# Modeling Interactions Across Skills: A Method to Construct and Compare Models Predicting the Existence of Skill Relationships

Anthony F. Botelho
Worcester Polytechnic Institute
100 Institute Rd.
Worcester, MA 01609-2280
abotelho@wpi.edu

Seth A. Adjei
Worcester Polytechnic Institute
100 Institute Rd.
Worcester, MA 01609-2280
saadjei@wpi.edu

Neil T. Heffernan
Worcester Polytechnic Institute
100 Institute Rd.
Worcester, MA 01609-2280
nth@wpi.edu

## ABSTRACT
The incorporation of prerequisite skill structures into educational systems helps to identify the order in which concepts should be presented to students to optimize student achievement. Many skills have a causal relationship in which one skill must be presented before another, indicating a strong skill relationship. Knowing this relationship can help to predict student performance and identify prerequisite arches. Skill relationships, however, are not directly measurable; instead, the relationship can be estimated by observing differences of student performance across skills. Such methods of estimation, however, seem to lack a baseline model to compare their effectiveness. If two methods of estimating the existence of a relationship yield two different values, which is the more accurate result? In this work, we propose a method of comparing models that attempt to measure the strength of skill relationships. With this method, we begin to identify those student-level covariates that provide the most accurate models predicting the existence of skill relationships. Focusing on interactions of performance across skills, we use our method to construct models to predict the existence of five strongly-related and five simulated poorly-related skill pairs. Our method is able to evaluate several models that distinguish these differences with significant accuracy gains over a null model, and provides the means to identify that interactions of student mastery provide the most significant contributions to these gains in our analysis.

## Keywords
prerequisite structures, skill relationships, feature selection, model comparison

## 1. INTRODUCTION
Many educational systems like ASSISTments and Khan Academy already implement a prerequisite structure as a suggested ordering in which skills should be presented to students. These structures are often developed by domain experts and teachers in the field of study, and are likely to hold ground-truth. It is clear, for example, that relationships can be identified by observing skills at the problem-level; by viewing the steps required for students to complete each item, it can be known that any skills required to complete such problems can be considered prerequisites. For example, Multiplying Whole Numbers may act as a prerequisite to Greatest Common Factors, as is used in our analysis. While causality suggests a strong relationship, it is possible for two skills to relate to each other in other ways. Such relationships are less intuitive, perhaps requiring a similar thought process or sequence of steps to solve, even if the content of such tasks differ. Many causal skill arches are identifiable by domain experts by observing content, but as described, other such relationships may be missed due to their non-intuitive structures. By observing strong skill relationships identified by domain experts, we construct a method of measuring the factors that are most predictive of their existence.

We also argue that identifying strong relationships is not enough for a method of prediction to be considered adequate. Such a method should also be able to identify weak or non-existent skill relationships. It is likely that while much attention and research is placed on structuring prerequisite links, some of these are false-positives. In other words, a skill may be listed as a prerequisite, but has no true relationship to its supposed post-requisite skill. In such a case there is little or no interactions of performance. Such links must also be identified and removed or reordered in learning platforms to benefit the students.

A significant amount of research has looked at measuring the strength of skill relationships [1],[4], and even the effects such relationships have on measuring student performance [3],[10], but without understood ground truths, it is difficult to compare across these methods. Furthermore, many of these methods represent similar conceptualizations of performance inherently, or through variations of representation such as aggregation or centering. For example, "student achievement" is likely a predictor of skill relationships (achievement on a prerequisite skill will likely influence achievement on a post-requisite skill), but can be represented as the percent of problems answered correctly, mastery speed (the number of items needed to complete an assignment as is commonly used in intelligent tutoring systems), or countless other combinations of features. It will be important to distinguish between these generalized components to avoid incorporating features that capture the same types of conceptualizations into predictive models.

This work provides a method to evaluate models that measure the strength of skill relationships, and with this model we attempt to identify which features best indicate a strong relationship between two skills. This analysis will incorporate a method of generalizing and distinguishing features that measure different aspects of learning and performance. With this methodology, we seek to answer the following two research questions:

1. What link-level features, expressed in this paper as interactions of performance between skills, are significant in predicting the existence or non-existence of skill relationships?

2. Which features are the strongest predictors of skill relationships, and does combining them make for a more accurate predictive model?

The next section of this paper will discuss some of the previous research performed on skill relationships and prerequisite structures. Then, we will discuss our theory and methodology to provide a baseline model of comparing methods of measuring skill relationships. Using this model, we then compare several commonly-used student-level features, and of the most accurate, compare several different representations of those features. Finally, we will discuss our findings and suggested future works.

## 2. PREVIOUS WORKS

The discovery and refinement of prerequisite skill structures has been an important research question in recent years. The impact of this research on educational systems cannot be overemphasized. Domain experts who design these structures need data centered methods to support the decisions they make; it is vital to have empirical data to support hypothesis regarding the order in which skills are presented as it can have a large impact on student achievement and either aid or impede the learning process. Additionally, identifying the best prerequisite skill structure will enhance student modeling; knowing a student's prior performance on prerequisite skills can help estimate that student's performance on the post-requisites. This can lead to earlier interventions for struggling students, or even help redefine mastery perhaps students who perform very well on a prerequisite requires less practice on a post-requisite, or can be given more advanced examples.

Tatsuoka, defined a data structure called the Q-Matrix, that represents the mapping of problems to skills: the rows of this matrix represent the problems, and the columns represent the skills [9]. Though the goal of the research was to diagnose the misconceptions of students, they set in motion a number of studies that have used this data structure as the first step to find prerequisite structures [2],[5],[8].

Desmarais and his colleagues developed an algorithm that finds the prerequisite relationship between questions, or items, in students' response data [6]. They compare pairs of items in a test and determine any interactions existing between each pair. Depending on the interactions and a set of interaction-related criteria, they determine whether the two items have a prerequisite relationship between them. This approach was applied by Pavlick, et al. to analyze item-type covariances and to propose a hierarchical agglomerative clustering method to refine the tagging of items to skills [7]. Brunskel conducted a preliminary study in which they use students' noisy data to infer prerequisite structures [4]. Further research by Scheines, et al. extended a causal structure discovery algorithm in which an assumption regarding the purity of items is relaxed to reflect real data and to use that to infer prerequisite skill structure from data [8].

## 3. DATASET

The dataset[1] used for this study consists of real-world student data from the ASSISTments online learning platform. The raw data contains student problem logs pertaining to ten math skills from the 2014-2015 school year. These ten skills represent five skill pairs, listed in Table 1, for which domain experts identified as having a strong prerequisite relationship. While we are not limiting the usage of our proposed baseline model to just prerequisite relationships, these are the most reliable to identify due to the causal effect of content (if problems in skill B require the use of skill A to complete, a strong relationship can be identified).

**Table 1: The strong skill pairs as determined by domain experts**

| Prerequisite | Post-requisite |
|---|---|
| Multiplication of Whole Numbers | Greatest Greatest Common Factor |
| Subtracting Integers | Order of Operations |
| Division of Whole Numbers | Dividing Multi-Digit Numbers |
| Volume of Rectangular Prisms Without Formula | Volume of Rectangular Prisms |
| Nets of 3D Figures | Surface Area of Rectangular Prisms |

In order to identify believable ground-truth skill pairs, a survey containing 24 skill pairs for which we had sufficient student data (greater than 50 student rows) was administered to 45 teachers and domain experts who use ASSISTments. Each was asked to rate on a scale of 1 to 7, indicating the perceived qualitative strength of the relationship of each skill pair. From the survey results, five skill pairs were selected to be the strongest related links with the smallest variance in opinion scores. As we are treating these links as truth, we wanted to be highly selective of these pairs.

The resulting dataset consists of 1838 total student rows from 896 unique students. This includes two rows of data per student for each of the five skill pairs included. The first row contains information of that student's performance on the pre- and post-requisite skills, while the second row contains student performance on the prerequisite and a simulated post-requisite described further in the next section.

[1]The full raw and filtered datasets are available at the following link: http://tiny.cc/veqg5x

For each student, a feature vector was selected using common performance metrics to compare within our model. This feature vector contained eight link-level features representing the interactions between student-level prerequisite and post-requisite performance metrics. The generated link-level features observed are as described below:

**Percent Correct**
The mean-centered[2] percentage of correct responses in the prerequisite skill multiplied by the mean-centered percentage of correct responses in the post-requisite skill.

**First Problem Correctness (FPC)**
The binary correctness of the first response in the prerequisite skill multiplied by the binary correctness of the first response on the post-requisite skill.

**Mastery Speed**
The mean-centered mastery speed of the prerequisite skill, defined as the number of problems required for each student to achieve three consecutive correct responses, multiplied by the mean-centered mastery speed of the post-requisite skill. In addition to centering, these values were also winsorized to make the largest possible value 10, chosen as this is often the maximum number of daily attempts allowed within ASSISTments. All centering and winsorizing occurred before multiplying the two values.

**Z-Scored Percent Correct**
The z-scored[3] value of mean-centered percentage of correct responses in the prerequisite skill multiplied by the z-scored value of mean-centered percentage of correct responses in the post-requisite skill.

**Binned Mastery Speed (Bin)**
The numbered bin of mastery speed as described in [3] of the prerequisite skill multiplied by the bin of mastery speed in the second skill. Students were placed into one of five bins based on mastery speed if the assignment was completed and based on percent correct if the assignment was not completed.

**Z-Scored Mastery Speed**
The z-scored value of mean-centered, winsorized mastery speed in the prerequisite skill, multiplied by the z-scored value of mean-centered, winsorized mastery speed in the post-requisite skill.

**Bin X FPC**
The binned mastery speed value in the prerequisite skill multiplied by the binary correctness of the first response in the post-requisite skill.

**Percent Correct X FPC**
The mean-centered percentage of correct responses in the prerequisite skill multiplied by the binary correctness of the first response in the post-requisite skill.

[2]All centering of features was performed at the skill-level.
[3]All z-scoring was performed at the class-level.

# 4. METHODOLOGY

The ultimate goal of this work is to provide the means of comparing models predicting the existence, or non-existence of skill relationships. Our approach to this is through the comparison and identification of features that most accurately predict these relationships. Using principal component analysis, we group similar features into more generalized conceptualizations to both compare which types of features matter when predicting relationships, but also to avoid problems of multicollinearity that may bias our estimates. Once this baseline model is established, we can construct new predictive models from the significant features and observe their accuracy in predicting the existence of skill relationships when compared to a simple null, or unconditional model.

In order to compare the usage of features against a weak or non-existent relationship, we simulated a new skill using students from the existing prerequisite skill by generating random sequences of responses. For each existing student, we randomly assign him/her a probability between 0.5 and 0.9 in order to create a random sequence of answers. For example, a student given a probability of 0.5 has a 50% chance of answering each given problem correctly. We simulate student answers until either mastery is achieved, defined as three sequentially correct responses, or the student reaches 10 problems without mastering; a value of 10 is chosen here, as many assignments in ASSISTments are given a daily limit of 10 problem attempts before asking the student to seek help or try again on another day. While we acknowledge there are many ways to accomplish this simulation step, we feel this simple method sufficiently creates a skill that has no relationship to the original prerequisite as intended. As our proposed method is intended to be used in the future to help identify undiscovered pre- or post-requisite links, we chose to use a simulated skill rather than a random existing skill to avoid the possibility of randomly selecting an undiscovered related skill. Again, we wanted to be highly selective and consider several such scenarios as we are attempting to create ground-truth values to which we can make our comparisons.

Using these two skill-pairs, one link representing a strong relationship while the other representing a non-existent relationship, we can calculate a feature vector for each student in the prerequisite skill with values from each skill-pair. We use a binary logistic regression with the existence of a relationship as the dependent variable and several link-level covariates to predict whether a skill relationship exists for each student row. The existence of a relationship can be determined then simply by majority ruling, but such calculation is not included in this work and instead observes accuracy at the student-level for a more accurate comparison.

We begin to compare commonly used student-level features in this study through two levels analysis. The first step attempts to compare groups of features, generalizing different representations of similar features into conceptual groupings. As such, we are able to view the predictive power of what we denote as initial performance, mastery, and correctness. The second experiment looks at the individual features as different representations of the overall group to compare

| | Component | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Percent Correct | | .821 | |
| First Problem Correctness (FPC) | | | .839 |
| Mastery Speed | .969 | | |
| Z-Scored Percent Correct | | .865 | |
| Binned Mastery Speed (Bin) | .972 | | |
| Z-Scored Mastery Speed | | | |
| Bin X FPC | | | .873 |
| Percent Correct X FPC | | .612 | |

**Figure 1: The results of the PCA analysis. All features except Z-Scored Mastery Speed mapped to one of three generalized components.**

these predictors at a closer level. We can take each factor of mastery, for example, and compare their usage in several models to determine which is the most accurate predictor of the existence of skill relationships.

## 4.1 Comparing Link-Level Features

In order to compare representations of student-level features, we must first be able to compare general conceptualizations of features to determine which provide more accurate predictions of the existence of skill relationships. We want to capture the true representations of each metric and attempt to interpret these generalizations as types of features. In order to accomplish this grouping of predictors, we use principal component analysis (PCA) to identify which student-level features correlate to and are representative of more generalized components. PCA is primarily used for dimensionality reduction as we are doing here and gives us the ability to create new variables from the component mappings. The resulting feature alignment can be seen in Figure 1. As is the case in our study, and was mentioned in the previous section, we have multiple metrics of mastery speed as well as several other features. As we can represent "mastery" in several ways, we want to know if the overall concept of mastery, as captured by the metrics used, is reliably predictive of the existence of skill relationships.

Creating a new set of predictors of these groupings, we are able to incorporate these into a binary logistic regression model to view the predictive power of each. While PCA groups similar features together based on their correlations, by viewing which features are grouped we are able to interpret and label each. From this process, we found that most of our features fell into three categories for which we have given the names "mastery," as this consists of representations of mastery speed, "correctness," as this consists of representations of the percentage of correct student responses, and "initial performance," as this consists of representations of student performance on the initial items of each skill. In addition to these three categories, we are also left with student mastery speed z-scored within student classes as a variable that did not fall under either of the three aforementioned categories; while a derivation of mastery speed, we believe that this did not correlate to the "mastery" category due to the method of standardization as it is capturing this metric in relation to students' peers. We will readdress this case in our section of discussion.

Once these predictors are identified and created, we construct a binary logistic regression model to predict, for each student row, whether a relationship exists or not. This model will give us a significance value and coefficient for each predictor in the model, as well as an overall predictive accuracy of the model which will be used more for the next analysis.

## 4.2 Comparing Feature Models

After being able to compare which generalized groups of features are significant predictors of the existence of skill relationships, we are able to compare the individual student-level features that fall into each category by incorporating them into separate models to observe predictive accuracy. The analysis of the first experiment is used to determine which categories are significant in predicting the existence of skill relationships. Using that information, we are able to focus on those groupings with significance to construct models that utilize factors from each grouping. The grouping of "mastery," for example contains the factors of mastery speed and binned mastery speed, so we can construct models using each to compare differences in predictive power. To avoid problems of collinearity, no single model contains more than one factor from a single grouping. This significantly reduces the number of combinations of features to test compared to running this experiment without first grouping like features and identifying those that are significant as we did in the first experiment.

Using the significant groupings, we are able to create 17 models consisting of single, pairs, and triplets of features. A logistic regression is run on each of these models to predict the existence of a skill relationship. Of the 17 models, 10 of them produce a statistically significant prediction when compared to a null model. Ideally, our null model should produce a 50% accuracy as there is an equal number of good and bad link rows in our dataset. This is not always the case, however, as depending on the feature observed, information may be missing for a particular student; mastery speed, for example, as the number of items attempted by a student before reaching 3 consecutive correct answers, would be missing for any student that did not complete the assignment. For this reason, the predictive power of each model is described as gains in predictive accuracy, or rather, the accuracy of each model minus the accuracy of the corresponding null model.

# 5. RESULTS

The results of the first analysis are expressed in Table 2. Each of the three feature groupings of Mastery, Correctness, and Initial Performance created using PCA in addition to the Z-Scored Mastery are compared within the same model, predicting the existence of a skill relationship. As these

**Table 2: The coefficients and significance values of the generalized components analyzed. From this we can focus on models that exclude features contained in the components with no significance.**

| Component | Coefficient Value (log-odds units) | Significance |
|---|---|---|
| Mastery | -.251 | <.001*** |
| Correctness | .015 | .802 |
| Initial Performance | .129 | .037* |
| Z-Scored Mastery Speed | -.129 | <.001*** |

again are link-level features describing interactions between student-level performance on prerequisite and post-requisite skills, it is difficult to draw tangible interpretations from the coefficient value, expressed in log-odds units. This coefficient, used in the logistic regression to make the predictions, describes each component's effect on the dependent variable. For example, for each unit increase in "Mastery," the probability that the link exists decreases. Again, as this component is an aggregation of interaction features, it is really describing an aggregation of differences of differences between student-level features making it difficult to make definitive claims regarding these values alone and were included purely to display a general trend of these components on the prediction.

From the table, we are able to determine the significance of each component on the overall prediction by viewing the corresponding p-values in the third column. Looking at these values, we can claim that the overall grouping of "Correctness" seems to have less of an impact on the predictive accuracy of the model. As this term is not significant, we can focus the remainder of our study on the remaining three components.

Table 3 illustrates the results of our second analysis comparing the models that we are able to construct with the remaining features once the "Correctness" grouping has been disregarded. This figure shows the comparative predictive accuracy of the 10 models that give statistically significant predictions as seen in Table 3. Again, these values are expressed as accuracy gains, or rather the percent accuracy increase over the null model run for each predictive model.

## 6. DISCUSSION

This work provides a baseline model of comparing student-level performance across skills to measure the strength of a skill relationship and compare the accuracy of both features and models that estimate this value. Such a model, in our experience, has not existed prior to this study. Our method attempts to identify not only the individual features that contribute to better predictions of these relationships, but also moves to generalize similar features into conceptualizations for comparison in order to minimize multicollinearity.

The principal component analysis step of our model found that all but one feature mapped to one of three components that we have interpreted as mastery, correctness, and initial performance. It was found the z-scored mastery speed, contrary to our intuition, did not map well to the grouping of mastery. We can speculate the reason for this occurrence by altering our interpretation of the feature. Mastery speed itself is an interesting metric as it attempts to capture two dimensions of performance: a level of understanding and a rate of learning. Also, to reiterate a prior distinction, these metrics are interactions of performance across skills. By z-scoring the metric, it is capturing a contextual effect of each student in comparison with other students in the class, a distinction that appears to have a significant effect.

Observing the resulting model components from the principal component analysis in Table 2, we were able to focus our attention to those components with significant values. Correctness was the only component of that model that was found to have no statistical significance on the dependent variable. This is certainly interesting, as percent correctness and other such measures are among the most common metrics of performance. Perhaps the interaction between pre- and post-requisite percent correct is losing some predictive power from when the metric is used for other predictions of performance.

This aspect illustrates one other important finding that the distinct representations of one metric or another each contribute differently to the predictive accuracy of the models studied. Models incorporating mastery speed, for example, had no significant accuracy gains over a null model, while mastery speed binning showed considerable gains as seen in Table 3. The baseline model of comparison proposed in this study provides the means to make that distinction regarding features contained within the same generalized component grouping. As is seen in that figure, combinations of features outperform any single feature, illustrating a more robust model by capturing multiple representations of performance.

## 7. FUTURE WORK

While we have shown that our model is able to compare and identify features that contribute to higher accuracy in predicting the existence of skill relationships, we also need to stress the importance of the usage of this information. The ability to compare features is only the first step of our model's goal. By identifying strong predictors of skill relationships that we know exist, we can apply it to other skills within ASSISTments and other systems to identify potentially new prerequisite arches, and also to better measure and predict long-term student performance, learning, and retention. Having an accurate estimate of skill relationships can help restructure prerequisite structures to provide skill sequences in an order that optimizes student learning and achievement.

The work in this paper incorporated several skills into a single dataset to make predictions. In this case, we wanted to create a method that is generalizable to some degree. While our selective skill set allows us to make some claims in terms of the accuracy these models over all skills, it may likely be the case that skill relationships are measurable in different ways for different skills. Further analysis could repeat the steps here on each one of the acquired skills in the dataset.

**Table 3: The models constructed from features in the significant generalized components. No one model contains more than a single feature from each generalized component.**

| Model | Null Accuracy | Model Accuracy | Accuracy Gain | Significance |
|---|---|---|---|---|
| Mastery Speed (MS) | 0.63 | 0.62 | 0.00 | 1.000 |
| Z-Scored Mastery Speed | 0.63 | 0.63 | 0.00 | 0.888 |
| First Problem Correctness (FPC) | 0.50 | 0.56 | 0.06 | <0.001*** |
| Binned MS | 0.50 | 0.69 | 0.19 | <0.001*** |
| Bin X FPC | 0.50 | 0.56 | 0.06 | <0.001*** |
| Bin, Z-Scored MS | 0.50 | 0.71 | 0.21 | <0.001*** |
| MS, FPC | 0.63 | 0.62 | 0.00 | 1.000 |
| MS, Bin X FPC | 0.63 | 0.62 | 0.00 | 1.000 |
| Bin, FPC | 0.50 | 0.69 | 0.19 | <0.001*** |
| Bin,Bin X FPC | 0.50 | 0.69 | 0.19 | <0.001*** |
| MS, FPC, Z-Scored MS | 0.63 | 0.63 | 0.00 | 0.754 |
| MS, Bin X FPC, Z-Scored MS | 0.63 | 0.63 | 0.00 | 0.979 |
| Bin, FPC, Z-Scored MS | 0.50 | 0.71 | 0.20 | <0.001*** |
| Bin, Bin X FPC, Z-Scored MS | 0.50 | 0.71 | 0.21 | <0.001*** |
| MS, Z-Scored MS | 0.63 | 0.63 | 0.00 | 0.843 |
| FPC, Z-Scored MS | 0.50 | 0.64 | 0.14 | <0.001*** |
| Bin X FPC, Z-Scored MS | 0.50 | 0.61 | 0.11 | <0.001*** |

While correctness was not significant in these results, perhaps it is significant when predicting certain types of skills. Perhaps, similar to our features, skills themselves could be generalized into conceptual types for different kinds of analysis pertaining to interactions of performance and their relationships.

The feature vectors generated for each student in our dataset captured many of the most common student-level metrics, but certainly not all of them. There are many other aspects that could be added including completion, measures of learning rate, time spent on the assignments, hint usage, and countless other variables. In addition, this study only observed interactions expressed as multiplications of these terms to describe them as link-level features. There are various other ways to represent interactions or other such transformations including differences of values, division of values, or just simply cross-feature interactions as was partially explored here by looking at Bin X FPC and Percent Correct X FPC. Such interactions model various other aspects of student performance and behavior that can be very useful in this type of relationship prediction.

The methodology presented observes models that predict the existence of skills as a binary outcome, while it can be modified to make comparisons on estimates of relationship strengths as a continuous outcome as well. The method observed model accuracy at the student level for better measurements, but it is a skill-level relationship that is being tested. One simple addition of future work could explore how to best combine the predictions at a student level to make a skill-level prediction. The methodology can then test relationships on the entire system skill structure.

# 8. ACKNOWLEDGMENTS

We acknowledge funding from the NSF (1316736, 1252297, 1109483, 1031398, 0742503, 1440753), the U.S. Dept. of Ed. GAAN (P200A120238), ONR's "STEM Grand Challenges," and IES (R305A120125, R305C100024).

# 9. REFERENCES

[1] S. Adjei, D. Selent, N. Heffernan, Z. Pardos, A. Broaddus, and N. Kingston. Refining learning maps with data fitting techniques: Searching for better fitting learning maps. In *Educational Data Mining*, 2014.

[2] T. Barnes. The q-matrix method: Mining student response data for knowledge. In *American Association for Artificial Intelligence 2005 Educational Data Mining Workshop*, 2005.

[3] A. Botelho, H. Wan, and N. Heffernan. The prediction of student first response using prerequisite skills. In *Learning at Scale*, 2015.

[4] E. Brunskill. Estimating prerequisite structure from noisy data. In *Educational Data Mining*, pages 217–222. Citeseer, 2011.

[5] Y. Chen, P.-H. Wuillemin, and J.-M. Labat. Discovering prerequisite structure of skills through probabilistic association rules mining. In *The 8th International Conference on Educational Data Mining*, pages 117–124, 2015.

[6] M. C. Desmarais, A. Maluf, and J. Liu. User-expertise modeling with empirically derived probabilistic implication networks. *User modeling and user-adapted interaction*, 5(3-4):283–315, 1995.

[7] P. I. Pavlik Jr, H. Cen, L. Wu, and K. R. Koedinger. Using item-type performance covariance to improve the skill model of an existing tutor. *Online Submission*, 2008.

[8] R. Scheines, E. Silver, and I. Goldin. Discovering prerequisite relationships among knowledge components. In *Educational Data Mining*, 2014.

[9] K. K. Tatsuoka. Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of educational measurement*, 20(4), 1983.

[10] H. Wan and J. B. Beck. Considering the influence of prerequisite performance on wheel spinning. In *Educational Data Mining*, 2015.