Open Virtual Mobility

O3-A2.1: Definite set of relevant matching criteria

to be used in preparation for

- (O3-A2.2: Matching tool service definition)

- **Final draft -**

Outcome 3 Activity A1.1 Competency Directory and Matching Tool	
Document submission and review information	
Declared due date of deliverable	30.03.2018
Reviewed due date of deliverable	30.03.2018
Actual submission date	16.05.2018
Organisation name of lead contractor	Beuth University of Applied Sciences, Berlin
Revision	1.1
Author and reviewer information	
Name of the author	Johannes Konert
Affiliation of the author	Beuth University of Applied Sciences, Berlin
Name of the reviewer	Diana Andone
Affiliation of the reviewer	UNIVERSITATEA POLITEHNICA TIMISOARA

Open Virtual Mobility

Erasmus+

The creation of these resources has been (partially) funded by the ERASMUS+ grant program of the European Union under grant no. **2017-1-DE01-KA203-003494**. *Neither the European Commission nor the project's national funding agency DAAD are responsible for the content or liable for any losses or damage resulting of the use of these resources.*

Imprint

Imprint: This publication is O3-A2.1 of the Open Virtual Mobility Erasmus+ strategic partnership founded by the European Commission 2017 - 2020 under **2017-1-DE01-KA203-003494,** URL: https://www.openvirtualmobility.eu/

This paper is to discuss the definite set of matching criteria to use for building groups of learners in the open virtual mobility learning hub. It is produced as part of Outcome O3 "Competency Directory and Matching Tool" and aims at providing an overview over current state of the art research results, algorithmic approaches and the set of matching criteria to be used in the project. Candidates for criteria are presented, but depend on other not yet finished outputs of the project.

PDF download

A full-text PDF of this report is available as a free download from:
https://www.openvirtualmobility.eu/

Social media

Find us on Twitter: @openVM_erasmus
Give us your feedback on social media with the following hashtag: #openvirtualmobility

Suggested citation

Konert, J. (2018) O3-A2.1: Definite set of relevant matching criteria. Open Virtual Mobility Erasmus+ (2017-2020). Berlin, Germany. Retrieved from https://www.openvirtualmobility.eu/topics/outputs

Corresponding author

Johannes Konert
Beuth University of Applied Sciences
Luxemburger Straße 10, 13353 Berlin, Germany
johannes.konert[at]beuth-hochschule[dot]de

Table of Contents

Executive summary

The achievement of virtual mobility skills is related to inter-personal, inter-cultural and inter-disciplinary teamwork competence. Thus with certain tasks of the online learning resources provides by the open virtual mobility project (openVM), learners (teachers, students) are expected to work collaboratively. Usually, group formation for collaborative course work is based on random selection or self-selection. Based on existing research in the field of algorithmic learning group formation, this text will define what is meant by successful group learning activity (the optimization goal). Before focusing on the openVM needs, the document will highlight the general scientific findings about important criteria for successful group work. It will become clear that several criteria, e.g. prior knowledge or motivation, are found to play important roles for the success of learning groups. In brief, criteria can be optimized to be heterogeneously spread among group members (to amend each other) of homogeneously (similar). Further the document will explain the modelling of such criteria as input vectors for the optimization algorithm. This includes nominal, ordinal and interval scale values. Afterwards, functionality of the existing Moodle plugin mod_groupformation is presented and specific requirements for the openVM learning hub will be discussed.
Specifically, for the process of defining the matching criteria to be used in the project, the draft will propose next steps like opinion collection after tasks are known/defined by output O6 as well as the technical aspects of adapting and integrating the plugin to the VMLH.

What are the objectives of this paper?

- goal of learning group formation and algorithmic optimization
- criteria found relevant in scientific literature
- modeling of criteria as number values to be applicable for optimization algorithms
- current state of the Moodle plugin mod_groupformation (to be used in the project openVM)
- requirements for the openVM learning hub (partly to be delayed until group tasks and content from output 6 is set)

Who is this paper for?

- Pedagogues and didacticians interested in optimized learning group formation
- Researchers of computer-supported learning group formation to get an overview of existing approaches
- users of openVM learning hub to read background information about the group formation

What topics are addressed in this paper?

definition of learning group formation, criteria types, personality traits, motivation, prior knowledge, level of knowledge, mod_groupformation plugin, requirements

Contributors

- Johannes Konert, Beuth University of Applied Sciences, Germany

Acknowledgements

intentionally blank

1. Aims and Scope

For matching learners in an online course into groups for collaborative tasks, a common pattern is either to assign them randomly or by self-selection. Both types have disadvantageous effects on individual learning outcomes from a didactical point of view. Thus, the aim is to find valuable criteria to be considered in learning group formation for the openVM Learning Hub (VMLH).
The first milestone of this activity was originally aiming for a final set of criteria to be used. During the project it became clear that the content of the VMLH and thus the collaborative tasks will not be quickly decided. After the tasks, which are performed by learners, are designed, then the skills can be defined which have to be represented in the learning groups.
This draft will therefore present the scientific ground knowledge, possibilities and limitations of an algorithmically optimized learning group formation. The text will

A. define what is meant by successful group learning activity (the optimization goal), it will
B. highlight the general scientific findings about important criteria for successful group work, it will
C. explain the modelling of such criteria as input vectors for the optimization algorithm, and finally it will
D. briefly introduce the requirements and functionality of the existing plugin mod_groupformation to be used with Moodle for VMLH.

Specifically for the process of defining the matching criteria to be used in the project, the draft will propose next steps like opinion collection after tasks are known/defined by output O6 as well as the technical aspects of adapting and integrating the plugin to the VMLH.

2. Background and rationale (State of the Art)

When learning groups are formed to work together towards a solution of one or several tasks, the learning situation is designed for a certain set of learning goals, which are achieved by performing several steps. These steps are usually defined and didactically pre-planned, but often are not completely fixed to allow some flexibility in order and chosen approach to problem solving. The field of research designing such learning scenarios for online collaborative learning is called computer supported collaborative learning (CSCL) (Amarasinghe et al. 2017). It is focused by Output 6, which covers the design of the tasks. Together with the tasks the desired learning outcomes and prerequisites will be delivered.
In all cases, the tasks require a certain combination of prerequisite traits, skills or competencies to be present to successfully solve the task as a team (Dillenbourg 1999). If all prerequisites are met by one person, they might solve the task alone without any group work (Michaelsen et al. 1997). If necessary prerequisites are missing in the group, the task might not be solvable.

Optimization goal

Thus, the goal of optimized learning group formation is to find the best combination of learners per group that they individually can maximize their learning improvements towards the learning goal and all learners have the same quality of learning group (Konert et al. 2014).

Relevant criteria

Concerning the relevant matching criteria, they first can be divided into stable and unstable criteria. The latter are derived from the task to be performed and may have different values for the same learner if the task changes (Harrison et al. 2002). For example, *level of expertise in the learning domain* can change drastically between tasks (if tasks are of different domains). Still, *prior knowledge* is found to be one strong factor for group performance, if mixed heterogeneously in the learning group (with the intention that learners benefit from each others knowledge) (Horwitz 2005, p. 230). Likewise, *number of comments* to a task or *amount of activity in the learning platform* are unstable criteria, which are easy to measure but hard to proof in their relevance for a task B, in case evidence has only been found for their positive impact on learning outcomes for task A. Still, machine learning algorithms are used to investigate such relevance and recommend learning groups for future tasks based on this (assuming the tasks are of similar kind and target domain) (Zheng & Pinkwart 2014).
Stable criteria are primarily personality traits, which depend on the learner's character and not primarily on the tasks. Examples of such criteria are level of *extraversion*, *team orientation*, or *motivational aspects*. Literature found their positive impact on learning outcome in several correlative studies and some experimental studies (e.g. Humphrey et al. 2007, Bell 2007, Nederveen Pieterse 2011). While the big five personality traits (*extraversion, conscientiousness, openness, neuroticism*, and *agreeableness*) are found to be relevant, among them *extraversion* and *conscientiousness* proved to be significantly relevant for better learning results and better group coherence, if combined as being heterogeneously represented in a group (Bellhäuser et al. 2018).
To optimize the distribution of criteria within a group for best group learning outcome, for each matching criterion the value per learner is collected and then groups are built by optimally combining these values. The combination can be done manually, but quickly this task is beyond human capabilities if more than two criteria have to be considered and more than 20 learners have to be combined (Konert et al. 2014).
Still, it is not enough to take care of representation of certain characteristics in learning groups, as for some requirements a similarity among group members is beneficial, and for some others a maximized heterogeneity. For which criterion the project desired which weighting factor and which type of matching will be decided later.
Considering more potential criteria, for the goal of improving virtual mobility skills it seems to be especially important to consider cultural background and language proficiency as stable criteria for the group formation. Language abilities should match homogeneously within the group, while cultural background (e.g. measured as country of residence or home organization) could be mixed heterogeneously within a learning group to maximize cultural knowledge gain and exchange.
Pragmatically, some related work includes as well overlapping in weekly schedules of learners to allow for synchronous online communication or face-to-face meetups (Cavanaugh 2004). At the moment it is unclear, whether or not, this is a requirement for the learning hub tasks.

Criteria modeling for algorithmic use

From algorithmic perspective the learning group formation problem is only solvable by approximation to find near optimal solutions (Henry 2013, Konert et al. 2014). Therefore, several approaches exist and were discussed (overviews in Konert et al. 2014 or Srba & Bielikova 2014). Semantic approaches have their strength in considering complex side conditions, but depend on an ontology of the target domain and involved entities (learners, tasks, goals). Machine learning approaches conclude relevant factors by observing former task solving processes and infer recommendations for future (similar) tasks. Their advantage is the continuous improvement (as long as tasks and learners do not change) without any needed for a-priori knowledge. Last, numerical approaches use vector representations of criteria per learner and calculate the near-optimal combinations at a certain point of time. The criteria values of learners are collected by questionnaires, assessment test or by data import and need to be present before first group formation. Such approaches are beneficial, if learners, tasks, or domain topics change during the course, but the used matching criteria remain valid. Additionally they build groups of high quality from first iteration, but are incapable of improving iteratively by considering intermediate group work results. As the GroupAL algorithm (Konert et al. 2014), underlying the Moodle plugin to be used within the project, is a numerical approach, some details about criteria types supported and supported value types are given in the following sections.

Criteria types and vector representation

A criterion to consider is represented per learner by an n-dimensional vector. If the criterion has only 1 value, the dimensionality is 1. All value types are considered to be interval scale values. It is recommended to normalize them to the interval of [0,1]. If this is not the case, the algorithm will do this automatically. If nominal values are desired to be used, they can be transformed to be a separate criterion vector with one dimension per possible nominal value. If the learner has this nominal value, the dimension is set to 1, otherwise to 0. If ordinal values are desired to be used, their order cannot be represented currently, but they can be considered as nominal values.
The algorithm can cope with any fixed number of criteria. Each criterion can be of any fixed dimensionality (which needs to be same only per criterion among all learners).
Weights are supported to adjust the impact of criteria to the optimization.
Criteria can be considered for optimization as either being as similar as possible for all learners of one group (homogeneous) or as maximization of the value distance among all group members (heterogeneous). A mixture of both criteria types in one optimization is possible.
Simulative experiments proved that the overall maximum group performance to be achievable depends on the value distribution of the cohort of learners and mostly on the targeted group size, if heterogeneous criteria are used (Konert 2014, p. 141ff). Thus, it is recommended to aim for rather small group sizes (3-5) and use a limited set of criteria (3-5) among which only 1 or 2 are heterogeneous.
The relevant criteria for optimal learning performance need to be selected by didactic experts designing the collaborative tasks. Beside this, the algorithm aims for building as similar qualities of all groups as possible to prevent one group to be really a “perfect match” while subsequent groups suffer from declining group formation quality. This is achieved by taking standard deviation of the calculated fitness value into account and minimizing it. For details about the optimization algorithm and the matching strategy consider Konert (2014, p. 137ff).

mod_groupformation characteristics

The plugin mod_groupformation is currently available in version 1.5 for Moodle 3.4. It uses the GroupAL algorithm with its group-by-group matching approach (without optimization iterations). It is an activity plugin which can be added in several instances to a course. It currently supports English and German language (Konert et al. 2017). It is expected that extra translation work by all partners is needed to provide the plugin in more languages, if desired by the project. The user interface currently does not allow to enter weights for criteria, but they can be set in the programming code. Likewise, the calculation of criteria vectors from questionnaire items is directly encoded into the programming code and needs to be changed there. Still, questionnaires are dynamically generated based on XML-file definitions which support multi-language translations easily (but only for the whole installation identically).

Consideration of other criteria than prior knowledge (by self assessment using a percentage slider), team orientation, motivation for course goal, personality traits (Big 5), and language spoken, require extra implementation work and testing (see Figure 2 for current state). It has to be discussed within the openVM project team whether prior knowledge values should be copied from assessment activities in Moodle (Output O4).

For the plugin to work properly, the cron tasks in Moodle need to be setup; ideally to start each minute.

Group formation is configured for a fixed size of groups to build (e.g. 5 learners). The formation process needs to be started manually by a user with course administration rights. The formation is then done within minutes, depending on the next cycle of cron to start. It is possible by Moodle settings to ask users for a self-invented pseudonym to be stored in database tables, if needed for later usage (scientific research, post-tests etc.). Build groups can be edited manually, before transferring them to Moodle. Afterwards the groups exist as regular Moodle groups. Learners can access their group members via the activity. More details about the plugin capabilities can be found at https://moodle.org/plugins/mod_groupformation .

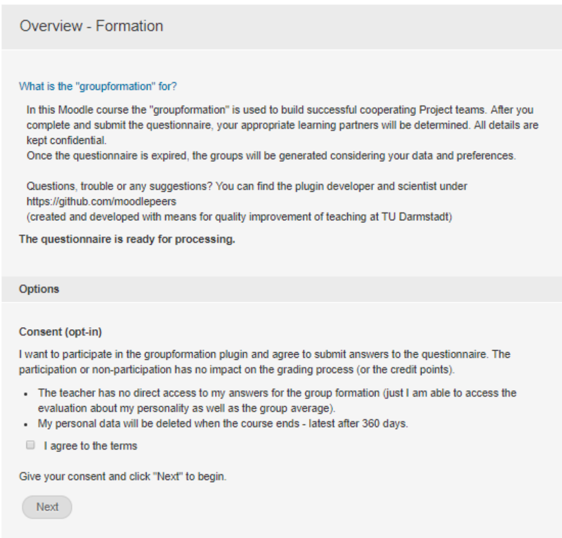

Figure 1: mod_groupformation opt-in dialog and explanation to users

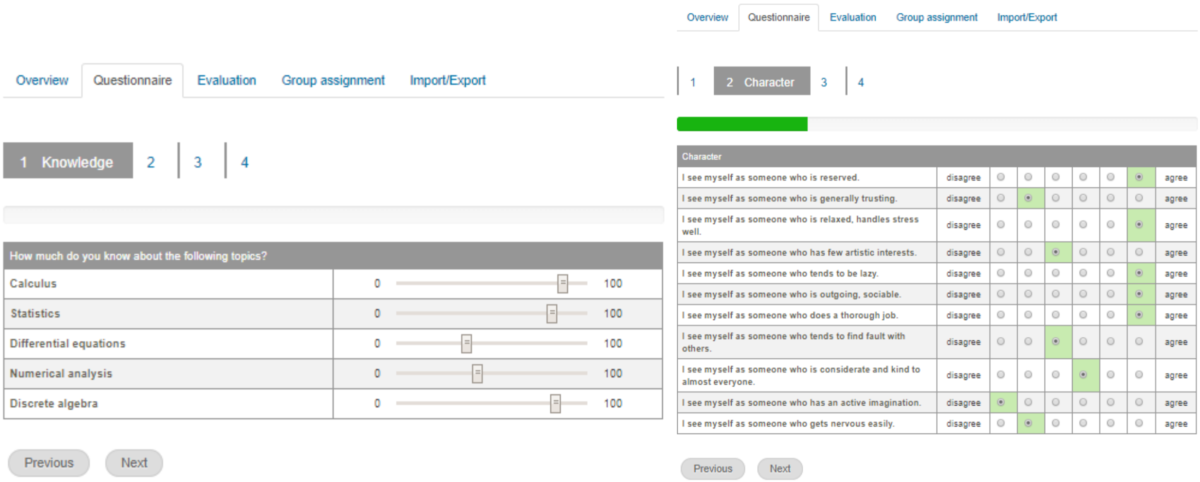

Figure 2: mod_groupformation examples of criteria data collection. left: prior knowledge self-assessment, right: Big5 personality traits as questionnaire

3. Methodology

Requirement analysis

Current state of the art in learning group formation is investigated and presented with three major perspectives. First, which criteria can be used and are found to be relevant for high-quality learning outcomes in group learning with collaborative tasks. Second, how can these criteria be represented mathematically and be used algorithmically and finally third, which algorithmic approach is used in the project. The findings are presented to the project partners for discussion and amendments. Such amendments are stimulated by three guiding questions

A. Which literature about relevant criteria for collaborative group work do you want to add to be considered in the project?
B. Which literature or study or source about relevant criteria to be represented in groups for learning VM skills do you want to add?
C. In your opinion, by which criteria should the learning groups be build for the open VM learning tasks? Please indicate for each proposal, whether the criterion should rather be heterogeneous (different) or rather similar (homogeneous) or rather something else (e.g. one of each kind) within the groups.

Partners are encouraged to invite external experts, like their organizations didactics center or pedagogical advisors to contribute to the discussion.
The new insights will be integrated into the document and assist the final decision of criteria to be considered.

Ethical issues and personal data collection

With the group formation tool, learners are put together in learning groups to communicate via the openVM LH and work collaboratively on tasks. For optimization a set of criteria will be used which are supported by literature that they were individually beneficial for better learning outcome and/or group performance in former learning scenarios. Still, there is no knowledge about the

benefits with the specific learners using the openVM LH, the specific combination and weighting of the criteria, and for the specific tasks to improve the VM skills. From an ethical perspective the algorithmic learning group formation should minimize the chance to expose learners to disadvantages they would otherwise not have. First, learners are informed about the intention and what will happen (criteria collection, group formation, group work). Second, learners can choose whether or not they want to take part in optimized group formation (opt-in). If they decline, they are matched randomly (see Figure 1). Moreover, learners are informed about contact possibilities to reach the developer team. The plugin integrates with the Moodle Privacy API to support requirements of European General Data Protection Regulation (GDPR) (European Commission, 2018; Moodle Community Team, 2018). Additionally, learners can – at all times during their participation – download all data stored about them by using a provided button of the plugin. Likewise, they can request deletion of their personal data. At the end of the online course, personal data is anonymized (pseudonymized) and extracted from the openVM LH for scientific studies.
As the openVM LH offers only voluntary online learning resources, no credits or marks are given. Thus, it can be justified to use the promising set of criteria and the optimization by algorithms, as the potential benefits for learners overweight the potential drawbacks for individual learning progress.
It is advised by the author of this document, to design a scientific user study with a questionnaire to be filled by participants. Its data is meant to investigate satisfaction with group members, quality of group work outcomes, and recommendations for the matching process.

Implementation

The plugin itself exists as a stable version with limited functionality as described above. It is available in Moodle plugin repository. It is expected that some adjustments have to be done to improve the applicability for the openVM LH. Such adjustments can be to the matching algorithm, the criteria to be used and their weighting as well as language support.
At current state it is unclear what are the must-have requirements for the group formation plugin. Thus the response to this document by project partners as well as the replies to the three questions are expected to assist in definition of the criteria set to be used.
The implementation is done by agile, iterative implementation based on user-stories. The plugin code currently exists as a Github repository which will be forked for the project adjustments. Open Source license GPL will be kept.

4. Intermediate results

The literature research, analysis and discussion revealed the following criteria to be relevant in general. The list does not contain VM skill specific criteria as the skill set is not yet decided (output O1) and responses to the questions A-C above are not yet collected.

Table 1: first candidate list of criteria to be used for openVM learning group matching (all weighted equally)

Criterion	Scale	Optimization goal	Data source
prior knowledge	Interval	heterogeneous	self-assessment
level of knowledge	Interval	homogeneous	self-assessment (mean value of prior knowledge fields)
extraversion	Interval	heterogeneous	questionnaire (big 5 items)
conscientiousness	Interval	heterogeneous	questionnaire (big 5 items)
language proficiency	Nominal (DE, EN, …)	homogeneous	questionnaire (MC)
cultural background/place of institution	Interval, if distance between learners Nominal, if asked by country or organization name	heterogeneous	City name (field) or questionnaire (MC)

Group size is not yet decided. The author of the document expects this information to be given by the designers of the collaborative tasks in Output O6.

An organizational account on Github has been established for openVM project (https://github.com/openVirtualMobility). Fork of the mod_groupformation will be done when implantation starts.

In contrast to the project proposal plan of milestones, it is evident that the final set of criteria cannot be decided a priori before the tasks and intended group work / learning outcomes are decided. Thus, the final decision needs to be done depending on Output 6 and Output 4 progress. The decision can then be simplified.

5. Discussion

The investigation and decision of most relevant criteria for learning group formation is a challenging research topic. While it has been around for some time, a lot of uncertainty still exists about the impact of certain criteria and the optimization approach to learning outcomes and coherence of the learning group. It is clear that optimization has an effect compared to pure random grouping. For the development and enhancement of virtual mobility skills it remains open which criteria exactly are the best. Thus, within the project a set of criteria (cf. Table 1) is used which is supported by prior scientific literature to be applicable for learning scenarios in general (without strong dependency on learning goals, tasks and scenario). If possible by Moodle integration, the project aims to use the country of origin and/or language proficiency as criteria as well. Using the existing mod_groupformation plugin as a basis for further development allows efficient integration into the process of the MOOC (Output 6). As the MOOC, tasks, and the open virtual mobility skills to be taught (Output 1) are still not developed yet, it remains a risk for Output 3 how well the learning group formation can be integrated and how important such cooperative tasks will be in the end. This will be named more clearly in the service description of the matching tool (O3-A2.2).

6. Conclusions

The document described the advantages of collaborative team work with tasks in a learning scenario to develop certain inter-personal skills, like the ones to be taught for open virtual mobility. The formation of such learning groups can be optimized to allow higher satisfaction of each group member with the group constellation and at the same time allows to optimize for better learning outcomes of everyone. To achieve this, certain personal characteristics need to be collected (by survey or from Moodle data), will be transformed into criteria vectors and then will be e.g. most similar (homogeneous) or amending another (heterogeneous) within a learning group to be built from all participants. Chapter 2 introduced the underlying research findings about relevant criteria, while chapter 3 outlined the methodology using an existing Moodle plugin (named mod_groupformation). Chapter 4 described the intermediate findings, mainly the criteria candidates to be used most probably. A final decision is made for next milestone O3-A2.2 which describes the final service characteristics of the group formation tool within the open virtual mobility project.

7. References

Amarasinghe, I., Hernandez-Leo, D., & Jonsson, A. (2017). Intelligent Group Formation in Computer Supported Collaborative Learning Scripts. Proceedings - IEEE 17th International Conference on Advanced Learning Technologies, ICALT 2017, (1), 201–203. http://doi.org/10.1109/ICALT.2017.62

Bell, S. T. (2007). Deep-level composition variables as predictors of team performance: A meta-analysis. Journal of Applied Psychology, 92(3), 595–615. http://doi.org/10.1037/0021-9010.92.3.595

Bellhäuser, H., Konert, J., Müller, A., & Röpke, R. (2018). Who is the Perfect Match? Effects of Algorithmic Learning Group Formation Using Personality Traits. I-Com Journal of Interactive Media, (4). http://doi.org/10.1515/icom-2018-0004

Cavanaugh, R., Ellis, M., Layton, R., & Ardis, M. (2004). Automating the Process of Assigning Students to Cooperative-Learning Teams. In Proceedings of the 2004 Proceedings of the 2004 ASEE Annual Conference. http://doi.org/10.1.1.65.682

Dillenbourg, P. (1999). What do you mean by Collaborative Learning? In P. Dillenbourg (Ed.), Collaborative-learning: Cognitive and Computational Approaches (pp. 1–15). Oxford: Elsevier.

European Commission. (2018). Key Changes with the General Data Protection Regulation – EUGDPR. Retrieved March 19, 2018, from https://eugdpr.org/the-regulation/

Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, teams, and task performance: Changing effects of surface- and deep-level diversity on group functioning. Academy of Management Journal, 45(5), 1029–1045. http://doi.org/10.2307/3069328

Henry, T. R. (2013). Creating effective student groups. Proceeding of the 44th ACM Technical Symposium on Computer Science Education - SIGCSE '13, 645. http://doi.org/10.1145/2445196.2445387

Horwitz, S. K. (2005). The Compositional Impact of Team Diversity on Performance: Theoretical Considerations. Human Resource Development Review, 4(2), 219–245. http://doi.org/10.1177/1534484305275847

Humphrey, S. E., Hollenbeck, J. R., Meyer, C. J., & Ilgen, D. R. (2007). Trait configurations in self-managed teams: a conceptual examination of the use of seeding for maximizing and minimizing trait variance in teams. The Journal of Applied Psychology, 92(3), 885–92. http://doi.org/10.1037/0021-9010.92.3.885

Konert, J. (2014). Interactive Multimedia Learning: Using Social Media for Peer Education in Single-Player Educational Games. Heidelberg, Germany: Springer.

Konert, J., Burlak, D., & Steinmetz, R. (2014). The Group Formation Problem: An Algorithmic Approach to Learning Group Formation. In C. Rensing, S. de Freitas, T. Ley, & P. J. Muñoz-Merino (Eds.), Proceedings of the 9th European Conference on Technology Enhanced Learning (EC-TEL) (pp. 221–234). Graz, Austria: Springer Berlin. http://doi.org/10.1007/978-3-319-11200-8_17

Konert, J., Röpke, R., & Bellhäuser, H. (2017). mod_groupformation: Moodle Plugin zur algorithmisch optimierten Lerngruppenbildung. In C. Igel, C. Ullrich, & M. Wessner (Eds.), Proc. der 15. E-Learning Fachtagung Informatik der g.i. (Bildungsräume DeLFI 2017) (pp. 399–401). Bonn, Germany: Bonner Köllen Verlag.

Michaelsen, L. K., Fink, L. D., & Knight, A. (1997). Designing Effective Group Activities : Lessons for Classroom Teaching and Faculty Development. To Improve the Academy, 385, 373–397.

Moodle Community Team. (2018). Privacy API - MoodleDocs. Retrieved May 15, 2018, from https://docs.moodle.org/dev/Privacy_API

Nederveen Pieterse, A., van Knippenberg, D., & van Ginkel, W. P. (2011). Diversity in goal orientation, team reflexivity, and team performance. Organizational Behavior and Human Decision Processes, 114(2), 153–164. http://doi.org/10.1016/j.obhdp.2010.11.003

Srba, I., & Bielikova, M. (2014). Dynamic Group Formation as an Approach to Collaborative Learning Support. IEEE Transactions on Learning Technologies, 8(99), 173–186. http://doi.org/10.1109/TLT.2014.2373374

Zheng, Z., & Pinkwart, N. (2014). Dynamic Re-Composition of Learning Groups Using PSO- Based Algorithms. Proceedings of the 7th International Conference on Educational Data Mining (EDM), (3), 357–358.

Attachments

intentionally blank