Tidsskriftet Læring og Medier

Dental and medical students' self-directed learning and motivation: An evaluation of two multiple-choice questions systems using machine learning

Emilie Leth Rasmussen, Centre for Educational Development, Aarhus University
Malthe Have Musaeus, It-University, Denmark
Mads Ronald Dahl, Centre for Educational Development, Aarhus University
Henrik Løvschall, Department for Dentistry and Oral Health, Aarhus University
Peter Musaeus, Centre for Educational Development, Aarhus University

Dansk abstrakt

Dette sammenlignende casestudie beskriver en undersøgelse, der ved hjælp af maskinlæring undersøger studerendes evaluering af multiple-choice spørgsmål (MCQ) som middel til læring. Med et fokus på selvstyret læring og motivation undersøger vi to systemer udviklet på Aarhus Universitet: "MED MCQ", der bruges af medicinstuderende, og "MCQ-anatomi", der bruges af tandlægestuderende. Med svar fra 126 medicinstuderende og 70 tandlægestuderende evaluerer vi to spørgeskemaer via SurveyXact. Ved anvendelse af topic modeling identificerer maskinlæringsmodellen to grupper af studerende, som på forskellig vis oplever systemet som motiverende og fremmende for deres læringsproces. De studerende oplever en øget selvstyret læring ved selv at have indflydelse på præsentation af spørgsmål og muligheden for at besvare spørgsmål uafhængigt af en underviser. Vi diskuterer, hvordan undervisere og udviklere kan bruge MCQ til at fremme studerendes læring. Artiklen diskuterer, hvordan man kan bruge maskinlæring til at analysere tekst-svar fra åbne spørgsmål.

Engelsk abstract

This comparative case study reports a study investigating student evaluation of Multiple-Choice questions (MCQ) through machine learning as a means of learning. The focus is on self-directed learning and motivation. The study evaluates two systems developed at Aarhus University: "MED MCQ" used by medical students, and "MCQ anatomy" used by dental students. The study evaluates two surveys in SurveyXact with responses from 126 medical students and 70 dental students. We use topic modeling over free text responses. The machine learning model identifies two groups of students who, in different ways, experience interacting with the system as motivating and facilitating their learning process. The students' experience increases self-directed learning by being able to choose the form of presentation of questions and answer questions independently of the instructor. The article discusses how educators and developers can use MCQs to promote student learning and how to analyze open-ended questions with machine learning.

ISSN: 1903-248X

Introduction

Multiple-Choice Questions (MCQs) represent a widely utilized assessment format characterized by a student selecting one or more answers from a predetermined list, as thoughtfully curated and presented by the test designer (Downing & Haladyna, 2004). Student learning can be facilitated in-class and out-of-class with MCQs and other written test questions (McDermott et al., 2014; Billings et al., 2021). Teachers and students evaluate MCQs positively, and research has demonstrated that they are a valid, economical, and effective means of summative assessment (Downing & Haladyna, 2004) and formative assessment, i.e., assessment for learning, perhaps because of their ability to cover a wide range of content in a relatively short period (Einig, 2013; Javaeed, 2018). However, there are conflicting views on whether MCQs mainly assess lower learning objectives or shallow learning.

MCQs and other formats of written test questions (Billings et al., 2021) have been extensively described in the context of health science education. Chandra et al. (2018) found that MCQs are a time-efficient assessment format displaying high levels of reliability in both undergraduate and postgraduate medical education. Some researchers have suggested that MCQs assess and produce superficial learning (Vansteenkiste, Simons, Lens, Sheldon, & Deci, 2004). In contrast, others have argued that MCQs assess deep learning and high-learning taxonomic goals, for instance, in dental and medical education, by including clinical vignettes (Billings et al., 2021; Palmer & Devitt, 2007).

Javaeed (2018) argued that MCQs could gauge medical students' higher-order thinking skills. The work by (Chandra et al., 2018) and (Javaeed, 2018) suggest that MCQs can assess medical students' understanding and knowledge across various educational levels. In anatomy, MCQs and computerized testing systems seem to be able to supplement rather than replace more traditional teaching methods such as lecturing and dissection (Tam et al., 2010; Durosaro et al., 2008). Furthermore, MCQ has been used to gamify anatomy teaching (Ang et al., 2018).

MCQs can be used as knowledge tests, formative assessments, and self-paced assessments (Say et al., 2022). It is an element of adaptability where the difficulty of the MCQ can be adjusted to the student's knowledge and skill level. It is developed in computer-adaptive testing (CAT) that employs item response theory to customize the difficulty of Multiple-Choice test questions according to students' abilities, as determined by their previous attempts (Bergstrom & Lunz, 1999). CAT for Certification and Licensure can be explored further in the context of computerized assessment innovations (Drasgow & Olson-Buchanan, 1999). Students can use MCQs outside the classroom to autonomously assess and enhance their knowledge and skills. This methodology not only serves as a testing tool but also a tool that facilitates the regulation of the student's behavior (Nicol, 2007).

Research has suggested that MCQs can enhance student memory (Little et al., 2012) and agency or self-directedness (Lee et al., 2021). In this context, we will explore explore two possible mediators of learning (memory and agency) in explaining how MCQ systems might benefit student learning.

First, the link between learning and memory is well established in the experimental psychological research field emanating from Bjork and Bjork (2019) and McDaniel et al. (2007). This line of research can guide us in designing MCQs to challenge students and enhance their memory (both memory retention and storage strength) and overall understanding and comprehension. MCQs can stimulate retrieval practice by presenting options and requiring the learner to select the correct answer (Little et al., 2012). It is a powerful method for strengthening memory and improving long-term retention. Thus, using MCQ, even if it involves the student making mistakes and experiencing initial difficulties, can contribute to deeper and more durable learning stemming from time on task or their potential for active recall during exam reading or better over the entire semester.

Second, there is merit in revisiting the notion of self-directed learning, which signifies students' active and autonomous position in higher education (Knowles, 1975; Merriam, 2001). Self-directed learning has been used to explain medical student learning (Mazmanian & Feldman, 2011) and dental student learning (Hendricson et al., 2006). Self-directed learning is not a unique feature of a few learning theories focusing on student agency; there are related studies on self-regulated learning (Zimmerman, 2002, 2008). Self-regulated learning can be conceived as a multifaceted process that incorporates cognitive, metacognitive, emotional, motivational, and behavioral strategies that students employ to set educational objectives and track their progress (Kizilcec, Pérez-Sanagustín & Maldonado, 2017; Zimmerman, 2002). Both theories acknowledge the importance of learners being proactive in setting educational goals, monitoring their progress, and employing various strategies to enhance their learning experience. However, self-directed learning theory is often associated with adult education and assumes that students are competent at taking responsibility for their learning, drawing on their life experiences (Edmondson et al., 2012). On the other hand, arguably, self-directed learning is a broader notion applicable to students' cognitive, metacognitive, emotional, motivational, and behavioral strategies.

Deci and Ryan's research on motivation points to the intricate connection between motivation and agency. According to Deci and Ryan (2012) and Ryan and Deci (2000), motivation can be conceptualized as the interplay between autonomy and control, profoundly intertwined with individuals' subjective sense of agency. This dynamic interplay acts as a potent mechanism for advancing students' educational achievements and enabling the cultivating of a distinct identity as proactive learners within a particular educational context. Guay, Ratelle, & Chanal's (2008) comprehensive literature review underscores the positive correlation between students' performance and their use of autonomous motivation types. Furthermore, fostering high autonomy and motivation yields positive learning outcomes (Guay et al., 2008). Studies have suggested that MCQ systems can foster self-directed learning in medical education (Pai et al., 2014; Zia et al., 2016). However, more is needed to know about how these systems mediate self-directed learning and motivation in a Danish context.

This article focuses on medical and dental students' evaluation of online MCQs, focusing on self-directed learning and motivation. We explore machine learning strategies for understanding student perceptions and study strategies. The study aims to explain to the university course administrator and teachers why MCQs might aid students' self-learning. The research aim is to explore the multidimensional nature of student evaluations through machine learning. The research question of the project is twofold. The first research question is: How can Uniform Manifold Approximation and Projection (UMAP) with Hierarchical Density-Based Spatial Clustering (HDBSCAN) uncover fine-grained patterns in students' evaluative responses? The research objective is to evaluate how students perceive MCQs and how they can stimulate self-directed learning. The second research question is how MCQ can be used pedagogically.

Materials and methods

Participants

126 medical students and 70 dental students at Aarhus University completed our surveys in SurveyXact. Their responses concerned evaluating their use of the subject-specific MCQ system. 87 dental students responded, but only 70 completed the questionnaire. 138 medical students responded to the survey, but only 126 completed the questionnaire.

Data set

In this paper, we used datasets of medical and dental students from Aarhus University. Specifically, we analyzed student evaluative responses collected from two university programs, the course in dental anatomy and five courses in the medical program: pathology (6th semester), medicine and clinical practice (7th semester), abdomen (8th semester), cardiovascular medicine (9th semester), and in gynecology, obstetrics, and pediatrics (11th semester). The dataset consisted of qualitative text feedback provided by students on their perceptions of course content and MCQs. We used sentiment analysis, an automated computational method for evaluating expressed opinions. The goal was to uncover sentiment patterns through machine learning. We wanted to identify different student groups and better understand student out-of-class study behavior in MCQs by clustering opinions.

We used Uniform Manifold Approximation (UMAP) and Projection with Hierarchical Density-Based Spatial Clustering (HDBSCAN) method. UMAP is a dimensionality reduction technique that maps high-dimensional data into a lower-dimensional space while preserving its underlying structure (McInnes, Healy & Melville, 2018). By visualizing complex relationships between data points, UMAP aids in identifying clusters and patterns that might not be apparent in the original data. Furthermore, it uses Hierarchical Density-Based Spatial Clustering (HDBSCAN). A density-based clustering algorithm can group data points based on their proximity and density distribution (dos Santos et al., 2019). It can uncover clusters of varying shapes and sizes, making it particularly useful for identifying fine-grained patterns within datasets (dos Santos et al., 2019; McInnes et al., 2018).

The study relied on a survey of dental and medical students' learning and motivation. The survey was conducted using the SurveyXact tool, comprising two questionnaires with 11 and 13 questions in medical and dental education, respectively (see Appendix-Table 4 and Table 5). The two questionnaires incorporated a combination of open-ended and closed-ended questions. Links to the surveys were mailed to dental students at Aarhus University from semesters 2-10, and another link was given to medical students who visited MCQ.AU.DK. Data was collected in the dental educational group over two weeks and three months for the medical. Table 1 compares the two surveys and samples on the structured part.

Table 1: Comparison of the two surveys

	Number of questions	**Survey duration**	**Participants (n)**	**University**	**Participants educational level**
Dental education/ dental students	13	Around 2 weeks	n= 70	Aarhus University	2nd semester to 10th semester
Medical education / Medical students	11	Around 3 months	n= 126	Aarhus University	6th,7th, 8th, 9th, and 11th semester

Design

In formulating our research design, we employed a comparative case study framework, outlined by Stake (2013). It is important to note that our study does not delve extensively into contextual elements, such as ethnographies of student usage within the systems. Alternatively, our approach could be characterized as a single case study with two illustrative instances, aligning with Yin's (2009) concept of case study research.

Evaluating the MCQ systems and characteristics of the MCQ systems

System 1 (Medical education): “MCQ.AU.DK”

MCQ.AU.DK is an online program intended to prepare medical students for exams. See Figure 1 for an example of how the program visually looks. This resource facilitates students' access to past exam papers from Aarhus University's Medical program, encompassing the five courses examined in this study. The system has an extensive database of 6266 questions. Students can tag each question as relevant or irrelevant and mark questions with comments for other students. For instance, they can comment on the relevance of the exams of a question. The analyst or teacher can use the system for data analytics while ensuring student anonymity. Teachers can integrate the system into their teaching, gaining insights into student behaviors and creating specialized question sets. In this way, it aims to provide tailored learning experiences to each student.

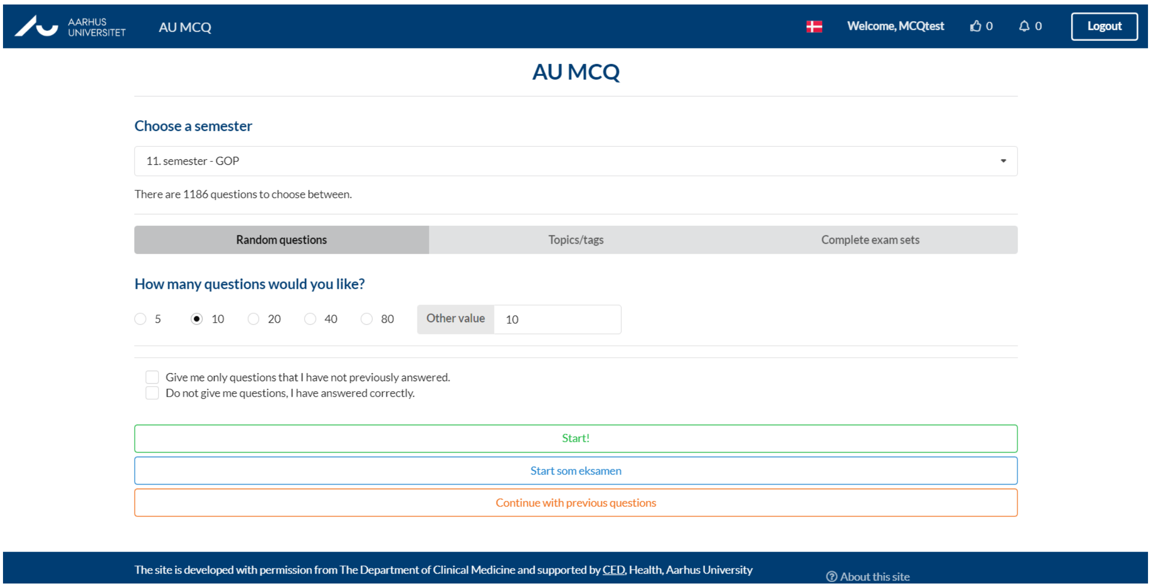

Figure 1: Screenshot of system MCQ.AU.DK

The online program was initiated around 2017 and developed by Sigurd Morsby Larsen, Mads Dahl, and Kasper Lauritzen at Aarhus University. In 2019, Centre for Educational Development took over the development and administration of the program. The program spans exam questions from fall 2011 to 2022, enabling students to engage with various medical topics.

Through the "www.mcq.au.dk/share" feature, teachers can curate specialized question sets tailored to specific topics or themes. These sets can be conveniently shared with students via learning programs, websites, or presentations. The program helps teachers foster engaging and efficient exam preparation scenarios because students can tag questions based on specialties, topics, and relevance, enhancing the program's utility. Additionally, users can provide comments and references for individual questions, fostering collaborative learning and knowledge sharing within the community. This interactive approach enhances the programs' question categorization and promotes a dynamic exchange of insights among students.

To maintain user privacy, MCQ.AU.DK utilizes Matomo. It is an open-source web analytics application that tracks online website visits. This technology respects browser opt-out preferences, avoids setting cookies, and anonymizes IP addresses. Consequently, users remain completely anonymous while benefiting from the programs' features. The usage data collected is stored securely at Aarhus University and is not shared externally.

The intention is that medical teachers can harness the capabilities of MCQ.AU.DK to enrich their teaching. The program enables educators to integrate old exam questions into their teaching materials, facilitating targeted learning experiences. Teachers gain insights into student-generated question tags, which can guide the creation of novel exam questions. Moreover, students and teachers can work collaboratively using tagged questions or sub-tags for focused preparation (see Figure 2).

In summary, the program exemplifies the synergy of technology and education by facilitating student engagement, question tagging, and teacher integration. It sets new standards for comprehensive knowledge assessment in the medical domain.

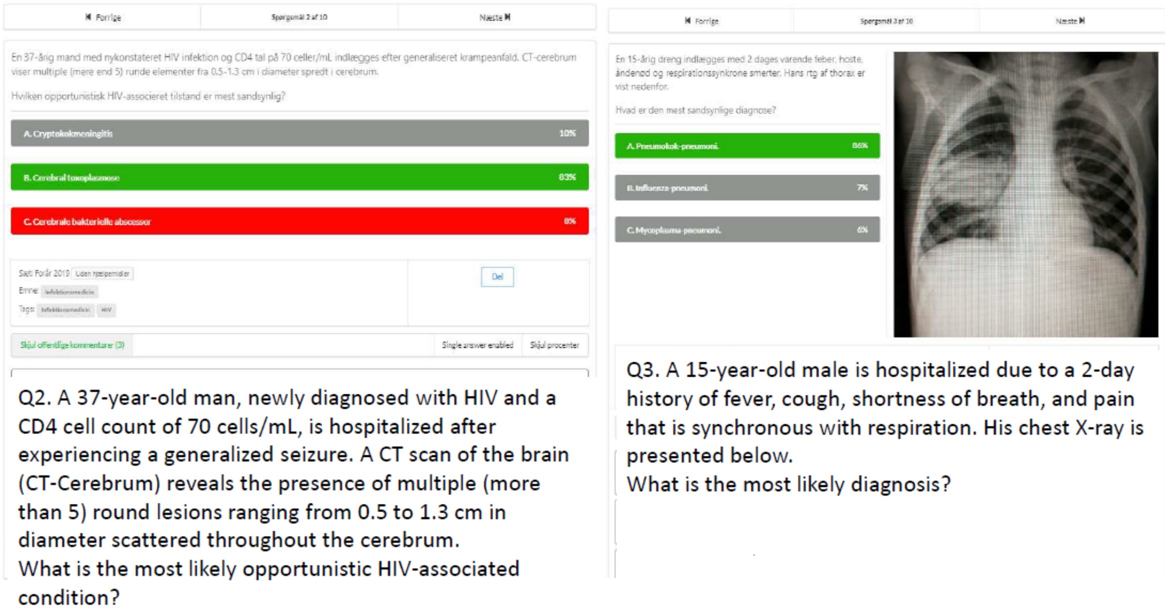

Figure 2: Screenshot of MCQ.au.dk and the provide feedback.

System 2 (dental education): "MCQ Anatomi"

The online MCQ anatomy program, developed in 2007 by Erik Christophersen and Henrik Løvschall at Aarhus University's Department for Odontology and Oral Health, aids students in anatomy training and self-assessment. It enhances anatomy learning alongside traditional teaching methods. The system has an extensive database of 5558 questions. See Figure 3 for an example of the subject overview.

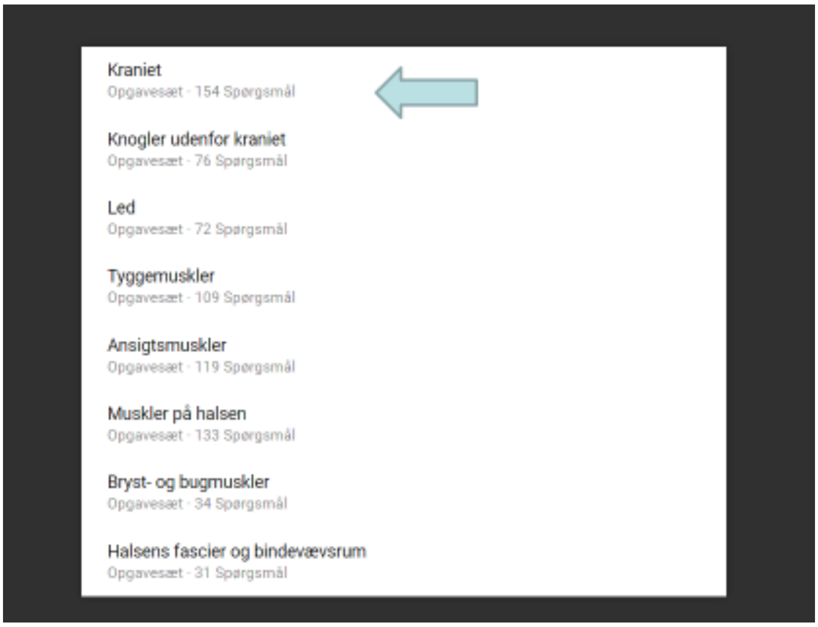

Figure 3: Screenshot of MCQ anatomy program.

The cranium: 154 questions - Bones outside the cranium: 76 questions - Joints: 72 questions - Masticatory muscles: 109 questions - Facial muscles: 119 questions - Muscles in the neck: 133 questions - Chest and abdominal muscles: 34 questions - Fascia and connective tissue spaces in the neck: 31 questions.

The MCQ collection covers the anatomy curriculum and mirrors exam question formulations. Continuously updated by Henrik Løvschall and Dennis Jensen, the program's structure follows textbook topics, allowing students to test their understanding after studying a section. MCQs have five options and a scoring range from zero to two points per question, with at least 70% needed to pass the exam.

Students receive individualized feedback and can access performance statistics. This self-assessment lets them see explanations for right and wrong answers to foster deeper tailored learning. Access codes, obtained when purchasing anatomy textbooks or eBooks, provide a year's access to the program, allowing progress tracking and self-evaluation. The MCQ program allows for individual feedback. In the MCQ program, the students can see statistics on how they are doing in their training topic. Visually, the program may show the score in the present and will be green when the student presents the level to pass the exam. The students can choose which answer options they want explained in each MCQ. In-depth explanations of both incorrect and correct answer options are available. The students can optionally use the in-depth description that belongs to each answer option. As a result, students can learn instantly or personalized according to their individual needs. See Figure 4 for an example of the provided feedback.

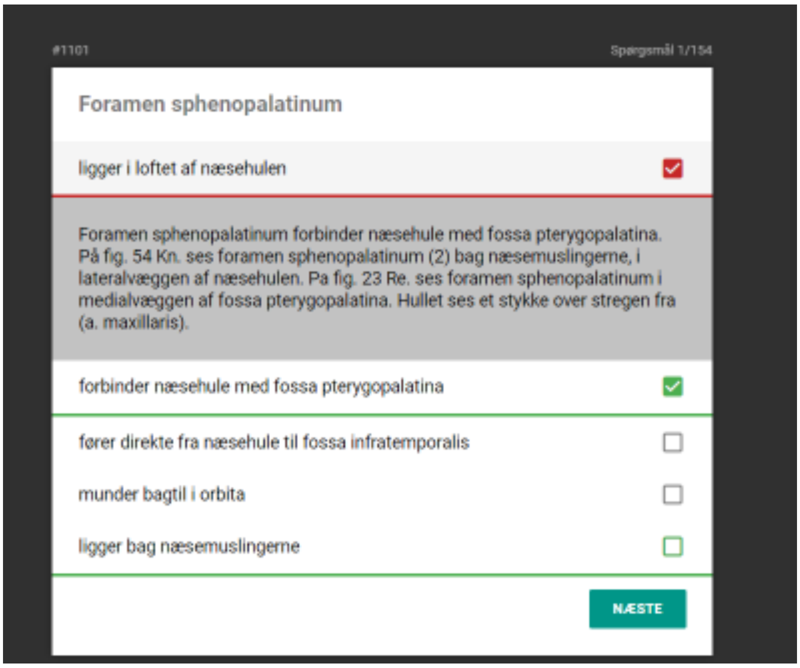

Figure 4: Screenshot of MCQ anatomy program and the provide feedback.

The foramen sphenopalatine is located in the ceiling of the nasal cavity, connecting it with the fossa pterygopalatine. In Figure 54, the foramen sphenopalatine (2) is situated behind the nostrils, within the lateral wall of the nasal cavity. In Figure 23, on the medial wall of the fossa pterygopalatine, the foramen sphenopalatine is positioned slightly above the line from the (a.maxillaries). This opening serves to connect the nasal cavity with the fossa pterygopalatine, leading directly from the nasal cavity to the fossa infratemporalis, located posterior to the orbit and behind the nostrils. Next.

In summary, the MCQ anatomy program has helped students test their knowledge of anatomy, a demanding subject requiring extensive memorization. The learning mechanism can be hypothesized to give students ownership of their studying pace and confidence in their learning. These mechanisms relate to self-directed learning and seem to have provided students with feedback on their progress. The feedback was adaptive because the knowledge tested was provided whenever students perceived a need for assessment and feedback.

Data analytic method

Using machine learning to analyze open-ended responses is a recognized method in the field of health education (Borakati, 2021). Using Python, we performed topic modeling over free-text responses, which we denote "documents." The initial step involved the generation of dense numerical representations for each document, a concept denoted as "embeddings." These embeddings served as vector portrayals of individual documents within a high-dimensional space, with the advantage of encapsulating semantic resemblances among the documents. Unlike a traditional Latent Dirichlet Allocation (LDA) approach (Yu & Xiang, 2023), the topic modeling algorithm employed for the analysis in this article is notable for its absence of assumptions regarding the intrinsic quantity of underlying topics.

Results

Spearman ranking correlation test for dentists' answers to question 1 (Q1)

Question 1 was: Describe how you use the MCQ program in your reading and exam preparation; in the original in Danish: "Beskriv hvordan du bruger MCQ program (anatomi) i din læsning og eksamensforberedelse"). The topic modeling technique for question 1 found two dominant topics. Topic 1=self- directed learning (n=25) and topic 2 = Rote learning (n=11). Most answers were classified as outliers (32 out of 68 documents), meaning that not enough responses were similar enough to be grouped. Representative documents in each cluster are shown below. Some people give detailed answers on their usage of the MCQ system, while others give short answers. Typical for the short answers in topic 2 is that respondents specify that they used the system to remember the questions by heart. On the other hand, the longer answers in topic 1 specify that students used the system to identify which topics they needed to study more (see Table 2).

The initial quote in Table 2 delves into the student's learning strategy, emphasizing self-directed learning through self-monitoring, goal setting, and planning. (See Table 2 in the appendix for the original Danish quotes.) Moreover, the student's motivation stems from a sincere interest in enhancing and mastering the subject matter, fostering a proactive and self-directed learning process. In contrast, the third quote in Table 2 illustrates that the student's self-directed learning was influenced by immediate exam requirements, leading to a constrained and somewhat superficial comprehension of the subject matter. The expressed regret reflects a yearning for a more effective and profound learning approach that extends beyond mere exam success. Furthermore, the third quote in Table 2 addresses self-directed learning because it showcases the learner's metacognitive awareness, monitoring of their study strategies, recognition of the need for better learning approaches, and a sense of responsibility for their learning. It is a reflection of the learner's self-awareness and a desire to enhance their learning experience.

Table 2: Examples response divided in the two topics funded in Question 1

Topic	Document
1	To figure out which subtopics I needed to improve on, I created an overview of all the things I made mistakes in. Then, I studied those extra before retaking the test. Additionally, I made a personal summary where I could track my performance, and based on that, I could establish a priority sequence for tackling the topics.
1	Used it to test myself in the topics I had reviewed
1	I used them because they were the ones I knew would appear on the exam. However, if they had just been supplementary to short-answer tasks, I definitely would not have used them as much. I could only remember that I had to choose the two that started with "connects," but I did not learn the actual material from them, only how to "survive" the exam. To this day, I regret that anatomy had not been approached differently so I would have been forced to learn it properly.
2	Mainly for memorizing the questions for the exam.
2	Only to memorize the questions for the exam.
2	To memorize the questions by heart.

We performed the Spearman ranking correlation test to test whether there was a correlation between the topic and sentiment for question 1. Utilizing a predetermined alpha level of α=0.05, our analysis revealed a statistically significant correlation, thus leading us to reject the null hypothesis, which posited the absence of a correlation between the topic and sentiment for question 1. The obtained p-value was 0.039, and Spearman's Rank Correlation Coefficient (rho) was calculated as 0.346. These results indicate a positive correlation between the topic and sentiment in question 1, suggesting that the greater the emphasis on self-directed learning within a response, the more positive the sentiment associated with that response. It could indicate that students in topic 2 were more neutral in assessing the program's value in preparing for the exam. On the contrary, students whose response was classified as topic 1 reported more positive perceptions (see Figure 5). Some students give detailed answers on their usage of the MCQ system, while others give short answers. Typical for the short answers in topic 2 is that respondents specify that they used the system to memorize the questions. The longer answers in topic 1 suggest that students in this group used the system to identify which topics they needed to study more.

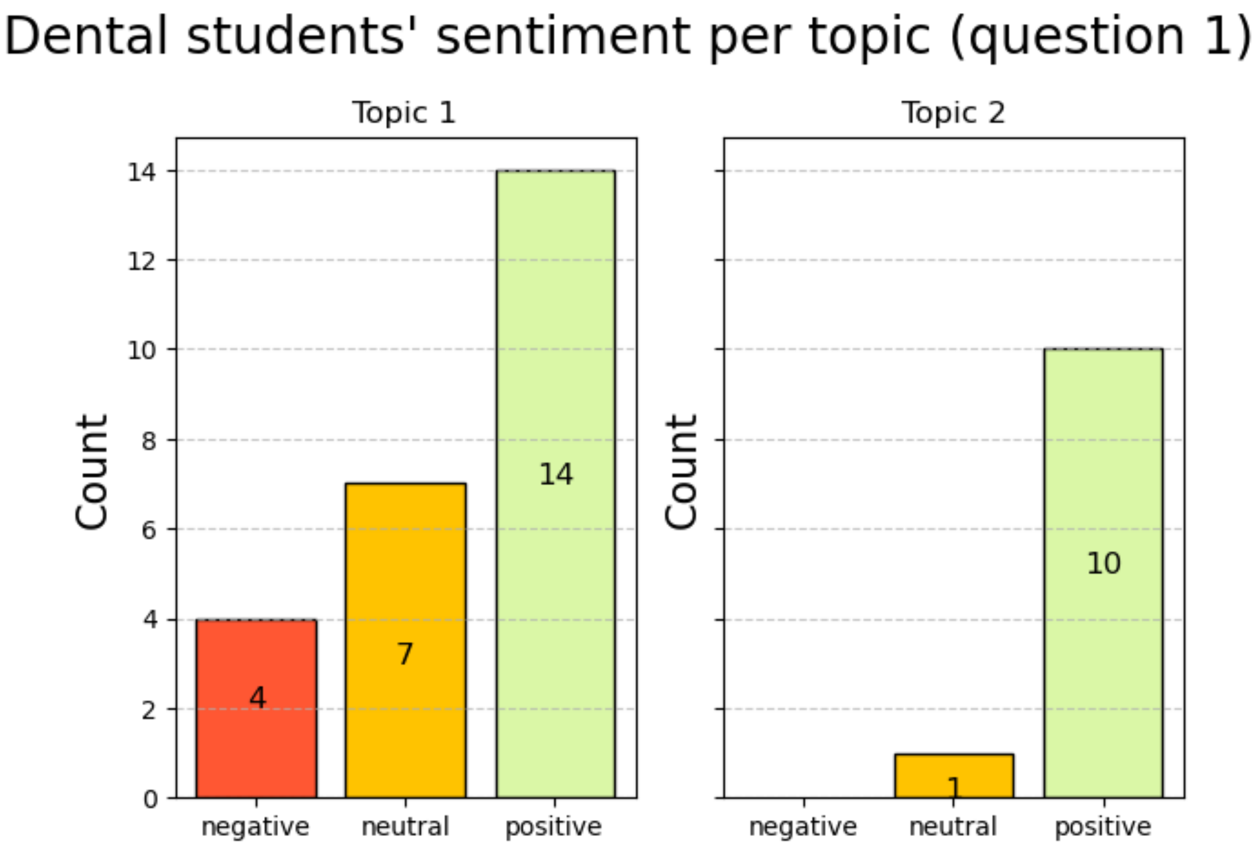

Figure 5: Distribution over sentiments in each of the two identified topics for dentist Q1.

C-TF-IDF stands for "Cluster-based Term Frequency-Inverse Document Frequency." C-TF-IDF operates at the cluster level, where topics group together similar documents. To do this, C-TF-IDF first combines all documents within a topic into one representative document. It then calculates how often specific words occur in that topic, considering the topic's size and differences from other topics. This results in a modified term frequency (TF) representation specific to that cluster. Next, it calculates an inverse document frequency (IDF) based on the average number of words in the topic and the frequency of each word across all topics. This combination of TF and IDF generates an importance score for each word within a cluster, allowing for a more accurate representation of topics. The C-TF-IDF scores range from 0 to 1, with higher values indicating greater importance and relevance of a word to a specific cluster or topic. A low C-TF-IDF score indicates that a particular word has relatively little relevance within a specific topic. Conversely, a high C-TF-IDF score signifies that a word is highly relevant and characteristic of that cluster or topic (see Figure 6).

Figure 6: A bar chart of C-TF-IDF with the top 5 words coming in each topic for Q1:
Topic 1: tried; topics; write; practice; used. Topic 2: questions; rote learning; understand; review; tasks

As introduced in the methods section, the LDA procedure was used for generating embeddings. They were facilitated by aggregating contextualized word embeddings extracted from pre-trained Bidirectional Encoder Representations from the Transformers (BERT) model. The resulting output from the pre-trained BERT model comprises 384-dimensional vectors for each word encompassed within a document. A singular vector representation is formulated for each document by pooling across these word vectors.

A dimensional reduction strategy is implemented by projecting the embeddings into a lower-dimensional space of five dimensions, employing the Uniform Manifold Approximation and Projection (UMAP) technique to enable the subsequent clustering of the generated embeddings. UMAP preserves local data structures throughout this reduction process (Ghojogh, Crowley, Karray, & Ghodsi, 2023).

The clustering is achieved using the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm. A notable attribute of HDBSCAN is its adaptability to identifying clusters characterized by varying densities, removing the necessity for a predetermined number of clusters (Malzer & Baum, 2020). Additionally, HDBSCAN demonstrates proficiency in recognizing outliers and documents that do not fit any specific cluster. These outliers may encompass isolated opinions expressed by one respondent only and are thus insignificant to make up their own topic.

To account for the fact that students wrote their answers in either English or Danish (and in a few cases Norwegian) in the documents, the embedding model we use was trained to align semantically similar documents to the same point in a vector space independent of the document language.

Sentiment analysis and dental students: Topic modelling over question (Q2) (motivation)

Q2 (question two) was: “Describe how the MCQ program (anatomy) motivates you to learn” (translated from Danish: “Beskriv hvordan MCQ program (anatomi) motiverer dig til at lære”). Running topic modeling over this question for the dentist students (n=70), we found two dominant topics in the responses. Topic 1 (appreciative of formative feedback group) with 22 responses includes responses where students find value in seeing which specific questions they answered incorrectly/correctly. The other discovered topic (topic 2= appreciative of summative feedback group), with 21 responses, includes responses where students express value in the quantitative answer percentages the MCQ system displays (appreciative of summative feedback). Table 2 shows examples of answers to Q2 from dental students. The remaining 27 responses were classified as outliers, meaning that the opinions that they each expressed were too small in numbers to make up their own topic.

In order to capture the nuances in sentiment within the text data, we will show the responses with three levels of granularity (See Table 3 in the appendix for the original Danish quotes.)

Table 3: Dental students - Sentiment analysis over Q2 with example responses

Negative	Neutral	Positive
In itself, it is not particularly motivating for learning, as you simply memorize answers to the questions posed. You really do not learn anatomy, but instead learn the familiar MCQ questions - in my view, this is a significant issue.	It motivated me in the sense that the answer did not always make sense, and therefore you had to try to find it in the book or Google it, as the reasoning was rarely formulated very well.	Getting a "score" at the end is motivating, so you know how many questions you answered correctly. It is easy to do many questions; they can be done in most places on the mobile phone and are representative of exam questions.
It does not.	You would like to have 100% correct, but it also quickly becomes rote learning.	It is an easy way to define the curriculum so that you get an impression of which parts of the book are most relevant.
It motivates more to become good at rote learning and not to understand the material.	It is condensed and organized into different topics, which makes it more manageable.	Being able to see how well I perform and concretely see where the focus is on the different topics.

Figures 7 and 8 show the distribution of sentiments for dentists and medical students' answers to Q2 ("Describe how the MCQ program motivates you to learn"/in Danish:"Beskriv hvordan MCQ program motiverer dig til at lære").

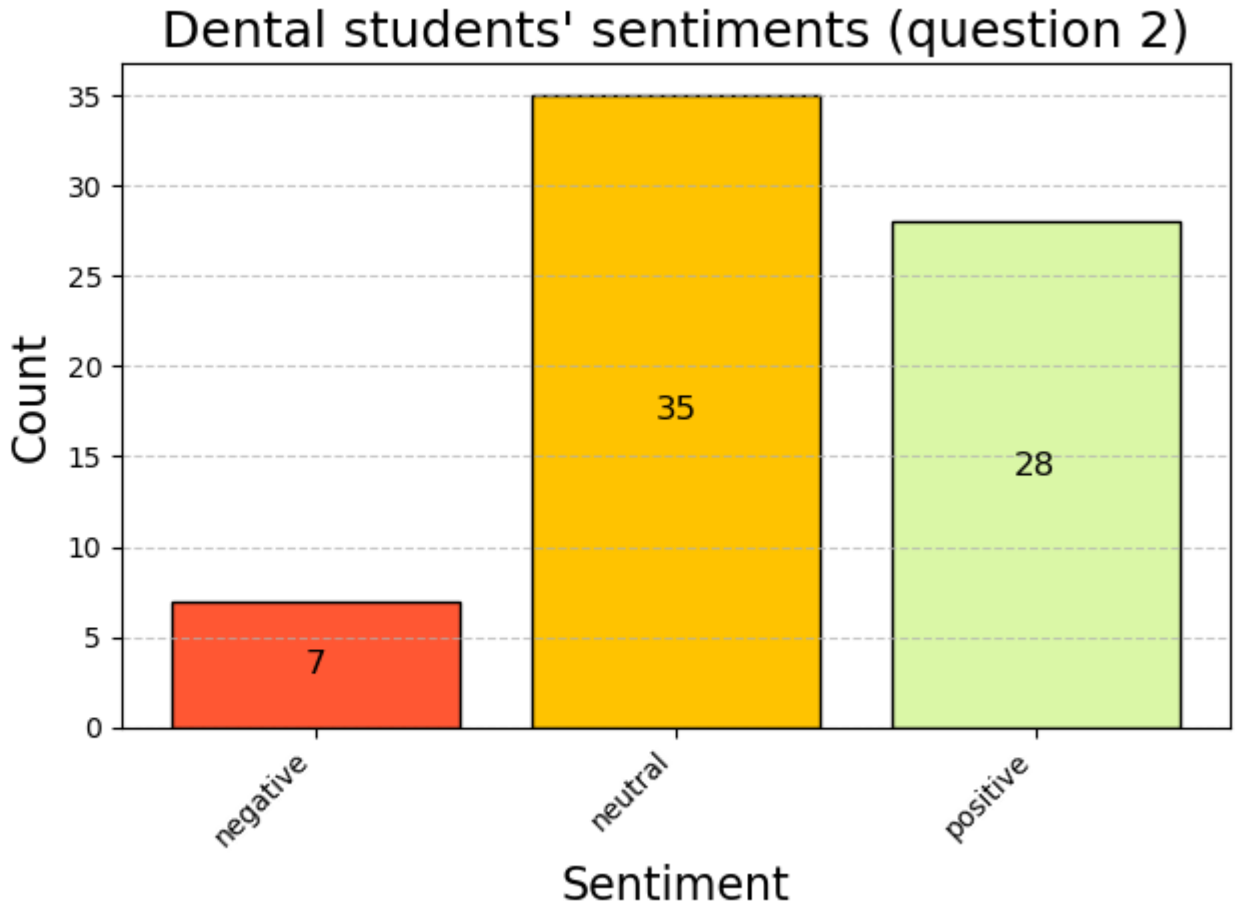

Figure 7: Dental students' sentiments (Q2)

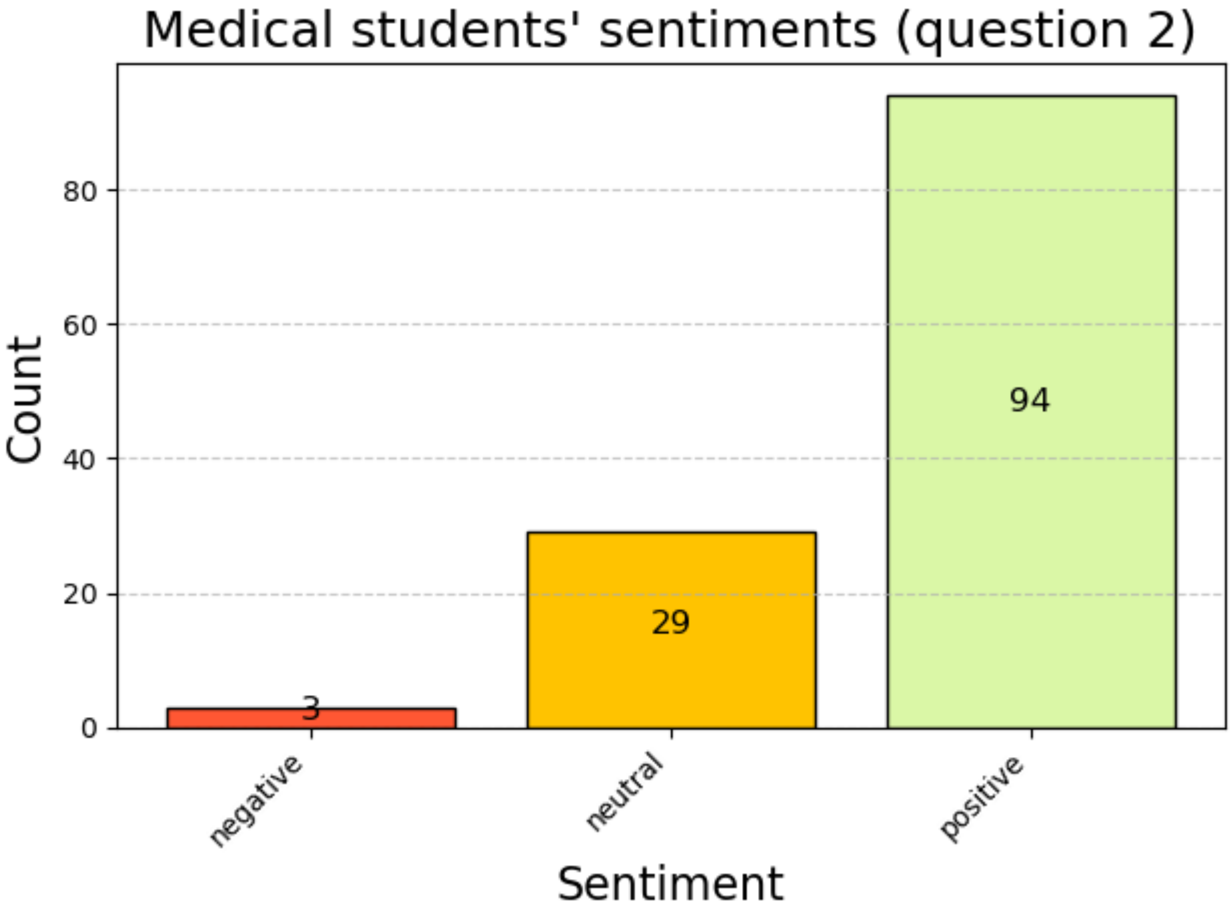

Figure 8: Medical students' sentiments (Q2)

Summary of questionnaire data

Based on the responses provided by dental and medical students regarding their utilization of the MCQ program and its impact on their learning, several notable trends emerged. When questioned about how the MCQ program motivates their learning, it was evident that a significant portion of students felt encouraged to engage in further study. They expressed a tendency to gravitate toward taking more MCQ tests and reading materials upon receiving their scores for various topics. 97% (n66/n68) of the dental participants either agreed or strongly agreed that it motivated them that the program assessed their performance through MCQ scores.

How did students integrate the MCQ program into their learning and exam preparation strategies? The responses indicated a common practice of students using the program to gauge their knowledge on specific topics. It served as a tool for them to measure their understanding and identify areas where additional study was needed to complete their curriculum. It is worth noting that 82% (n56/n68) of the dental students either agreed or strongly agreed that the MCQ program allowed them to manage their learning process effectively. During the exam period in January, 67% (n54/n81) of the participants devoted more than 20 hours to the MCQ program.

Then, we focus on medical students' motivation. Their answers demonstrated a tendency to perceive the MCQ.AU.DK program as tailored for their exam preparation and as a concrete means of mastering the curriculum in the few subjects mentioned above. 72% (99n/n138) of medical students reported using the MCQ.AU.DK over 100 times, with 86% (n119/138n) dedicating an average of more than 15 minutes per session.

When inquiring about the reasons behind their choice of MCQ order and their motivations for learning strategies, medical students typically favored selecting questions by topic during the semester. However, as the exam date approached, many opted for a change in strategy, often choosing random questions or referring to past exam sets. This shift was driven by a desire to minimize guesswork throughout the semester and to conduct self-assessments of their learning. As a result, during the exam period in January and June, students sought to consolidate their knowledge across the entire curriculum.

Both dental and medical students commonly expressed satisfaction with the questions' difficulty level, deeming it adequate and aligned with the curriculum. Moreover, a substantial majority of participants (90% (n61/n68) of dental students and 87% (n118/n136) of medical students) reported using the MCQ programs independently. Additionally, a strong consensus existed among both groups, with 98% (n135/n138) of medical students and 90% (n61/n68) of dental students agreeing or strongly agreeing that the program motivated them to seek explanations for specific answers. Notably, 65% (n90/n138) of medical students preferred using the program with questions organized by topic, while 87% (n59/n68) of dental students opted to complete the program by self-testing on specific subjects.

Discussion

This study explored dental and medical students' motivation for self-directed learning. The study demonstrated that medical and dental students were motivated by getting formative and summative feedback on their performance. Students reported that the MCQ program provided them with scores as a motivational component. By tracking their advancements, they discerned areas of improvement and identified domains warranting further practice. This evaluative feedback aids in calibrating their level of readiness for forthcoming examinations. However, as this study also explores, it is fundamental that the MCQ program is developed to provide this feedback to get the students to self-assess and take an active part in self-directed learning.

Initially, dental and medical students access the MCQ program to engage in the review and practice of questions about specific dental or medical subjects. The structured nature of the questions is particularly beneficial, as they encompass diverse facets within the discipline. Following this, students systematically work through the questions presented. When faced with uncertainties regarding answers, they take note of such instances and subsequently consult their core textbooks or supplementary study materials to facilitate a more comprehensive comprehension of the underlying concepts. This approach facilitates the consolidation of understanding, transcending mere rote memorization.

In the following, we will discuss how well the machine learning techniques help us analyze the complex dataset and throw light upon whether the two MCQ systems produced a learning environment for learning. In other words, did the students perceive that the MCQ system was a relevant means of formative assessment? Them set relevant learning goals, monitor their learning progress, and adapt their study strategies to achieve such goals? First, we will discuss machine learning as an analytic strategy, and second, we will discuss how MCQ can be used pedagogically.

Considering machine learning as a strategy

Machine learning techniques like UMAP and HDBSCAN offer a novel way to analyze and interpret student evaluative responses. This paper used UMAP for dimensionality reduction and HDBSCAN for density-based clustering. Thus, UMAP helped transform the dataset into a more manageable form while ideally retaining as much relevant information as possible.

The fine-grained patterns, sentiments, and themes enabled us to investigate the motivational impact of MCQs on students' learning experiences. Analyzing responses from 70 dental students, the study identifies two dominant topics. The first topic, consisting of 22 responses, pertains to the appreciative of formative feedback group. These students value the program for its ability to reveal the correctness of their answers, allowing them to focus on specific areas of improvement. The second topic, comprising 21 responses, encompasses the appreciative of summative feedback group, wherein students appreciate the system's display of answer percentages. The HDBSCAN helped us detect outliers in the responses. This group, consisting of 27 responses, falls into the outlier category because it represented opinions too scarce to form distinct topics. There is a danger in relying unthinkingly on one analytic technique, for instance, machine learning.

Due to practical considerations such as cost, time, and staff availability, evaluators often face the challenge of choosing between quantitative and qualitative data. They frequently need to pay more attention to the potential of incorporating carefully selected open-ended questions within quantitative instruments. Despite respondents' investment of time and effort in providing this data, open-ended responses often need more attention, relegated to a word cloud or a table in the report's appendix. Open-ended responses can illuminate the field from a broad perspective by combining qualitative and quantitative data.

As university teachers, educational evaluators, and faculty developers, we may use open-ended survey responses. However, we often need more explicit analytical strategies or expertise in analyzing and presenting text data beyond qualitative methods (Decorte et al., 2019). When open-ended data is presented alongside quantitative results, the approaches can be superficial, such as displaying a few quotes or using word clouds, which diminishes the value and depth of the information, thus undermining the effort respondents put into responding (Eleanor & Mick, 2017). It is important to note that this issue is not solely attributed to (the limitations of individual researchers), but rather stems from the field. Thus, it has been argued that the development of qualitative visualization techniques compared to quantitative analysis methods has been slower (Evergreen & Metzner, 2013).

MCQs as formative assessment

Alignment with the curriculum is pivotal when the teacher intends to use the designed MCQ question and the whole system, including how the teacher uses MCQs in and out of class (Kanzow et al., 2023; Schuwirth & Van Der Vleuten, 2004). MCQ items must continuously align with evolving learning objectives, ensuring their relevance. When students experience an alignment between a learning test format and the final test format, be it MCQs as prospective exam questions, this seems to motivate students to engage with the given test format (Little et al., 2012; McDaniel et al., 2007).

We will point to five strategies when using MCQs. First, strategic integration refers to teachers using pre-reading assignments with accompanying MCQs. It might foster proactive engagement and in-class discussions of MCQ answers to address misconceptions, promoting metacognition. Second, timely feedback after MCQ assessments and immediate feedback within the system aid comprehension assessment and improvement identification (Nicol, 2007). Adaptive learning systems enhance self-directed learning by personalizing recommendations based on strengths and weaknesses. Third, ways of using automation features such as automating grading and ensuring standardized evaluation. Fourth, strategies of collaboration. To foster self-directed learning, self-assessment MCQ sets can identify areas requiring attention. Collaborative learning via MCQ-based group discussions enhances comprehension and engagement. Fifth, the design of MCQs and their pedagogical integration demand a focus on authenticity. MCQs offer self-testing potential and exam preparation, often mirroring real-life scenarios. They also serve as valuable resources for out-of-class learning and collaborative group work, promoting agency and active learning. Even though this study suggests that students generally prefer to work individually with the MCQ, it might benefit them to participate in collaborative learning. Collaborative work with the MCQ might enhance deeper learning through group discussions, sharing perspectives, and a deeper understanding of the materials. When participating in collaborative learning, they will practice professional teamwork and critical thinking skills throughout the problem-solving of the MCQ. Despite the benefits, there are also challenges due to working collaboratively with the MCQ, such as time management in the group, the easiness of sitting at home when it fits the students, and ensuring that all the students benefit equally from the collaborative learning and not just let the other in the group answer the MCQ. Whether it is the challenge of time management within the group that leads students to prefer working independently, it could be one factor due to a tight schedule.

Although these and other studies have pointed to MCQs' convenience, reliability, and efficiency, other studies have pointed to their shortcomings associated with test pressure. In an experimental investigation involving 5th-grade students, Grolnick & Ryan (1987) found that student retention of material and learning was influenced by the degree of autonomy experienced by students. It was found that students in the experimental group with MCQ testing exhibited more surface learning with rote memorization. Thus, MCQs seemed to lead students to engage in low-taxonomic learning objectives and poor conceptual understanding when contrasted with their peers who engaged in learning without test-related pressures. This finding was repeated by Vansteenkiste et al. (2004), who found that autonomous motivation mediates first-year college student teachers' self-disclosed inclinations toward surface-level and in-depth cognitive processing.

However, this may not be generalizable to health science or other university students. University students might be more autonomous than younger students when selecting means of studying and assessment for learning. However, the cited experiments raise the question of whether the MCQ per se produces a negative pressure rather than the context of using MCQs. For instance, if students perceive the MCQs without the pressure to pass an MCQ exam, this might lead to a different learning outcome. Furthermore, critiques of MCQs have pointed out that they foster shallow thinking and surface-level assessment; this raises the need for supplementary evaluation methods (Palmer & Devitt, 2007). Considering the hypothesis that if students perceive MCQs without the pressure to pass an exam, this positively affects self-directed learning, such an approach might encourage a shift from a surface to a deeper learning strategy. If students view MCQs as a tool for self-assessment and learning rather than a means to pass an exam, they might engage more deeply with the material. It aligns with the overall theme of the passage, emphasizing the importance of the purpose and stakes associated with assessments in influencing students' approaches to learning. Addressing these concerns involves enriching MCQs' contextualization and feedback provision. When thoughtfully integrated within pedagogical approaches, MCQs can benefit student motivation and self-directed learning. To ensure that the MCQ can be used as a pedagogical tool for providing self-directed learning. It is essential that MCQ can provide immediate feedback to the students. MCQs provide instant feedback, letting students know whether their answers are correct or incorrect. This immediate feedback is crucial for self-directed learning as it helps students promptly understand their strengths and weaknesses, which are aligned with the participants' answers, where the feedback impacts the motivation for learning and their self-directed learning. The provided feedback also allows the students to self-assess their performance in the MCQ program. Here, the MCQs allow students to assess their understanding of the material. By answering questions, students can gauge their knowledge and identify areas requiring further study. Furthermore, MCQs allow for repeated practice, reinforcing learning through retrieval. Recalling information from memory strengthens knowledge retention, contributing to long-term learning.

While MCQs demonstrate value in self-directed learning, understanding their role in promoting motivation and deep learning requires consideration. To enhance the effectiveness of MCQs in medical and dental education, aligning with motivational theories and exploring alternative assessment methods should be priorities. Deci and Ryan's motivational theory (Deci & Ryan, 2012; Ryan & Deci, 2000) highlights the importance of meaningful learning experiences. MCQs' alignment with this theory rests on their capacity to convey the purpose and meaning to students. Alternative assessment strategies might complement MCQs to foster deep learning.

In this paper, datasets of medical and dental students from Aarhus University are being used. We uncovered sentiment patterns through machine learning. As we consider the future directions for research in this domain, several avenues emerge from the findings of our analysis:

1. Resampling Designs for Improved Confidence

Our study reveals a notable limitation related to class imbalance: uneven distribution of samples across different classes or categories within the dataset. Future research could explore the application of resampling designs to enhance the robustness and generalizability of our findings. By leveraging resampling techniques, other researchers might seek a design that mitigates the impact of skewed class distributions.

2. Longitudinal Analysis

Building upon the insights gained from our analysis of student sentiments, a valuable direction for future investigation involves applying a longitudinal study design. By capitalizing on the dynamic structure in the phenomenon (and ideally in the dataset) of how students develop self-directed learning over time, researchers could use a longitudinal perspective to reveal patterns, trends, and potential adaptations that went unnoticed in our analysis.

3. Generalizability

To ascertain the broader applicability of our sentiment analysis method, future studies should consider applying the approach to other subject areas in higher education. Exploring the pros and cons of MCQs in different subjects and countries will contribute to a better understanding of the generalizability and robustness of the MCQs as resources or barriers to student self-directed learning.

Despite the limitations, this study indicates that well-developed MCQ programs can provide a basis for developing students' motivation and self-directed learning. In addition, MCQ can be used as a learning tool to provide self-directed learning by providing feedback on the MCQ and making it possible for the students to use the MCQ in the way the students find it suitable for their learning. The study shows a tendency towards machine learning, and topic modeling can complement qualitative data analysis.

4. Designing better MCQs

Future studies might also investigate the importance of better use of student self-directed learning in groups and terms of all cyclical phases of learning. Thus, the MCQ systems could be designed to use student reflection more. Students could be prompted to anticipate, execute, and evaluate their learning experiences during this process. By integrating these phases, individuals can enhance their cognitive processes and achieve more effective outcomes (Zimmerman & Schunk, 2011).

Conclusion

This study utilized machine learning techniques to analyze student evaluations of MCQs, focusing mainly on the perspectives of dental and medical students about their self-directed learning and motivation. In the study, we tried to exemplify this approach in augmenting student knowledge and motivation while thinking about ways more teachers and students can use MCQs for self-directed learning. The application of UMAP-HDBSCAN demonstrated promise in extracting valuable insights from student feedback, shedding light on potential areas for improvement. However, it is crucial to acknowledge the need for further investigations to validate its efficacy across diverse educational scenarios and datasets. Moreover, exploring alternative machine learning methodologies and natural language processing techniques could offer a more comprehensive understanding of student sentiments. The system's ability to categorize questions into thematic domains proves advantageous, allowing targeted learning during study sessions and directing efforts toward specific subjects that need reinforcement.

The study emphasizes students' agency through MCQs for self-directed learning, enabling them to control their academic progress. The immediate feedback MCQs provide fosters motivation and self-directed, helping students identify their strengths and areas for improvement. Its informed approach assists them in strategically allocating efforts for further enrichment. Integrating motivation and self-directed learning into freely practiced university-level MCQs presents a complex challenge but holds potential benefits. Despite concentrating on dental and medical education, the study suggests that findings on the motivational potency of MCQs and the crucial role of feedback may resonate universally. However, further investigation is needed to confirm these insights and address potential biases like information and selection bias.

Although our study concentrated exclusively on two MCQ programs tailored for dental and medical students, it is reasonable to suggest that our findings regarding the motivational potency of MCQs and the crucial role of feedback may resonate universally across a broader range of educational contexts. However, due to the diverse subjects and practices inherent in dental and medical education, further investigation is necessary to confirm these insights and discern between subjects, with a particular focus on information bias, selection bias, and other potential biases. Because there is selection bias in evaluating only two relatively similar systems, further exploration in more diverse subjects within dental or medical education is encouraged.

Besides the limitations mentioned above, the study has limitations, including a small sample size not associated with big data and machine learning. The study's focus on machine learning and topic modeling introduces limitations related to data quality and algorithm choices, potentially introducing bias but also giving a time-efficient way to deal with open-ended questions.

The two MCQs examined in this study should ideally adopt a dynamic approach. This approach empowers students to expand their knowledge by acquiring enhanced storage of anatomical information and refining their ability to retrieve and apply these concepts. An ideal system should allow students to shape their own learning experiences. Teachers play a pivotal role in this context by designing MCQs that focus on facilitating choices rather than merely presenting content. Recognizing that testing and assessing one's knowledge catalyzes learning, transcending a mere demonstration of content mastery is essential. Through well-designed MCQs, students are encouraged to interact with learning resources and strategies and are provided with a platform to explore diverse possibilities. Students can effectively leverage the system to tailor their learning needs and methods independently or collaboratively with teachers.

Our study suggests that it is necessary to shift the emphasis in health science education from knowledge testing to the development of diagnostic reasoning and self-directed learning strategies. If MCQs could be used to promote cognitive flexibility and motivation for self-directed learning, it would have implications for improved clinical and scientific reasoning. While our study contributes to the evolving field of self-assessment and machine learning in evaluation research, acknowledging the study's strengths and limitations, future investigations should explore correlations between motivation, self-directed learning sub-processes, and student outcomes regarding memory. Although students could tailor some features of the MCQs, a more differentiated approach might also be able to distinguish between low versus high-performing students to guide effective educational interventions.

References

Ang, E. T., Chan, J. M., Gopal, V., & Li Shia, N. (2018). Gamifying anatomy education. *Clinical Anatomy, 31*(7), 997-1005.

Armstrong, R. D., Jones, D. H., Koppel, N. B., & Pashley, P. J. (2004). Computerized adaptive testing with multiple-form structures. *Applied Psychological Measurement, 28*(3), 147-164.

Bergstrom, B. A., & Lunz, M. E. (1999). CAT for certification and licensure. *Innovations in computerized assessment*, 67-91.

Billings, M., DeRuchie, K., Hussie, K., Kulesher, A., Merrell, J., Morales, A., Paniagua, M.,

Sherlock, J., Swygert, K., & Tyson, J. (2021). NBME item writing guide. Constructing written test questions for the health sciences. *Philadelphia: National Board of Medical Examiners.*

Bjork, R. A., & Bjork, E. L. (2019). Forgetting as the friend of learning: Implications for teaching and self-regulated learning. *Advances in Physiology Education, 43*(2), 164-167.

Borakati, A. (2021). Evaluation of an international medical E-learning course with natural language processing and machine learning. *BMC Med Educ, 21*, 1-10.Chandra, S., Katyal, R., Chandra, S., Singh, K., Singh, A., & Joshi, H. (2018). Medical education/original article creating valid multiple-choice questions (MCQs) Bank with Faculty Development of Pharmacology. *Indian J Physiol Pharmacol, 62*(3), 359-366.

Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. *Handbook of theories of social psychology, 1*(20), 416-436.

Decorte, T., Malm, A., Sznitman, S. R., Hakkarainen, P., Barratt, M. J., Potter, G. R., Werse, B., Kamphausen, G., Lenton, S., & Asmussen Frank, V. (2019). The challenges and benefits of analyzing feedback comments in surveys: Lessons from a cross-national online survey of small-scale cannabis growers. *Methodological innovations, 12*(1), 205979911982560. https://doi.org/10.1177/2059799119825606

dos Santos, J. A., Syed, T. I., Naldi, M. C., Campello, R. J., & Sander, J. (2019). Hierarchical density-based clustering using MapReduce. *IEEE Transactions on Big Data, 7*(1), 102-114

Downing, S. M., & Haladyna, T. M. (2004). Validity threats: overcoming interference with proposed interpretations of assessment data. *Med Educ, 38*(3), 327-333.

Drasgow, F., & Olson-Buchanan, J. B. (1999). *Innovations in computerized assessment*. L. Erlbaum Associates, Publishers.

Durosaro, O., Lachman, N., & Pawlina, W. (2008). Use of knowledge-sharing web-based portal in gross and microscopic anatomy. *Annals Academy of Medicine Singapore, 37*(12), 998.

Edmondson, D. R., Boyer, S. L., & Artis, A. B. (2012). Self-directed learning: A meta-analytic review of adult learning constructs. *International Journal of Education Research, 7*(1), 40-48.

Einig, S. (2013). Supporting Students' Learning: The Use of Formative Online Assessments. *Accounting education (London, England), 22*(5), 425-444. https://doi.org/10.1080/09639284.2013.803868

Eleanor, S., & Mick, P. C. (2017). Some Methodological Uses of Responses to Open Questions and Other Verbatim Comments in Quantitative Surveys. *Methoden, daten, analysen, 11*(2), 115-135. https://doi.org/10.12758/mda.2017.01

Evergreen, S., & Metzner, C. (2013). Design Principles for Data Visualization in Evaluation: Design Principles for Data Visualization. *New directions for evaluation, 2013*(140), 5-20. https://doi.org/10.1002/ev.20071

Ghojogh, B., Crowley, M., Karray, F., & Ghodsi, A. (2023). Uniform Manifold Approximation and Projection (UMAP). I B. Ghojogh, M. Crowley, F. Karray & A. Ghodsi (red.), *Elements of Dimensionality Reduction and Manifold Learning* (s. 479-497). Springer International Publishing. https://doi.org/10.1007/978-3-031-10602-6_17

Grolnick, W. S., & Ryan, R. M. (1987). Autonomy in children's learning: an experimental and individual difference investigation. *Journal of personality and social psychology, 52*(5), 890.

Guay, F., Ratelle, C. F., & Chanal, J. (2008). Optimal Learning in Optimal Contexts: The Role of Self-Determination in Education. *Canadian psychology = Psychologie canadienne, 49*(3), 233-240. https://doi.org/10.1037/a0012758

Hendricson, W. D., Andrieu, S. C., Chadwick, D. G., Chmar, J. E., Cole, J. R., George, M. C., Glickman, G. N., Glover, J. F., Goldberg, J. S., Change, A. C. o., & Education, I. i. D. (2006). Educational strategies associated with development of problem-solving, critical thinking, and self-directed learning. *J Dent Educ, 70*(9), 925-936.

Javaeed, A. (2018). Assessment of higher ordered thinking in medical education: multiple choice questions and modified essay questions. *MedEdPublish, 7*, 128.

Kanzow, P., Schmidt, D., Herrmann, M., Wassmann, T., Wiegand, A., & Raupach, T. (2023). Use of Multiple-Select Multiple-Choice Items in a Dental Undergraduate Curriculum: Retrospective Study Involving the Application of Different Scoring Methods. *JMIR Med Educ, 9*, e43792. https://doi.org/10.2196/43792

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education, 104*, 18-33. https://doi.org/https://doi.org/10.1016/j.compedu.2016.10.001

Knowles, M. S. (1975). Self-directed learning: A guide for learners and teachers.

Lee, N. W., Shamsuddin, W. N. F. W., Wei, L. C., Anuardi, M. N. A. M., Heng, C. S., & Abdullah, A. N. (2021). Using Online Multiple Choice Questions with Multiple Attempts: A Case for Self-Directed Learning among Tertiary Students. *International Journal of Evaluation and Research in Education, 10*(2), 553-568.

Little, J. L., Bjork, E. L., Bjork, R. A., & Angello, G. (2012). Multiple-choice tests exonerated, at least of some charges: Fostering test-induced learning and avoiding test-induced forgetting. *Psychological science, 23*(11), 1337-1344.

Malzer, C., & Baum, M. (2020). A Hybrid Approach To Hierarchical Density-based Cluster Selection. Ithaca.

Mazmanian, P., & Feldman, M. (2011). Theory is needed to improve education, assessment and policy in self-directed learning. *Med Educ, 45*(4), 324-326.

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European journal of cognitive psychology, 19*(4-5), 494-513.

McDermott, K. B., Agarwal, P. K., D'Antonio, L., Roediger III, H. L., & McDaniel, M. A. (2014). Both multiple-choice and short-answer quizzes enhance later exam performance in middle and high school classes. *Journal of Experimental Psychology: Applied, 20*(1), 3.

McInnes, L., Healy, J., & Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.

Merriam, S. B. (2001). Andragogy and self-directed learning: Pillars of adult learning theory. *New directions for adult and continuing education, 2001*(89), 3.

Murray, J. H. (1997). *Hamlet on the holodeck : the future of narrative in cyberspace*. MIT Press.

Nicol, D. (2007). E-assessment by design: using multiple-choice tests to good effect. *Journal of further and higher education, 31*(1), 53-64. https://doi.org/10.1080/03098770601167922

Pai, K. M., Rao, K. R., Punja, D., & Kamath, A. (2014). The effectiveness of self-directed learning (SDL) for teaching physiology to first-year medical students. *The Australasian medical journal, 7*(11), 448.

Palmer, E. J., & Devitt, P. G. (2007). Assessment of higher order cognitive skills in undergraduate education: modified essay or multiple choice questions? Research paper. *BMC Med Educ, 7*(1), 1-7.

Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *The American psychologist, 55*(1), 68-78. https://doi.org/10.1037/0003-066X.55.1.68 (Positive Psychology)

Say, R., Visentin, D., Cummings, E., Carr, A., & King, C. (2022). Formative online multiple-choice tests in nurse education: An integrative review. *Nurse education in practice, 58*, 103262.

Schuwirth, L. W. T., & Van Der Vleuten, C. P. M. (2004). Different written assessment methods: what can be said about their strengths and weaknesses? *Med Educ, 38*(9), 974-979. https://doi.org/10.1111/j.1365-2929.2004.01916.x

Stake, R. E. (2013). *Multiple case study analysis*. Guilford press.

Tam, M., Hart, A., Williams, S., Holland, R., Heylings, D., & Leinster, S. (2010). Evaluation of a computer program ('disect') to consolidate anatomy knowledge: A randomised-controlled trial. *Med Teach, 32*(3), e138-e142.

Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K. M., & Deci, E. L. (2004). Motivating learning, performance, and persistence: the synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *Journal of personality and social psychology, 87*(2), 246.

Yin, R. K. (2009). *Case study research: Design and methods* (årg. 5). sage.

Yu, D., & Xiang, B. (2023). Discovering topics and trends in the field of Artificial Intelligence: Using LDA topic modeling. *Expert systems with applications, 225*, 120114. https://doi.org/10.1016/j.eswa.2023.120114

Zia, S., Jabeen, F., Atta, K., & Sial, N. A. (2016). Self-directed learning (SDL), an effective method for teaching physiology to medical students. *PJMHS, 10*(3), 699-702.

Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory into practice, 41*(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

Zimmerman, B. J. (2008). Investigating Self-Regulation and Motivation: Historical Background, Methodological Developments, and Future Prospects. *American educational research journal, 45*(1), 166-183. https://doi.org/10.3102/0002831207312909

Zimmerman, B. J., & Schunk, D. H. (2011). Self-regulated learning and performance: An introduction and an overview. I *Handbook of self-regulation of learning and performance*. (s. 1-12). Routledge/Taylor & Francis Group.

Appendix

Table 2: Example responses divided in the two topics funded in Question 1 with original quotes in Danish.

Document
Til at finde ud af, hvilke delemner jeg skal få bedre styr på. Jeg lavede en oversigt med alle de ting, jeg fik fejl i, og så læste jeg ekstra på dem, før jeg tog testen igen. Derudover lavede jeg en oversigt for mig selv, hvor jeg kunne holde styr på, hvordan jeg klarede mig, og derudfra kunne lave en prioritetsrækkefølge med, hvilke emner, der skulle "tages" først.
Brugte det til at teste mig selv i de emner, som jeg havde læst op på.
Jeg brugte dem, fordi det var dem, jeg vidste der ville komme til eksamen. Men hvis det bare havde været supplement til kortsvarsopgaver, havde jeg helt sikkert ikke brugt dem lige så meget. Kunne kun huske, jeg skulle vælge de to, der startede med, "forbinder", men lærte ikke selve stoffet af dem. Kun at "overleve" eksamen. Fortryder den dag i dag, anatomi ikke havde været på en anden måde, så jeg var tvunget til at lære det ordentligt.
Primært til at lære spørgsmålene udenad til eksamen
Kun til at lære spørgsmålene udenad
Til at lære spørgsmålene udenad.

Table 3: Dentist - Sentiment analysis over Q2 with example responses with original quotes in Danish.

Negative	Neutral	Positive
Det er i sig selv ikke særligt motiverende til læring, da man blot indlærer svar til de stillede spørgsmål. Dermed lærer man sådanset ikke sin anatomi, men i stedet de kendte MCQ-spørgsmål - i min optik er dette et stort problem.	Det motiverede mig i den forstand, at svaret ikke altid gav mening, og man derfor måtte prøve at finde det i bogen eller google det, da begrundelsen sjældent var formuleret vildt godt.	Det er motiverende at få en "score" til sidst, så man ved hvor meget man har svaret rigtigt. Spørgsmålene er nemme at lave mange af, kan laves de fleste steder på mobiltelefon, og er repræsentative for eksamensspørgsmål.
Det gør det ikke	Man vil gerne have 100% [korrekt], men det bliver også nemt udenadslære	Det er en let måde at afgrænse pensum, så man får et indtryk af hvilke dele af bogen som er mest relevante.
Det motiverer mere til at bliver dygtigt til udenadslære, og modsat ikke til at forstå stoffet.	Det er komprimeret og opsat i forskellige emner, hvilket gør det mere overskueligt	At kunne se hvor godt jeg klare mig og se konkret hvor fokus er på de forskellige emner

Table 4: Questionnaire dental students – MQC program (anatomy)

What semester are you in? 1: 2: 3: 4: 5: 6: 7: 8: 9: 10:
How much time have you spent on the MCQ program (anatomy) before the exam period? 0 (no time) Up to 1 hour Over 1 and up to 10 hours Over 10 and up to 15 hours Over 15 and up to 20 hours Over 20 hours
How much time did you spend on the MCQ program (anatomy) during the exam period? Up to 1 hour Over 1 and up to 10 hours Over 10 and up to 15 hours Over 15 and up to 20 hours Over 20 hours
In my experience, the questions are up to date and in line with the curriculum Strongly agree: Agree: Disagree: Strongly disagree:
In my experience, the questions have an adequate level of difficult Strongly agree: Agree: Disagree: Strongly disagree:
It motivates me to get explanations to each question after answering Strongly agree: Agree: Disagree: Strongly disagree:
It motivates me that I can continuously see, how I manage Strongly agree: Agree: Disagree: Strongly disagree:

The MCQ program's questions allow me to control my own learning process Strongly agree: Agree: Disagree: Strongly disagree:
I have mostly used the MCQ program (anatomy) Based on randomly selected questions From topics or tags I could choose
Describe how you use the MCQ program (anatomy) in your reading and exam preparation
Describe how the MCQ program (anatomy) motivates you to learn
I have used MCQ au.dk Mostly individually/alone: With my peer study group:
Other comments on your learning or ideas to improve MCQ program (anatomy) as help for dental students' learning?

Table 5: Questionnaire medical students – mcq.au.dk

At which university are you a medical student at: Aarhus University: Aalborg University: Copenhagen University: SDU Odense University: Other:
How often have you used MCQ.AU.DK? This is my first time: 2-5 times: 5-20 times: 20 times: More than 100 times:
On average I spend the following time on MCQ.AU.DK Up to 5 minutes: From 5 to 10 minutes; Between 10 and 15 minutes: More than 15 minutes:
In my experience, the questions are up to date and in line with the curriculum Strongly agree: Agree: Disagree: Strongly disagree:
In my experience, the questions have an adequate level of difficult Strongly agree: Agree: Disagree: Strongly disagree:
It motivates me to get explanations to each question after answering Strongly agree: Agree: Disagree: Strongly disagree:
Which feature do you use the most? Random choice of questions: From topics or tags: From the whole set of exams:
Describe your reasons for this choice of feature (random choice, from tags or whole set of exams), your motivations or learning strategies?
What motivates you to use MCQ au.dk? Describe your learning goals when using MCQ au dk?
I have used MCQ au.dk Mostly individually/alone: With my peer study group:
Other comments on your learning or ideas to improve MCQ au.dk as help for medical students' learning?

ISSN: 1903-248X

Forfattere

Forfatter Emilie Leth Rasmussen

Titel Cand odont, Ph. D studerende

Institution Centre for Educational development, Aarhus Universitet

Forfatter Malthe Have Musaeus

Titel Bachelor student in Data Science

Institution It-Universitet København

Forfatter Mads Ronald Dahl

Titel MSc biotechnology, Ph.D, Master in Informatics, Special konsultent.

Institution Centre for Educational development, Aarhus Universitet

Forfatter Henrik Løvschall

Titel Cand odont, Ph. D, Lektor.

Institution Deparment for Dentistry and Oral Health

Forfatter Peter Musaeus

Titel Cand psych, Ph. D, Lektor

Institution Centre for Educational development, Aarhus Universitet