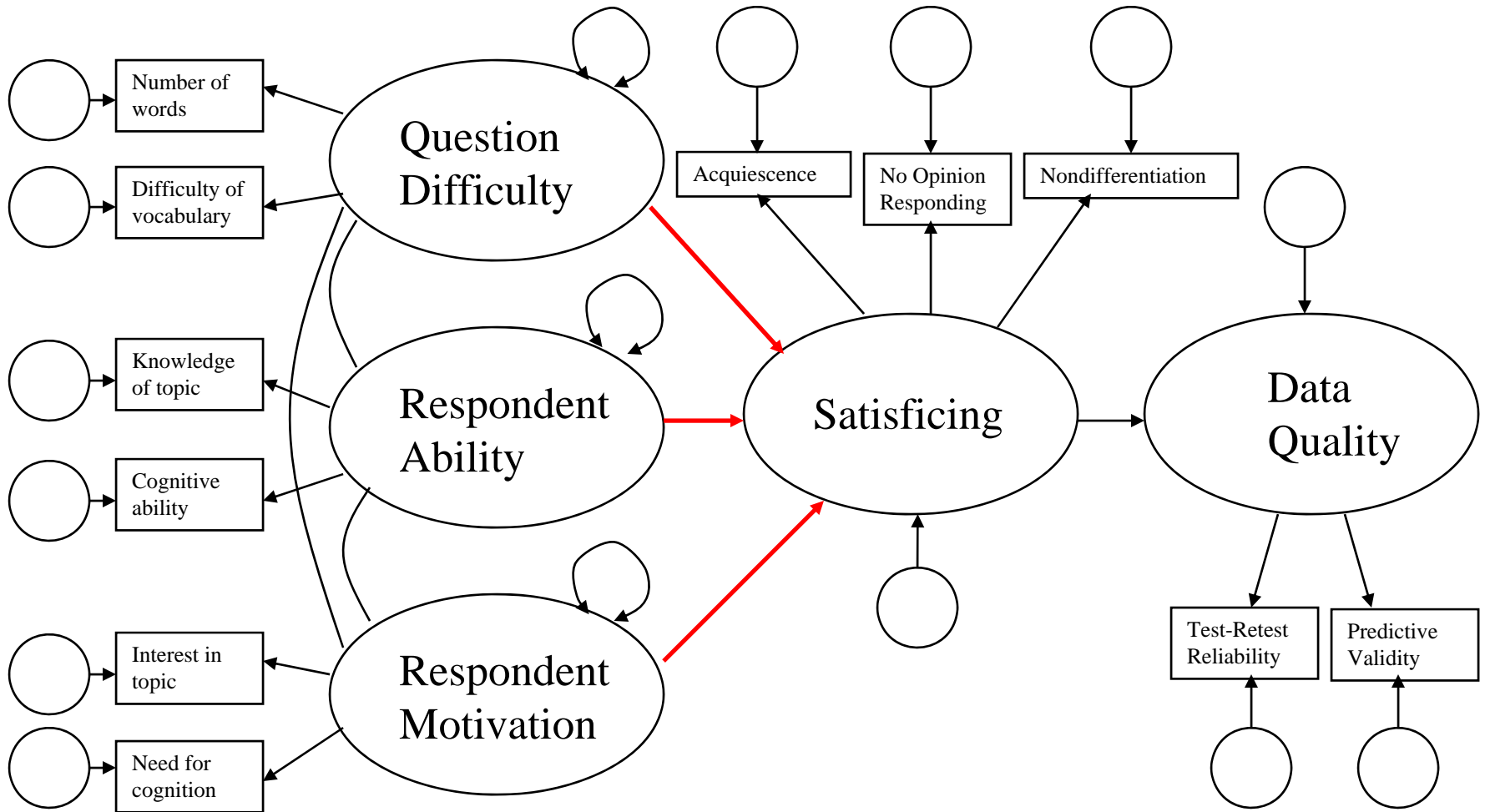
Example CSM: LX



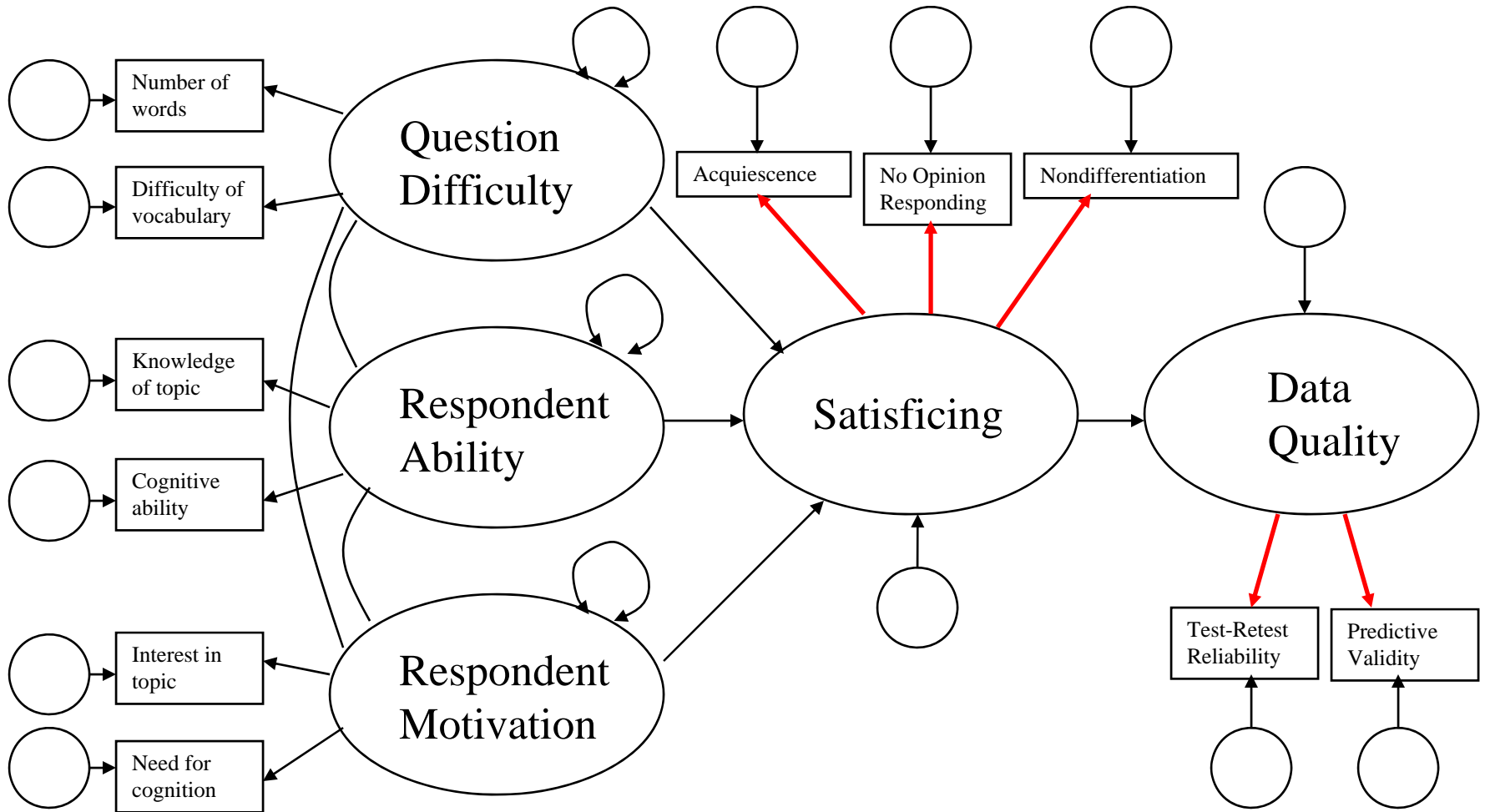
Example CSM: PH



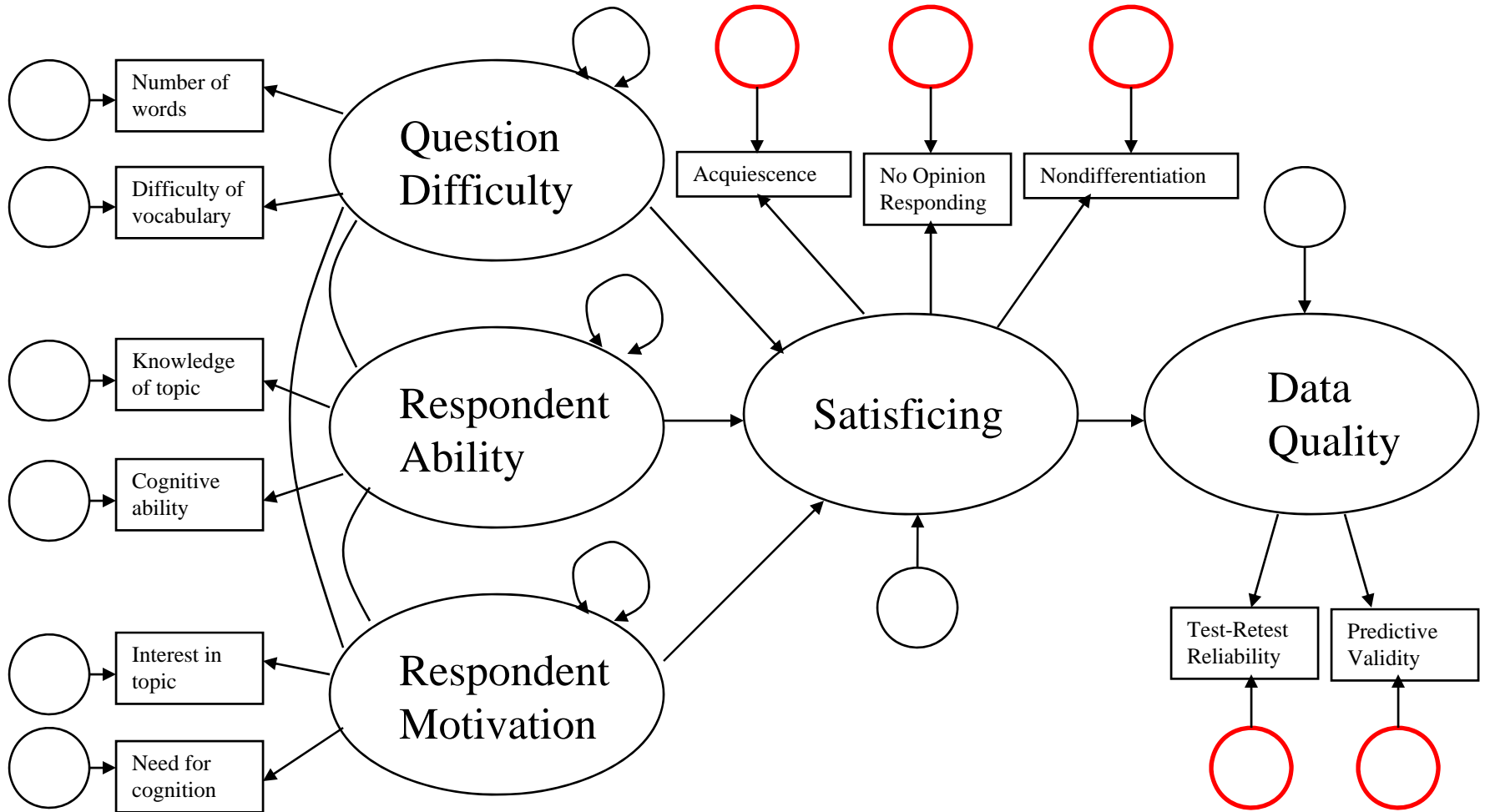
Example CSM: GA



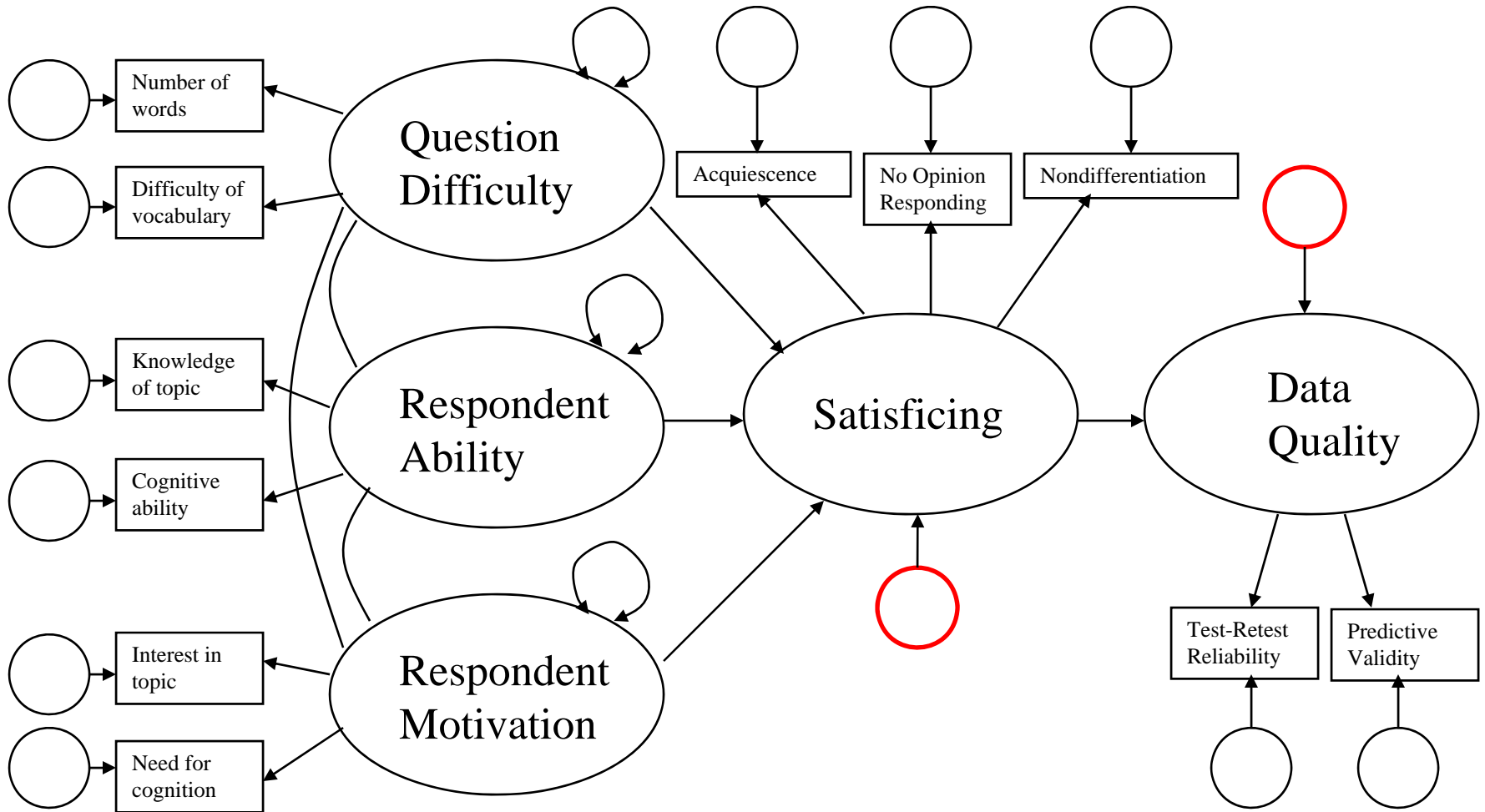
Example CSM: LY



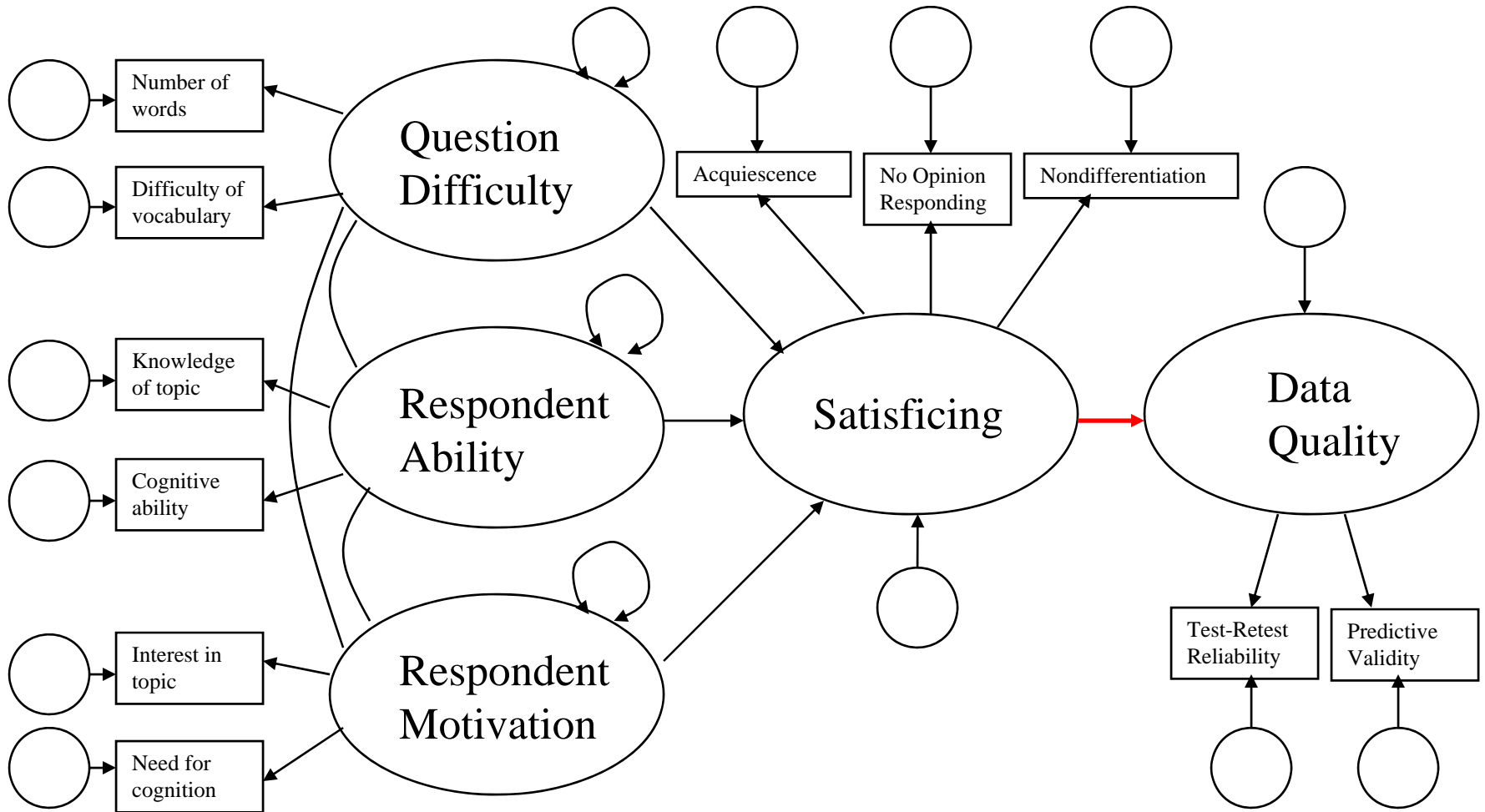
Example CSM: TD



Example CSM: PS



Example CSM: BE



Output

- Unstandardized estimates and standard errors for each estimated parameter (test of significance)
- R^2 for each endogenous LV
- Model fit
 - χ^2 fit
 - RMSEA
 - Goodness of fit index (GFI)
 - Adjusted GFI
 - Advice on indicators of model fit

Some other applications of CSM

- Composite LVs (result of causal MVs rather than cause)
 - Conceptual decision – somewhat unusual to use composite LVs
 - Makes MVs exogenous (no error terms, can covary)
 - e.g., SES
- Modeling means as well as relationships
- Multi-group analyses
 - E.g., model in men versus women
 - Role of constrained parameters (and comparing nested models)
- Longitudinal analyses
- Use of method factors to account for measurement error

Things CSM can't do or doesn't do well:

- Dichotomous manifest variables as predictors of latent variables – polychoric and tetrachoric correlations, but...
- Interactions between continuous variables
 - Interaction of each MV with each other MV
 - Estimation problems

Things to beware of

- Sample size
 - Chi-squared significance (RMSEA)
 - Large sample sizes are necessary, especially when models are complex and when effects are small
- Modification indices
 - Take advantages of chance variance/covariance
 - Replicate with new data
 - Theoretically sensible (covariances between error terms often not)

Things to beware of (2)

- Equivalent models
 - Models that would have identical fit, the only difference is the direction of a causal relationship or causal/not causal nature of a relationship
 - No universally agreed upon rules for generating equivalent models (E.g., Lee-Hershberger “Replacing Rule” – applies to relationships among LVs or stand-alone MVs)
 - MacCallum (1993) – takes model published by Sidanius (1988) and shows at least 51 equivalent models

Summary/Conclusion

- CSM is a very flexible tool
- Particularly useful when you have multiple indicators of variables
 - Models measurement error (unlike some other methods)
 - Allows flexibility in model to be tested
 - Allows estimation of overall model fit
- Also a complex tool – LISREL manuals are great references
- As always, results may be influenced by
 - Quality of input data
 - Decisions of researcher – particularly true here