

# Potential Applications for Multidimensional Item Response Models

Michael C. Edwards

May 24, 2009

# Today's Talk

- Current state of affairs
- Item response theory and factor analysis
- Potential uses for multidimensional models

For a long time IRT has been dominated by two critical assumptions:

- 1 There is only one construct being measured
- 2 That construct is normally distributed

# As we stand today

As you've heard from Carol and Li's talks, those aren't necessary assumptions anymore. Developments are in progress that should make these new methods widely available in user-friendly software packages.

# Back together again

Focusing on the dimensionality aspect of these developments, we see that one of the long-standing differences between how IRT operates in the world and how FA operates in the world no longer exists. It seems reasonable to shed some of the existing distinctions. In this light, several folks have started using the phrase “item factor analysis” to highlight that these models are, after all, just factor analytic models for item-level data.

# Beginning at the beginning

Once upon a time, item response theory (IRT) and factor analysis (FA) were developed out of a common basis in the Thurstonian common factor model.

# Beginning at the beginning

Once upon a time, item response theory (IRT) and factor analysis (FA) were developed out of a common basis in the Thurstonian common factor model.

Despite this shared origin, the two methodologies (or more appropriately two different frameworks) diverged onto two separate tracks for about forty years.

# The good, the bad, and...

This bifurcation has had good and bad consequences.

# The good, the bad, and...

This bifurcation has had good and bad consequences.

One consequence was that development in each framework happened more or less independent of the other. IRT, by virtue of many of its developers being in education, became very concerned with issues of scoring.

# What we've learned along the way

As the FA and IRT frameworks merge back together, it is useful to consider what each camp has learned in their solitary wanderings over the years.

# What we've learned along the way

As the FA and IRT frameworks merge back together, it is useful to consider what each camp has learned in their solitary wanderings over the years.

As I mentioned before, IRT folks have long been preoccupied with scoring and it may be time to start thinking more about item factor analysis models in terms of the meaning we would like to attach to scores.

# What now?

I am reminded of the famous saying by statistician George Box that all models are wrong, but it is only necessary that they be useful.

# What now?

I am reminded of the famous saying by statistician George Box that all models are wrong, but it is only necessary that they be useful.

This brings me to my main theme for today: How can applied researchers benefit from the use of item factor analysis?

# Choosing a model

In the FA tradition, model choice has been guided mostly by the various and sundry fit indices that exist and are provided by most SEM-based software packages.

# Choosing a model

In the FA tradition, model choice has been guided mostly by the various and sundry fit indices that exist and are provided by most SEM-based software packages.

I think it might be time to worry a little less about fit and a little more about function.

# Function over fit

I'm certainly not arguing that we adopt models which do not provide reasonable approximations to our observed data.

# Function over fit

I'm certainly not arguing that we adopt models which do not provide reasonable approximations to our observed data.

However, I'm not aware of any research which can, in concrete terms, tell me why I should strongly prefer a model with an RMSEA of 0.05 over one with an RMSEA of 0.08.

# An example

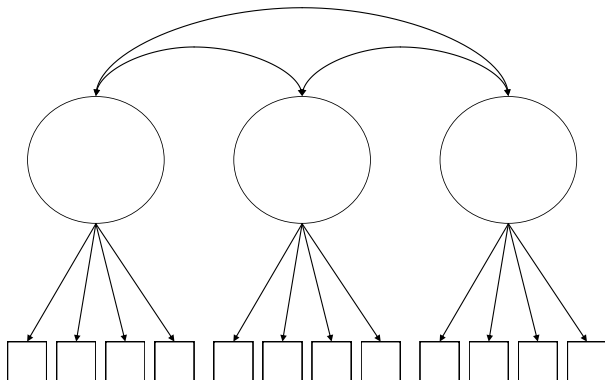
In the slides that follow I examine three different models which could be used to model three factors. If they provide reasonable and similar approximations to the observed data, my belief is that we should choose the model which best suits our current purposes.

# An example

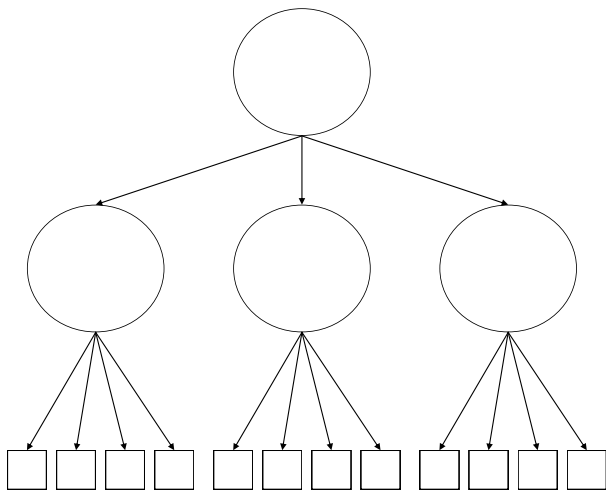
In the slides that follow I examine three different models which could be used to model three factors. If they provide reasonable and similar approximations to the observed data, my belief is that we should choose the model which best suits our current purposes.

Remember, they are all wrong, so we should focus on choosing the most useful model - and this will depend upon the goals of the researcher.

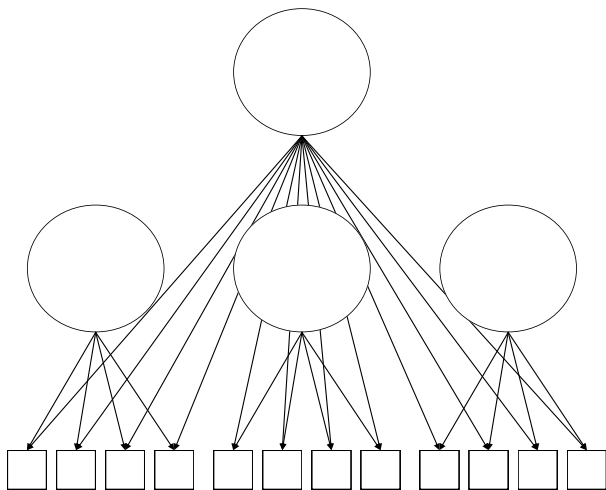
# Three “primary” factors



# Three lower order factors



# Three “cluster-specific” factors



# Meaningful scores

As we've seen, the interpretation attached to the scores produced by these three models can be quite different from a substantive standpoint, even if the models may provide similar (or identical) fit from a statistical one.

# Conclusion

It is an exciting time in psychometrics, as new developments have greatly expanded the range of plausible models. To make maximum benefit of these methodological extensions, we must also consider in what contexts these tools can best be used.

# Thank you

edwards.134@osu.edu