Predicting Learning Performance in Serious Games

Michael D. Kickmeier-Rust(✉)

Institute for Educational Assessment, University of Teacher Education, St. Gallen, Switzerland
michael.kickmeier@phsg.ch

Abstract. The prediction of learning performance is an important task in the context of smart tutoring systems. A growing community from the field of Learning Analytics and Educational Data Mining investigates the methods and technologies to make predictions about the competencies and skills, learners may reach within a specific course or program. Such performance predictions may also enrich the capabilities and the effectiveness of serious games. In game-based assessment, predictions add a novel dimension for the personalization and adaption in games for which these functions may provide a valuable data basis. The Learning Performance Vector (LPV) allows utilizing information about the learning domain (i.e., the competencies and the structure of competencies) and log file information from games to make performance predictions. In a simulative study based on existing datasets, we explored the characteristics of the approach and compared it to a linear regression model. The results indicate that the LPV is a promising method, specifically in data rich game-based scenarios with limited external information.

Keywords: In-game assessment · Performance prediction · Learning analytics Competence-based Knowledge Space Theory

1 Introduction

The assessment of learning performance plays a crucial role in many serious games. Specifically adaptive games require a certain understanding of competencies of the learners and their learning progresses. Accordingly important is a sound, valid, and theoretically grounded assessment. The evidences, thereby, may be divided into performance related aspects, emotional-motivational as well as personality related aspects [1]. The performance related aspects include measuring, gathering, analyzing, and interpreting scores, task completion rates, completion times, success rates, success depths (the quality or degree to which a ask has been accomplished), etc. [2]. The approaches to in-game assessment, stealth assessment, and non-invasive adaptation of games have been refined significantly over the past decade [3, 4]. There exist structural models (related to KST and micro-adaptivity [5]), cognitive diagnosis models [6, 7], Bayesian approaches [8], latent variable models [9] and methods from the field of learning analytics (LA) research [10]. A concept that is not as popular in the context of serious games as assessment approaches are, is the prediction of learning performance

S. Göbel et al. (Eds.): JCSG 2018, LNCS 11243, pp. 133–144, 2018.
https://doi.org/10.1007/978-3-030-02762-9_14

in games. Performance prediction has a long(er) tradition in the context of Learning Analytics, for example.

In the context of serious games, prediction of learning performance may be important in two areas. The one is game-based assessment, the assessment of certain performance constructs on the hand of games or simulations. A prominent example is the National Observational Teaching Exam (NOTE) by the Educational Testing Service (ETS) [11]. NOTE is a test instrument for teacher's abilities based on simulated classroom scenarios, accredited in the USA. Meanwhile already a number of commercial psychometric games exist. Prediction of achievements may be a valuable dimension of such games and simulations. The second area is assessment for games, for example to inform personalization and adaption of games. An example is the approach of micro-adaptivity [5], which is a probabilistic, non-numerical framework to build believe models about learner performance on the basis of fine grained activities in the game. This paper describes an extension of the micro-adaptivity concept, aiming at the prediction of a so-called learning horizon of a learner. This concept refers to the likelihood with which a particular learner will achieve the learning goals in the domain of the learning game. This approach is specifically interesting for serious games because with each action of a learner in the game, the prediction model can be updated and the prediction gets more accurate. By this means, the game may predict possible achievements already at a comparably early stage and the right didactic consequences can be drawn (e.g., an adaptation of the game at an early stage).

2 Predictive Learning Analytics

The prediction of academic success has a longer tradition, for example in the context of university entry exams, which in the end aim to predict the performance and the chances to graduate. This research basically focuses on two types of predictors: cognitive ability or traditional measures, and non-cognitive, affective or non-academic factors. Cognitive factors usually refer to measures such as high school grades and standardized test scores whereas non-cognitive measures are related to psychological factors, like social support and academic related skills [12]. Of course, there are mixed approaches as well [13]. Often, very simple measures – such as engagement – predict study success best [14]. In general, one has to distinguish the attributes and variables on which predictions are based and, second, the methods how these variables are processed. The most frequent methods to process variables are classifications, regressions, and categorizations. In a review [15] list and describe the following methods: Decision Tree, Artificial Neural Networks, Naive Bayes, K-Nearest Neighbor, and Support Vector Machine. These authors conclude that Neural Network and Decision Tree approaches have the highest prediction accuracy. [16] provide an overview of approaches over the past fifteen years. These authors also demonstrate the effectiveness and the limitations of four approaches (Logistic Regression, Naïve Bayes, Support Vector Machine) in the context of an early alert system. An interesting comparison of eight methods (ranging from K-Nearest Neighbor to Decision Tree algorithms) along a variety of learning factors was published by [17] who found only fair prediction accuracy (60 to 80%). Also, the various algorithms did not differ substantially. [18]

compared different methods from data mining and from the field of recommender systems such as Bayesian Probabilistic Matrix Factorization and Bayesian Probabilistic Tensor Factorization and could demonstrate that the methods are, in principle, equally accurate [19] demonstrated, that prediction accuracy can be improved when binary regression algorithms are extended to partial credit models and when the algorithms includes penalties for hints and attempts.

Using multiple and continuously changing sources of data are the basis for predictions, which perfectly suits the nature of serious games. [20] discuss how the predictive capacity of different sources of data changes as the course progresses and also how a student's pattern of behavior changes during the course, which in turn affects predictions. [16] conclude that prediction and risk detection approaches do work, however, they have their strengths in large lecture-style electronic courses. It remains unclear though, to what extend these methods are helpful in smaller, perhaps more limited games.

In conclusion, an overview of the literature indicates that a number of sophisticated prediction models do exist and that the accuracy of the methods is widely acceptable. The different prediction models and methods appear to have a lower impact on the accuracy in comparison to the underlying data basis (the variables and attributes of students). A critical factor, obviously, is the settings within which the methods can be applied. Only few studies outside "ideal" settings such as (i) a general forecast of academic success (likelihood of completing a course or school) or (ii) as MOOCs or distance learning scenarios report a practical success. This argument is mirrored by studies that yielded that conventional methods could not predict student success (e.g., [20–22]). A number of researchers argue that further work is needed to investigate the applicability of methods in small scale, heterogeneous scenarios with incomplete data basis (e.g., [16]). The literature indicates that a differentiating factor is whether predictions are made over a long period (e.g., by predicting college success at the time of the enrollment) or on a short scale (e.g., a course or a game, cf. [20]). In general, such settings reveal the limitations of prediction methods in general. With this paper, I want to introduce a different approach of predicting student performance, originating from the community of probabilistic and combinatorial test theory.

3 Competence Spaces

We developed a combinatorial approach to educational personalization in games, which is called micro-adaptivity. In projects such as ELEKTRA and 80Days (www.eightydays.eu) we introduced and evaluated the usefulness of this approach. The goal was to complement the widely bottom-up driven, data mining and statistics focused methods of assessment and adaptation with a top-down approach, driven by psycho-pedagogical theories. At the same time, we attempted to work towards solutions for the areas, within which typical methods have certain weaknesses (as discussed above). One direction, micro-adaptivity pursued was Competence-based Knowledge Space Theory (CbKST), which is an extension of Knowledge Space Theory (KST) established by [23, 24]. KST is a set-theoretic framework for addressing the relations among problems (e.g., test items). It provides a basis for structuring a domain of knowledge and for representing the knowledge based on prerequisite relations. Similar to Item Response Theory (IRT), KST attempts to order test items and problems. As opposed to IRT, which establishes linear orders, KST allows for multiple dimensions. It establishes a Knowledge Structure by identifying relationships between the items. While KST focuses only on the items – or rather whether learners are able to master the item (performance), CbKST introduces a separation of observable performance and latent, unobservable knowledge and competences, which determine the performance [25]. Very briefly, the fundamental idea of CbkST is to assume a set of atomic competencies and a so-called Prerequisite Relation between them. Such relation is, in fact, a pedagogical model that explains the course of learning and development in a specific domain and the structural relations in the domain. As an example, one such relationship is to assume that adding integers is a prerequisite to learn multiplying integers. An individual learner can have none, all, or a specific set of competencies of a domain (e.g., being able to add, subtract, and multiply integers) – this is called the learner's Competence State. By a combinatorial permutation, the Prerequisite Relation induces a so-called Competence Space, the collection of all possible Competence States (cf. Fig. 1). Due to the pedagogical model, not all possible combinations of competencies are meaningful states; for example being able to multiply integers but not to add them is not reasonable state. This happens on a latent, conceptual level; the knowledge, the competencies and skills, the aptitude of a learner cannot be observed directly. CbKST now links the performance to the competence level on a stochastic level by so-called Representation Functions. Concrete test items and problems serve as behavioral indicators. Mastering an item increases the probabilities of all those Competence States that include the associated competencies. By this means, the probability distribution over the Competence Space is updated on the basis of a continuous interpretation of all sorts of behavioral indictors. Each gaming activity, each achievement, each learning activity contributes to CbKST's believe model. [5] demonstrate that this approach can be broken down to a very fine granularity and that it can be utilized in the context of serious games.

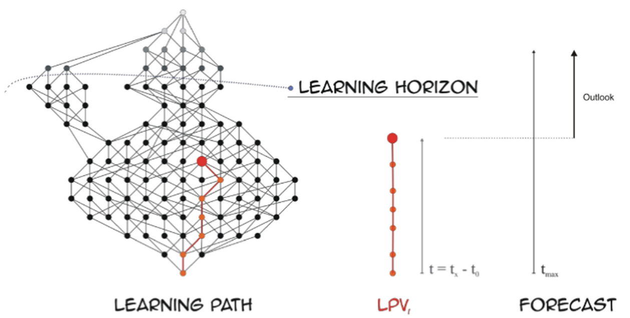

Fig. 1. The left part shows a Competence Space including the admissible learning paths. The red path may be that of a particular student. The left part of the figure illustrates the prediction principle of the LPV. (Color figure online)

4 The Learning Performance Vector

In the game context, it might be important to estimate the competence states a learner may reach within a reasonable time span of gaming. This helps, for example, preventing the learner to fail in reaching the goals of a game. As outlined initially, a good portion of the existing methods are statistics-based data mining techniques. These perform well on a general, statistical basis however have clear weaknesses when operating on a level of individual learners. Also, many statistical approaches build upon a set of (at least) debatable statistical assumptions and decision criteria. In this paper, we introduce a first evaluation study, comparing a simple statistical prediction method with a CbKST-based one, we term Learning Performance Vector (LPV). The purpose is to elucidate primarily the prediction characteristics of the method.

The origin of the prediction algorithm is a Competence Space. This space gives us a model of the learning domain, starting from having no competencies in a domain, leading to the complete mastery. This allows us to identify the progress of a particular learner given the timeline of a course. Mathematically speaking, we have the set of all admissible learning paths. This indicates the average learning efforts, given that transitions have specific difficulties or weights. We have a set of competencies $Q = \{a, b, c, \ldots\}$ with a relationship c $\leq$ c' among the competencies, which establishes the Competence Space. The sum of the resulting Competence States is $\sum(|Q|r)$. Given that the transitions from one competence state to another has a difficulty parameter, which in turn is the average of the difficulty parameters of the competencies being a part of the state, we have a set of tuples of the initial state, the end state, and the difficulty $\tau = [s_1, s_2, w]$. This results in a set of such tuples for the entire Competence Space $T = \sum(r|Q)$. In addition, we have a set of indicators providing evidences for competencies: I = {e_i, {c} * w}, with a given weight w. Based on the evidences we can estimate the likelihood of each competency. The probability of a Competence State is the average of its competencies $p(s) = \sum(\pi)/n$. To identify the learning path of a person, we identify the states with the highest probability at each assessment point. For each step, we compute the difficulty (as a value between 0 and 1). The sum of the values gives us an indicator of the efforts a student spent on her learning history (the individual learning path). In a next step, given the concrete Competence State of the learner, we have to identify the possible paths towards the final learning goal, which is a (rather small) subset of all possible paths. Equal to the computation of the difficulty to reach the current state, we can compute the potential difficulties of all possible paths towards the goal. This now is an indicator for the efforts that are necessary for an individual learner to reach the learning goal.

As illustrate in Fig. 1, considering the progress of a student within a given span of time, we can make a prediction about how far a student can come within the remaining time (of a course, for example). So, as a final step, we can identify exactly those states (and therefore the competencies) a particular learner will be able to reach within the time limits. We call the set of the reachable Competence States the learning horizon.

5 Identifying the Prediction Characteristics

The purpose of this study was to investigate the characteristics of the prediction method as opposed to an existing, well-elaborated approach. To judge the accuracy, it is necessary to compare the predictions with – what often is called – "ground truth". Therefore, we simulated the learning performance of excellent, medium, and poor learners. On this basis, we made systematic comparisons. Since no particular examples exists for the prediction with games, we used a conventional test data set.

5.1 Data Set

The first step for this study is to select an appropriate data set. To build upon a realistic data we selected a data set from Carnegie Mellon's DataShop. It is a data set of "Assistments Math 2004–2005", data set id 92 (accessible at pslcdatashop.web.cmu.edu/DatasetInfo?datasetId=92). This data set covers mathematics (which offers an easy 'playground' because it is a well-defined domain) and includes the data of 912 students. The data set is based on in total 80 competencies (knowledge components). For the simulation study, we selected a subset of 11 competencies and established a competence model (see Fig. 2). The weights are derived from the inverse solution frequencies of the real data set. Furthermore, we selected 12 item types and 111 items from the data set. These cover one or more of the selected competencies, partially also other competencies (Fig. 2). Based on the real data set, we simulated prototypical learners, taking the characteristics of 912 students and the item solution frequencies into account. The ability parameter was defined on a scale from 1 to 10, while 1 means no knowledge in the domain and 10 means having all competencies. The parameters were simulated based on a normal distribution, assuring the medium level abilities are most common and extreme position rather seldom. Finally, because this study is about prediction, we simulated 9 time points with the assumption that in the time intervals learning occurs, depending on the student abilities. In summary, we simulated the answer patterns of 15 students across 9 time points in 111 fictitious test items, covering 12 competencies. The simulated data set consists of 1665 data points. The following chart shows the prototypical simulated results of an excellent learner (squares), a medium learner (circles), and a poor learner (diamonds) (Fig. 3a). The values show the relative increase in correctly solved items over the 9 time intervals. The bold black diagonal indicates the optimal increase, so that with each of the 9 points in time 1/9 of the items is solved correctly – or in other terms, 1/9 of the competencies have been acquired.

The results show that the increase is determined by the student abilities, due to error rates (lucky guesses and careless errors) we see that the optimal learner is a bit below the ideal diagonal while thee poor learner still shows a slight increase.

5.2 A Simple Linear Prediction Model

To evaluate the characteristics of the LPV, we established a baseline prediction model. The model is a simple linear regression model based on a retrospective view of a particular student's performance. The model considers the performance of a particular student and predicts the future performance on the basis of the slope the general

Competencies

ID	Competence	Weight
1	Addition	0.10
2	Subtraction	0.15
3	Multiplication	0.27
4	Division	0.40
5	Fraction	0.45
6	Division /w decimals	0.55
7	Fraction multiplication	0.66
8	Fraction division	0.70
9	Fraction percent	0.80
10	Fraction /w decimals	0.89
11	Comb. Equation	0.93

Test Items

Item Type	Competencies
1	1
2	1,2
3	1,3
4	2
5	1,2,3
6	1,2,3,4
7	1,2,3,4,5
8	1,2,3,4,6
9	1,2,3,4,5,7
10	1,2,3,4,5,6,7,8,9
11	1,2,3,4,5,6,7,8,10
12	1,2,3,5

Students

Student ID	Ability Level (1 – 10)
1	2
2	9
3	3
4	3
5	4
6	5
7	5
8	6
9	7
10	7
11	8
12	8
13	9
14	5
15	4

Competence Model

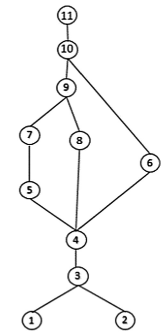

Fig. 2. From left to right: the selected competencies, the selected test items, the simulated students, and the derived competence model (prerequisite relation).

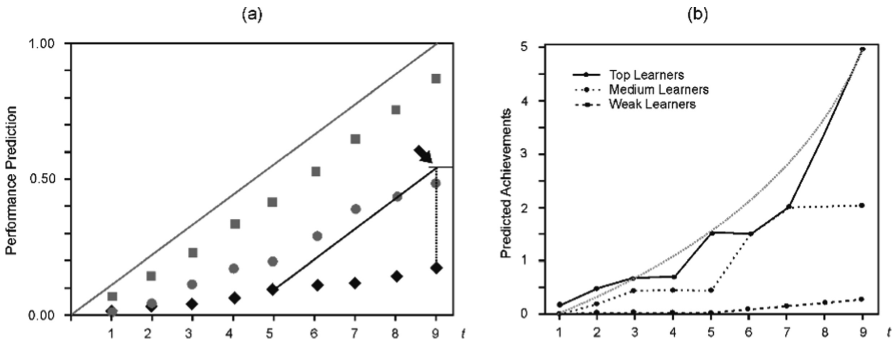

Fig. 3. Panel (a) shows the simulated results of three prototypical students as opposed to the ideal learning performance. Panel (b) shows the prediction results of the CbKST approach for a good (bold line), medium (dotted line) and poor (dashed line) student.

regression lines. This is demonstrated in Fig. 3. The low performing (diamonds) student reached a solution frequency of 0,054 at the end of interval 4. The model prediction is indicated in the grey line. This, however, is a significant overestimation of a student's abilities. There is a strong discrepancy between the results of a student and such estimations (dotted line in the figure). Figure 4 reports the predictive power of this approach over time. The left panel shows the predicted final achievements over the time intervals for the good (bold line) and the poor (dashed line) students. The right panel shows the accuracy (difference of simulated end values and predictions) of the approach. It is evident that the method overestimates the achievements by far, even for a nearly optimally performing student. This optimal and average linear increase is a problematic approach, obviously.

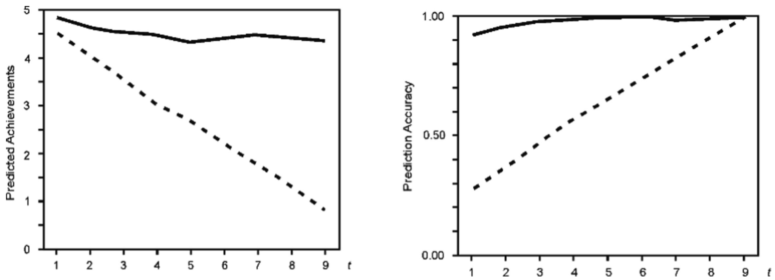

Fig. 4. The left panel shows the predicted achievements over time; the right panel shows the accuracy of the prediction method over time. The bold line refers to the LPV, the dashed line to the linear model.

5.3 The CbKST-Based LPV

The regression model, of course, is over simple. Methods that are more sophisticated are available, as introduce in this paper. These methods use more information about student performance and establish more complex, non-linear models. The contribution of CbKST is to use the multidimensional domain and learner models (the Competence Spaces) to add information about the nature of a learning domain to the model. This information includes the number and complexity of competencies as well as the relationships between them. Moreover, since the Competence Space is composed of the admissible Competence States, the lines between the states indicate the set of different learning paths that are possible – starting from having none of a domain's competencies to having all (cf. Fig. 1, left part).

The prediction logic of the LPV is to assume a finite number of learning paths leading from the trivial Competence State of having no competencies (the empty set) to the trivial state of having all competencies (the full set). We assume a well-graded space, claiming that in each step in the learning paths only one competency is acquired. The set of learning paths a learner is on can be identified on the basis of the current and past answer patterns (i.e., which item types were mastered and which not). The various paths can be characterized by their complexity, which is determined by the weights of the individual steps, which in turn result from the item solution frequencies in the original data set. The advantage is that we have a specific instance of the prediction model for each individual learner and her specific learning paths. In other words, mastering many easy items at an early stage is a weak indicator because major challenges are still ahead for the student. In turn, mastering highly complex items (with high weights and perhaps a larger number of prerequisites) is a very strong indicator because all prerequisite items are assumed to be possessed by the learner. The following figure shows the prediction results for the same simulated data set and the same students. In this example, the sum of weights assigned to the competence structure is 4.98 (the grey curve in Fig. 3b indicates the average prediction of this approach, contrasting the linear approach we described above). The curves show the performance of the three prototypical students across the 9 time intervals.

Figure 5(a and b) illustrates the final values predicted at each point in time. For the low performing student (dashed line) we obtain a clear overestimation of achievements but this overestimation is decreasing very quickly; after time interval 5, the prediction is very low – which is an accurate prediction. In case of a high performer (bold line) we have similar predictions as for the low performer and we see the same decrease in the predicted achievements. This decrease is much smaller and after time interval 5 the prediction becomes quite accurate also for the high performer. Figure 5b shows the accuracy, defined as the difference of predicted and actual achievements. When analyzing the different answer patterns (which items were processed in which order) for the 12 item types in our data, we found that the order of item types strongly influences the prediction characteristics (the predictive accuracy) of the LPV approach. If items that are more difficult are presented already at an early stage, the accuracy of the LPV can be increased, while such item order effects do not affect the accuracy of a linear model. Figure 5c illustrates the accuracy depending on the item order. With difficult items in the beginning, the accuracy of the LPV is very high already after 3 time intervals and superior to the linear model (Fig. 5c). In turn, the linear model has higher accuracy only when the order of item presentation strictly follows the assumption of an evenly distributed linear increase of item difficulty (Fig. 5d).

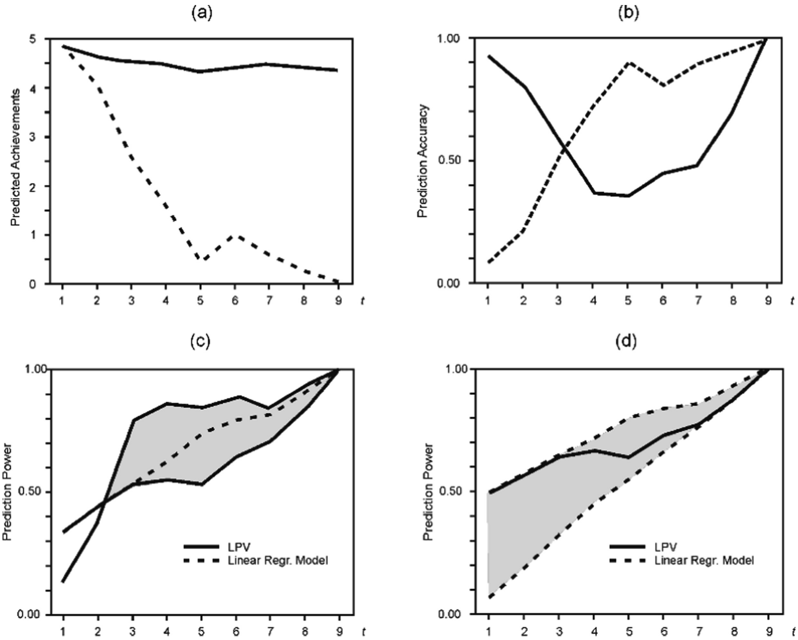

Fig. 5. Panel (a) shows the predicted performance over time, panel (b) the method's accuracy. The bold line displays an optimal student, the dashed line a poor student. Panels (c) and (d) show a comparison of predictive power (LPV vs. linear).

6 Discussion

The aim of this simulation study was to identify the characteristics of the LPV and if and to what extent the LPV is a suitable method to predict learning performance in serious games. The simulation-based study, developed on the basis of a real data set, allowed us to explore the characteristics and dependencies of the method in its application to various data characteristics. In this paper, we described the main findings.

As benchmark, we used a simple, linear regression model. Although there are much more sophisticated methods available, for the prediction of performance within a single course and without additional information about the learner, only few suitable methods exist [20]. Considering the development of the LPV as a robust prediction method, at this stage of our research, we avoided to involve too many student attributes. And certainly, one of the main goals of the LPV is to provide a performance prediction method that operates in scenarios with little to no background data available, with a shallow and incomplete data basis, and that allows a continuous monitoring of performance. Insofar, we obtained promising results. The information added to the prediction model – the domain structure and the weights of competencies – allows a more accurate performance prediction and the predictions converge quicker to a reasonable accuracy.

A critical aspect is the weighting process for the competencies and indirectly the test items (by associating items and competencies). This process clearly has a strong influence on the predictions. KST as well as CbKST describe a number of approaches for the structuring and weighting of a domain, ranging from data mining approaches to expert decisions (see [25] for details). With respect to the weighting process, a simple but practical method is a manual assignment of weights by teachers. This, however, bears the peril of an arbitrary and unfounded weighting. On the other hand, the strength of this approach could be that the weights would be grounded on the very concrete and practical experiences of a teacher. A second and more data driven approach is to refer to the solution frequencies of items in large data sets. This is the method we used in this study. If items are solved with a high frequency, we can assume a low difficult of the competencies covered by the item and also a low predictive power in terms of CbKST-type prerequisites between the competencies. A third method we will explore in future work is the so-called Component Attribute Approach [26]. This theoretical approach describes test items by components and their attributes. Components are major characteristics, for example, which algebraic operations are included in a math item. The attributes describe the individual components, for example, which types of numbers are part of the item. It was shown, that a Competence Space can be derived due to mathematical set inclusion. In our context decomposing and analyzing the components and their attributes can support the domain analysis and the weighting process because for typical course settings, usually elaborated curricula are available. A forth method is to analyze existing test items on the basis of their cognitive depth. This refers back to the famous taxonomy of Benjamin Bloom, revised by [27]. In the so-called Concept – Action Verb approach. Bloom proposed six such levels. An example would be "understand that a house has windows and apply this understanding in a new situation".

The taxonomy also separates the knowledge dimensions factual, conceptual, procedural, and metacognitive knowledge, which in the end established a two-dimensional hierarchy. In our context, this taxonomy provides a scaffolding to analyze the items, to identify the covered competencies, and to rank the competencies according the taxonomy – which in the end specifies the weights.

Certainly, the LPV presented in this paper stands in close relationship to other game analytics and game learning analytics (GLA) solutions [28]. Gaming Learning Analytics refers to analyzing gaming behaviors of students to obtain the relevant information about the learning process of the student. The goal is to understand how a student learns something new and how to help students achieve a higher outcome of their interaction with it – either in form of new designs or in-game adaptions. Several complex frameworks have been proposed and realized [28]. The aspect of the prediction of learning performance, however, was not necessarily in the focus of the systems. Instead they rather provide the information necessary for the predictions to stakeholders (students, teachers), for example in form of dashboards. Future steps will demonstrate the use of the LPV (and other predictive analytics approaches) in the context of a math game, developed by the Technical University of Graz. Moreover, we will formally extend the micro-adaptivity concept for serious games by performance predictions. In conclusion, the recent initiatives of introducing assessment, prediction, and adaptation methods from fields such as psychometrics and Learning Analytics should be intensified to make games stronger and more reliable means of educational assessment.

References

1. Kickmeier-Rust, M.D., Albert, D. (eds.): An Alien's Guide to Multi-Adaptive Educational Games. Informing Science Press, Santa Rosa (2012)
2. Wiemeyer, J., Kickmeier-Rust, M.D., Steiner, C.M.: Performance assessment in serious games. In: Dörner, R., Göbel, S., Effelsberg, W., Wiemeyer, J. (eds.) Serious Games: Foundations, Concepts and Practice, pp. 273–302. Springer, Berlin (2016). https://doi.org/10.1007/978-3-319-40612-1_10
3. Bellotti, F., Kapralos, B., Lee, L., Moreno-Ger, P., Berta, R.: Assessment in and of serious games: an overview. Adv. Hum.-Comput. Interact. (2013)
4. Shute, V.J., Ventura, M.: Stealth Assessment. Measuring and Supporting Learning in Video Games. The MIT Press, Cambridge (2013)
5. Kickmeier-Rust, M.D., Albert, D.: Micro adaptivity: protecting immersion in didactically adaptive digital educational games. J. Comput. Assist. Learn. **26**(2), 95–105 (2010)
6. Heller, J., Stefanutti, L., Anselmi, P., Robusto, E.: Erratum to: on the link between cognitive diagnostic models and knowledge space theory. Psychometrika **81**, 250–251 (2016)
7. Heller, J., Stefanutti, L., Anselmi, P., Robusto, E.: On the link between cognitive diagnostic models and knowledge space theory. Psychometrika **80**, 995–1019 (2015)
8. Käser, T., et al.: Design and evaluation of the computer-based training program Calcularis for enhancing numerical cognition. Front. Psychol. **4**, 289 (2013)
9. Mislevy, R.J.: Evidence-centered design for simulation-based assessment. Mil. Med. **178** (10), 107–114 (2013)

10. Kickmeier-Rust, M.D. (ed.): Learning analytics for an in serious games. In: Proceedings of the Joint workshop of the GALA Network of Excellence and the LEA's BOX project at EC-TEL 2014, 17 September 2014, Graz, Austria (2014). http://css-kmi.tugraz.at/mkrwww/leas-box/ectel2014.htm
11. Sandberg, H.: The Good and the Bad of Game-Based Assessment. ETS, Focus on R&D, Issue 1, April 2016. https://www.ets.org/research/policy_research_reports/focus_on_rd/issue1
12. Ma, Y., Liu, B., Wong, C., Yu, P., Lee, S.: Targeting the right students using data mining. In: Proceedings of the 6th International Conference on Knowledge Discovery and Data Mining (KDD), Boston, Massachusetts, USA, 20–23 August, pp. 457–464 (2000)
13. Shahiria, A.M., Husaina, W., Rashid, N.A.: A review on predicting student's performance using data mining techniques. Procedia Comput. Sci. **72**, 414–422 (2015)
14. Christian, T.M., Ayub, M.: Exploration of classification using nbtree for predicting students' performance. In: Proceedings of the International IEEE Conference on Data and Software Engineering (ICODSE), pp. 1–6 (2014)
15. Romero, C., López, M.I., Luna, J.-M., Ventura, S.: Predicting students' final performance from participation in on-line discussion forums. Comput. Educ. **68**, 458–472 (2013)
16. Jayaprakash, S.M., Moody, E.W., Lauría, E.J.M., Regan, J.R., Baron, J.D.: Early alert of academically at-risk students: an open source analytics initiative. J. Learn. Anal. **1**(1), 6–47 (2014)
17. Gray, G., McGuinness, C., Owende, P., Hofmann, M.: Learning factor models of students at risk of failing in the early stage of tertiary education. J. Learn. Anal. **3**(2), 330–372 (2016)
18. Sahebi, S., Huang, Y., Brusilovsky, P.: Predicting student performance in solving parameterized exercises. In: Trausan-Matu, S., Boyer, K.E., Crosby, M., Panourgia, K. (eds.) ITS 2014. LNCS, vol. 8474, pp. 496–503. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07221-0_62
19. Ostrow, K., Donnelly, C., Heffernan, N.: Optimizing partial credit algorithms to predict student performance. In: Proceedings of Educational Data Mining (EDM) 2015 (2015)
20. Wolff, A., Zdrahal, Z., Herrmannova, D., Knoth, P.: Predicting student performance from combined data sources. In: Peña-Ayala, A. (ed.) Educational Data Mining. SCI, vol. 524, pp. 175–202. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-02738-8_7
21. Strang, K.D.: Can online student performance be forecasted by learning analytics? Int. J. Technol. Enhanc. Learn. **8**(1), 26–47 (2016)
22. Papamitsiou, Z., Economides, A.: Learning analytics and educational data mining in practice: a systematic literature review of empirical evidence. Educ. Technol. Soc. **17**(4), 49–64 (2014)
23. Doignon, J.-P., Falmagne, J.-C.: Spaces for the assessment of knowledge. Int. J. Man-Mach. Stud. **23**, 175–196 (1985)
24. Doignon, J.-P., Falmagne, J.-C.: Knowledge Spaces. Springer, Berlin (1999). https://doi.org/10.1007/978-3-642-58625-5
25. Albert, D., Lukas, J. (eds.): Knowledge Spaces: Theories, Empirical Research, and Applications. Lawrence Erlbaum Associates, Mahwah (1999)
26. Heller, J., Steiner, C., Hockemeyer, C., Albert, D.: Competence-based knowledge structures for personalized learning. Int. J. E-Learn. **5**, 75–88 (2006)
27. Anderson, L.: Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Pearson Education, Boston (2013)
28. Freire, M., Serrano-Laguna, Á., Iglesias, B.M., Martínez-Ortiz, I., Moreno-Ger, P., Fernández-Manjón, B.: Game learning analytics: learning analytics for serious games. In: Spector, M., Lockee, B., Childress, M. (eds.) Learning, Design, and Technology. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-17727-4_21-1