On the Dependence Structure between Learners' Response-time and Knowledge Mastery: If Not-linear, then What?

Zacharoula Papamitsiou*
Norwegian University of Science and
Technology
Trondheim, Norway
zacharoula.papamitsiou@ntnu.no

Kshitij Sharma*
Norwegian University of Science and
Technology
Trondheim, Norway
kshitij.sharma@ntnu.no

Michail N. Giannakos*
Norwegian University of Science and
Technology
Trondheim, Norway
michailg@ntnu.no

ABSTRACT

Popular approaches in learner modeling explore response-time as observational data supplemental to response correctness, to enrich the predictive models of learner knowledge. It has been argued that the relationship between response-time and knowledge mastery is non-linear. Determining the degree of association (dependence structure) between those two observations is an open question. To address this objective, we propose an approach based on copulas, i.e., a statistical tool suitable for capturing dependence structure between two variables. All of the information about the dependence structures can be estimated by copula models separately, allowing for the construction of more flexible joint distributions than existing multivariate distributions. This paper puts into practice a two-step pipeline for building the analytical models. Specifically, we propose a flexible copula-based approach that describes the dependence structure between students' response-time and mastery, in learning and testing contexts, and apply the methodology on four datasets. The two datasets are coming from Intelligent Tutoring Systems and are shared via an online repository, and the other two were collected during the validation of an (adaptive) assessment system. The results reveal five generic patterns of associations *across-datasets*, for various types of activities, domains and learner characteristics (i.e., not across-contexts). We elaborate on those findings and on the implications of our approach for adaptive systems.

KEYWORDS

Copula Theory; dependence structure; tail dependence; response-time; knowledge mastery; learner models

---

*All authors contributed equally to the paper

---

1 INTRODUCTION

Adaptivity and adaptive learning environments are in the epicentre of the technology enhanced learning research community. The 2019 NMC Horizon Report characterized adaptive learning as a "breakthrough teaching model of the future" that needs to be scaled to its potential [1]. The tailored support provided by those systems is usually delivered as adaptive course content or activities [46], as individualized (or group) recommendations [10, 36], as insightful analytics dashboards and open learner models [2, 37], or by adjusting the learning design to meet the learners' abilities [47, 55]. In all cases, in order to adaptively scaffold the learners throughout the learning process, it is a prerequisite for the learning systems to deeper "learn" and "understand" the learners [5], i.e., to make decisions by considering the learner models. Essentially, the learner models are *estimations* of learners' current states and *approximations* of next states, based on the available observational data from their activity and behaviour within a learning environment.

Learner modeling involves diverse aspects and decisions ranging from what to model (e.g., knowledge, behaviour, affects), to whether it concerns learning or assessment processes, to what data to include in the models, and to how exactly to model them [12, 16, 43].

Regarding modeling learner knowledge in particular, i.e., a latent construct that we try to quantify based on available observational data, the integration of *different types* of observational data is an open question, among others [43]. For instance, if the model is used for mastery detection, it is more important *what data* are used for modeling than the exact details of models, e.g., the impact of incorporating response-times to mastery criteria is higher than using a different model of learning [43]. Across numerous studies, researchers have explored the use of response-time (as a task attribute and an indicator of learners' engagement and mastery) as supplemental information to usual correctness of answer, targeting to enrich and improve their models [11, 26, 29, 44, 62]. In many cases, it was shown that incorporating students' response-time outperformed the original learner knowledge models [11, 29], while in others, the results could not demonstrate a clear trend between response-time and correctness or mastery [8, 62], or concluded that response-time does not always predict performance per step [26]. Overall, prior studies suggest that the relationship between those two variables is non-linear [3, 9, 11, 23], attempt to model it using parametric approaches [11, 29, 30, 56] or non-parametric approaches [4, 9, 25], and argue that using response-time as a predictor for success is a challenging open question [3, 9, 11].

A step towards that direction is to determine the nature of the relation between response-time and mastery, i.e., the dependence structure between those two variables. Learning more about the

shape of this relationship, allows one to study in more detail the impact that the time-spent on the response has on the response accuracy and performance, which, in turn, can be informative for the possible underlying processes. However, determining the dependence structure between response-time and correctness of responses is not a trivial task. That is because contemporary analytical modeling techniques (either parametric, e.g., quadratic, log-normal, polynomial, or non-parametric, e.g., Generalized Additive Models - GAM) have specific limitations. On the one hand, the common parametric approaches are not data-driven, having underlying assumptions on the distributions (e.g., log-normal assumes that $\ln(x)$ is normally distributed; polynomial has assumption on the degree), while the data-driven non-parametric ones have overfitting problems [33, 49], describing random error rather than a true underlying relationship. Although modeling approaches are powerful, they may not give a full and detailed view of the relationship and cannot always handle the full complexity of empirical relationships. This paper contributes to the field in two ways: (a) it proposes a flexible approach based on copulas, i.e., a parametric, yet data-driven approach with no assumptions about the distributions [17, 31]; and (b) it identifies five generic patterns of dependence structures by applying this methodology on four datasets, two from two Intelligent Tutoring Systems and two from an (adaptive) assessment system.

2 RELATED WORK

In this section, first we provide a generic overview of models for learners' knowledge mastery, next we explain the approaches on modeling the relation between response-times and mastery, and we end-up with the motivation of the research and research question.

2.1 Modeling learners' knowledge mastery: a brief overview and approaches

The basic goal of modeling learners' mastery in knowledge or ability (the two terms will be used interchangeably in this paper) is to estimate the current knowledge state and to predict future mastery based on past observational data. The literature distinguishes two contexts of learner knowledge modeling, i.e., the learning and the testing/assessment contexts. In the first case, the learner models are used to capture changes in knowledge and skills and predict the next learner states *per task*, where the tasks are directly relevant with each other (e.g., they are sub-components of the same skill). In the second case, where knowledge and skills are not expected to change during a test, the learner models are used to estimate the learners' *overall* ability by continually re-calibrating it based on learners' responses. As such, here we present related work on the approaches for both contexts in parallel (i.e., not across-contexts).

The prevailing approaches for modeling learner knowledge mastery can be distinguished in two generic families, i.e., knowledge tracing, and logistic models [43]. The knowledge tracing approaches – popular in learning contexts – either model learners' knowledge as a binary hidden state with a level of uncertainty attached to it (e.g., Bayesian Knowledge Tracing - BKT [13]) or model learners' knowledge with distributed continuous hidden states that are updated in non-linear, deterministic ways (e.g., Deep Knowledge Tracing - DKT [45]). The other family of methods for modeling learners' knowledge mastery, i.e., logistics models, typically include an item difficulty parameter and a logistic function for mapping

a difference between a skill and an item difficulty into the probability of a correct response. Such models are intensively used in Item Response Theory (IRT - [58]), popular approach in testing settings, whereas logistic models commonly employed in learning settings include the Performance Factor Analysis (PFA)[41], the Additive Factors Model (AFM) [7], and the Elo Rating System [42]. For comprehensive reviews of those methods, please see [15, 43].

2.2 Utilizing response-time to improve the models

In the "The measurement of intelligence", it is stated that "*other things being equal, if intellect A can do at each level the same number of tasks as intellect B, but in a less time, intellect A is better*" [54, p.33]. This statement synopsizes that in any ability or mastery measure, both the *result* of interacting with a task and *how long* it took to reach the result need to be considered, i.e., observed response and response-time jointly affect mastery. The idea of incorporating response-time as supplemental observational information to improve the accuracy of the learner models is not new. A wide range of approaches have been suggested to conceptualize and model the latent structure underlying mastery (or ability) and response-times both in the field of Intelligent Tutoring Systems (ITS) and learning and in the field of Psychometrics and testing (e.g., [4, 11, 26, 57]).

In learning settings (e.g., ITS), most of the approaches combine response accuracy and response-time into a single knowledge mastery measure, aiming to improve the accuracy of the learner models (e.g., [26, 48, 60]), without, exploring their between dependence structure. For instance, BKT was extended by adding a binary variable ("quick/slow") to describe the learner's response-time, depending on whether it is below or above the median response-time [26]. It was found that response-time can be a good predictor of post-test scores, but it does not always predict mastery in individual steps. Instead of binary values, first response-time data were discretized into four categories to predict students' correctness of the next response [60]. The categories were defined based on assumptions on what the response-times might indicate with respect to knowledge mastery; next, the authors computed the percentage of students who answered the current question correct, when the previous response fell into the corresponding category, and utilized a linear regression model to extent the existing knowledge tracing model.

The early attempts to study the relation between response-time and knowledge mastery in learning settings, utilized students' previous response-times for directly predicting the correctness of their next actions. For instance, the z-scored response-time was used to identify where the predictive power of response-time may come from and where the variance lies, but no clear trend between response-time and correctness was identified [62]. The Elo function was also used to estimate ability, utilizing response accuracy and response-times data based on the *High Speed High Stakes* rule, i.e., for both correct and incorrect answers, increasing response-time results in the score going towards zero [23]. The idea was grounded on the hierarchical framework from educational measurement [57] (discussed later here). In another approach, the authors' hypothesis that rapid responses on easy items may imply fluent knowledge was implemented by utilizing a log function of response-times for correctly answered items (response-times for wrong answers were not considered meaningful) [29]. In a recent approach, the use of

response-time as a predictor of student mastery was explored in two ways: by modeling response-time as a linear parameter and by modeling response-time as a quadratic parameter [11]. The authors found that including response-time as a quadratic parameter improves the model's performance both with respect to goodness-of-fit and with respect to predictive accuracy on unseen steps.

In educational measurement (i.e., testing), one of the most popular approaches for the joint modeling of item response-accuracy and response-time is the hierarchical framework [57]. The assumption in this model is that response-times follow a log-normal distribution, and as such, they are incorporated as a log-normal transformation of the original values. A major assumption of the model is the conditional independence of response accuracy and response-time, given the overall speed and ability. Although the correlation between response-time and ability can be interpreted *within* a test-taker (or item), however, the assumption of conditional independence might be violated in the correlation between responses and response-times *across* test takers (or items). To address this issue, it has been proposed to explicitly model the residual dependence between time and accuracy by incorporating the effects of the residual response-time [3]. Another, purely data-driven approach explored the residual relationship of response-time and accuracy using plots of residual log-transformed response-times against proportions of correct responses [9]. The exploratory findings from this study revealed a curvilinear dependence, further confirmed in a fine-grained approach [4], and that there is a very stable relationship between response-time and accuracy when the residual dependencies are considered. However, this approach appears to have overfitting problems, describing random error rather than a true underlying relationship. Furthermore, the residual dependency between accuracy and response-time has been found to be negative for most items in both approaches [4, 9]. Yet, the rapid guessing mixture model cannot explain these results because it implies a positive dependency (slower responses are more correct) [61]. Moreover, a generalized linear modeling framework has been proposed to model differences in different forms of cross-relations linking separate measurement models for item responses and response times [30]. However, this approach is a parametric approach in the sense that it requires an assumption about the distribution of the response-times.

Overall, previous works converge on the fact that the dependence structure between response-time and knowledge mastery needs to be determined, since it can be utilized for predicting one variable from the other. The literature strongly indicated that this relation is non-linear [4, 9, 11]. Yet, the popular approaches in literature either were non data-driven, employing analytical methods with inherent assumptions on the distributions (e.g., parametric models), or were suffering from overfitting issues (e.g., non-parametric models).

2.3 Motivation of the Research and Research Objective

The current research clearly demonstrates that response-times are a useful source of information. Learning more about the shape of the dependence structure between response-time and knowledge mastery, allows one to study in depth the impact of response-time on the response accuracy (correctness), and the possible underlying processes. Determining the nature of this relationship and precisely *how* this is to be exploited, awaits further research. In line with

the above, the objective of this study is to explore the degree of association between those two observations, in learning and in assessment settings. Thus, the research question is as follows:

Research Question (RQ): *Of what kind is the relationship between learners' response-time and knowledge mastery?*

We propose an approach based on copulas, i.e., a statistical tool suitable for capturing dependence structure between two variables. All of the information about the dependence structures can be estimated by copula models separately, allowing for the construction of more flexible joint distributions than existing multivariate distributions. This paper puts into practice a two-step pipeline for building the analytical models, and applies the methodology on four datasets: the two datasets are coming from an Intelligent Tutoring System and are shared via an online repository, and the other two were collected during the validation of an adaptive assessment system.

3 THE COPULA-BASED APPROACH

3.1 Copulas: definitions and families

The dependence between two random variables X_1 and X_2 is contained in their joint distribution, and its study needs to go beyond common measures of linear dependence [18]. For instance, one drawback of the Pearson's correlation is that it depends on the marginal distributions F_1 and F_2 of the two random variables X_1 and X_2. This explains why it is not invariant on monotonic increasing transformations of X_1 and X_2, an undesirable property for a dependence measure [34]. Copulas model the dependence structure independently of the marginal distributions; from a distributional viewpoint, copulas can be utilized to identify a dependence structure capturing scale invariant dependencies not affected by the marginal behaviours of the considered variables. The idea behind the concept of copulas is to separate a joint distribution function into the dependence structure and the marginal behaviour [31].

Decomposing the multivariate distribution into marginal distributions and a copula, allows for the construction of better models of the individual variables than if only explicit multivariate distributions were considered. Copulas allow the construction of joint distributions with arbitrary margins and address the dependence structure of multidimensional random variables; they are functions that "couple" the marginal distributions to the corresponding joint distribution functions [31], i.e., a copula is a multivariate distribution function $C : [0, 1]^d \rightarrow [0, 1]$ with standard uniform margins [22, 31]. Any multivariate distribution function can serve as a copula. The theoretical foundation for the application of copulas is provided in Sklar's theorem [51]. In brief, Sklar's theorem claims that any multivariate joint distribution can be written in terms of univariate marginal distribution functions and a copula function that describes the dependence structure between the variables. In this paper, considering only two variables, i.e., response-time and knowledge mastery, we restrict ourselves to the bivariate case ($d = 2$). In this case, Sklar's theorem is written as follows:

Sklar's Theorem: Let F be a joint distribution function with margins F_1, F_2. There exists a copula such that for all x_1, x_2 in $[-\infty, \infty]$

$$F(x_1, x_2) = C\{F_1(x_1), F_2(x_2)\}$$

If the margins are continuous, then C is unique. Conversely, if C is a copula and F_1, F_2 are univariate distribution functions, then F defined above is a bivariate distribution function with margins F_1, F_2.

Most of the copulas are parametric and there are basically two types; the copulas of normal mixture distributions and the Archimedean copulas. The first type refers to implicit copulas (elliptical), such as the Gaussian copula and the Student's t copula (for simplicity, t copula), which are implicit in well-known multivariate distributions; the Gaussian copula supposes (X_1, X_2) being the standard bivariate normal with correlation matrix Σ and is defined as:

$$C_\Sigma^{\text{Ga}}(u_1, u_2) = \{X_1 \leq \Phi^{-1}(u_1), X_2 \leq \Phi^{-1}(u_2)\}. \tag{1}$$

The t copula, with ν degrees of freedom, is implicitly built as:

$$C_{\Sigma,\nu}^{\text{t}}(u_1, u_2) = \{X_1 \leq t_\nu^{-1}(u_1), X_2 \leq t_\nu^{-1}(u_2)\}. \tag{2}$$

In (1) and (2), the $\Phi^{-1}(\cdot)$ and $t_\nu^{-1}(\cdot)$ are the inverse of the standard univariate Gaussian distribution function and the standard univariate student-t-distribution with ν degrees of freedom, respectively.

Another type of parametric copulas is the Archimedean family, in which copulas have simple closed forms. The Gumbel copula is

$$C_\beta^{Gu}(u_1, u_2) = \exp\left[-\left\{(-\log u_1)^\beta + (-\log u_2)^\beta\right\}^{1/\beta}\right], \tag{3}$$

with $\beta \geq 1$. When $\beta = 1$, we get independence and $\beta \to \infty$ provides perfect positive dependence.

Another Archimedean copula is the Clayton copula

$$C_\beta^{Cl}(u_1, u_2) = \left(u_1^{-\beta} + u_2^{-\beta} - 1\right)^{-1/\beta}, \tag{4}$$

with $\beta > 0$. When $\beta \to 0$, it tends to independence, whereas $\beta \to \infty$ gives perfect positive dependence. Other Archimedean copulas and other families of copulas exist as in [28] and the references therein. This study, the selection of the specific copulas was grounded on the following facets: (a) they are the most commonly met in datasets, and (b) they are found to be the most robust for modeling both the positive and negative dependencies in bivariate distributions [6, 40]. More details on Copula theory can be found in [20, 32, 52].

3.2 Copula-based dependence measures

If the joint distribution of the random variables is well represented by an elliptical distribution, e.g., Gaussian or t, their dependence structure is linear, and Pearson's correlation coefficient is a meaningful measure of dependence. However, as it has been argued, between response-time and knowledge mastery there is a non-linear dependence. Outside the world of elliptical distributions, the use of the linear correlation coefficient may induce misleading conclusions about the dependence. There exist standard and useful dependence measures solely related to the copula and not to the margins. These are, for instance, rank correlations like Kendall's tau $\tau(X_1, X_2)$ and Spearman's rho $\rho_S(X_1, X_2)$ or tail dependence coefficients. The tail dependence coefficients are measures of extremal dependence that quantify the dependence in the upper and lower tails of a bivariate distribution of two random variables X_1 and X_2 with continuous marginal distributions F_1 and F_2. The coefficients are defined in terms of quantile exceedences, and when the limit exists, the coefficient of upper tail dependence is

$$\lambda_u(X_1, X_2) = \lim_{q \to 1}\{X_2 > F_2^{-1}(q) \mid X_1 > F_1^{-1}(q)\}. \tag{5}$$

Analogously, the coefficient of lower tail dependence is

$$\lambda_\ell(X_1, X_2) = \lim_{q \to 0}\{X_2 \leq F_2^{-1}(q) \mid X_1 \leq F_1^{-1}(q)\}. \tag{6}$$

Convenient mappings between such measures and the parameters of common copulas often exist, using for instance Kendall's tau and Spearman's rho. In that terms, the tail-dependence coefficients (5) and (6) are conveniently functions of the copula. More precisely,

$$\lambda_u = \lim_{q \to 1} \frac{1 - 2q + C(q, q)}{1 - q},$$

$$\lambda_\ell = \lim_{q \to 0} \frac{C(q, q)}{q}.$$

Therefore, $\lambda_u \in [0, 1]$ and $\lambda_\ell \in [0, 1]$. When $\lambda_u > 0$, we get upper tail dependence. If $\lambda_u = 0$, we get asymptotic independence in the upper tail. When $\lambda_\ell > 0$, we reach lower tail dependence and if $\lambda_\ell = 0$ we obtain asymptotic independence in the lower tail. For elliptical copulas, $\lambda_u = \lambda_\ell$ and the Gaussian copula is asymptotically independent for $|\rho| < 1$, so that for the Gaussian copula $\lambda_u = \lambda_\ell = 0$. When the copula shows tail dependence (lower or upper or both), the tail dependence can be explicitly mapped to the copula parameters (β). Let us denote the mapping

$$\lambda = g(\beta). \tag{7}$$

The t copula is tail dependent when $\rho > -1$ and the mapping is

$$\lambda = 2\bar{t}_{\nu+1}\left(\sqrt{\nu + 1}\sqrt{1 - \rho}/\sqrt{1 + \rho}\right), \tag{8}$$

where $\bar{t}$ denotes the survival of the t distribution and ν its degrees of freedom. The Gumbel copula is upper tail dependent for $\beta > 1$ and the Clayton is lower tail dependent for $\beta > 0$, with mappings:

$$\lambda_u = 2 - 2^{1/\beta}. \tag{9}$$

and

$$\lambda_\ell = 2^{-1/\beta}. \tag{10}$$

The Gumbel and Clayton copulas can only capture one side of the tail dependence, and cannot display negative tail dependence; once the dependence between the variables is negative, these copulas would not fit. Therefore, these copulas can be "rotated" (90 degrees, 180 degrees [i.e., survival], 270 degrees) and applied again (Fig. 1).
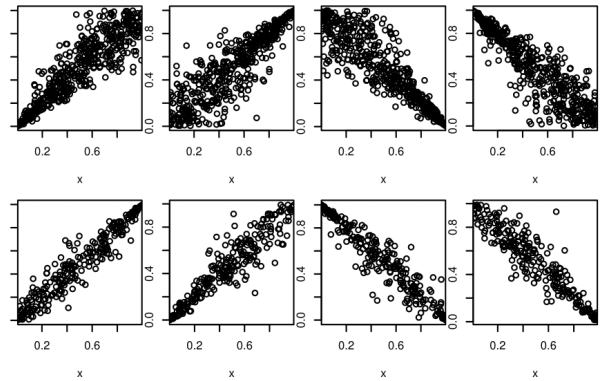

Figure 1: (from left to right) Top: Clayton – normal, survival, rotated 90 degrees, rotated 270 degrees. Bottom: Gumbel – normal, survival, rotated 90 degrees, rotated 270 degrees.

Although some of the pairs across the two generic families in Figure 1 look similar, Clayton family is still characterised by λ_l (10) and Gumbel family is still characterised by λ_u (9). The parameters of the rotated copulas correspond to Kendall's tau τ value 0.5 for positive dependence and -0.5 for negative dependence.

In the next section, we put into practice the copula-based approach for determining the dependence structure (i.e., what kind is the relationship) between response-time and knowledge mastery, in learning and assessment (testing) settings. The approach presented here is not an alternative for modeling, i.e., the aim is not methodological. Instead, we utilize a purely data-driven method, as a means to empirically explore the relationship between these two observations in an exploratory way that can lead to new models. We use this method to address the open problem: to both characterise the nature of the relation between these two variables (i.e., the copula) and to measure its strength (i.e., the correlation coefficient).

4 EMPIRICAL STUDY

In this section, first we describe the details of the datasets and the pre-processing applied on them to prepare them for analysis, and next we explain the empirical study setup, i.e., the copula-based pipeline, along with the measures computed at each step.

4.1 Datasets

To illustrate how the dependence between response-time and knowledge mastery can be investigated using the copula-based approach, we applied the method to four datasets, two from Intelligent Tutoring Systems (referred to as "learning data" from this point on), and two from an adaptive assessment system (referred to as "assessment data" from this point on). The four datasets contain learners' interaction logs in various types of activities and disciplines.

The learning data are shared via an online repository [24]. In particular, we used the following datasets:

- "*Algebra 2005-2006*" (or "Algebra") collected with the Carnegie Learning Algebra System; this dataset has been used in several educational data mining analyses/student modeling studies (e.g., [14, 27]) and contained 813,661 steps of 575 students;
- '*Bridge to Algebra 2008-2009*" (or "Bridge") collected with the Carnegie Bridge to Algebra system; this dataset was also used in previous student modeling studies (e.g., [53, 63]) and had 20,012,499 steps of 6,043 students.

From these datasets, we only used steps that had information about the learners' response-times (duration) and were identified as correct or incorrect (i.e., we excluded hints and unidentified steps).

The assessment data were collected with a widely used assessment system that supports both fixed and adaptive assessment procedures, as multiple-choice quizzes, for different disciplines and courses [35]. Specifically, here, we used the following datasets:

- "*Computers II 2014*" (or "Computers II"); this dataset contained 2,397 steps of 259 students;
- '*Microeconomics II 2016*" (or "Micro II"); this dataset had 7,654 steps of 452 students.

The aforementioned datasets have been used in former studies (e.g., [38, 39]). Here, to identify the kind of the relationship between response-time and mastery, we used the correctness of response (knowledge mastery) and response-times from all the participants on the those datasets. Note that, in this study, we put in practice the copula-based approach *across-datasets*, i.e., not across-contexts.

4.2 Study setup

4.2.1 Pre-processing: We initially plotted all original datasets to depict the obtained response-times (y-axis) in relation to the respective knowledge mastery (x-axis), in order to explore the data and gain an intuitive insight about the observations, their distribution, and the existence of extreme values (outliers). In this study, in the pre-processing of the datasets (explained above), we did not consider the difficulty of the tasks/items in terms of a baseline response-time they might required on average to read and understand. As a result, the observed response-times and mastery may have been induced by the students' individual differences or by the tasks'/items' difficulty, or possibly by both. To address this issue, for each dataset, we divided the student populations into smaller datasets of sub-populations (i.e., "*blocks*"), in order to "dive-into" the data-points and explore the bivariate dependence structure for each sub-population separately. In this study, we sliced the data using the 25 and 10 percentiles of both the response-time and knowledge mastery, but any other fine-grained splitting can be utilized. Note that, in the joint model, response-time is treated as a continuous variable, so as there is no loss of information in the measurement.

4.2.2 Two-steps pipeline and measures: Modeling the bivariate dependence involves quantifying dependence and correlation. For the first one, the copula family needs to be decided upon a goodness-of-fit criterion (e.g., Akaike information criterion - AIC), and for the second one, the tail-dependence coefficients need to be computed to show the strength of the relationship. Accordingly, for each block, we fitted the ten *theoretical* models (i.e., four Clayton, four Gumbel, Gaussian, and t - explained in Section 3.1), and concluded to the *empirical* models, choosing the ones with the lowest AIC, in a purely data-driven fashion. Furthermore, the strength of the copula between the two variables is measured by the respective tail-dependence coefficient, i.e., λ_u for Clayton types, λ_l for Gumbel types, Kendall's τ for Gaussian, and ρ_S for t.

The bivariate dependence between the variables of interest (i.e., response-time and knowledge mastery) as a function that best fit the original data, and the observed data, were also graphically displayed using Contour plots, for the copulas detected in the datasets. These plots show explicitly how the copula, as a function of the two variables, fits the original data, with respect to tail dependence, in a bivariate "loss function" manner, for each sub-population (block).

It should be noted that, for the analysis, none of the response-time outliers were removed. In the copula-based models, the tail dependence describes the concordance between extreme values of random variables X_1 and X_2, and the tail-dependence coefficient is a measure of the extremal dependence between the variables [19].

5 RESULTS

5.1 Plotting of the datasets

The results from the exploratory scatter plots of the datasets are illustrated in Figures 2 and 3 for the learning data and the assessment data respectively. These results show that, in each context (i.e., pairs of plots), the structures of the joint distributions look different, but they share similar characteristics in the shape, with the existence of outliers/extremes in all cases. As such, it is expected that *generic patterns of the dependence structure between response-time and mastery exist*, and need to be identified, isolated and explained.

5.2 Dependence – Selection of copulas

To better and deeper explore the nature of the underlying relationship between response-time and mastery, we performed an analysis
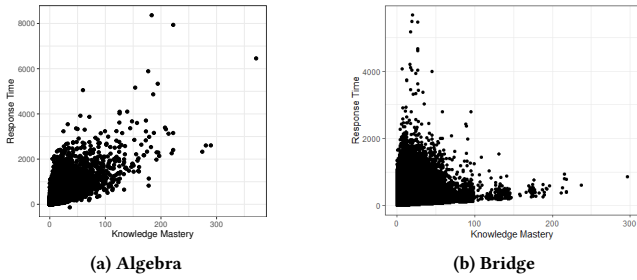
(a) Algebra

(b) Bridge

Figure 2: Plots of the original datasets for learning data

(a) Computers II

(b) Micro II
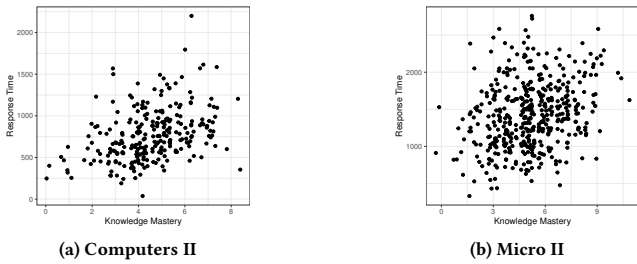
Figure 3: Plots of the original datasets for assessment data.

(a) 25 percentile

(b) 10 percentile

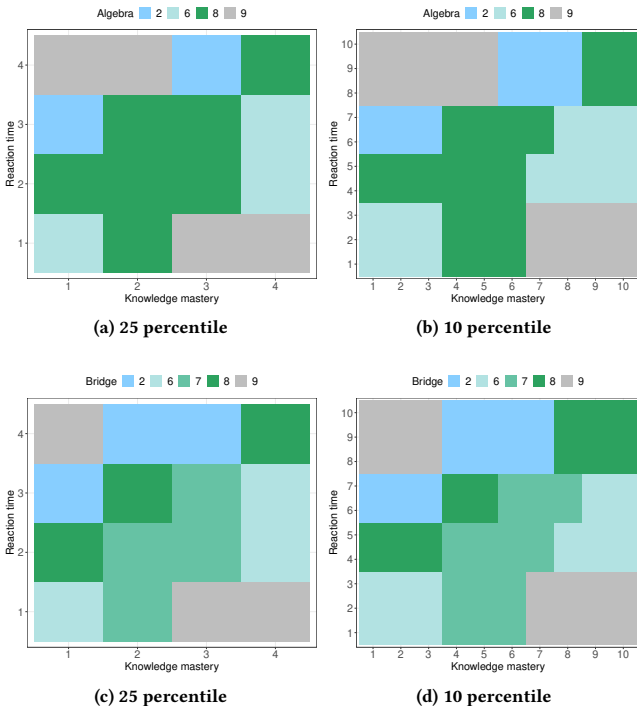(c) 25 percentile

(d) 10 percentile

Figure 4: Block-based Copula families for the learning data. In all figures: 2–survival Clayton; 6–survival Gumbel; 7–rotated Gumbel (90); 8–rotated Gumbel (270); 9–Gaussian.

based on 25-percentiles of the population and a more granular analysis based on 10-percentiles. We investigated the potential to fit the *theoretical* copula models on each of the sub-populations, and studied each block separately, as explained in the previous section.

Figures 4 and 5 depict the fitting result for all datasets. For concluding to the empirical (i.e., data-driven) copula family for each block, we employed the lowest AIC value. In the figures, each color corresponds to one of the detected copula families.

(a) 25 percentile

(b) 10 percentile

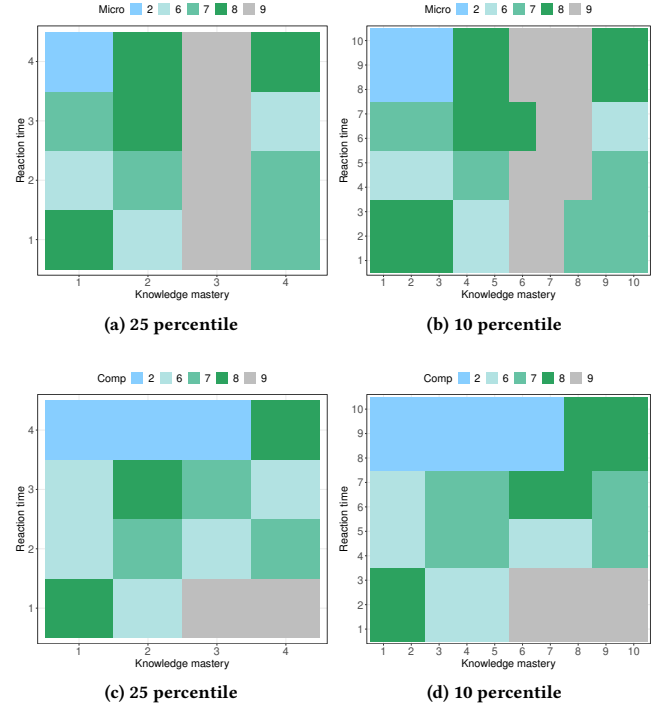(c) 25 percentile

(d) 10 percentile

Figure 5: Block-based Copula families for the assessment data. In all figures: 2–survival Clayton; 6–survival Gumbel; 8–rotated Gumbel (270); 9–Gaussian.

Overall, out of the ten (10) theoretical copulas, five (5) of them were empirically detected across datasets. From the elliptical copulas (that imply a linear dependence structure), the t is not detected in the data and the Gaussian is detected in specific blocks (i.e., high response-time – low knowledge mastery; low response-time – high knowledge mastery). Hence, the nature of the underlying relation between the variables of interest is non-linear, but some linearity is detected, as blocks in specific percentiles follow linear behaviour.

Furthermore, as it can be observed from the figures, the empirical copulas follow generic patterns for each context separately. For the learning data, the overlap of the detected copulas across datasets is 93.8% for the 25 percentile and 91% for the 10 percentile, when we only consider the generic copula family, and 68.8% for the 25 percentile and 68% for the 10 percentile, when we take into account both the direction of the dependence and copula family. The respective overlap for the assessment data is 62.5% for the 25 percentile and 49% for the 10 percentile, despite the dependence direction. This finding indicates that the *generic patterns of associations across-datasets exist*, confirming the initial exploratory insights from the scatter plots regarding the shared similarities of the distributions.

In the rest of the paper, we explain the results for the 25 percentiles, yet, the results are similar for the 10 percentiles. The method can be scaled-up for finer-grained splitting, if the sub-populations (blocks) sizes allow for the copula estimation.

5.3 Strength of relations

Table 1 summarizes the results for the tail-dependence coefficients for each block for all datasets in the 25 percentile. In addition, in this table, the number of students in each block is included to showcase that even for smaller partitions of learners, the copula-based method can map a bivatiate function for the joint dependence between response-time and knowledge mastery. It can be observed that for most of the learning data, the detected copula reveals a strong or moderate dependence structure, i.e., the tail-dependence coefficient is above 0.5; for the assessment data, most of the dependencies are moderate. An advantage of the method is that it can detect the dependence structure even for extreme values in the dataset.

Table 1: The tail-dependence coefficients and the number of students in each sub-population (block) in the 25 percentile.

| Block[1] | Tail-dependence Coefficient (# of students) | | | |
	Algebra	Bridge	Computers II	Micro II
1-1	0.97 (110)	0.45 (1320)	0.81 (54)	0.78 (53)
1-2	0.96 (24)	0.44 (170)	0.45 (30)	0.81 (44)
1-3	0.74 (6)	0.37 (11)	0.48 (22)	0.17 (29)
1-4	0.21 (4)	0.15 (8)	0.30 (20)	0.28 (28)
2-1	0.89 (24)	0.55 (176)	0.49 (33)	0.96 (51)
2-2	0.71 (80)	0.49 (1050)	0.03 (34)	0.13 (46)
2-3	0.72 (30)	0.46 (270)	0.07 (35)	0.46 (38)
2-4	0.17 (6)	0.47 (9)	0.32 (36)	0.41 (42)
3-1	0.43 (6)	0.09 (5)	0.23 (8)	0.03 (48)
3-2	0.77 (34)	0.47 (277)	0.10 (20)	0.01 (45)
3-3	0.52 (94)	0.44 (981)	0.68 (23)	0.18 (50)
3-4	0.70 (15)	0.45 (250)	0.09 (21)	0.07 (47)
4-1	0.32 (7)	0.17 (4)	0.30 (12)	0.21 (26)
4-2	0.43 (8)	0.44 (5)	0.13 (36)	0.12 (45)
4-3	0.78 (10)	0.41 (260)	0.49 (44)	0.11 (56)
4-4	0.77 (120)	0.63 (1251)	0.15 (46)	0.40 (56)

[1] The blocks correspond to the respective ones in Figures 4 and 5, from bottom-up and left-to-right, e.g., Block 1-1 corresponds to "*Low response-time – Low knowledge mastery*" and Block 4-4 corresponds to "*High response-time – High knowledge mastery*". The in-between states are marked as "*Moderate High/Moderate Low*" both for response-time and mastery.

5.4 Goodness-of-fit

If no prior knowledge about the dependency structure, e.g., prevalence of asymptotic or tail dependencies, is available, it is common to employ goodness-of-fit tests or measures of tail dependency in order to choose an appropriate model. Fitting the detected copulas to the datasets can be visualized using contour plots. For instance, Figure 6 illustrates how the detected copulas change as a function of the response-times and knowledge mastery, for the different partitions of the populations (blocks) in which they are detected, in the Algebra dataset. The contour plots show how well the copulas fit the existing data-points. In the figures, the inner rounded curves indicate stronger relation between response-time and mastery captured by the copula. When the contours become very tall and narrow, only the edges are visualized across the plotted range, i.e., not the rounded ends. In Figure 6a, the Gaussian copula has been detected in the block 4-1 (i.e., high response-time – low knowledge mastery), and the respective contours visualize the underlying structure and the fitting on the data points (i.e., the seven students). The moderate strength of the detected dependence in this block – based on the tail-dependence coefficient reported in Table 1 – is

also shown in the contour plot, as a second inner contour is also detected (upper right corner). Similarly, the strong dependence in block 4-4, in which the Rotated Gumbel copula has been detected, is shown in Figure 6d. The results for all other datasets are analogous.
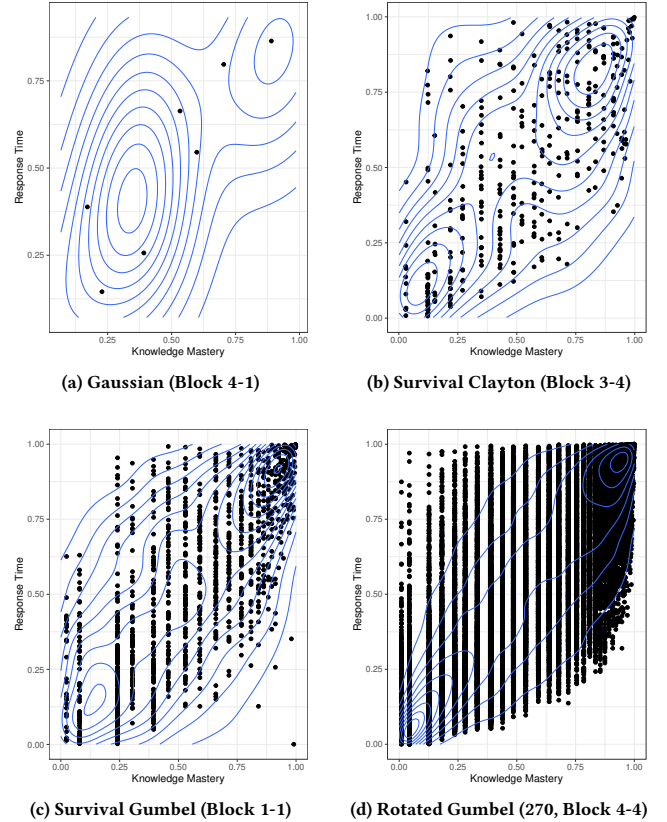

(a) Gaussian (Block 4-1)

(b) Survival Clayton (Block 3-4)

(c) Survival Gumbel (Block 1-1)

(d) Rotated Gumbel (270, Block 4-4)

Figure 6: Contour plots for the learning data (Algebra dataset) with the respective copulas.

6 DISCUSSION, IMPLICATIONS AND CONCLUSIONS

Current research on the improvement of learners' knowledge mastery models focuses on incorporating in the models supplemental observational data from learners' activity, beyond response-accuracy or correctness alone. The most commonly used, yet promising data is learners' response-times. The topic is of high interest for both the Intelligent Tutoring Systems (i.e., learning) and the Psychometrics (i.e., testing) communities. Previous research attempted to shed light to the underlying relationship between response-time and knowledge mastery [9, 11, 26, 56]. Determining the shape of the dependence structure between those two variables, is expected to allow one to study in depth the impact that response-time has on response correctness and knowledge mastery; in turn, this can be informative for the possible underlying processes and for refining the learner models and guiding the design of adaptive systems.

Previous studies have shown that this relation is non-linear [4, 9, 11, 23]. An in-depth investigation of this relationship is still missing, as *discovering* this relationship is not a trivial task *per se* – contemporary analytical modeling approaches are either based on

arbitrary assumptions about the distributions, or they may describe the error instead of the true relationship. In this study, we proposed an empirical, purely data-driven approach based on copulas. To determining the kind of the relationship between response-time and knowledge mastery, the idea is to separate the joint distribution into the dependence structure and the marginal behaviour [31]; here, we characterise both the nature of their dependence (i.e., the copula family) and measure its strength (i.e., the tail-dependence coefficient). In other words, this information reveals both the pattern of the underlying relationship and its strength for each data-point in a dataset, and as such, it enables the development of new models, and allows for revisiting the adaptation decision for each data-point accordingly. We empirically explored the relationship between these two observations using datasets from diverse disciplines, contexts and learners. The copulas selected for this study have been found to be the most common and robust for modeling both the positive and negative dependencies in bivariate distributions [6].

There are numerous interesting observations one can make from the findings of the empirical study. Firstly, figures 4 and 5 show that five generic patterns of dependence structure were detected across-datasets (i.e., not across-contexts). In each context, the overlap of the detected empirical patterns between the different datasets was high (up-to 68.8% for the learning data when considering the sign of the dependence, and 93.8% when considering only the copula families; and up-to 62.5% for the assessment data, regardless of the directionality of dependence). Interestingly, although the approach is purely data-driven, it does not overfit the data, and is not affected by the differences in the margins, in contrast to [4]. Furthermore, both Clayton and Gumbel are for the one side of the tail-dependence – in a way they model either upper or lower tail dependence. In the terms of knowledge mastery and response-time, those copulas show that the co-movement of these variables in the tails of the distribution follows specific patterns, and the above-mentioned copulas are utilized to model those patterns. The amount of time used by the learners to respond to a task, and their knowledge mastery is reflected in the patterns: according to the pattern detected, the learners are located in one of the blocks (see Figures 4 and 5) that corresponds to that pattern, and we can learn the mapping function between their response-time and mastery. Learning this function, allows us to drive adaptation, e.g., to select the task that has the greatest expected reduction in entropy, i.e., that better fits the learner's mastery class.

In addition, the detection of the Gaussian copula in blocks with "high response-time – low knowledge mastery" or "low response-time – high knowledge mastery" is in line with the *High Speed High Stakes* rule used in [23], and based on the hierarchical framework that is extensively used in testing settings [56]. This is an intriguing finding because, in our study, the Gaussian copula was detected mostly in the learning data (see Figure 4). Although the framework has recently received some criticism [3], the results indicate that for some learners' sub-populations, it can model well the relationship between response-times and knowledge mastery.

Furthermore, the tail-dependence coefficients (Table 1) revealed the nonlinear nature of the underlying relation between the variables of interest. Indeed, the nonlinearity is further confirmed both in the learning and in the assessment context, although the two

contexts are not compared with each other. For most of the sub-populations in all datasets, the copulas that describe the dependence structure do not imply linear structures. In other words, the nonlinear trend is met across datasets, disciplines, learners, and contexts.

The approach allows for nonlinear dependence detection and characterization, the identification of the exact underlying patterns of dependence, and captures both positive and negative dependencies. In a sense, the rotation of copulas that do not display negative dependencies, can be utilized to explain rapid guessing using the model suggested in [61]. This model (i.e., from [61]) implies only positive dependencies, and could not be used to explain results from previous approaches that detected negative relations (e.g., [4, 9]). The copula-based approach goes a step beyond this limitation, using the rotation of the copulas to address directionality issues.

6.1 Implications

Previous studies have shown how to model the dependence structure between two random variables as a function of an arbitrary set of exogenous predictors (covariates) [50, 59]. Specifically, those studies built-upon the GAM [21] that link the mean behaviour of a random variable X with a set of covariates W. GAM provides a flexible model for a univariate response. In a bivariate context, if we are interested in the dependence between X_1 and X_2 that both can be individually explained by covariates of W, it is likely that their dependence structure also depends on some of the covariates in W. This problem has been addressed in [59], and GAMs have been extend to a tail dependence structure between random variables in [50]. Practically, a GAM for tail dependence coefficients in the presence of covariates can be utilized to explore how the joint response-time and knowledge mastery depends on covariates. Considering other meaningful attributes of the learners' (e.g., self-efficacy) or the tasks' (e.g., difficulty) as covariates has the potential to explain the detected dependence structures, and justify the relationship between response-times and knowledge mastery.

Furthermore, one of the advantages of the method is that it can handle and detect the dependence structure even for extreme values in the datasets. Considering the contemporary massive open online courses, in which the populations at the tails of the distributions might be thousands, the copula-based approach can be scaled-up and allows for detecting the nature of association of response-time and knowledge mastery, for those students, and opens the path for designing adaptive systems for such environments.

6.2 Limitations and Future work

The strength of the relationships was low in some cases – mostly in the assessment data (Table 1). This result might have been due to not considering other characteristics of the tasks (e.g., difficulty), that can catalyze both learners' response-times and accuracy. For instance, we did not examine random guessing or "gaming the system" behaviours; this is within our future work plans.

Another limitation of this study is that the assessment datasets were relatively small, and some of the dependencies were weak. Bigger datasets might had provided more granular information.

Exploiting the analytical models and the knowledge gained about the dependence structure between response-time and knowledge mastery is expected to also improve the predictive models. The next step is to incorporate this information in the learner models to drive adaptation, and to evaluate the accuracy of the new models.

REFERENCES

[1] Bryan Alexander, Kevin Ashford-Rowe, Noreen Barajas-Murph, Gregory Dobbin, Jessica Knott, Mark McCormack, Jeffery Pomerantz, Ryan Seilhamer, and Nicole Weber. 2019. *EDUCAUSE Horizon Report 2019 Higher Education Edition*. Technical Report. EDU19.

[2] Robert Bodily, Judy Kay, et al. 2018. Open learner models and learning analytics dashboards: a systematic review. In *8th Int. Conf. on Learning Analytics & Knowledge*. ACM, 41–50.

[3] Maria Bolsinova, Paul de Boeck, and Jesper Tijmstra. 2017. Modelling conditional dependence between response time and accuracy. *Psychometrika* 82, 4 (2017), 1126–1148.

[4] Maria Bolsinova and Dylan Molenaar. 2018. Modeling nonlinear conditional dependence between response time and accuracy. *Frontiers in psychology* 9 (2018), 1525.

[5] Peter Brusilovsky, Sibel Somyürek, Julio Guerra, Roya Hosseini, Vladimir Zadorozhny, and Paula J Durlach. 2015. Open social student modeling for personalized learning. *IEEE Transactions on Emerging Topics in Computing* 4, 3 (2015), 450–461.

[6] Piyapatr Busababodhin and Pimpan Amphanthong. 2016. Copula modelling for multivariate statistical process control: a review. *Communications for Statistical Applications and Methods* 23, 6 (2016), 497–515.

[7] Hao Cen, Kenneth Koedinger, and Brian Junker. 2006. Learning factors analysis–a general method for cognitive model evaluation and improvement. In *International Conference on Intelligent Tutoring Systems*. Springer, 164–175.

[8] Shu-Ren Chang, Barbara S Plake, Gene A Kramer, and Shu-Mei Lien. 2011. Development and application of detection indices for measuring guessing behaviors and test-taking effort in computerized adaptive testing. *Educational and Psychological Measurement* 71, 3 (2011), 437–459.

[9] Haiqin Chen, Paul De Boeck, Matthew Grady, Chien-Lin Yang, and David Waldschmidt. 2018. Curvilinear dependency of response accuracy on response time in cognitive tests. *Intelligence* 69 (2018), 16–23.

[10] Yunxiao Chen, Xiaoou Li, Jingchen Liu, and Zhiliang Ying. 2018. Recommendation system for adaptive learning. *Appl. Psychol. Meas.* 42, 1 (2018), 24—41.

[11] Irene-Angelica Chounta and Paulo F Carvalho. 2019. Square it up! How to model step duration when predicting student performance. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*. 330–334.

[12] Konstantina Chrysafiadi and Maria Virvou. 2013. Review: Student Modeling Approaches: A Literature Review for the Last Decade. *Expert Syst. Appl.* 40, 11 (2013), 4715–4729.

[13] Albert T Corbett and John R Anderson. 1994. Knowledge tracing: Modeling the acquisition of procedural knowledge. *User modeling and user-adapted interaction* 4, 4 (1994), 253–278.

[14] Ryan SJ d Baker, Albert T Corbett, and Vincent Aleven. 2008. More accurate student modeling through contextual estimation of slip and guess probabilities in bayesian knowledge tracing. In *International conference on intelligent tutoring systems*. Springer, 406–415.

[15] Paul De Boeck and Minjeong Jeon. 2019. An overview of models for response times and processes in cognitive tests. *Frontiers in psychology* 10 (2019), 102.

[16] Michel C Desmarais and Ryan S Baker. 2012. A Review of Recent Advances in Learner and Skill Modeling in Intelligent Learning Environments. *User Modeling and User-Adapted Interaction* 22, 1-2 (2012), 9–38.

[17] Paul Embrechts, Filip Lindskog, and Alexander McNeil. 2001. Modelling dependence with copulas. *Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich* 14 (2001).

[18] Paul Embrechts, Alexander McNeil, and Daniel Straumann. 2002. Correlation and dependence in risk management: properties and pitfalls. *Risk management: value at risk and beyond* 1 (2002), 176–223.

[19] Gabriel Frahm, Markus Junker, and Rafael Schmidt. 2005. Estimating the tail-dependence coefficient: properties and pitfalls. *Insurance: mathematics and economics* 37, 1 (2005), 80–100.

[20] Christian Genest and Jock MacKay. 1986. The joy of copulas: bivariate distributions with uniform marginals. *The American Statistician* 40, 4 (1986), 280–283.

[21] Trevor J Hastie. 2017. Generalized additive models. In *Statistical models in S*. Routledge, 249–307.

[22] Harry Joe. 1997. *Multivariate models and multivariate dependence concepts*. CRC Press.

[23] Sharon Klinkenberg, Marthe Straatemeier, and Han LJ van der Maas. 2011. Computer adaptive practice of maths ability using a new item response model for on the fly ability and difficulty estimation. *Computers & Education* 57, 2 (2011), 1813–1824.

[24] Kenneth R Koedinger, Ryan SJd Baker, Kyle Cunningham, Alida Skogsholm, Brett Leber, and John Stamper. 2010. A data repository for the EDM community: The PSLC DataShop. *Handbook of educational data mining* 43 (2010), 43–56.

[25] Amar Lalwani and Sweety Agrawal. 2019. What Does Time Tell? Tracing the Forgetting Curve Using Deep Knowledge Tracing. In *International Conference on Artificial Intelligence in Education*. Springer, 158–162.

[26] Chen Lin, Shitian Shen, and Min Chi. 2016. Incorporating student response time and tutor instructional interventions into student modeling. In *Proceedings of the 2016 Conference on user modeling adaptation and personalization*. 157–161.

[27] Noboru Matsuda, William W Cohen, Jonathan Sewall, Gustavo Lacerda, and Kenneth R Koedinger. 2007. Evaluating a Simulated Student Using Real Students Data for Training and Testing. In *International Conference on User Modeling*. Springer, 107–116.

[28] Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. 2005. *Quantitative risk management*. Princeton University Press, Princeton, NJ. xvi+538 pages. Concepts, techniques and tools.

[29] Everett Mettler, Christine M. Massey, and Philip J. Kellman. 2011. Improving Adaptive Learning Technology through the Use of Response Times. In *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*. 2532–2537.

[30] Dylan Molenaar, Francis Tuerlinckx, and Han LJ van der Maas. 2015. A bivariate generalized linear item response theory modeling framework to the analysis of responses and response times. *Multivariate Behavioral Research* 50, 1 (2015), 56–74.

[31] Roger B. Nelsen. 1999. *An introduction to copulas*. Lecture Notes in Statistics, Vol. 139. Springer-Verlag, New York.

[32] Roger B Nelsen. 2007. *An introduction to copulas*. Springer Science & Business Media.

[33] Robert Nisbet, Gary Miner, and Ken Yale. 2018. Chapter 8 - Advanced Algorithms for Data Mining. In *Handbook of Statistical Analysis and Data Mining Applications (Second Edition)* (second edition ed.), Robert Nisbet, Gary Miner, and Ken Yale (Eds.). Academic Press, Boston, 149 – 167. https://doi.org/10.1016/B978-0-12-416632-5.00008-6

[34] Harry H Panjer. 2006. *Operational risk: modeling analytics*. Vol. 620. John Wiley & Sons.

[35] Zacharoula Papamitsiou and Anastasios A Economides. 2014. Temporal learning analytics for adaptive assessment. *Journal of Learning Analytics* 1, 3 (2014), 165–168.

[36] Zacharoula Papamitsiou and Anastasios A Economides. 2018. Motivating students in collaborative activities with game-theoretic group recommendations. *IEEE Trans. Learn. Technol.* (2018).

[37] Zacharoula Papamitsiou, Anastasios A Economides, and Michail N Giannakos. 2019. Fostering Learners' Performance with On-demand Metacognitive Feedback. In *European Conference on Technology Enhanced Learning*. Springer, Cham, 423–435.

[38] Zacharoula Papamitsiou, Anastasios A Economides, Ilias O Pappas, and Michail N Giannakos. 2018. Explaining learning performance using response-time, self-regulation and satisfaction from content: an fsQCA approach. In *Proceedings of the 8th international conference on learning analytics and knowledge*. 181–190.

[39] Zacharoula Papamitsiou, Eirini Karapistoli, and Anastasios A Economides. 2016. Applying classification techniques on temporal trace data for shaping student behavior models. In *Proceedings of the sixth international conference on learning analytics & knowledge*. 299–303.

[40] Andrew J Patton. 2012. A review of copula models for economic time series. *Journal of Multivariate Analysis* 110 (2012), 4–18.

[41] Philip I Pavlik Jr, Hao Cen, and Kenneth R Koedinger. 2009. Performance Factors Analysis–A New Alternative to Knowledge Tracing.. In *Proceedings of the 2009 Conference on Artificial Intelligence in Education*. 531–538.

[42] Radek Pelánek. 2016. Applications of the Elo rating system in adaptive educational systems. *Computers & Education* 98 (2016), 169–179.

[43] Radek Pelánek. 2017. Bayesian knowledge tracing, logistic models, and beyond: an overview of learner modeling techniques. *User Modeling and User-Adapted Interaction* 27, 3 (2017), 313–350.

[44] Radek Pelánek and Petr Jarušek. 2015. Student modeling based on problem solving times. *International Journal of Artificial Intelligence in Education* 25, 4 (2015), 493–519.

[45] Chris Piech, Jonathan Bassen, Jonathan Huang, Surya Ganguli, Mehran Sahami, Leonidas J Guibas, and Jascha Sohl-Dickstein. 2015. Deep knowledge tracing. In *Advances in neural information processing systems*. 505–513.

[46] KR Premlatha and TV Geetha. 2015. Learning content design and learner adaptation for adaptive e-learning environment: a survey. *Artif. Intell. Rev.* 44, 4 (2015), 443–465.

[47] Bart Rienties and Lisette Toetenel. 2016. The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules. *Computers in Human Behavior* 60 (2016), 333–341.

[48] Jirí Rihák and Radek Pelánek. 2017. Measuring Similarity of Educational Items Using Data on Learners' Performance. *International Educational Data Mining Society* (2017).

[49] Guillaume Rochefort-Maranda. 2016. Simplicity and model selection. *European Journal for Philosophy of Science* 6, 2 (2016), 261–279.

[50] Kshitij Sharma, Valérie Chavez-Demoulin, Pierre Dillenbourg, et al. 2018. Nonstationary modelling of tail dependence of two subjects' concentration. *The Annals of Applied Statistics* 12, 2 (2018), 1293–1311.

[51] A. Sklar. 1959. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris* 8 (1959), 229–231.

[52] Abe Sklar, A SKLAR, and CA Sklar. 1959. Fonctions de reprtition an dimensions et leursmarges. (1959).

[53] Nguyen Thai-Nghe, Lucas Drumond, Tomás Horváth, and Lars Schmidt-Thieme. 2012. Using factorization machines for student modeling.. In *UMAP Workshops*.

[54] Edward L Thorndike, Elsie Oschrin Bregman, Margaret Vara Cobb, and Ella Woodyard. 1926. The measurement of intelligence. (1926).

[55] Brendon Towle and Michael Halm. 2005. Designing adaptive learning environments with learning design. In *Learning design*. Springer, 215—226.

[56] Wim J van der Linden. 2006. A lognormal model for response times on test items. *Journal of Educational and Behavioral Statistics* 31, 2 (2006), 181–204.

[57] Wim J van der Linden. 2007. A hierarchical framework for modeling speed and accuracy on test items. *Psychometrika* 72, 3 (2007), 287.

[58] Wim J van der Linden and Ronald K Hambleton. 2013. *Handbook of modern item response theory*. Springer Science & Business Media.

[59] Thibault Vatter and Valérie Chavez-Demoulin. 2015. Generalized additive models for conditional dependence structures. *Journal of Multivariate Analysis* 141 (2015), 147 – 167. https://doi.org/10.1016/j.jmva.2015.07.003

[60] Xu Wang, Diyi Yang, Miaomiao Wen, Kenneth Koedinger, and Carolyn P Rosé. 2015. Investigating How Student's Cognitive Behavior in MOOC Discussion Forums Affect Learning Gains. *International Educational Data Mining Society* (2015).

[61] Steven L Wise and Lingyun Gao. 2017. A general approach to measuring test-taking effort on computer-based tests. *Applied Measurement in Education* 30, 4 (2017), 343–354.

[62] Xiaolu Xiong, Zachary A Pardos, and Heffernan Neil T. 2011. An analysis of response time data for improving student performance prediction. *KDD 2011 Workshop: Knowledge Discovery in Educational Data, Held as part of 17th ACM SIGKDD Conference on Knowledge Discovery and Data Miningr* (2011).

[63] Michael Yudelson and Kenneth Koedinger. 2013. Estimating the benefits of student model improvements on a substantive scale. In *Educational Data Mining 2013*.