Development of a student-centered assessment: Students formulate, solve and evaluate their own questions

Cristina Manatschal

Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

SoTL Project CAS Hochschuldidaktik

ABSTRACT

Questions are central to thinking, since they enable us to understand how the answers are connected, and give rise to the next questions. Likewise, several studies indicate that engaging students into formulating good questions increases their depth and breadth of knowledge and fosters their intrinsic motivation to learn. We have therefore asked biochemistry students to develop exam questions about the topics of lectures of a third-year bachelor course. The questions had to be technically correct, unambiguously formulated, should assess understanding and help their peers improve learning. In contrast to previous studies, where these questions were mainly used as a self-assessment tool, students also had to systematically evaluate their questions in a student-conducted assessment based on the defined criteria giving rise to their final grades for the course. We accompanied the entire process and evaluated the students' motivation level and learning outcome. We found that formulating and solving questions of peers was generally very motivating, while the task of ranking questions of peers is less motivating and was perceived to be very challenging. The majority of students felt they learned at least as much about the topics as compared to learning for a traditional exam, and more than half of the students felt they even learned more. Interestingly, many students also stated that they additionally learned to think critically and see other points view, to provide own insights and solutions into a group, communicate and collaborate with others and to think more deeply and with greater complexity, all of which are key competences for a successful professional development inside and outside academia.

Keywords: student-centered assessment, questions, deep learning, motivation

INTRODUCTION

The framework of a didactical setting should enable students to achieve a deep understanding of the topics of their studies (Pellegrino & Hilton, 2013). To do so, students need to go way beyond memorizing and recalling facts; they need to understand and apply their knowledge, to be able to connect it with their previous knowledge and finally learn to independently acquire and assess new knowledge. Whereas memorizing and recalling facts is often accompanied by so called surface learning strategies, that are rather passive and include little reflection, deep learning strategies include more active and conceptional thinking approaches with a strong intention to reach a high level of understanding. In addition, students should move away from an extrinsically motivated and targeted way to learn only what is required to pass a defined course, but instead be stimulated to work in an autonomous and intrinsically motivated way.

It has long been recognized that teaching students to ask good questions is very beneficial to positively influence student's learning (Bowker, 2010) (Marbach-Ad & Sokolove, 2000). Many studies describe and evaluated settings where students developed questions during classes, and most of them confirm the positive effects. It has been shown that the task to develop questions improves student's performance on examinations, that the task is perceived as very motivating and that the students' questions provide a rich source of material to detect and discuss common misconceptions (Kerkman et al., 1994) (Sircar & Tandon, 1999) (Kolluru, 2012) (Poot et al., 2017). Furthermore, students report that the task to develop questions had changed their learning strategies towards a higher degree of reflection, and that it increased their depth and breadth of knowledge (Baerheim & Meland, 2003) (Rhind & Pettigrew, 2012).

Finally, the type of assessment can influence the students' way of learning and their future learning strategies (reviewed in (Struyven et al., 2005)). For instance, it was shown that if students expect a multiple-choice examination they perform poorly in an assay-type examination as compared to students that actually prepared themselves for an assay test. In contrasts, students preparing for an assay exam performed equally well in multiple-choice tests (Traub, R. E. & MacRury, 1990). In line with these results, many studies showed that the learning environment can stimulate students to apply deep learning approaches (reviewed in (Entwistle, 1991)).

In light of these considerations, we decided to completely replace the final written exam in a biochemistry course of the third bachelor year with an alternative assessment type where students develop their own exam questions (**Figure 1)**. The course consists of an introduction

week with lectures that provide the necessary theoretical knowledge for the practical experiments by recapitulating important concepts and introducing new methodology (**Figure 1A, week 1**). In the second week, the students formulated their own exam questions (**Figure 1B, week 2**). The questions had to be technically correct, unambiguously formulated, should assess understanding and help their peers improve learning. We believe that the task to formulate such exam questions promotes students learning in an ideal way. To develop questions that assess the understanding and not only pure factual knowledge, it is necessary to deeply engage with the study material and finally design it in a clear, unambiguous and interesting way. In most studies, the student generated questions are used as self-assessment tools. However, we considered it as very important to engage students in solving questions of their peers and provide a platform for discussion among students and with lecturers. Solving questions of peers might most likely arise additional questions regarding the understanding of the topics and also arise the need to discuss the quality of the developed questions. We therefore asked students to solve and evaluate the questions of their peers based on the previously defined criteria (**Figure 1C, week 3)**. Questions were subsequently discussed in a symposium (**Figure 1D**). The final evaluation of the questions resulted in the final grade for the course (**Figure 1E, week 4**). Student questions were also rated by the lecturers and the ratings of the lecturers accounted for one third of the grade. During the time of formulating, solving and evaluating the questions, lecturers were available for questions and discussions upon request. The students worked in small self-organized groups that got scrambled for the tasks of solving and evaluating questions.

AIM OF THE STUDY AND METHODS

In this present study, we assessed if our setting is motivating for students, and explored if we can detect motivational differences regarding the tasks of formulating, solving and evaluating questions. In a second part, we assessed which kind of strategies students apply to formulate and solve questions and how the learning in this setting compares to learning for a traditional exam. Finally, we asked students what they liked and disliked most and what they missed in the present setting and whether they would have preferred a written exam about the topics. We assessed all these aspects using an online questionnaire at the end of the course (**Appendix 1).**

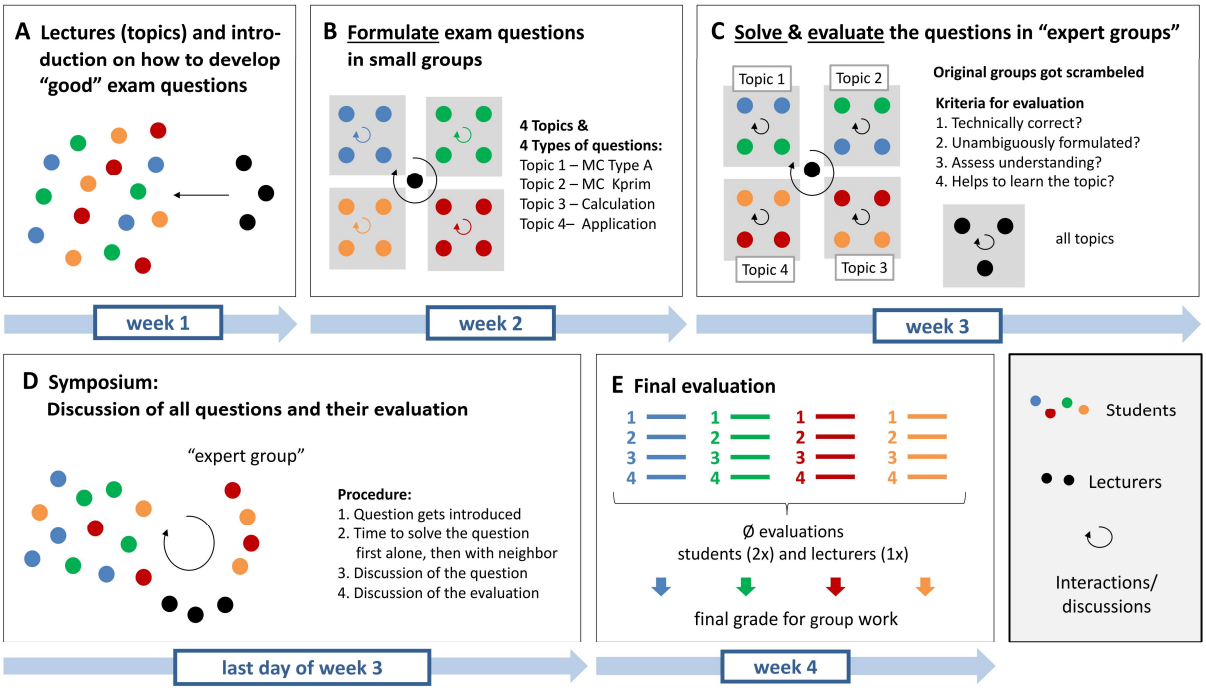

Figure 1: Graphical illustration of the described didactical setup. To illustrate the different composition of the groups for the different tasks, the students are depicted in different colors, lecturers correspond to black dots. (**A**) In the new setting, the theoretical lectures were conducted the same way as in previous years and aspects of question design were introduced by discussing examples of different types of exam questions (**B**) To formulate exam questions, students were assigned into small groups of 4-6 students. For each topic of the lecture, a specific type of question had to be developed, either a multiple-choice question (MC) of type A or K-prim, a calculation or application questions. All developed questions became available for all students after the second week. (**C**) In a later week during the semester, students were re-assigned to new groups (pairs of 2-3 students stayed together) and asked to solve and evaluate all designed questions of a defined topic (expert groups). The evaluation of the questions had to be done based on the previously defined criteria - technical correctness, unambiguity, assessing understanding and help to improve learning. For each criterion, students had to rate the questions with 0-2 points (0 = not fulfilled, 1 = partially fulfilled, 2 = fully fulfilled), the total amount of points for all questions than translated to the final grade for this part of the course. (**D**) At the end, all questions and their evaluations were discussed in a symposium. (**E**) After this discussion, the groups had the possibility to revise their initial rating of the questions.

Before surveying the students self-perceived motivation for formulating, solving and evaluating questions, we assessed motivational criteria according to the motivational expectancy-value and self-determination theories (reviewed in (Wigfield, 1994) and (Niemiec & Ryan, 2009)). The expectancy-value theory initially developed by Atkinson back in the 1950s (Atkinson, 1957) states that the motivation of an individual for a task is a positive function of the own perception

to perform well as well as to appreciate the value of the task. The self-determination theory (SDT) builds on work from Richard M. Ryan and Edward L. Deci at the University of Rochester who investigated how initially extrinsic motives can be actively assimilated into personally endorsed values resulting in high intrinsic motivation and self-determination. To do so, three psychological needs are of key importance, the needs for competence, autonomy and social relations (relatedness).

In the first four questions of the questionnaire, we therefore asked whether, after the course introduction, students felt they are able to perform well in the setting (competence) and if they felt the tasks are useful to better understand the topics of the lecture (task value). In addition, we asked whether the setting enabled enough autonomy (regarding group-work, time and self-organization) as well as enough possibilities to clarify questions with the supervisors (relatedness).

RESULTS

In general, the response rate of the questionnaire was very high (**Figure 2A**). 42 students signed in for the course and 39 students complete the course in a regular way (93%). 35 students finished the questionnaire (90%). The answers of the two students that did not complete the questionnaire (stop after one or two questions) were not included in the statistics.

Perceived competence, task value, autonomy and relatedness

For each motivational criterium, we asked the students to rate them on a five-level scale (**Figure 2B, C, D and E**). After the course introduction, the majority of students thought they will perform well in the setting (competence) ("rather yes", 51.4% and "yes", 28.5%) and only 8.6% thought they will rather not perform well (**Figure 2B).** Equally, a great majority of students felt they had sufficient autonomy and that there were sufficient possibilities to clarify questions with the lecturers (relatedness), with a slight trend to more negative answers for solving and evaluating questions compared to the task to formulate questions (**Figure 2D and E**). Regarding the question whether the students thought that the tasks will be useful to better understand the topics (prospective task value), the results of the survey are more variable (**Figure 2C**). Most students thought formulating questions will be most useful ("rather yes", 28.6% and "yes", 60%), followed by solving questions of their peers ("rather yes", 45.7% and "yes", 37.1%) and

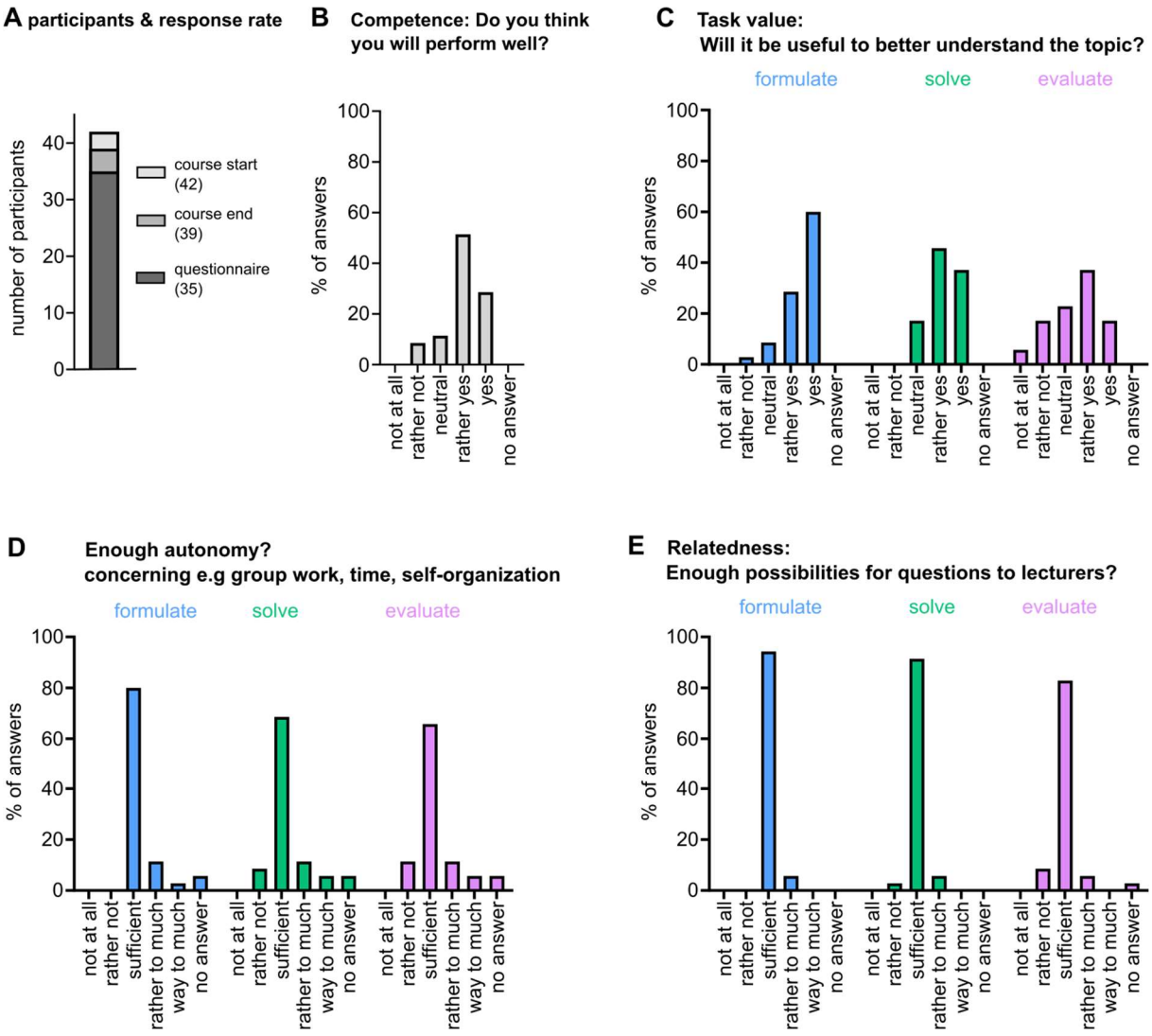

Figure 2: Perceived competence, task value, autonomy and relatedness for formulating, solving and evaluating questions. (A) Overview of course participation and completion of the questionnaire. Students answers concerning **(B)** competence, **(C)** prospective task value, **(D)** autonomy and **(E)** relatedness on a five-level answer scale for the entire course (grey) and for the tasks to formulate (blue), solve (green) and evaluate (purple) questions.

evaluating the questions ("rather yes", 37.1% and "yes", 17.1%). Regarding the task of evaluating the questions, a significant number of students also stated that they think it will not or not at all be useful ("rather not", 17.1% and "not at all", 5.7%).

In conclusion, the students rated all motivational criteria based on the expectancy-value and self-determination theory in a predominately positive way, especially the task to formulate questions. However, ratings were consistently slightly more negative for the tasks to solve the questions, and evaluating the questions was rated least positive.

Motivation level and retrospective task value

We next assessed the self-perceived motivation level of the students for the different tasks (**Figure 3A, B, E, F, I and J).** To also get information regarding the evolution of their motivation with time, we asked them to rank their motivation level for the different tasks before they did it and while they did it. The overall statements of all students are depicted in **figures 3A, E and I** and in the figures on their right (**Figures 3B, F and J**) the data of the students' individual evolution of motivation over time is summarized. In general, the students' motivation level before doing the task was highest for formulating the questions ("high/very high", 37%) followed by solving questions of their peers ("high/very high", 22.6%) and evaluating the questions ("high/very high", 20%) (**black bars in Figure 3A, E and I**). Interestingly, while doing the task, the overall motivation level significantly increased for the tasks to formulate questions ("high/very high", from 37%-before to 74.3%-while) and for solving the questions ("high/very high", from 22.6%-before to 42.9%-while) (**compare black and blue bars in Figure 3A and E**). In contrast, the motivation for evaluating the questions stayed on a similar level ("high/very high", from 20%-before to 22.8%-while) (**compare black and blue bars in Figure 3I**). We observed the same trends when evaluating the data on an individual level (**Figure 3B, F and J**). For the task to formulate questions, compared to their reported motivation before doing it, the majority of students reported an increase of their motivation level while doing it (60%, green bar) among which, 47.6% initially clicked "medium" (**Figure 3B**). Only 11.4% of the students reported a decrease of motivation level (orange bar), among which 50% initially clicked "high" (**Figure 3B**). In contrast, for evaluating the questions, the majority reported an equal motivation level (48.6%, blue bar), only 25.7% an increase (among which 55.6% initially clicked "low") and 25.7% a decreased level (among which 55.6% initially clicked "medium") (**Figure 3J**). The task to solve the questions of their peers, was rated in a slightly more negative manner as compared to formulating questions and slightly more positive as compared to the task to evaluate the questions (**Figure 3E, F**).

Regarding the degree of motivation change, in most cases the great majority reported one degree of motivation change (e.g. from medium to high, from high to very high etc.). Only for one case, a significant number of students reported more than one degree of motivation change, namely for the task of formulating questions. In this case, 22.8% of all students reported an increased motivation by two degrees (corresponding to 38% of the ones that reported an increased motivation).

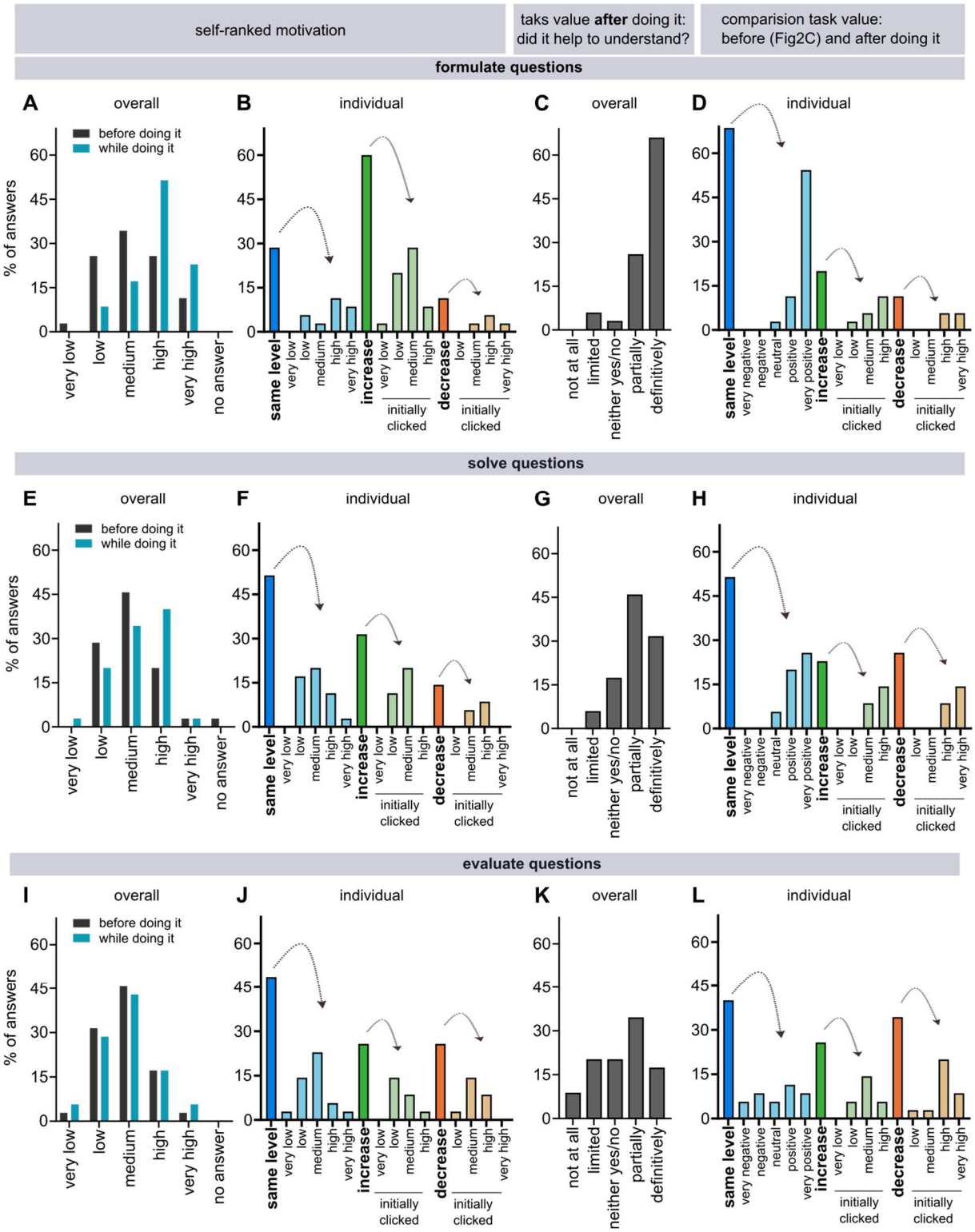

Figure 3: Motivation level and retrospective task value. Students self-ranked overall motivation before doing and while doing the specific tasks is depicted in panels **A**, **E** and **I**. The evolution of the motivation over time resulting from an analysis on an individual level is depicted in panels **B**, **F** and **J**. Results regarding the perceived retrospective task value for the respective tasks, and the evolution of the task value over time on an individual level are depicted in panels **C** and **D**, **G** and **H**, **K** and **L**.

After the students ranked their motivation level, in a later question we asked whether the tasks actually helped to better understand the topics of the lecture (retrospective task value, "Did the task help to better understand the topics", **Figure 3C, G and K**). The distribution of the answers for all three tasks is very similar to the prospective task value question (Figure 2C, "Do you think it will be useful to better understand the topics"). The rankings are highest for the task of formulating questions ("definitively", 65.7%), followed by solving questions ("definitively", 31.4%) and evaluating questions ("definitively", 17.1%). We also analyzed, the time evolution of the task value on an individual level (**Figure 3D, H and L**). Whereas it mostly stayed on the same level for the tasks of formulating and solving questions (68.6% and 51.4%, blue bars, **Figure 3D and H**), the number of students that reported a decreased perceived task value is highest for the task of evaluating the questions (34.3%, orange bar**, Figure 3L**). Among those students, 58.3% initially clicked a high value. Regarding the degree of change of task value, in most cases the great majority reported one degree of change. Again, only for one case, a significant number of students reported more than one degree of change, this time for the task of evaluating questions. Here, 11.4% of all students reported a decrease by two degrees (corresponding to 33.3% of the ones that reported a decreased task value).

Together, both, the motivation level as well as the perceived task value followed the same trends; for formulating questions the rankings were highest and also showed the greatest positive trend concerning its evolution over time. Similarly, the rankings for evaluating the questions were lowest and also showed the greatest negative evolution trend over time. Solving the questions of peers resulted in intermediate rankings.

Correlation analysis

To analyze potential correlations between the criteria of the expectancy-value and self-determination motivational theory (competence, task-value, autonomy, relatedness) and students reported self-ranked motivation level more precisely, we performed a Spearman cross-correlation analysis (**Table 1**). There was no statistically significant correlation between the criteria "competence", "autonomy", "relatedness" and the reported self-ranked motivation level of the students (data not shown). Interestingly however, there was a positive correlation between the perceived task value and the motivation level for the respective tasks (**Table 1**). Regarding the task of formulating questions, the positive cross-correlation is statistically significant for the prospective as well as the retrospective task value, while for solving and evaluation of questions, this is only true for the retrospective task value.

Table 1: Spearman correlation coefficients		Self-ranked motivation level					
		formulate (Fig2A)		solve (Fig2E)		evaluate (Fig2I)	
		before	*while*	*before*	*while*	*before*	*while*
Prospective task value (Fig1E)	formulate	0.46 (0.52%)	0.40 (1.8%)	0.36 (3.4%)	0.04	0.13	-0.06
	solve	0.13	0.27	0.29	0.24	0.00	-0.22
	evaluate	0.10	0.01	0.12	0.31	0.22	0.21
Retrospective task value (Fig2C, G, K)	formulate	0.51 (0.19%)	0.46 (0.53%)	0.30	0.17	0.05	0.17
	solve	0.09	0.05	0.49 (0.3%)	0.44 (0.8%)	0.13	0.13
	valuate	0.19	0.24	0.29	0.42 (1.2%)	0.43 (1%)	0.63 (0.004%)

Table 1: Spearman correlation coefficients between the prospective and retrospective task value and the self-ranked motivation level of the students before and while doing the tasks. The analysis was performed with Graph Pad Prism 8.4.2 (www.graphpad.com). The five-level answer scale was converted to numbers in the following way: for the prospective task value, "not at all" = 1, "rather not" = 2, "neutral" = 3, "rather yes" = 4 and "yes" = 5 and for the retrospective task value, "not at all" = 1, "limited" = 2, "neither yes nor no" = 3, "partially" = 4 and "definitively" = 5. For the motivation levels, "very low" = 1, "low" = 2, "medium" = 3, "high" = 4 and "very high" = 5. Correlation coefficients are color coded (positive = blue and negative = red) and the p-values are shown in parentheses in case of values below 5%.

Together, this underlines the significance of the trends described above, namely, that the more the students felt, the task will be, or was useful to understand the topic, the higher they ranked their motivation level for the task.

Applied strategies while working on tasks and comparison to learning for exam

In the second part of the questionnaire we intended to get an overview of the strategies that students applied while working on the different tasks and asked them to compare their learning experience with learning for a traditional exam (**Figure 4**). To explore the level and frequencies of applied problem-solving strategies, we provided them with a list of strategies we had observed in the course and asked them to click on all statements that apply (**Figure 4A**). For both, formulating and solving questions, a great majority of students stated that they "studied the lecture material" (formulate 97.1%, solve 88.6%), "thought for themselves" (formulate 85.7%, solve 88.6%) and "discussed with peers from their group" (formulate 94.3%, solve 85.7%). In contrast, for both tasks, only a comparably minor part of the students "discussed with peers from other groups" (formulate 20%, solve 17.1%). For other strategies that probably

require more proactivity, the variation for the different tasks was bigger. Interestingly, students stated that they applied these strategies more often while formulating questions as compared to solving questions, namely “study materials from previous courses” (formulate 88.6% versus solve 51.4%), “discussing with lecturers” (formulate 60% versus solve 45.7%) and “study literature found autonomously” (formulate 54.3% versus solve 25.7%). While solving the questions of peers, google was consulted more often (62.6%) as compared to while formulating the questions (45.7%). Moreover, only 40% of the students stated that they “looked at the solutions provided by their peers” while solving the questions.

In a second series of questions, we explored how the learning experiences compare to learning for a traditional exam (**Figure 4B, C**). Concerning the expert group topics, the majority of students stated that they think they learned more (45.7%) or way more (20%), resulting in total number of “more/way more” statements of 65.7%. Only a minority stated they learned less (5.7%) about their expert group topics. Concerning the other topics, although more students stated they learned less (22.9%), still a considerable number of students stated they learned more (22.9%) or even way more (11.4%) (resulting in total number of “more/way more” statements of 34.3%) and 31.4% of the students stated that they learned the same. As expected, also a great majority of students stated they learned to discusses with others more (40%) or way more (42.9%). A considerable number of students reported positive effects regarding learning to organize themselves (learned more or way more, 34.3%). Interestingly, almost half of students felt they were able to rate their understanding more (31.4%) or way more (11.4%) (total more/way more, 45.7%) as compared to the situation of having a traditional exam about the topics. Finally, many students stated they additionally learned to think critically (55.9%), to provide own insights and solutions into a group and argue (50%), to communicate and collaborate with others (41.2%) and to think more deeply and with greater complexity (38.2%) (**Figure 4C**). In this case, students could choose from a list of 10 statements, but click on maximum 3 statements (the ones that applied most) and the majority did so (33 clicked on 3 items, 1 clicked at one item and only 1 student did not click anything).

When performing cross-correlation analysis between approaches taken to formulate and solve the questions and the perceived learning effects compared to learning for a traditional exam, there were some interesting statistically significant positive correlations. Students that stated they discussed with peers from their group while formulating questions, often also stated they learned a lot more about their expert group topic (correlation coefficient 0.38, p-value 2.9%) and stated they were able to better rate their understanding about the topics (correlation

coefficient 0.4, p-value 2.6%). In addition, students that stated they used additional material found autonomously for formulating the questions, often also stated that they learned a lot more to discuss with others (correlation coefficient 0.54, p-value 0.17%). No statistically significant correlations were found regarding approaches takes for solving the questions.

Overall, this survey confirmed our impressions that the majority of students very actively participated and engaged in formulating and solving questions and that, while doing it, interactions among all participants (including lecturers) were a lot higher as compared to previous years, were students performed a final examination at the end of the course.

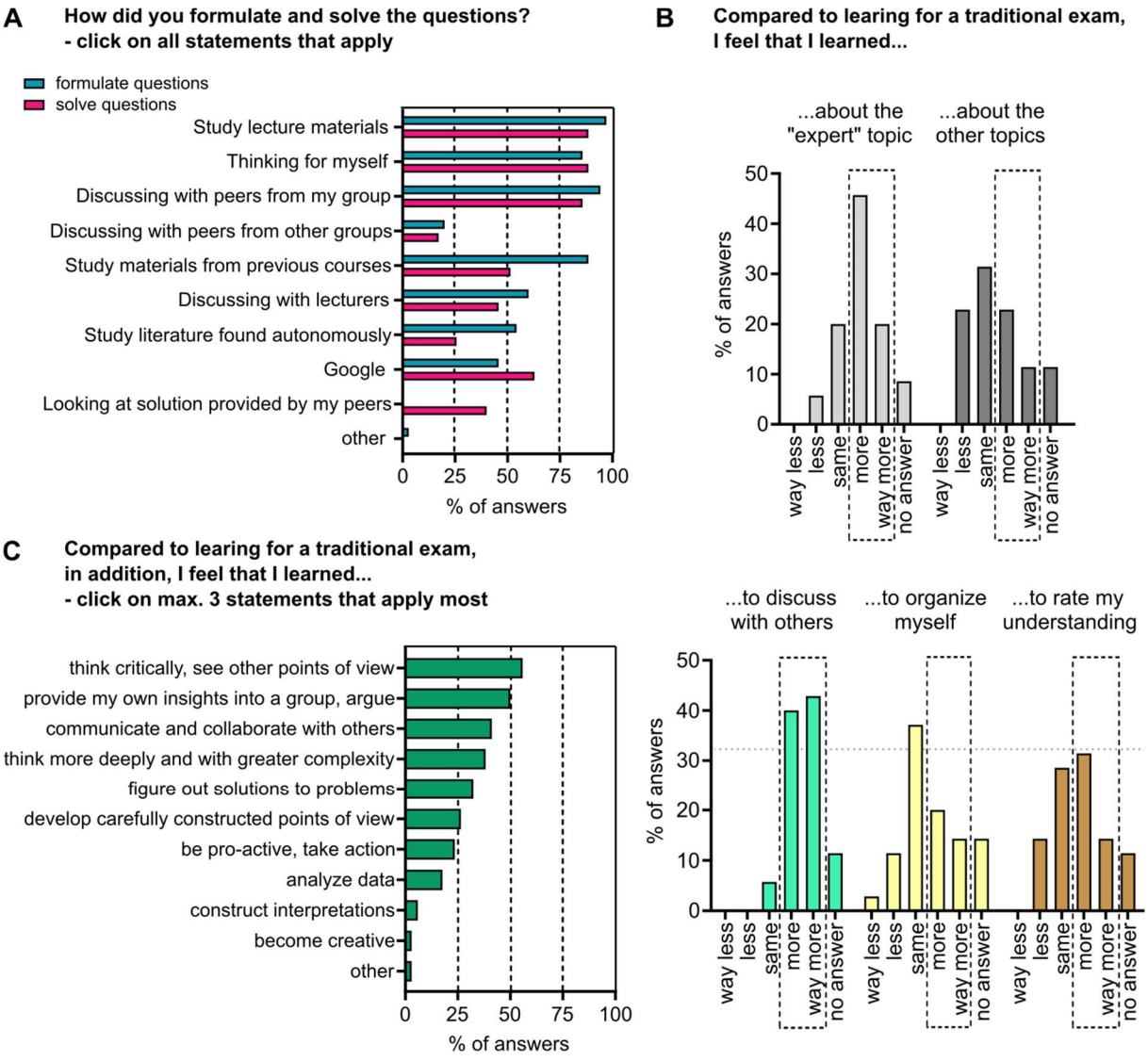

Figure 4: Applied strategies while working on depicted tasks and comparison to learning for a traditional exam. (A) Approaches taken to formulate and solve questions. In panels **(B)** and **(C)** the data regarding the students' comparison of their learning compared to a traditional exam is represented.

This is also reflected in the fact that we got more than 70 written additional comments for the separate questions and more than 20 general comments and suggestions for improvement of this format (listed in the appendix). The great majority of the comments and suggestions are very elaborated and purposeful and some of them will be considered in more detail in the discussion.

Overall feedback

To collect more target-oriented feedback on the format, at the end of the survey, we openly asked the students what they "liked most", "disliked most" and "missed" during this course. For all three questions, the response rate was very high (28/35 answers for "I liked most", 26/35 answers for "I disliked most" and 16/35 answers for "I missed". The statements are listed in the appendix and many concerned similar issues. Among the statements "**I liked most**", many concerned social aspects (relatedness) (42.9%), among which the majority specifically mentioned the group work and discussions (83.3%) and the others interactions with the lecturers (16.7%). Here, many students also stated aspects regarding their autonomy (25%) and the task of formulating questions (25%). Some appreciated the fact that it led to a better understanding (task value) (14.3%) and the new approach by itself (7.1%). Among the statements "**I disliked most**", the majority concerned the process of evaluating the questions (34.6%) as well as the discussion day (26.9%). Interestingly, 26.9% of all "I disliked most" statements were also related to social aspects (relatedness) (among which again, most of them concerned the group work), and 11.5% to aspects regarding the autonomy. Among the statements "**I missed**", most concerned guidance (68.9%), among which the majority stated the evaluation of the questions (43.8%) and some the formulation of the questions (18.8%).

Comparison of evaluations by students and lecturers

In line with the many good impressions of the lecturers during the question formulation process, in general, the evaluations were very good (on average, more than 5 out of 8 points for all type of questions) and were very similar for the students and lecturers (**Figure 5A**). The multiple-choice questions of both types were evaluated in a worse way by the lecturers, mostly regarding the criteria "ambiguity", "understand & apply" and "helped to learn" (**Figure 5B**). Among all the evaluations, the criteria "understand & apply" and "helped to learn" for the multiple-choice questions of type A (one answer is correct), generated the biggest difference between the evaluation of the students and the lecturers.

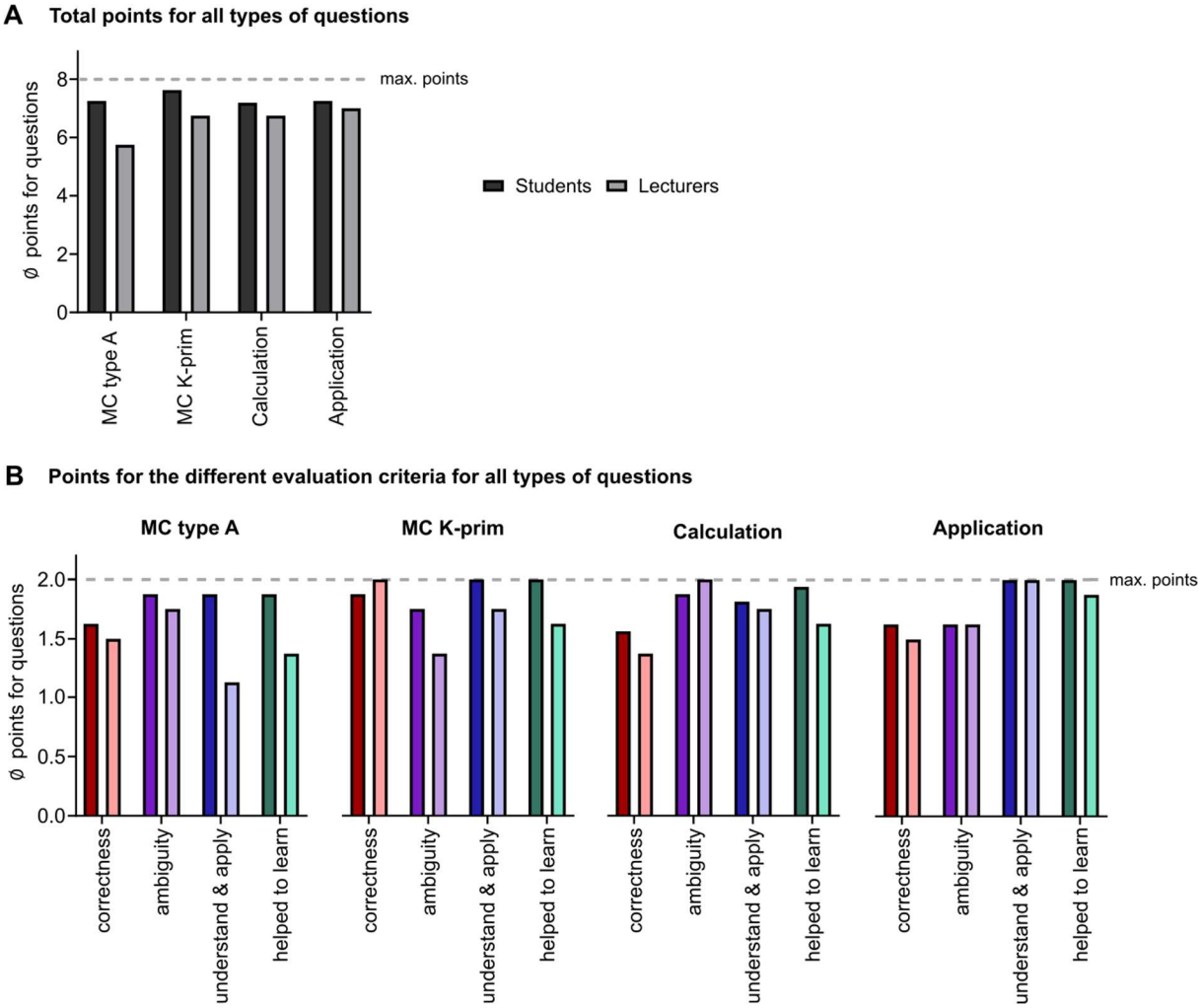

Figure 5: Comparison of evaluations by students and lectures. (A) Average number or points for the different type of questions. **(B)** Average number of points for each criterion for each type of question. Dark colors correspond to evaluations by the students and light colors to evaluations by the lecturers. Evaluation scheme for all criteria was the following: 0 points = not fulfilled, 1 point = partially fulfilled, 2 points = fully fulfilled.

DISCUSSION AND CONCLUSIONS

In general, our setting to formulate, solve and evaluate questions was perceived as motivating and the corresponding motivational criteria were rated in a positive way (**Figures 2 and 3**) and a great majority of students would not have preferred a written exam ("Would you have preferred a written exam": no = 80%, yes = 8.6%, no answer = 11.4%). Amongst the different tasks, the task to formulate questions was rated most positively, followed by the task to solve questions of peers and the task to evaluate the questions. We detected a positive correlation of the perceived task value with students' motivation to perform the respective task (**Table 1**). This is in line with results from a recent study that assessed the motivational behavior of

students to voluntary generate multiple choice questions for an online self-assessment test (Poot et al., 2017). The authors of this study reported that students that participated explained that the task value was the most important reason to participate. Interestingly, in our study, the majority of statements about what students liked most concerned social aspects regarding the group work and only few statements concerned the task value. Consistently, we observed that students intensively interacted and discussed within the groups and it is not surprising that more than 80% of the students stated they learn more or way more to discuss with others as compared to learning for a traditional exam (**Figure 4B**).

More than 60% of all students stated they felt that they learned more or way more about their expert group topic as compared to learning for a traditional exam. Additionally, more than 60% of all students stated they learned the same or more about the other topics (**Figure 4B**). This is also reflected in the fact, that all lecturers of the course perceived many aspects of this setting in a very positive way: The quality of the developed question was very high. Questions gave rise to discussion about aspects of the topics that went well beyond the usual discussion with students on this level. Since many students stated they learned way more about their expert group topic, in an attempt to improve the setting, it could be considered, to enable students to work on more than one topic while solving questions of their peers or to emphasize that all students solve all questions before the day of discussion. Some students' comments and suggestions for improvement also went in these directions.

The task to evaluate questions was rated the least positive, and students perceived this task as least useful to better understand to topics of the lecture. Consistently, when asking what students disliked most and missed most in this setting, the majority of comments concerned the process of evaluating the questions and stated that guidance was missing. Even among the lecturers, rating of questions was difficult and led to many interesting discussions. Often, this concerned issues regarding the ambiguity of statements and formulations in the questions of the students and in the symposium the corresponding discussion were sometimes highly controversial. Students also stated that the pressure, that questions will be ranked based on the defined criteria (technical correctness, unambiguity) lead to develop questions about topics that were already understood well and that the things that were unclear until this point were not touched. It might be worth to introduce an additional reward or evaluation criterion that emphasizes questions that go beyond things discussed in the lecture, that are very creative or involve very difficult aspects of the topics.

In conclusions, although the evaluation of the questions was perceived as most difficult, we

believe that this aspect likely also contributed to the very encouraging fact, that a significant number of students stated that they additionally learned to think critically, provide their own inside into a group and to think more deeply and with greater complexity (**Figure 4C**).

ACKNOWLEGEMENTS

I would like to thank the entire teaching team of the Department of Biochemistry (C. Berger, F. Borquin, O. Büchi, S. Gloor, D. Grossenbacher, and S. Lobet) for their support and help at all stages of the project and S. Rast for his technical support to set up the online questionnaire.

REFERENCES

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*. https://doi.org/10.1037/h0043445

Baerheim, A., & Meland, E. (2003). Medical students proposing questions for their own written final examination: Evaluation of an educational project. *Medical Education*, *37*(8), 734–738. https://doi.org/10.1046/j.1365-2923.2003.01578.x

Bowker, M. H. (2010). Teaching Students to Ask Questions instead of Answering them. *Thought & Action*, 127–134.

Entwistle, N. J. (1991). Approaches to Learning and Perceptions of the Learning Environment : Introduction to the Special Issue Author (s): N . J . Entwistle Source : Higher Education , Vol . 22 , No . 3 , Approaches to Learning and Perceptions of the Published by : Springer St. *Higher Education*, *22*(3), 201–204.

Kerkman, D. D., Kellison, K. L., Piñon, M. F., Schmidt, D., & Lewis, S. (1994). The Quiz Game: Writing and Explaining Questions Improve Quiz Scores. *Teaching of Psychology*, *21*(2), 104–106. https://doi.org/10.1207/s15328023top2102_11

Kolluru, S. (2012). An active-learning assignment requiring pharmacy students to write medicinal chemistry examination questions. *American Journal of Pharmaceutical Education*, *76*(6). https://doi.org/10.5688/ajpe766112

Marbach-Ad G., & P.G Sokolove (2000). Good Science Begins With Good Questions. *Journal of College Science and Teaching*.

Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom:Applying self-determination theory to educational practice. *Theory and Research in Education*, *7*(2), 133–144. https://doi.org/10.1177/1477878509104318

Papinczak, T., Peterson, R., Babri, A. S., Ward, K., Kippers, V., & Wilkinson, D. (2012). Using student-generated questions for student-centred assessment. *Assessment and Evaluation in Higher Education*, *37*(4), 439–452. https://doi.org/10.1080/02602938.2010.538666

Pellegrino, J. W., & Hilton, M. L. (2013). Education for life and work: Developing transferable knowledge and skills in the 21st century. In *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. https://doi.org/10.17226/13398

Poot, R., de Kleijn, R. A. M., van Rijen, H. V. M., & van Tartwijk, J. (2017). Students generate items for an online formative assessment: Is it motivating? *Medical Teacher*, *39*(3), 315–320. https://doi.org/10.1080/0142159X.2017.1270428

Rhind, S. M., & Pettigrew, G. W. (2012). Peer generation of multiple-choice questions: Student engagement and experiences. *Journal of Veterinary Medical Education*, *39*(4), 375–379. https://doi.org/10.3138/jvme.0512-043R

Sircar, S. S., & Tandon, O. P. (1999). *Involving Students In Question Writing*. *22*(1), 84–91.

Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment and Evaluation in Higher Education*, *30*(4), 325–341. https://doi.org/10.1080/02602930500099102

Traub, R. E. & MacRury, K. (1990). Multiple-choice vs. free response in the testing of scholastic achievement. *In: K. Ingenkamp & R. S. Jager (Eds) Test Und Tends 8: Jahrbuch Der Pädagogischen Diagnostik (Weinheim Und Base, Beltz Verlag)*, 128–159.

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review*, *6*(1), 49–78. https://doi.org/10.1007/BF02209024

Appendix 1: Questionnaire

We thank you for your feedback on the BCH303 part about question formulation/evaluation/discussion. Your feedback is voluntary. This questionnaire is anonymous and will not have any influence on your mark in this course. We might use the data of this course evaluation for a publication.

QUESTION 1

After the course introduction, I felt confident that I can perform well in this setting

QUESTION 2

After the course introduction, I felt that the following tasks are useful for me to better understand the topics of the lectures (formulate questions, answer questions of my peers, evaluate questions of my peers)

QUESTION 3

Did you had enough autonomy (e.g. concerning group work, time, self-organization) to perform the following tasks (formulate questions, answer questions of my peers, evaluate questions of my peers)

QUESTION 4

Were there enough possibilities to clarify questions with the lecturers to formulate questions, answer questions of my peers, evaluate questions of my peers.

QUESTION 5

Please rank your motivation level for the following tasks and timepoints **-** formulate questions, answer questions of my peers, evaluate questions of my peers – each "before doing it" and "while doing it".

QUESTION 6/7

How did you approach the task to formulate the questions/to solve the questions of your peers - click on all statements that apply.

QUESTION 8

After finishing the following tasks, did it helped you to better understand the topics of the lectures? (formulate questions, answer questions of my peers, evaluate questions of my peers)

QUESTION 9

In this setting *(lecture - formulate/solve/evaluate questions)*, compared to a traditional one *(lecture – written exam)*, I feel that…

… I learned about my expert group topic … I learned about the other topics

… I learned to discuss with others … I learned to organize myself … I was able to rate (german = einschätzen) my understanding about the topics

QUESTION 10

In this setting *(lecture - formulate/answer/evaluate questions)*, compared to a traditional one *(lecture – written exam)*, I feel that additionally I learned to… (choose from the following list up to 3 statements that apply most)

QUESTION 11

In this setting, I…"liked most"…"disliked most" …"missed"

QUESTION 12

Would you have preferred to have a written exam about the topics of the lecture?

QUESTION 13

Do you have additional comments/remarks/suggestions about this format?

Appendix 2: Students' answers on the Question 11 (liked, disliked, missed)

Grammatical errors were corrected. German answers that were translated to English are marked with quotation marks (“).

I liked most …	classification
Independence	autonomy
Working in groups, have presentations in-front of small audience	relatedness (group work, presentations)
We were free to organize ourselves. We had the chance to take the initiative	autonomy
Not so much pressure with the grading.	relatedness (grading)
Learning new aspects about the topics	task value -better understanding
prepare the questions in the group	relatedness (group work)
The creative and organizational freedom (a hate-love relation)	autonomy
the discussions (in the group of 4 or six and in the group with everyone)	relatedness (discussion in group)
elaborating something (not only simple studying and written exam) and rethink it again, again. Discuss topics personally with teaching assistants	autonomy, relatedness (lecturers)
to formulate and to discuss the questions.	formulating questions
A new approach to learning about a topic	new approach
the interaction within/between the group(s)	relatedness (group work)
the open frame within which we could develop our own questions	autonomy
The discussions which happened during formulating the questions, because it led to a better understanding	relatedness (discussions), task value -better understanding
formulating questions	formulating questions
Getting to work with new people and thinking more deeply about the lecture materials	relatedness (group work), task value-better understanding
Question Design	formulating questions
the self-organization and many opportunities for discussions	autonomy, relatedness (discussion)
The question design	formulating questions
group works	relatedness (group work)
Formulate Questions	formulating questions
Great teamwork,	relatedness (group work)
one specific lecturer, students profit enormously - very clear, emphatic, motivational and caring about to share knowledge)	relatedness (lecturers)
working independently	autonomy
Deepen knowledge and gaining a new perspective on topics learned before. "Trivial" knowledge got new depth by approaching from another angle.	task value - better understanding
formulating questions	formulating questions
try to think from another perspective when writing the questions	new approach, other perspective
generating the questions	formulating questions

	#	%
1. Relatedness - group work, discussions -lecturers	12 10 2	42.9
2. Autonomy	7	25.0
3. Formulating questions	7	25.0
4. Task value - better understanding	4	14.3
5. New approach	2	7.1

I disliked most …	**classification**
Grading	evaluating questions, grading
Go over the same topics over and over again	repeating
That we had to repeat every theme several times, gets boring from time to time	repeating
It was exhausting to concentrate the whole day and participate concentrated in the discussion.	last day
Focusing during the expert group time on only one topic	only one expert group topic
the evaluation of the questions in the expert group on Friday	evaluating questions, last day
The creative and organizational freedom (a hate-love relation)	autonomy
the time we needed in the small groups for minor formulating stuff	discussions, minor formulating stuff
presentation (I do not like to present), competition in group	presentation, relatedness (competition in group)
The group work and the evaluation process	relatedness (group work), evaluation
the focus on criteria since day 1. Suggestion: limit the question designing (with calculator, like in exam) and criteria only given just before evaluating	evaluation process, criteria
being afraid that others could misunderstand something that was extremely clear to us	relatedness (others could misunderstand…)
The rating of the questions	evaluation of questions
evaluation questions of peers	evaluation of questions
Sometimes no one in the group took a ‚lead', as in group work no one feels responsible (or wants to feel responsible). I feel like it is tempting to put as little effort in as possible. While this was not the case for the construction of the questions, whilst evaluating I felt like there is a rather rushed atmosphere	autonomy (nobody takes lead), relatedness
The discussion about the questions	last day
the unfortunate aspect of group work that not everyone contributes and understands equally	relatedness (group work)
the rating system	evaluation of questions, criteria
vague focus during question design and evaluation	autonomy (focus)
Presentation and evaluation of questions	presentation), evaluation of questions
answering of the questions	solving the questions
some group members	relatedness (group members)
Grading seemed random from a grading and graded perspective as unclear on how to grade and no theory on good question design was available. A scheme/questionnaire with how to weigh aspects of question design would decrease randomness and social pressure of giving points as you would have something concrete to argue why points were given/not given. Grading own questions feels ethically wrong as such most abstained anyways..	evaluation of questions (criteria)
the discussion of the questions in the two blocks	last day
grading system (some students were harsher with their grading than others)	evaluation of questions, relatedness (peer evaluation)
presentations which were too stiff	presentations

	#	%
1. Evaluation of questions	9	34.6
- criteria	3	
2. Relatedness	7	26.9
- group work, group members	5	
- others could misunderstand, some groups harsher with grading	2	
3. Last day, presentations	7	26.9
- presentations	3	
4. Autonomy	3	11.5
5. Discussing about minor formulating stuff, repeating	3	11.5
6. Only one expert group topic	1	3.8
7. Solving questions	1	3.8

I missed …	**classification**
Direction	too much autonomy
lecturer's evaluation about own questions	evaluation by lecturers
the evaluation of the profs	evaluation by lecturers
some variation in the topics evaluated.	variations in questions/topics
creativity	
guidance on preparing proper presentation, requirement to dig deeply in all topics	too much autonomy, guidance evaluation
Studying by myself by solving provided exercises previously to constructing my own	too much autonomy, guidance for formulating
some clearer instructions regarding "target audience"/level of the questions	too much autonomy, guidance for formulating
The four criteria for evaluation were too little to make a proper grading scheme.	too much autonomy, guidance evaluation
a better explanation of the evaluation criteria and of the time schedule	too much autonomy, guidance evaluation
A smoother, faster transition between the topics.	variations in questions/topics
Get deep reflexive by heart knowledge on all topics	
good opportunity to attend seminars (or just one) and discuss it together. (maybe as part of the introduction)	too much autonomy, guidance evaluation
clearer rating scale for the evaluations	too much autonomy, guidance evaluation
Guidance on constructing exam questions. Some points mentioned during discussion e.g. common mistakes and things you as lecturer keep in mind when constructing questions would have been helpful before constructing questions. No points for constructing a challenging/creative question.	too much autonomy, guidance for formulating
the discussions during the presentations between the peers	not enough discussions

	#	**%**
1. guidance evaluation	7	43.7
- evaluation by lecturers	2	
- guidance for evaluation	5	
2. guidance for formulating	3	18.7
3. variations, questions, change between topics	2	12.5
4. guidance in general	1	6.25
5. not enough discussions of peers during presentations	1	6.25

Appendix 3: Students' comments on all Questions

Grammatical errors were corrected. German answers that were translated to English are marked with quotation marks (“).

Question1: After the course introduction I felt confident that I can perform well in this setting
For topics like buffers and enzymology it was only a short recap which was enough since the topic was already discussed in detail in a lecture. On the other hand, things that were not clear from the lecture so far, were also not clear after this short recap. But with chromatography it was only a crash course on a topic we have not heard much about it so far. This made it rather hard to create and evaluate questions.
“One realized only later, how exhausting and time consuming it is to formulate questions.”
I found a refreshing very useful. I thought it was a little bit to basic for the topics of BIOCHEMIE 1 “(module of a previous semester with similar topics)”, I would prefer more information about chromatography.
The lectures in the beginning helped for the design the questions and it was also a repetition and a further learning. For example we learned how to design a protein purification strategy with the help of protein characteristics.
If you mean overall BCH303, then: Well, everything was pretty clear, but since there not every experiment is graded to the same weight, a table with the composition of the final grade would have been helpful. If you mean Question Designing: The Task was clear, but the limitation were not - Do I have to design a question like in an exam (with/without calculator) or more like an exercise for the module BCH I “(module of a previous semester with similar topics)”?
The introduction was good, but we have never really worked in the lab for a long time. Half of the methods were still unknown at the beginning of the block course.
I think it is good practice, since we don’t have many opportunities to present or design questions in other modules.
The introduction felt reassuring. I was excited for the question design.
the introduction could be shorter. Short introduction what the experiments are and with who we are doing it (supervisors). -Implement it the very first day. I knew my performance the day before the experiment. (That's ok!)

Question 2: After the course introduction, I felt that the following tasks are useful for me to better understand the topics of the lecture
Mostly the questions prepared covered the topics that were already understood well and the things that were unclear until this point were not touched.
“It was interesting to solve the questions, since one realized what was important for the others or how they thought differently about things. The evaluation was based on the criteria and there was little relation to the topics. “
I thought that answering the questions will probably help me most, because this will cover the most topics. By formulating questions, we only focus on one aspect (like enzymology, only solving one equation)
The evaluation day was very long and intense, maybe it would make sense to split it up.
It was clear what was expected. I found it good that we could consult afterwards individually, because the most problem are specific when we did elaborate the questions.
The task splitting, which happens automatically is problematic. It enables students to avoid topics that they do not like or are not good at. By this, it is highly probable that in the end not all students learnt about all topics the same. An exam would still be the most accurate assessment.
Questions were important from 1. semester, but I did not notice it. I understand topics more when I formulate and try to answer the questions about the topic and the understanding improve when I try to design a more precise question during evaluate it.
The pressure of evaluating the question with fair grading was slowly becoming important during evaluation. We started off very well, but somewhere the criteria become more important.
it is way more intense to evaluate the questions than to simply answer them and one gains a deeper insight into the topic, which is helpful!
Evaluation did not really help as the questions were designed some months ago and a lot of theoretical details were forgotten in the meantime.
Please define course introduction better. For me this was the first week, not the answer and evaluation of the questions.
I think it would be good if the work in the lab or other contributions counted towards the grade. As is the issue with group work, some members work very hard and some contribute less. Participation, preparation and understanding during the labs is very important during this course and should count in some way.
I profited the most from the question formulation.

Question 3: Did you have enough autonomy (e.g. concerning group work, time, self-organization) to perform the following tasks?
Enough time to prepare the questions and task was quite clear. For answering the question there was enough time, every question could be discussed in detail in the groups Evaluating the questions was rather hard and the group presentation and discussion took rather longer than expected and not even all questions could have been discussed
The criteria on which the evaluation based, could have been improved, i.e. level of difficulty of the question etc.
“I think it depended on the group. If someone was there to set the pace, things could proceed fast. “
I really liked that we had so much time that we could organize ourselves. I liked that we could work in different groups to get a lot of different inputs. The evaluation day was a bit long, I had a hard time fallowing the whole day and I got a bit confused in the end.
One day is more than enough to answer and evaluate the question type we had.
I would prefer to have a little bit more time to answer the question which were not of my "expert" topic. That was probably a problem of myself, as I only thought about the "expert" questions before Friday. It would good to refresh the other question as well.
Topics and question types were complex to a different manner. By this a lot of time was needed to answer and evaluate the questions. It would be an advantage to have an additional half day to discuss the evaluation, uncertainties and also the presentation with a teaching assistant. Resolving the exercises by all students before the presentation day would increase understanding a lot.
We had time to discuss the questions and to improve them.
Regarding enzymology it was rather easy. But I think if I has the topic ligand and receptor then it would have been more difficult.
if the dynamics in the group is difficult, for me it gets hard to not be very critical against ideas of others, I tend to not like them. In our group, our subgroup (team of 3) ended up writing the majority of the questions while the others got stuck up with one theme. But in the end, I think we were all happy with the outcome
Very good time management
Distinguish between knowledge of the topic and conceptual designing of questions
I felt that we could have used more time for the groupwork, since it was sometimes very hard to discuss the questions

Question 4: Were there enough possibilities to clarify questions with the lecturers for the following tasks?
Lecturers were always around when questions appeared and they were always happy to help!
“We knew where to find you, but could also work alone.”
I think this was very good. There was always someone available to ask if there were insecurities.
There was always someone around to help with issues.
Very helpful advices were given. Always available when needed.
It was pretty easy to find someone, everyone was very motivated and explained everything clear and good
Teaching assistants were available for questions and took their time.
We were independent and if we needed help there was always someone from the teaching team. There was no need for writing emails and for waiting a long time for answers. This was very good! Also, the IT team was very helpful!
I liked that by time someone was coming to our "room/table", looked after us and tried to clarify our problems. “We did not have the feeling to be thrown into cold water.”
We did not approach our lecturers that often for the question design. We could have done this better.
More concepts for how to design questions and more specific evaluation criteria.
The presence of the teaching team was very helpful and very balanced

Question 5: Please rank your motivation level for the following tasks and timepoints:
Too much peer evaluation. Teachers grades should carry more weight
Before and while answering my motivation was rather low since we had to get back into the topic again. And then many questions very similar and things had to be discussed over and over again. Evaluating all the questions was exhausting after a while.
I realized while formulating the questions, that it is not as easy as it seems. And after 2 days of formulating I had my fill. “I had enough of our questions and did not want to look at them again.”
I thought that the evaluation was the least important part. It's also a bit unpleasant - I don't want to give my colleagues a bad grade. Therefore, in the evaluation we were biased in my opinion. But it's also clear to me that it's like a bonus system and nobody can get a grade below 4. However still you always think about how you could give maximal points, and in my opinion, time is wasted on that.
Low due to some complex questions and the discussions that ensued when evaluating.
Personally, I found the questions to solve the most interesting part. When formulating the question, I think the most time went into find a good spelling without any misunderstandings. I think even I learned there also, I think at a certain point it was just formulate again and again, with low learning effect, specially at the multiple/single choice. I thought at open question you focus more at the topic and less on the formulation and I think I learned more at the evaluation and at the formulating.
I prefer acquiring knowledge on my own and assessment with exams.
It was pretty hard to exclude the criteria while doing the questions. We were on some point more concentrated to fulfill the criteria that we missed the basic (labeling of axis).
Evaluating the questions was very difficult especially with the evaluation criteria. The evaluation criteria were actually good (a little bit inaccurate) but the criticism of teaching assistants was sometimes very good but often too detailed.
answering questions during the course (e.g set of questions has to be solved via Olat “(online learning platform)” in groups) what evaluation is, was vague formulated.

Question 8: Did the following tasks help you to better understand the topics of the lectures?
In chromatography it helped since this topic was not covered before but in enzymology every group mostly covered the same part of the topic (the part that was understood the best)
I don`t think evaluating helped. It was just nice to compare the different questions from the different groups.
Formulating the questions definitely helped to have a look at the topic from another perspective than I would do to learn for an exam. The only point is that maybe it helps to understand a topic in a very specified topic (example: I formulate a question about competitive inhibition; I don't learn much about uncompetitive etc. inhibition). By answering the questions, the field is broader and more topics are covered; very helpful for understanding. Evaluating the questions didn't really help with understanding the topic, however it helped to think about formulation of the questions, and if they are misleading, what could be meant by the person who wrote it.
When answering the questions that all were similar one gains little new knowledge...
Teamwork may enable students to rely on peers and not be sufficiently motivated to really study the theory.
It helped also to see what I do not understand yet.
I don't think evaluating helps with understanding as the focus often is more on if the question was clear. To me a big part was making sure the question couldn't be misunderstood which took away from learning about the actual topic. I don't think it makes sense for us to learn how to write very precise question since it does not seem to be essential for this stage in our studies. While writing the questions I also felt we were only focused on the topic of our question and all other aspects of a topic were neglected and so I only came into contact with them again while solving the other's questions.
Evaluation/answering questions of peers should rather be done directly after creating the questions in order to have the same level of knowledge.
solutions were already given and topics weren't that clearly present anymore.

Question 10: In this setting (lecture - formulate/answer/evaluate questions), compared to a traditional one (lecture – written exam), I feel that additionally I learned to… (choose from the following list up to 3 statements that apply most)
I think, it was relaxing that there was less pressure than in an Exam. It was nice to do it in a group and share the "burden". But I think it could depend on the composition of the different personalities in the group.
I think that by learning for a written exam, I learn more by myself and also, I learn more things, because I have more time to learn complex things at my pace. But this way, I looked at the topics from another point of view. Also, it was more fun to discuss with my colleagues than learn for an exam by myself.
What I found the coolest way to learn new things was the purification of the protein. With an open task like this you can be creative, critical thinking learn to research and you read about many different topics and papers.
Pressure in teams may also have negative effects: Students may be hesitating to ask thing that are unclear to them, competition.
Exams we know from school and from the exam of the driving license. We know how to prepare for an exam and to learn definitions and facts. The most what I learned for exams I forgot after a while. The preparation for an exam is not attractive and only what all students want is to write it and to go home. It was more attractive to try to understand a topic without a pressure of a preparation for an exam time. By every question also if we don't write it down for the presentation I learned or I realized something. The topic was more attractive and e.g. buffers become interesting. The topics become something which I want to improve the skills for that, like a sport which we want to be pros. For example, I want to improve my skills for chromatography to separate a protein to 99.999% even not possible. With an exam I would learn what is SEC, which column material is needed for which protein sizes and after 3 weeks I would forget it. In the next practical course, I would have to look in the scripts from the last courses and may I would make all mistakes which I've made by the first time. Also the work with a team is very important. Exams you write alone. A team is important because all have skills and you learn also from that. In an exam if you don't know something then the question is not answered and you get 0 points. This you also forget after the exam and the question is still not answered.
more knowledge was acquired than with studying for exams. it helps to find out what is relevant to really understand the topic and explain it. when the group dynamic is good, the benefit is huge. (questions to ask: how to establish a good environment in advance? = still get a benefit even the group dynamic is suboptimal)

General comments
It was good to work in groups and learn to present results in front of other peoples. I would have probably understood more if i just had recaptured all the lecture slides, like this i only recaptured on part of a topic that i already understood quite well again. but of course one gets a more critical thinking when discussing the topics in more details rather than just learn things by heart.
I think I would have preferred to focus on multiple topics during the expert group time. Maybe split the four topics on the different expert groups. Because the discussion happened 1.5 months after generating the questions, I didn't remember very much not even about our own questions. It would give a larger insight into all topics if the four topics would be split in the expert groups.
To vary the question type per evaluation... - To not just evaluate one single topic, give each expert group two „exams" of 2 questions on each topic. Each question comes from a different group. (E.g. For Groups1-8 and Questions A-D, that would give an evaluation asset for Expert1 composed of 1A+3B+5C+7D and 8A+6B+4C+2D and so on...) - Only suggest the Type of question for each topic. „Make 4 Question on xyxy, use one of each of the following types: KPrim,TypA etc."
I think it makes no sense to have a exam about this topics, as it is almost the same test material as in BCH1. Even you have a lot of creativity by formulating a question it is stuff about the material we had and no independent thinking. No research about a new topic or like in the chromatography find out a new purification way (or something like that).
I think a written exam assesses more accurately the performance of the individual student and not of a group as a whole.
"For the individual understanding an exam would have been better. The exam would have generated a lot of time pressure, due to another exam in another module. Therefore, an exam would have been counterproductive. The concept of the questions is good, but requires improvement. Maybe less questions, and everyone needs to solve all questions. Like this, the final discussion makes more sense and one would achieve more understanding with less time pressure. The best of two worlds. "
better concept needed, how to organize the expert group.

Suggestions for improvements
This format showed me how to think more critically about topics and not just accept things that I am told!
I liked it. I learned/repeated a lot, which I definitely would not have done on my own. In addition, this could be made possible also without an exam.
I think it was demanding, but we learned a lot, especially for our expert topic.
I could feel that it was the first time the lecturers made this format with students. It was interesting and I learned a lot, about the topics but also about discussing with others and thinking about other opinions. Something I don't really do by studying for a written exam. However, as I said before I think it would be better to split the four topics in the expert groups. I would have learned more this way, than focusing only on buffers (which was my topic). Generally, go on like this! It was fun and interesting.
I really liked the format and I think I learned more compared to a normal exam, especially while preparing the questions. The evaluation on Friday was a bit messy and intense but I think you should do it again next year :)
1. Make an intermediary evaluation step by a partner team and then let the students rethink/reformulate their questions based on the feedback. Then make a final evaluation through an expert group OR just let the group reformulate/rethink the questions based on the final feedback from the experts. We had too little time to consult other peers on how they think they would solve the question and were sometimes lost in irrelevant discussions about how to formulate the question even better. It would help me more to understand each topic if you have something like a finalizing step/finished product. So, I think having to correct/reformulate your own questions + answers would be favorable. 2. If you continue with this type of examination (1Expertgroup per Topic) then following criteria would be interesting. - Was the question creative/innovative? A harsh and broad criteria, but many questions were similar and pretty simple. If students try to get the „creativity bonus“, then maybe the evaluation will not be that „monotonous“ for some topics.
I found it very interesting and definitely learned a lot. I would never thought that much if I just had an exam. Still I think if I compare the chromatography questions and the purification strategy I think I learned more (or I am more motivated) with the open task to develop a strategy.
Split up presentation day in two days and add theory parts. Students should at least try to solve all questions in advance.
I enjoyed the idea of the questions. I think some more theory could have been provided, especially more examples of purification strategies in chromatography as it was difficult to construct them ourselves since we lack practical knowledge. We also did not discuss this topic as much in previous lectures as for example enzyme kinetics. Personally, I think I would have learned more in all topics by writing an exam. However, I do think working in a group - as frustrating as it was at times - was beneficial, especially for such a task, since it animated many discussions about aspects that I would have otherwise not noticed.
I like this format a lot. On an exam, there are rarely chances to discuss and look at different viewpoints. There is also no opportunity to ask questions and present new ideas. This format feels like more of a preparation for real life, and helped my understanding more than an exam would.
This was a great experience!
“The final discussion could be improved. More interaction and another concept would be desired.”
thank you to the assistants for their effort! most important module.
I enjoyed the format and think I gained a lot from it. However, the input from the lecturer during question design were essential to get on the right track and rethink theoretical concepts. I think some of the time during question design and evaluation could be used to look at the theory of question design. Not the definition of the types of questions but rather what makes a good and unambiguous question. What is good and bad complexity. How to guide someone to the right idea. The concept of the most correct answer. Common errors during question design.