

Team formation instruments to enhance learner interactions in open learning environments

Citation for published version (APA):

Spiegelstra, H., Van Rosmalen, P., Houtmans, T., & Sloep, P. (2015). Team formation instruments to enhance learner interactions in open learning environments. *Computers in Human Behavior*, 45, 11-20. Advance online publication. <https://doi.org/10.1016/j.chb.2014.11.038>

DOI:

[10.1016/j.chb.2014.11.038](https://doi.org/10.1016/j.chb.2014.11.038)

Document status and date:

Published: 01/04/2015

Document Version:

Early version, also known as pre-print

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

<https://www.ou.nl/taverne-agreement>

Take down policy

If you believe that this document breaches copyright please contact us at:

pure-support@ou.nl

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 26 Sep. 2023

Open Universiteit
www.ou.nl



Team Formation Instruments to Enhance Learner Interactions in Massive Open Online Courses

Howard Spoelstra, Peter van Rosmalen

(Welten Institute, Open University of the Netherlands, Heerlen, the Netherlands
{howard.spoelstra, peter.vanrosmalen}@ou.nl)

Tilly Houtmans

(Faculty of Psychology and Educational Sciences, Open University of the Netherlands, Heerlen, the Netherlands
tilly.houtmans@ou.nl)

Peter Sloep

(Welten Institute, Open University of the Netherlands, Heerlen, the Netherlands
peter.sloep@ou.nl)

Abstract

Current open learning environments such as Massive Open Online Courses often show a lack of learner collaboration possibilities and high levels of drop-out. Introducing project-based learning can enhance learner collaboration and motivation. Project-based learning requires extensive support from expert teachers and therefore does not easily scale up into Massive Open Online Courses. Team formation instruments are introduced, aimed at supporting teachers and learners in defining and staffing projects. These consist of team formation principles and algorithms to form productive, creative, or learning teams. They use data on the project and on learner knowledge, personality and preferences to propose teams. A study was carried out to validate the team formation principles and the results from the algorithms. The data were provided by Bachelor students Psychology and master students Learning Sciences (n=168) and processed by the algorithms. By means of a survey among human assessors (n=56), the instruments were validated. The principles for learning teams and productive teams were accepted, while the principle for creative teams was not. The algorithms were validated using team classifying tasks and team ranking tasks. Human assessors classify and rank small productive, creative and learning teams in accordance with the algorithms. This indicates that the algorithms differentiate effectively and in line with human assessors between teams with high or low fit to a team formation principle. Results also shows that forming teams quickly becomes complex when team size and the number of topics in a project increase. The article closes with a discussion of the results, conclusions, and directions for future research.

Keywords: Open learning environments, MOOC, Social Learning Network, Project-based learning, Project team formation.

1. Introduction

Open learning environments, such as Massive Open Online Courses (MOOCs), currently attract large bodies of learners. Initially these environments were envisioned to provide learning settings based on the pedagogical vantage point of networked learning, with a strong emphasis on learner self-direction and learner contribution. Downes (2006) and Siemens (2004) coined the term “connectivism” to label such learning settings. In parallel a different kind of MOOC rose to attention, one that builds on behaviourist rather than social-constructivist principles. Reports, however, from both learners and MOOC providers indicate that dropout rates from MOOCs are massive, and that they offer limited opportunities for learner collaboration. (Daniel, 2012; Edinburgh University, 2013, Morrison, 2013; McGuire, 2013). While there are many reasons for drop-out rates being high, these effects can at least also partly be explained by learning settings that do not motivate learners. In the up till now smaller scale connectivist MOOCs learners are expected to be self-directing, which can present learners with difficulties (Kop et al., 2011). In the current large scale behaviourist MOOCs, scaffolding, teacher-learner contacts and collaborative learning opportunities are limited, which leads to sub-optimal learning (Daniel, 2012; Edinburgh University, 2013). To address some of these issues, recent initiatives (Stanford University, 2012; NovoEd, 2014) seek to support collaboration between learners by offering learner group formation support and group facilities in MOOCs. This does in effect introduce small-scale (group-based) learning settings in large-scale MOOCs. In e.g., NovoEd (2014), learners can self-select or receive recommendation for a small set of co-learners, based on e.g., geographical locations and common language.

In this article we discuss an approach to team formation that aims to go beyond self-selection or relatively simple criteria such as location and language. It builds on extensive research in team formation and team-based learning (Spoelstra et al., 2013; Spoelstra, van Rosmalen & Sloep, 2014) and aims to be utilizable in a large scale environment such as a MOOC. We propose to implement the well-researched team-based learning settings of project-based learning (PBL) in MOOCs. This will address several of the issues outlined above: First, besides playing a role in learner retention (Dahms and Stentoft, 2008; Fisher and Baird, 2005), PBL improves the learners' motivation, so that learners are more inclined to deal with hard, complex problems and spend more time studying (Johnson, Johnson, Stanne and Garibaldi, 1990; Marin-Garcia and Lloret, 2008). Second, PBL blends learning and working, thereby creating realistic (inter-professional) learning experiences (Springer, Stanne and Donovan, 1999; Felder, Felder and Dietz, 1999) which prepare learners for real life working conditions. And third, generally speaking, collaboration between learners as envisioned in PBL has been shown to lead to an increase in learning outcomes, when compared to individual learning (Hsiung, 2010). Implementing PBL in traditional educational settings requires expertise from teachers for defining project tasks and staffing them. However, as in large scale MOOCs staff burden needs to be kept down, we propose that learners themselves will play an active role in defining projects for PBL. Learners who are enabled to *self-define tasks* develop a motivating sense of ownership and responsibility for their learning processes. At the same time, however, *self-selection of teams* ought to be discouraged. Fiechtner and Davis (1985), Oakley, Felder, Brent and Elhadj (2004) hold that for teams to be effective, team formation should be performed by experts. These experts use knowledge of the project tasks and of the prospective team members to form teams (Oakley, Felder, Brent and Elhadj, 2004; Obaya, 1999). In large scale MOOCs however, a complicating factor is that these experts will most probably not be available. Therefore we argue that if large groups of learners in MOOCs are to be enabled to self-define project tasks and to receive effective team formation suggestions, we need to develop *automated support services*. These need to mimic expert behaviour in assessing whether projects fit with the MOOC learning materials and form teams based on task and team member characteristics, beyond language and geographical location. We therefore develop an *automated service for learner-induced project definition and team formation in MOOCs*. This service is based on an analysis of team formation expert behaviour and provides intelligent team formation principles. In earlier work we already inferred several team formation principles from team formation theory and developed the corresponding team formation algorithms to implement these principles (Spoelstra et al., 2013). It is our future goal that these instruments will be able to assess whether suggested projects qualify for execution inside a MOOC and to form effective project teams based on important factors in team formation, such as knowledge, personality, and preferences (Oakley, Felder, Brent and Elhadj, 2004; Obaya, 1999).

In this article, however, we focus on the *validation* of the set of instruments we developed. First, we aim to validate the team formation principles we inferred. Second, we aim to validate their implementation in algorithms, using real world learner data for their input.

The remainder of the article is structured as follows: In section 2 we present a team formation model, which uses learner knowledge, personality and preferences to suggest teams fit for executing a project. In section 3, we present the research questions and hypotheses, on the basis of which we aim to validate the team formation instruments. Section 4 describes the materials and methods we used to test the hypotheses. In section 5, the results are presented. Section 6 provides an extensive discussion of these results, while in section 7 we draw conclusions and suggest future research.

2. A team formation model

The automated service builds on earlier work in which we introduced a team formation model for use in open learning environments such as MOOCs, Social Learning Networks, and in more traditional learning settings. The model was constructed based on a review of team formation literature. It aims to mimic the behaviour of team formation experts (i.e., use knowledge on task and team members to form teams fit for various tasks) (See Figure 1) (Spoelstra et al., 2013).

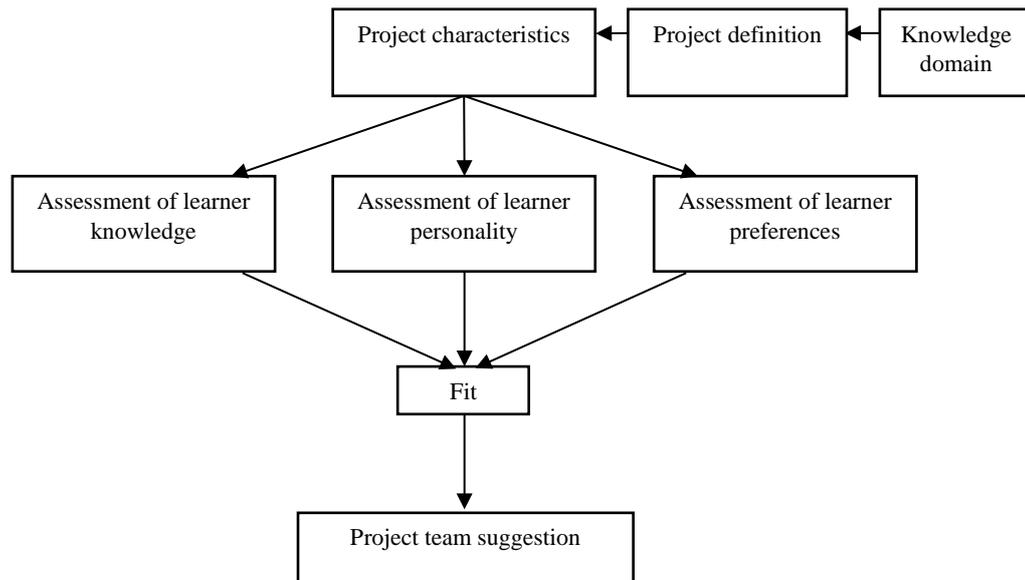


Figure 1: The team formation model

This model describes the definition of a project (a task addressing multiple topics carried out by multiple learners) in a knowledge domain, after which assessments of learner *knowledge*, *personality* and *preferences* occur, in order to determine a fit-value for a team of learners for the project. The factor preferences denotes conditions for collaboration (e.g., available time slots or languages spoken). These determine whether a project can take place at all with a particular team of learners. From this it follows that the *first* step in the chronology of the team formation process is finding overlapping sets of preferences by comparing the project characteristics and learner preferences. This limits the number of learners from which teams can be formed. In the *second* step, the assessment of knowledge is used to match the knowledge required for executing the project to the knowledge the prospective team members can provide. The factor personality is represented by the personality trait “Conscientiousness”, which is assessed with the Big Five personality test (Barrick and Mount, 1991). This particular trait is chosen because of its ability to predict job performance (Jackson et al., 2010; George and Zhou, 2001; Barrick & Mount, 1991) and because of its ability to predict a person’s creativity. (George and Zhou, 2001; Wolfradt and Pretz, 2001). This enabled us to use just two dimensions (topic knowledge and conscientiousness) to form teams fit for various team tasks. In the *third* step the resulting data are combined to determine the fit of a team of learners to a task and to suggest project teams. We discern between three common types of project tasks: 1) expertly and productively working on a project, 2) creatively solving a project problem, and 3) sharing knowledge (teach and learn) with fellow team members while solving a project problem. Based on team formation theory, in earlier work (Spoelstra et al, 2013), we inferred three team formation principles that vary on the dimensions knowledge and conscientiousness. Each principle is directed at optimising the team formation process toward one of these three types of tasks. In the next subsection we present these team formation principles.

2.1 Team formation principles

The team formation principles we aim to validate are the following:

- 1.) The team formation principle for productive teams: *"Productivity in a team is fostered when team members have high scores on knowledge of the project topics and the team members show high, homogeneous levels of conscientiousness"*.
- 2.) The team formation principle for creative teams: *"Creativity in a team is fostered when team members have differentiated scores on knowledge of the project topics and the team members show low levels of conscientiousness."*
- 3.) The team formation principle for learning teams: *"Learning in a team is facilitated when knowledge on the project topics is distributed over the members (allowing each member to learn and teach). However, the differences in knowledge should not be too high, and the team members should show high levels of conscientiousness."*

These team formation principles, through their corresponding expressions (see Appendix A), were implemented in computer algorithms. The implementation of the algorithms for productive and creative teams was rather straightforward: the algorithm for productive teams favours teams with members that all have high knowledge scores and high conscientiousness scores. The algorithm for creative teams favours teams in which topic knowledge is maximally diversified over team members, who also have low conscientiousness scores.

The team formation algorithm for learning teams, however, is more complex as:

1) It models one of the aspects from Vygotsky's principle of "zone of proximal development": *difference in knowledge between learners*. This aspect is expressed in the parameter "zpd", which puts a limit on the knowledge differences allowed between team members. From this difference it calculates teaching and learning effectiveness between each team member inside each project topic. It currently follows a 10-point grading system (grades range between 1 and 10, with 10 being the highest possible grade while 6 is considered to be the passing mark). For the current experiment the value of the parameter "zpd" is set to 3.

2) It implements a minimum knowledge level, below which teaching is assumed to be unwanted, as the member considered for the teaching role is assumed not master the topic sufficiently well. This value is set to 6.

Please note that this means that the algorithm assumes that learner/peer-tutor pairs with topic knowledge grades of 7 and 10, 6 and 9, 5 and 8, 4 and 7, and 3 and 6 can all learn effectively, provided they also have high scores on conscientiousness. In practice a pair with a smaller knowledge difference may be selected if the optimal learner/peer-tutor pairing is not available.

As indicated in the introduction, our current focus is on the validation of the team formation principles and the results from their implementation in algorithms. Therefore, in the next section we present our research questions and hypotheses.

3. Research questions and hypotheses

As we are focussed on the question whether the principles and algorithms correspond with human judgement, our two main research questions are:

(R1) *Are the team formation principles for forming productive, creative and learning teams in alignment with the opinions and experiences of practitioners from the educational field about how such teams should be formed?*

(R2) *Are the results from the computer algorithms in alignment with the results of practitioners performing the same task?*

To answer research question 1 (R1), we put forward the following hypothesis:

(H1) *Practitioners from the educational field agree that the three individual team formation principles for productive, creative and learning teams lead to the formation of teams fit for their associated tasks.*

We consider H1 to be accepted when the practitioners agree that each of the three individual team formation principles lead to the formation of teams fit for their associated tasks (i.e. when practitioners answer "agree" or "strongly agree" on a five-point Likert scale).

Research question 2 (R2) will be answered by evaluating the results (i.e. team formation suggestions) of an implementation of the team formation expressions in algorithmic form, applied to real world learner data. We draw up the following connected hypotheses:

(H2a) *Given the same data as used by the algorithms, practitioners classify the teams in accordance with the algorithms.*

(H2b) *Given the same data as used by the algorithms, the practitioners rank the teams in accordance with the algorithms.*

H2a will be accepted when the practitioners classify teams in accordance with the team formation principles, while H2b will be accepted when the practitioners rank teams in accordance with the team formation principle. Due to the complexity of the tasks, we assume that human performance will be effected negatively when the tasks get more complex. Hence, related to R2, we explore whether the performance of practitioners on

classification and ranking is effected when we increase both the size of the teams and the numbers of topics the project addresses.

4. Materials and Method

For the experiment a representative set of real world learner data on knowledge, conscientiousness and preferences was required. This set was gathered by means of an online survey which is described in subsection 4.1. The survey was taken from learners from the School of Psychology and from the Master Educational Sciences of the Open University of the Netherlands. It was conducted in the Dutch language. The data gathered could then be processed with the team formation algorithms. Their output is described in subsection 4.2. The method applied for the experiment is described in subsection 4.3.

4.1 Learner data

Part 1 of the learner inquiry gathered learner demographics (gender, age, etc.). In total, 168 complete responses were gathered. Of the respondents, 31 participants were male, 137 female. Stratification over age groups was as follows: 20-29 (24), 30-39 (46), 40-49 (58), 50-59 (37) and 60-69 (3). Of these, 121 learners studied at the Psychology faculty, while 47 learners studied Learning Sciences.

In part 2 we determined learner *conscientiousness* scores. To that end we presented the learners with a Big Five personality test (Barrick and Mount, 1991), containing 44 questions. We used the Dutch translation of the test (Denissen, Geenen, van Aken, Gosling, Potter, 2008). The test assessed all Big Five personality aspects (Extraversion, Neuroticism (vs. Emotional Stability), Conscientiousness, Agreeableness, and Openness to Experience). The learners' conscientiousness scores varied between 2.00 and 4.56. The reliability scores (rounded to two significant decimals) for the five factors of the test were: Extraversion (.81), Agreeableness (.74), Conscientiousness (.84), Neuroticism (.85), and Openness (.86). These results are fully in line with an earlier validation of the BFI in the Dutch language.

In part 3 of the survey we asked learners to self-rate their *knowledge* on four topics that were addressed in courses on research methods and techniques. The topics were: 1) Defining research questions and theoretical designs for a study, 2) Gathering data, 3) Analysing data, and 4) Discussing and concluding on results. Following a 10-point grading system (with grades ranging from 1-10, with 10 as highest grade), the self-reported scores on the topics ranged between 3-9, 1-10, 1-10 and 1-9, respectively.

Part 4 asked the learners about their project work *preferences*, such as their preferred collaboration languages, the time slots in which they were available for collaboration (in the morning, and/or in the afternoon, and/or in the evening for every week day and the weekend as a whole) and the total number of hours they had available for collaboration weekly. Additionally, learner time zone information was gathered to be able to adjust for time zone related availability mismatches. As indicated in Section 2, learner preferences effectively filter the number of possible team members for any project. For the current experiment we filtered using the data on availability. As criterion we used learner availability on the separate days of the week and on the weekend as a whole. This resulted in 8 groups of learners. The numbers of learners available in these groups were as follows: Monday (27), Tuesday (23), Wednesday (30), Thursday (34), Friday (29), Saturday (33), Sunday (29), and the whole weekend (29). Please note that learners could be available on multiple days.

4.2 Team formation algorithms output

The knowledge and conscientiousness scores of the 8 groups of learners were processed by the team formation algorithms to form project teams with 2 members covering 2 topics (using the topic knowledge grades on topics 1 and 2), to form project teams with 3 members covering 3 topics (using the topic knowledge grades on topics 1, 2 and 3), and to form project teams with 4 members covering 4 topics (using the topic knowledge grades on topics 1, 2, 3 and 4). This resulted in a total of 8 (one for each availability slot) times 3 lists (one for each team size). Each of the 24 lists contained the fit values for the task types productive, creative, and learning.

Please note that for the remainder of this article we will refer to these teams as *2x2 teams*, *3x3 teams*, and *4x4 teams*, respectively. The number of team members and the number of project topics were chosen for the purpose of the present experiment only. The team formation algorithms did not impose these choices. The robustness of the algorithms was tested by inputting the data of all learners (n=168) and calculating of fit values for all possible 2x2, 3x3 and 4x4 teams. This resulted in text files containing team formations and fit values for 14,028, 776,216, and 32,018,910 unique teams respectively, with sizes of 1Mb, 73Mb, and 3.3Gb. On a machine with an Intel i7 CPU and with 4Gb of internal memory, the algorithms completed successfully.

4.3 Method

The validation of the hypotheses described in Section 2 was conducted by means of an online survey. We invited all members ($n=405$) of the teaching staff of our university to participate on a voluntary basis. In total 56 respondents completed the survey. Of these, 26 were female, while 30 were male. Of the participants 14 had experience in forming teams. The distributions over age groups was: 20-29 (4), 30-30 (9), 40-49 (10), 50-59 (20), 60-69 (12), 70-79 (1).

In order to test hypothesis 1 (H1: team formation principles), in three separate questions the respondents were presented with the three team formation principles. They were asked whether they agreed whether applying the principle would lead to the formation of teams fit for the type of task the principles described. (cf. R1). The questions could be answered on a 5-point Likert scale, with the answer options “strongly disagree”, “disagree”, “neither agree nor disagree” “agree”, and “strongly agree”.

In order to test hypothesis 2a (H2a: classifying teams) we first presented the participants with a preparatory question with a near perfect 2x2 team for each of the 3 team formation principles, isolated from each other. The examples elaborately explained the application of the team formation principles and asked the participants to classify the teams. Next, from each of the calculated lists of teams for each task type (productive, creative, and learning) and each team size (2x2, 3x3, and 4x4) we selected the three highest scoring teams. We randomly grouped the teams of equal size into 3 sets of 3 teams. The 9 sets (each containing 3 samples) were presented sequentially to the respondents in order of increasing team size. The respondent’s task was to classify the teams as examples of either productive, creative or learning teams. The participants were instructed to only give “No answer” for their answer if they could not decide on one type.

In order to test hypothesis 2b (H2b: ranking teams), we first presented the respondents with an example of the task. Next, from each of the calculated lists of teams for each task type (productive, creative, and learning) and each team size (2x2, 3x3, and 4x4) we selected the teams with the three highest, three most average and three lowest fit values on productivity, creativity, and learning. We randomly ordered the teams of equal task type and equal size into 9 sets of 3 teams. The 9 sets were presented separately to the respondents. We started with all 2x2 teams of the individual task types (productive, creative, and learning) and then, while keeping this order, increased team size to 3x3, and finally to 4x4. The respondent’s task was to rank the teams in accordance to their assessment of the level of adherence of the teams shown to the current task type. The final question in the survey invited the participants to comment on the survey and their tasks.

5. Results

The results of the survey among teaching staff ($n=56$) are presented in the order of the research questions as stated in Section 3.

5.1 Team formation principles

Our participants were asked to indicate their level of agreement on whether each of the team formation principles would lead to the formation of teams fit for the task type. The results are presented in Figure 2.

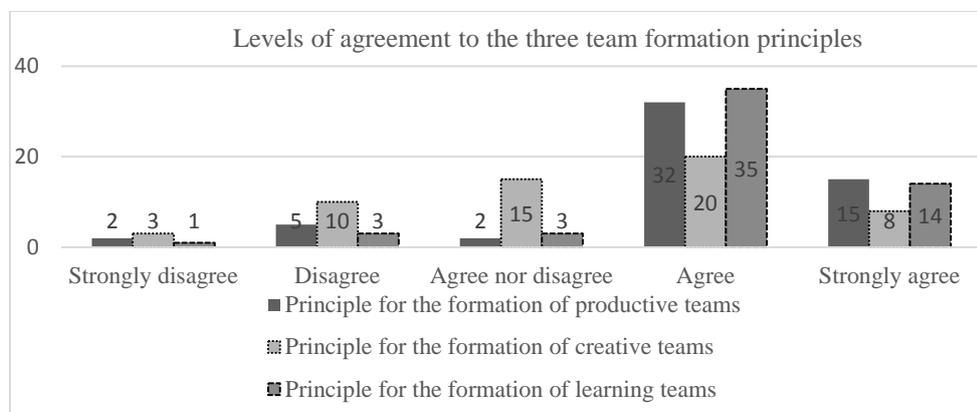


Figure 2: Levels of agreement to the three team formation principles on a 5-point Likert scale.

The number of participants indicating agreement and strong agreement to the principles for the formation of productive, creative, and learning teams totalled to 47 (84%), 28 (50%) and 49 (88%), respectively.

5.2 Classifying teams

As indicated in subsection 4.3 above, the actual task of classifying teams was preceded by a preliminary question. It presented constructed examples adhering well to the team formation principles. (See Tables 1, 2 and 3.)

Tables 1, 2, and 3: Examples of productive, creative and learning teams with two members (L1 and L2), each having two scores on knowledge topics (T1 and T2) and one score on conscientiousness (Cons).

Productive	T 1	T 2	Cons	Creative	T 1	T 2	Cons	Learning	T 1	T 2	Cons
L 1	8	9	4.32	L1	8	4	1.67	L 1	9	6	4.33
L 2	9	8	4.78	L2	3	9	2.11	L 2	7	9	4.45

For these teams the team formation algorithms calculated fit values of 0.880, 0.493, and 0.939, respectively. The numbers of participant's classifying these team in line with the algorithms were: 52 (93%), 54 (96%), and 48 (86%), respectively. In 4 cases the productive team was alternatively classified as a learning team. In 2 cases the creative team was alternatively classified as a learning team. The learning team was in 2 cases alternatively classified as a creative team, and in 6 cases as a productive team.

The next 3 questions asked participants to classify teams of size 2x2. The cumulative results from these tasks are shown in Figure 3. The labels on the vertical axis indicate for which classification these team had the highest fit values. As all three types of teams were shown 3 times, the total number of answers on any type of team is 168 (3x56).

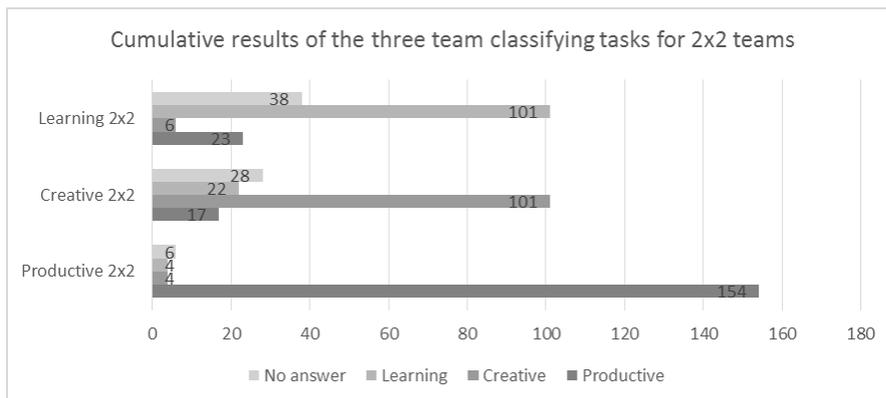


Figure 3: Cumulative results of classifying three 2x2 teams of each type of team.

The participants classified these 2x2 teams in accordance with the team formation algorithms as follows: Productive 154 (92%), Creative 101 (60%), and Learning 101 (60%).

Additionally, we explored to what extent this task becomes more complex when both team size and number of topics addressed in the project increased. Therefore the next 6 questions we asked to classify 3x3 and 4x4 team respectively. Figures 4 and 5 depict the results of these tasks.

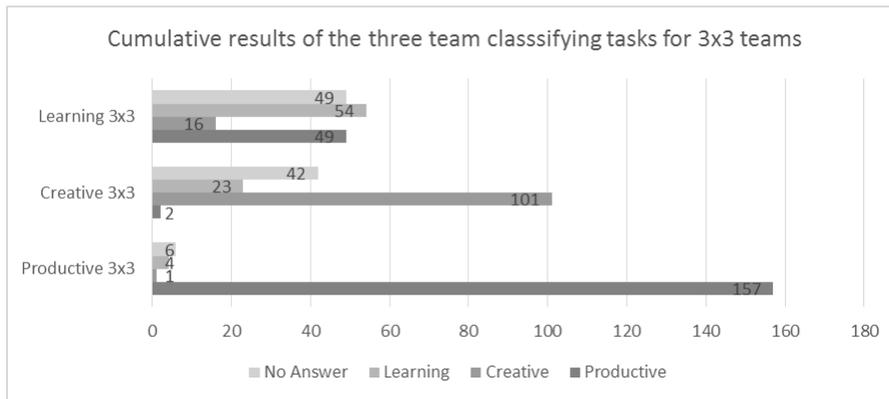


Figure 4: Cumulative results of classifying three 3x3 teams of each type of team.

The participants classified these 3x3 teams in accordance with the team formation algorithms as follows: Productive 157 (93%), Creative 101 (60%), and Learning 54 (32%).

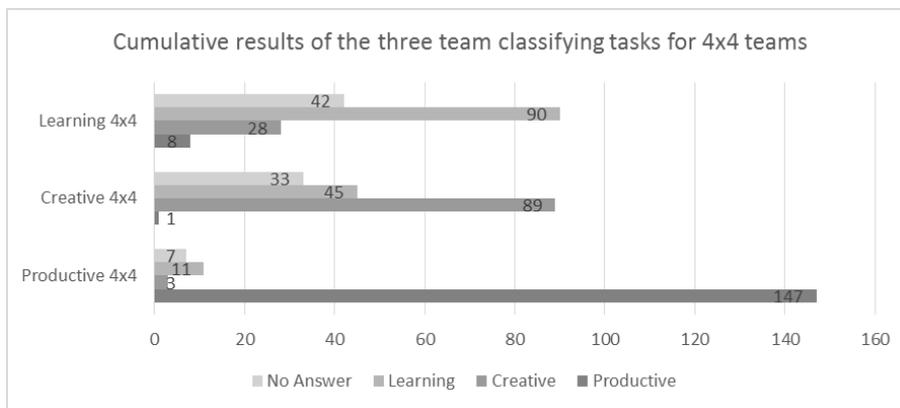


Figure 5: Cumulative results of classifying three 4x4 teams of each type of team.

The participants classified the 4x4 teams identical to the team formation algorithms as follows: Productive 147 (88%), Creative 89 (53%), and Learning 90 (54%). To allow for easy comparing, the collective results of the 2x2, 3x3 and 4x4 classifying tasks are displayed in Table 4.

Table 4: Percentages of classifications identical to the team formation algorithms for three types of teams of sizes 2x2, 3x3, and 4x4, including “no answers”

	Productive	Creative	Learning
2x2	92 %	60 %	60 %
3x3	93 %	60 %	32 %
4x4	88 %	53 %	54 %

In Table 5 we present a breakdown of these results into team sizes, numbers of classifications identical to the algorithm results and numbers and kinds of alternative classifications, including the number of no-answers. The cells in the diagonal from upper left to lower right for each team size represents the classification that aligns with the one calculated by the algorithm, while the other cells in each row represent the type and number of the alternative classifications.

Table 5: Numbers of identical and alternative classifications, and no answers for the three types of teams of size 2x2, 3x3, and 4x4.

n=168	Productive	Creative	Learning	No answer
Productive 2x2	154	4	4	6
Creative 2x2	17	101	22	28
Learning 2x2	23	6	101	38
Productive 3x3	157	1	4	6

Creative 3x3	2	101	23	42
Learning 3x3	49	16	54	49
Productive 4x4	147	3	11	7
Creative 4x4	1	89	45	33
Learning 4x4	8	28	90	42

Table 6 shows the percentages of identical classifications, excluding the “no answers”.

Table 6: Percentages of classifications identical with the team formation algorithms for three types of teams of sizes 2x2, 3x3, and 4x4, excluding “no answers”

	Productive	Creative	Learning
2x2	92 %	72 %	78 %
3x3	97 %	80 %	45 %
4x4	91 %	66 %	71 %

5.4 Ranking teams

Our participants ranked teams based on how well they adhered to each individual team formation principle. They did this for 9 sets of 3 teams in the order productive, creative, and learning and with increasing team size. Figure 6 shows the results of these ranking tasks.

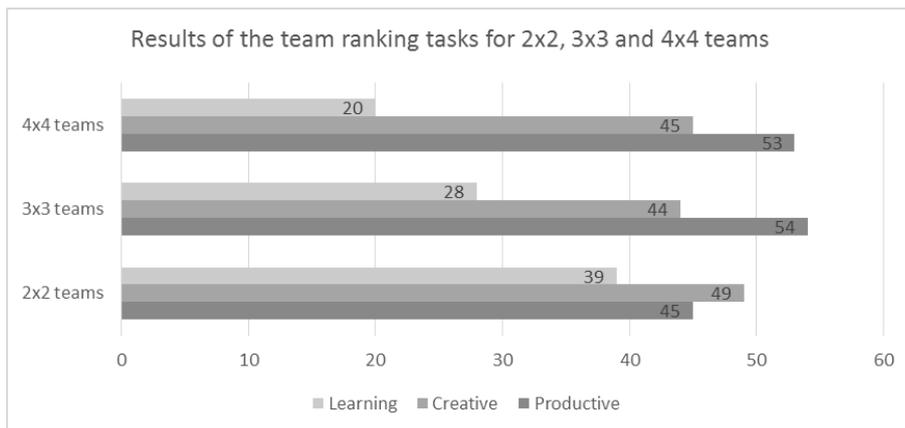


Figure 6: Numbers of rankings of team sizes 2x2, 3x3, and 4x4 in accordance with the ranking from the team formation algorithms for each type of team.

The numbers and percentage of rankings matching the ranking from the team formation algorithms for the 2x2 productive teams were 45 (80%), for the creative teams 49 (88%), and for the learning teams 39 (70%).

The collective results are displayed in Table 7, which allows easy comparing.

Table 7: Numbers and percentages of rankings of three types of teams in line with the ranking from the team formation algorithms for team sizes 2x2, 3x3 and 4x4.

n=56	Productive	Creative	Learning
2x2 teams	45 (80 %)	49 (88 %)	39 (70 %)
3x3 teams	54 (96 %)	44 (79 %)	28 (50 %)
4x4 teams	53 (95 %)	45 (80 %)	20 (36 %)

Our exploration into the complexity of the team formation task showed that the numbers of rankings identical to the results from the team formation algorithms for the 3x3 productive teams were 54 (96%), creative teams 44 (79%), and learning teams 28 (50%). The results for the 4x4 productive teams matching the rankings from the team formation algorithms were 53 (95%), creative teams 45 (80%), and learning teams 20 (36%). Both these sets of results are also shown in Figure 6 and Table 7. We observe that while the numbers of rankings of productive and creative teams in accordance with the rankings from the team formation algorithms remained

roughly on the same level, the number of matching rankings of learning teams showed a considerable drop (from 70 % to 36 %).

5.5 Comments on the survey

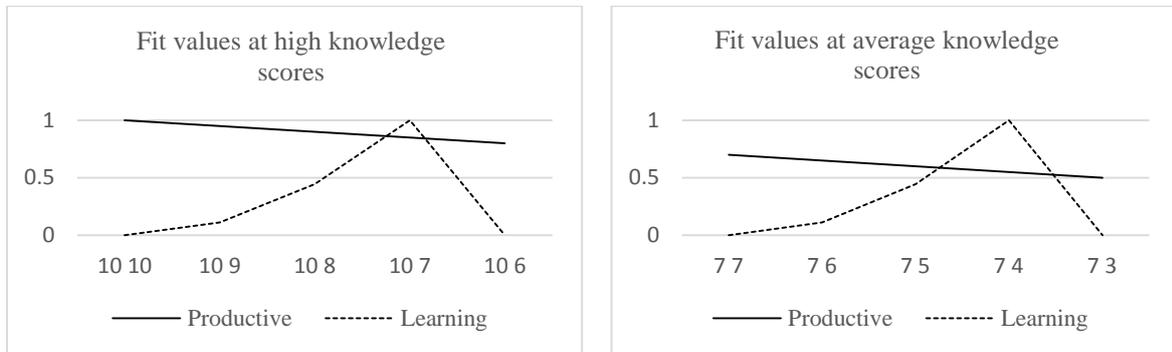
The most relevant comments given were as follows: One participant noted that using only conscientiousness as a personality factor would underrepresent personality in the team formation principles. More specific, in line with the low acceptance rate of the principle for forming creative teams, some participants remarked that creativity cannot effectively be described with only knowledge and conscientiousness as factors. Another participant remarked that the survey was difficult to answer, mostly because of having to take into account the use of Vygotsky' zone of proximal development in the team formation principle for learning teams. More specific, several participants remarked on the heavy mental load the survey put on them.

6. Discussion

Given the results of the validation of our first hypothesis (84%, 50% and 88% of the participants agreed with the proposed principles for the formation of productive, creative, and learning teams, respectively), the hypothesis was accepted for productive and learning teams. The principle for the formation of creative teams was not accepted. Likely, this is due to the use of only two factors in the team formation principles: knowledge distributions and conscientiousness levels. This can also be surmised from the comments on the survey. Even though there is evidence for a relationship between conscientiousness and creativity (see e.g., Robert & Cheung, 2010, who show that there is a significant negative relationship between group conscientiousness and group performance on a creative task), this apparently is a too narrow a basis for the formation of creative teams. It may be necessary to include additional personality factors, for instance based on Barrick and Mount (1993), who find a relation between creativity, openness to experience (one of the Big Five personality factors) and job performance. One may also have to take into consideration the sub-factors of which conscientiousness is made up. Research by e.g., Barrick, Mount and Strauss (1991) and Reiter-Palmon, Illies and Kobe-Cross (2009) indicates that conscientiousness consists of two components: an achievement component (consisting of the facets competence, achievement striving and self-discipline) and a dependability component (consisting of the facets order, dutifulness and deliberation). Reiter-Palmon et al. (2009) and Kaufman (2011) argue that the achievement component is related to creative job performance, while the dependability component is not. While some researchers suggest that all humans have creative ability, but with different styles and levels (e.g., Kirton, 2003), others search to define factors above and beyond conscientiousness and openness to experience that determine a person's creativity (Cambridge Handbook of Creativity, 2010). Both of these approaches can inform future research for a better delineation of what makes a team creative and, in case, how to form a creative team. The validation of the implementation of the principles in algorithms was conducted by means of classifying and ranking tasks. The classifying task aimed at confirming hypothesis 2a (Given the same data as used by the algorithms, practitioners classify the teams in accordance with the algorithms). The results from the preparatory question with well-formed examples show that participants were able to perform these tasks well when isolated from each other. As the principle for the formation of creative teams was rejected, we focus this part of the discussion on the classification and ranking of productive and learning teams. The results obtained from the actual classifying tasks show that participants could classify productive teams of all sizes successfully. With regard to the learning teams, the results show a fair amount, i.e. 60%, of classifications align with the algorithm for 2*2 teams. At the same time it shows that the participants find this task difficult, and even more so for the 3*3 and 4*4 teams. Table 5 shows that both the numbers of alternative classifications and no answers increase when the team size and the number of topics addressed in the project rise. This can be explained by e.g., the phenomenon of bounded rationality, in which time constraints and limited human power of abstraction hinder the rational decision making process (Gigerenzer and Selten, 2002). Several aspects of the classifying task relate to this phenomenon:

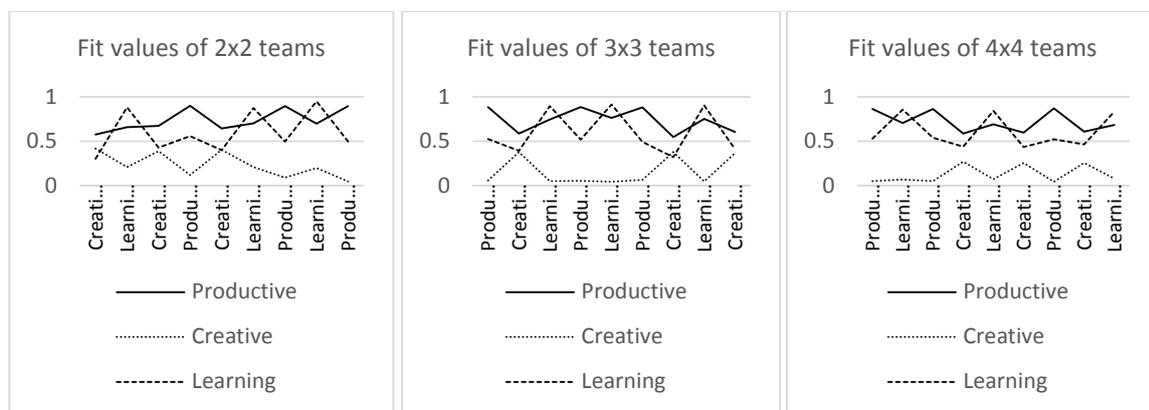
1. The participants had to mentally apply the variables from all three team formation principles to the teams shown and consider each result to come to a classification. For both productive and creative teams only two basic rules applied, while for learning teams considerably more rules had to be taken into account to decide on a classification.
2. The number of team members and topics addressed in the project increased. This had the effect that overall the numbers of classifications in accordance with the team formation algorithm results declined (table 1).

3. The participants were allowed to indicate that they could not come to a conclusion by selecting “no answer”. As can be seen in Table 5, the numbers of “no answers” were highest for learning teams of any size. From Table 5 it can also be observed that overall the numbers of “no answer” increased as soon as the team size was larger than 2x2.
4. In specific circumstances the difference in fit to a principle between productive and learning teams is minimal: The productive teams receive high fit values when knowledge scores are high, while the learning team receive high fit values when knowledge scores have an optimum difference, so also when knowledge scores are high. For the algorithms even a very small difference is sufficient to make a distinction, obviously this is not the case for the participants. This effect is demonstrated in the Figures 7 and 8.



Figures 7 and 8: Fit values from the productive and learning algorithms for pairs of knowledge scores of two team members on one topic, for high and average knowledge scores.

It then depends on the data available whether the examples drawn from it provide sufficient basis for clearly distinctive teams. As is shown in the Figures 9, 10, and 11 the differences between productive and learning teams was often limited, specifically with the 3x3 teams.



Figures 9, 10 and 11: Fit values for the productive, creative and learning teams shown in the classifying part of the survey.

This effect is reflected in the classifications given by the participants for 3x3 learning teams (see Tables 4, 5, and 6). With respect to hypothesis 2a (Given the same data as used by the algorithms, practitioners classify the teams in accordance with the algorithms) this indicates that by and large human assessors do classify teams identical to the classifications from the algorithms.

The ranking task showed similar, but not as many, complicating aspects. In this task only one team formation principle had to be considered at any time, but the team size and numbers of topics increased. The results presented in Table 7 reflect the complexity of the tasks, as the numbers of rankings identical to the ranking from the team formation algorithms for the principle for productive and creative teams remained roughly on the same level, while the numbers of rankings of the learning teams dropped from 70% to 50% to 36% when the team size went up. With respect to hypothesis 2b (Given the same data as used by the algorithms, the practitioners rank the teams in accordance with the algorithms) this indicates that human assessors rank teams in accordance with the team formation principles. But again, for learning teams, human performance drops considerably when the task complexity rises.

7. Conclusions and directions for future research

This article discussed and investigated an automated project-based learning and team formation service for learners in open learning environments, such as MOOCs. It provided several reasons why this can be beneficial for both learners and support staff. Among them are the benefits of collaborative learning with respect to motivation, and therewith drop-out. As providing support for such learning settings in open access learning environments can be resource-intensive for staff, we introduced several instruments to implement automated project definition and team formation. These instruments exist of team formation principles and algorithms for the formation of productive, creative, and learning teams. The algorithms use data on learner knowledge, personality and preferences to form teams. It was stressed that first and foremost such an implementation requires validated instruments. We therefore presented an experiment which validated the principles and the results from the team formation algorithms. Participants were recruited from practitioners in educational field. The results from an acceptance test of the three team formation principles demonstrated that the team formation principles for the formation of productive teams and for the formation of learning teams received wide support. We discussed the possible steps to take to further refine the principle for the formation of creative teams. In order to make sure the team formation algorithms could perform in settings with large groups of learners we gathered a large amount of real world data on learner knowledge, personality and preferences ($n=168$). The processing of these data proved the algorithms' robustness. The validation of the results of the team formation algorithms followed a dual approach. Based on their understanding of how the team formation principles work, practitioners first classified teams into three classes (productive, creative and learning teams). Next, they ranked teams within one type of team. The results of these classifying and ranking tasks showed that the participants classified small productive, creative and learning teams largely in accord with how our algorithms judged these teams. However, with increasing task complexity, especially when learning teams were concerned, increased divergence occurred between classifications and rankings resulting from the application of the algorithms and classifications and rankings performed by human assessors. As the team formation principles for productive and learning teams were accepted and thus form a validated basis for team formation, we take this as an indication of the usefulness of automating the team formation task. Our overall conclusion is that we believe we found clear support for both the team formation principles for productive and learning teams and the correct implementation of these principles in our team formation algorithms.

In order to service the learning aspects of MOOCs, our future research will first focus on the implementation of the project-based learning and team formation instruments. Personality tests suitable for our purpose are publicly available and we can easily implement a preferences filter. We therefore focus on the automated assessments of both the fit of proposed projects to the knowledge domain, and of the knowledge available with learners. As we aim to implement the service in learning settings, we will restrict ourselves to forming learning teams only. We will investigate whether the formation of teams following the principle for learning teams can be proven to foster learning.

References

- Barrick, M. R., Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis, *Personnel Psychology*, 44,, 1–26.
- Barrick, M. R., Mount, M. K., Strauss, J. P. (1993). Conscientiousness and performance of sales representatives: Test of the mediation effects of goal setting. *Journal of Applied Psychology*, 78, 715 – 722.
- Cambridge handbook of creativity (2010). James C. Kaufman, J.C., Sternberg, R.J. (Eds). Cambridge University Press, New York.
- Dahms, M-L., Stentoft, D. (2008). Problem Based Learning in Engineering Education: a Development Option for Africa? Paper presented at The 4th African Regional Conference on Engineering Education (ARCE-2008): 'Capacity Building in Engineering Education for Sustainable Development', Dar es Salaam, Tanzania, United Republic of.
- Daniel, J. (2012). Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility, *Journal of Interactive Media in Education*, www-jime.open.ac.uk. Retrieved June 3, 2013, from <http://www-jime.open.ac.uk/jime/article/viewArticle/2012-18/html>
- Denissen, J. J. A., Geenen, R., van Aken, M.A., Gosling, G., Samuel, D., Potter, J. (2008). Development and Validation of a Dutch Translation of the Big Five Inventory (BFI), *Journal of Personality Assessment*, 90:2, 152 – 157.
- Downes, S. (2006): Learning networks and connective knowledge. Retrieved June 3, 2013, from <http://it.coe.uga.edu/itforum/paper92/paper92.html>

- Edinburgh University (2013). MOOCs @ Edinburgh 2013 - Report #1. Retrieved June 5, 2013, from <http://www.era.lib.ed.ac.uk/bitstream/1842/6683/1/Edinburgh%20MOOCs%20Report%202013%20%231.pdf>
- Felder, R. M., Felder, G. N., Dietz, E. J. (1999). A longitudinal study of engineering student performance and retention v. comparisons with traditionally-taught students, *Journal of Engineering Education*, 87, 4 (1999), 469–480.
- Fiechtner, S. B., Davis, E. A (1985). Why some groups fail: A survey of students' experiences with learning groups, *The Organizational Behavior Teaching Review*, 9(4), 75–88.
- Fisher, M., Baird, D.E., (2005). Online learning design that fosters student support, self-regulation, and retention. *Campus-Wide Information Systems*, 22, 2, pp.88 - 107
- George, J. M., Zhou, J. (2001). When openness to experience and conscientiousness are related to creative behavior: An interactional approach, *Journal of Applied Psychology*, 86, 3, 513–524.
- Gigerenzer, G., Selten, R. (2002). *Bounded Rationality*. Cambridge: MIT Press.
- Hsiung, C. M. (2010). An experimental investigation into the efficiency of cooperative learning with consideration of multiple grouping criteria, *European Journal of Engineering Education*, 35 (6), pp 670–692.
- Jackson, J., Wood, D., Bogg, T., Walton, K., Harms, P., Roberts, B. (2010). What do conscientious people do? Development and validation of the Behavioral Indicators of Conscientiousness (BIC), *Journal of Research in Personality*, 44, 4, 501–511.
- Johnson, D.W., Johnson, R., Stanne, M.B. Garibaldi, A. (1990): Impact of group processing on achievement in cooperative groups, *The Journal of Social Psychology*, 130, 4, 507–516.
- Kaufman, S.B. (2011). Is conscientiousness compatible with creativity? *Psychology Today*, August. Retrieved November, 19 2013 from <http://www.psychologytoday.com/blog/beautiful-minds/201108/is-conscientiousness-compatible-creativity>
- Kirton, M. (2003). *Adaption—Innovation in the context of diversity and change*, London: Routledge.
- Kop, R., Fournier, H., & Mak, J. S. F. (2011). A pedagogy of abundance or a pedagogy to support human beings? Participant support on massive open online courses. *International Review of Research on Open and Distance Learning (IRRODL)*, 12(7). Last accessed on 09-06-2014 from <http://www.irrodl.org/index.php/irrodl/article/view/1041/2025>
- Marin-Garcia, J.A., Lloret, J. (2008): Improving teamwork with university engineering students: the effect of an assessment method to prevent shrinking, *WSEAS Transactions on Advances in Engineering Education*, 5, 1, 1–11.
- Morrison, D. (2013). How NOT to Design a MOOC: The Disaster at Coursera and How to Fix it. Retrieved online on 09-11-2013 from <http://onlinelearninginsights.wordpress.com/2013/02/01/how-not-to-design-a-mooc-the-disaster-at-coursera-and-how-to-fix-it/>
- McGuire, R. (2013). Building a senses of community in Moocs. Retrieved online on 09-11-2013 from <http://online.qmags.com/CPT0813/default.aspx?pg=31&mode=1#pg31&mode1>
- NovoEd (2014). <https://novoed.com/how-it-works>
- Oakley, B., Felder, R.M., Brent, R., Elhajj, I. (2004). Turning student groups into effective teams, *Journal of Student Centered Learning*, 2, 1, 10–34.
- Obaya, A. (1999). Getting cooperative learning, *Science Education International*, 10, 2, 25–27.
- Reiter-Palmon, R., Illies, J. J., & Kobe, L. M. (2009). Conscientiousness is not always a good predictor of performance: The case of creativity. *The International Journal of Creativity & Problem Solving*, 19(2), 27-45.
- Siemens, G. (2004). Connectivism: A learning theory for the digital age. *International Journal of Instructional Technology and Distance Learning*. Retrieved on June, 3rd 2013, from http://www.itdl.org/Journal/Jan_05/article01.htm
- Spoelstra, H., Van Rosmalen, P., van de Vrie, E., Obreza, M., Sloep, P.B. (2013). A Team Formation and Project-based Learning Support Service for Social Learning Networks. *Journal of Universal Computer Science*, 19, 10, 1474–1495.
- Spoelstra, H., Van Rosmalen, P., Sloep, P.B. (2014). Toward Project-based Learning and Team Formation in Open Learning Environments. *Journal of Universal Computer Science*, 20, 1, 57-76.
- Springer, L., Stanne, M.E., Donovan, S.S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology, *Review of Educational Research*, 69, 1, 21–51.
- Stanford University. (2012). Retrieved online on 29-11-2013 from <http://news.stanford.edu/news/2012/september/venture-lab-platform-091712.html>. VentureLab has since developed into NovoEd.
- Vygotsky, L. S. (1978). *Mind in society: the development of higher psychological processes*, edited by Michael Cole (et al.) Harvard University Press, Cambridge.
- Wolfradt, U., Pretz, J. E. (2001). Individual differences in creativity: personality, story writing, and hobbies, *European Journal of Personality*, 15, 297–310.

Appendix.

1. The team formation expression for productive teams:

$$FitP_i = W_K * \frac{Avg_K_i}{Max_K} + W_C * \frac{Avg_C_i}{Max_C}$$

Expression 1: Team formation expression for productive teams

Explanation of the terms used:

$FitP_i$: The level of adherence of team_i to the team formation principle for productive teams

Avg_K_i : The average of the knowledge scores of all team members over all topics addressed in the project.

Avg_C_i : The average of the conscientiousness values of all team members.

Max_K : The maximum knowledge score of a team member. Following a 10-point grading system, this value set to 10.

Max_C : The maximum conscientiousness score, calculated from the Big Five test. The maximum is 5.

W_K : The weight of the factor knowledge in the team formation expression.

W_C : The weight of the factor conscientiousness in the team formation expression.

Both weights add up to 1.0, which guarantees that the $FitP_i$ value always varies between 0 and 1. For the experiments weights were set to 0.5 each, so knowledge was of equal importance in the calculation of $FitP_i$ as was conscientiousness.

2. The team formation expression for creative teams:

$$FitC_i = W_K * \frac{\sum_j DifK_j}{n * Max_K} + W_E * \frac{\sum_t DifK_t}{k * Max_K} + W_C * \frac{Max_C - Avg_C_i}{Max_C}$$

Expression 2: Team formation expression for creative teams

Explanation of the terms used:

$FitC_i$: The level of adherence of team_i to the team formation principle for creative teams

$DifK_j$: The sum of the differences between the highest and next highest score over all members inside the respective project topics.

n : The number of members in the team.

Max_K : The maximum knowledge score of a team member. Following a 10-point grading system this is set to 10.

$DifK_t$: The sum of the difference between the highest and next highest score over all topics inside the respective member' scores.

k : The number of topics in the project

Max_C : The maximum conscientiousness score from the Big Five test. The maximum is 5.

Avg_C_i : The average of the conscientiousness values of the members of team_i.

W_K : The weight of the factor knowledge in the team formation expression.

W_E : The weight of the factor expertise in the team formation expression.

W_C : The weight of the factor conscientiousness in the team formation expression.

The three weights add up to 1.0, which guarantees that the $FitC_i$ value always varies between 0 and 1. For the experiments they were set to 0.33 each, so that knowledge inside a topic, knowledge over topics and conscientiousness were of equal importance in the calculation of $FitC_i$.

3. The team formation expression for learning teams:

$$FitL_i = W_K * \frac{\sum_t \sum_l \sum_j |DifK_{tjl}|}{d_{jt} \cdot zpd \cdot n \cdot k} + W_C * \frac{Avg_C_i}{Max_C}$$

Expression 3: Team formation expression for learning teams.

Explanation of the terms used:

$|DifK_{tjl}|$: The absolute difference between two learners' (j and l) scores inside a topic. These are summed up over all pairs of learners $_{j,l}$ inside a topic and over all topics $_t$ in the project. Topic scores can vary between 1 and 10, following a 10-point grading system.

d_{jt} : the difference between the number of times a member has a higher score and a lower score when compared to other members (i.e., the number of times a member can act as a peer-tutor or as a learner).

zpd : (zone of proximal development) (Vygotsky, 1978). The maximum difference in knowledge between learners. This value is set to 3 grade points for the current experiment.

n : the number of members in the team

k : the number of topics in the project

Avg_C_i : The average of the conscientiousness values of the members of team $_i$.

Max_C : The maximum conscientiousness score, calculated from the Big Five test. The maximum is 5.

W_K : The weight of the factor knowledge in the team formation expression.

W_C : The weight of the factor conscientiousness in the team formation expression.

The expression describes teams whose members can teach and learn to and from each other inside each topic, while having a high score on conscientiousness. It optimises the match between peer-tutors and learners in the team by modelling one of the aspects of Vygotsky's principle of "zone of proximal development": *difference in knowledge* between learner and peer-tutor. The parameter "zpd" is used to calculate teaching and learning effectiveness for the team on a topic. The algorithm implementing this expression adds two exemptions to the rule: If the difference between two topic scores is higher than the value of the parameter zpd , or when a peer-tutor has a score on a topic lower than a set minimum grade (currently set to 6), teaching and learning effectiveness for that peer-tutor/learner pair is set to be 0.

The weights W_K and W_C can be set to stress the importance of knowledge over conscientiousness in the team formation, or vice versa. The scores from the first and second part are multiplied by their weights (W_K and W_C) separately and then summed. As the two scores each result in a value between 0 and 1 and the sum of the weights always is 1, this results in a measure of fit for each team considered ($FitL_i$) between 0 and 1.