

## **Title:** Diagnostic Modeling with R and Stan

### **Abstract:**

Diagnostic classification models (DCMs; also known as cognitive diagnostic models) have gained interest in the fields of educational assessment in recent years due to their ability to provide fine-grained actionable feedback while keeping test lengths short. In this workshop, we will cover how to easily estimate and evaluate DCMs with Stan and the R package `measr`. The `measr` package includes convenience functions for creating Stan scripts, estimating models, and extract relevant pieces of model output. The workshop will include hands-on examples of defining a DCM, estimating the model, and evaluating the fit (e.g., test- and item-level fit, classification accuracy and consistency, etc.).

The goal of the workshop is to enable participants to implement DCMs in their own work, and thus is intended for anyone who uses, or would like to use, DCMs for applied or research uses (e.g., psychometricians, faculty, applied researchers, graduate students). Although not necessary, prior experience with R will be helpful. All workshop materials, including slides, examples, and solutions will be available on a workshop website. Participants should have access to a laptop they can bring to the workshop in order to follow along with the examples. Instructions for installing any necessary software will be provided.

### **Summary:**

Diagnostic classification models (DCMs) are a powerful tool for analyzing assessment data (e.g., dichotomous or polytomously scored items) to understand human cognition and behavior. These models provide actionable feedback by reporting specific skills a respondent has mastered, or which traits are present. DCMs are intended to be multidimensional, allowing many attributes to be measured on a single assessment. For example, an educational assessment may assess student proficiency on several standards. DCMs can be used to report a status (i.e., proficient/not proficient) for each student and assessed standard. Additionally, the reported proficiency results can be used to make inferences about the order in which student typically acquire the knowledge required for each standard.

Due to the categorical nature of the latent constructs, reliable results can be achieved with much shorter assessments than with traditional scale score assessments. Despite these benefits, DCMs are not often used in applied settings. One reason for this is that methods for estimating and evaluating DCMs are often presented in papers, but applied practitioners are left to implement the methods on their own. Additionally, software that does exist for DCMs is often complicated, under-documented, or limited to specific DCM subtypes. In this tutorial, we will showcase how DCMs can be easily estimated and evaluated using Stan and a new R package, `measr`, which is designed to be user friendly, with extensive and accessible documentation.

We will begin the workshop with a high-level introduction to DCMs for participants who may have a limited familiarity with this class of models; however, this workshop is not intended to provide a deep technical understanding of DCMs. Rather, the majority of the session will focus on learning tools (i.e., Stan and `measr`) for implementing these models in the participants' own work. Participants will first learn how to estimate DCMs using Stan and `measr`. Our focus will then turn to interpreting the model output to evaluate respondent attribute profiles and

understand the relationships between the attributes. We'll then shift to the evaluation of DCMs, including assessing how well the model fits the data, investigating item performance, and reliability.

The presenter has extensive experience with DCMs and R and is the author of the `measr` package. As the Assistant Director of Psychometrics for Accessible Teaching, Learning, and Assessment Systems (ATLAS) at the University of Kansas, Dr. Thompson oversees operational scoring and reporting for a large-scale accountability assessment, the Dynamic Learning Maps Alternate Assessment, which uses DCMs to report results. He also leads the DCM research program for ATLAS, including the creation of the `measr` R package. Thus, he is not only skilled in using R to estimate and evaluate DCMs, but also has extensive experience in using DCMs in applied or operational settings. Dr. Thompson also has experience teaching workshops using R and RStudio for analyzing data. An example illustrating the accessibility of his workshop can be found at <https://tidyds-2021.wjakethompson.com>. A similar website will be created for this workshop.