Questions and Sample Answers for Chapter 5

Section 1 – Questions

Question 1

What are key conceptual and practical differences as well as similarities between multidimensional item response theory models and DCMs?

Question 2

DCMs have much in common with factor analysis (FA) and item response theory (IRT) models. Which of the following describes a critical difference between these three families of latent variable models?

a. Unlike IRT and FA models, which directly model response vectors of respondents, estimation routines for DCMs utilize summary statistics (i.e. tetrachoric and polychoric correlations).

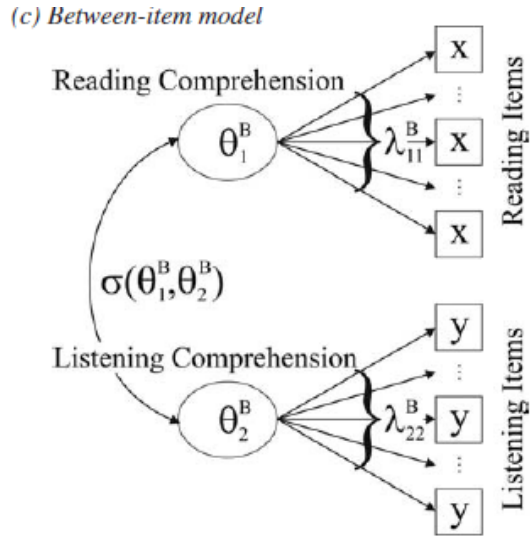
b. The interpretational focus of DCMs is on how and to what extent different attributes relate to one another, while the latent variables in IRT and FA models typically represent components of a much more narrowly defined construct.

c. IRT and FA models are typically used for norm-referenced interpretations about respondents relative to a particular population whereas DCMs support instead criterion-referenced interpretations about absolute or partial mastery / possession of attributes.

d. DCMs cannot assign attribute profiles with absolute certainty and so cut-points for mastery are assigned consensually by a panel of experts, while IRT and FA models statistically assign attribute profiles and their corresponding mastery states to respondents directly.

Question 3

In their work on large-scale assessment of foreign language ability, Hartig and Hoehler (2008) propose and then compare several different IRT models to describe students' performance on items designed to measure foreign language reading and listening comprehension. One model they propose is a two-dimensional between-item model in which listening and reading comprehension are conceptualized as two distinct abilities and student performance on the reading and listening items depends, respectively, on these two separate dimensions. This design is shown in the following figure:



Using the conventions of the graphical representations introduced in section 5.1.6, how would you represent this two-dimensional MIRT model as a two-dimensional confirmatory CFA model with a complex loading structure such that the first dimension (reading comprehension) represents the abilities common to all items and the second dimension (listening comprehension) represents only those abilities specific to the listening items? How would you represent this as a two-dimensional confirmatory DCM with the same complex loading structure? Draw the graphs that represent each model and describe the statistical features that distinguish the models from one another.

Reference

Hartig, J., & Hoehler, J. (2008). Representations of competencies in multidimensional IRT models with within-item and between-item multidimensionality. *Journal of Psychology, 216*, 89-101.

Question 1

DCMs are conceptually related to traditional multidimensional IRT models. DCMs are also multidimensional measurement models and directly estimate the correlations of latent variables, based on which we infer how theoretical constructs are related to each other. Both modeling families use information contained in item responses to estimate item parameters, which, in turn, are used to estimate respondent parameters. DCMs also resemble multidimensional IRT models in that they "break down" a coarsely defined construct into subcomponents that are represented by separate latent variables in the model. Both modeling families can be used to specify and estimate measurement structures with between-item multidimensionality or within-item multidimensionality.

In terms of the scales of observed response variables, DCMs are similar to multidimensional IRT models in that they are typically used when the response variables are dichotomous or polytomous. Yet, neither modeling family is technically limited to categorical item responses. This has implication for the estimations procedures that are used to estimate DCMs and multidimensional IRT models, which are full-information estimation approaches that utilize the information that is contained in the unique item response vectors to estimate model parameters.

One additional similarity is the manner in which latent variables are combined. In multidimensional IRT models and DCMs latent variables can be combined in the form of a sum or a product, which is often associated with the terms "compensatory models" and "non-compensatory models", respectively, even though such a terminological distinction is not statistically clean.

A difference between the two modeling families is the scale of the latent variables. Whereas latent variables in IRT models are continuous resulting in norm-referenced interpretations about respondents, latent variables in DCMs are categorical and typically support multiple criterion-referenced interpretations.

Another distinction between multidimensional IRT models and DCMs is the confirmatory or exploratory nature of the models. DCMs are exclusively confirmatory in nature as they are used to confirm or invalidate a hypothesis about the relationships between item responses and the underlying mental mechanism. The use of DCMs thus requires that a loading structure be specified a priori in the form of a Q-matrix. Multidimensional IRT models, in contrast, can be either exploratory or confirmatory: items can be allowed to load freely on all latent variable or specific items can be specified to load on specific latent variables.

Question 2

a. Both IRT and DCMs utilize the full information contained in the data matrix for estimating unknown parameters; all unique response vectors contribute to the estimation of parameters in these models. Factor analysis models, in contrast, use summary statistics that include covariances / correlations and means appropriate for continuous observed variables to estimate unknown parameter values.

b. The interpretational focus of factor analysis models is typically on the relationships between latent variables in the model (i.e., on the general model structure), which are often more broadly defined than the latent attributes found in a DCM. For DCMs, as for IRT models, the interest is typically in scoring respondents even though model-data fit procedures give some insight into the dimensionality and, thus, structure of the model as well.

c. Correct answer.

d. It is true that DCMs cannot assign profiles with absolute certainty. However, it is DCMs that classify respondents directly into mastery states using respondents' estimated probability of mastery whereas IRT and FA models will first and foremost create proficiency estimates from a continuous population distribution. These estimates can then be used to rank-order respondents relatively finely and would only lead to classifications of respondents if an a priori cut score were used on the resulting scale. Such a cut score is not developed statistically but, rather, through a blend of consensual and, sometimes, additional post-hoc quantitative analyses.

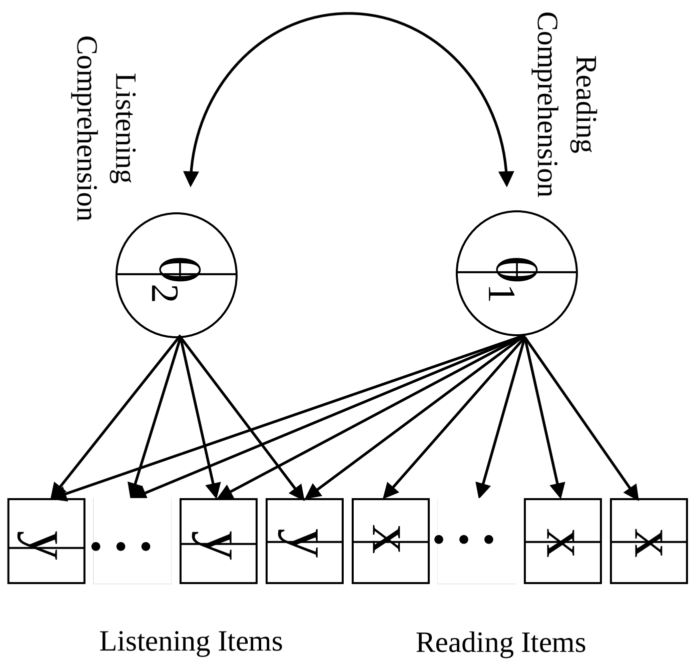
Question 3

The graphs of a two-dimensional CFA model and a two-dimensional DCM are shown in the figure on the next page.

For both the CFA model and the DCM, the latent attribute variables that represent listening and reading comprehension are represented as circles while the observed item score variables are represented as squares. Both drawings also have appropriate loading structures to represent within-item multidimensionality, with reading comprehension (θ_1) mapping onto all of the items and listening comprehension (θ_2) mapping only onto the listening items. The arrows are pointing from the latent variables to the outcome variables, the *X's* and *Y's*, as variation in the outcome variables is predicted by variation in the latent variables (i.e., since these models are akin to systems of simultaneous multiple regression models with latent predictor variables). In both diagrams a two-sided arrow representing the correlation between the latent variables is included. The features that graphically distinguish DCMs from CFA models are (1) the addition of vertical bars through both the observed response and latent variables for the DCM graph; and (2) the absence of arrows pointing toward the observed response variables to represent individual error terms in the DCM graph.

As these graphs underscore, although a CFA Model and a DCM are structurally similar in some respects, there are also key differences between the models which can be communicated through these graphical representations. A CFA model is designed to generate continuous multidimensional profiles of respondents, mapping continuous latent variables onto continuous observed response variables; tetrachoric or polychoric correlation matrices would have to be used for discrete variables in general. In contrast, DCMs classify respondents directly according to their probability of mastery on one or more latent variables; both the latent and observed variables are categorical variables and they are connected via an appropriate non-identity link function.

Two-dimensional DCM with complex loading structure



Two-dimensional CFA Model with complex loading structure