A newly designed comprehensive Science Learning Activation Lab instrument based on the multidimensional structure of dispositions, skills, and knowledge theorized to comprise science activation was examined. Two methods of investigating the psychometric properties are typically used: confirmatory factor analysis (CFA) and Item Response Theory (IRT). The relationship between CFA and IRT is well established:  
- CFA is better at modeling the items to the underlying latent trait(s) relationship  
- IRT is better at modeling the items to persons relationship.  

However, using a combined approach can lead to development of stronger social science instruments and better evaluation of those instruments.  

### Research Design & Methods

To provide dimension scores and an overall activation score, CFA and IRT were iteratively used to refine the final item set to include items that:  
- contributed significantly to their dimension  
- minimized cross-loadings

#### IRT analysis:  
- IRTPRO 2.11  
- Maximum likelihood estimator (MLE)  
- Graded response model (GRM) for polytomous items  
- A 2-parameter logistic (2PL) model for dichotomously scored items. Model fit indices showed that slopes were significantly predicting, so the 1PL was insufficient for this instrument.

#### CFA analysis:  
- MPlus v.6.19  
- Items designated as categorical  
- Weighted least squares means and variances (WLSMV) estimator

#### What We’ve Learned So Far

<table>
<thead>
<tr>
<th>IRT</th>
<th>CFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item difficulty &amp; person ability independent</td>
<td>Item difficulty &amp; person ability independent</td>
</tr>
<tr>
<td>Item fit: weighted &amp; unweighted mean square errors</td>
<td>Instrument Dev.</td>
</tr>
<tr>
<td>Item performance independent of other items</td>
<td>Scale fit</td>
</tr>
<tr>
<td>S.E. sample independent</td>
<td>Item performance dependent on other items</td>
</tr>
<tr>
<td>S.E. varies across sample</td>
<td>S.E. averaged across sample</td>
</tr>
<tr>
<td>S.E. varied across sample</td>
<td>CFA</td>
</tr>
<tr>
<td>Item fit: factor loadings, error variance, communalities</td>
<td></td>
</tr>
<tr>
<td>Instrument Dev.</td>
<td>Scale fit</td>
</tr>
<tr>
<td>Item performance</td>
<td></td>
</tr>
<tr>
<td>single fit indicator</td>
<td></td>
</tr>
</tbody>
</table>

M2011M was retained after combining categories based on the IRT results and CFA category frequencies. Both methodologies in subsequent analysis of the sense-making scale showed good item fit for the revised item.

### Recommendations

We offer the following observations for researchers desiring to use both methodologies:  
1. IRT can identify items as misfitting if the underlying model is not specified correctly. While the underlying structure is generally based on theory, best practice dictates confirming the dimensional structure prior to IRT estimation using CFA.  
2. Items with response categories that are unendorsed or rarely endorsed result in inconsistent item fit statistics.  
3. Due to model specifications, CFA and IRT cannot determine if items are loading across dimensions. While some cross-loading items were flagged by IRT analysis as misfitting, not all items identified by the EFA structural analysis were flagged by the IRT analysis.  
4. Polytomously-scored items can result in good item fit statistics during CFA, however, review of the IRT analysis for the same items can show poor item fit when thresholds are not consistent with the expected ordering of responses.  
5. Individual factor scores obtained from the CFA and IRT dimension scores (thetas) should be highly correlated. However, when items in violation of underlying assumptions of either methodology (items with negative factor loadings or items with inconsistent thresholds values) are included in the item set, the correlation between scores becomes less than desirable.  
6. Theory is essential, but insufficient to determine if items are functioning as expected. Items with the best fit and alignment with theory will provide the best possible information.

### References