COACHING AND THE FLEMISH MEDICAL ENTRANCE EXAM: EFFICACY AND SELF-SELECTION

Lars VANDEKERCKHOVE[1]

Abstract: *This paper is the second part of a dissertation, investigating the effect of coaching and online coaching on the result of the Flemish Medicine Admission Test (FMAT). The dissertation also examines the self-selection variables into coaching, as individual differences between coached and uncoached participants could potentially mask the coaching effect. Firstly, a brief introduction refreshes the main topics of the first paper. Then, the used method of research is discussed, with attention for the sample, the content of the admission test, the content of the questionnaire and the used variables. The next part displays the results of the statistical analyses. Finally, the discussion interprets the results through the initial hypotheses before stating a few limitations and attention points for further research.*

Key words: *coaching, online coaching, e-learning, admission tests, self-selection.*

1. Introduction

Twice a year, in July and August, the Flemish government organizes an admission test to the study of medicine and dental medicine. With this admission test, the Flemish parliament tries to avoid that students had to study for 5 or 7 years after which they could possibly be denied to practice their preferred profession, as there is a limitation on the enrollment of new physicians and dentists in Flanders. [9]

Research showed that participants prepare in a variety of ways for academic admission tests [4], [5]. One preparation activity that got special attention in previous studies was coaching, as it mostly is a paying activity in which a private tutor or tutoring agency promises a gain in results by attending their program.

However, these promises have often a weak empirical base. When investigating coaching effect, it is important to separate this effect from the individual differences that self-select participants into coaching programs. Otherwise, the perceived coaching effect is confounded with the effects of these individual differences on the test result. Eight possible self-selection variables were carefully selected from previous studies, which led to four hypotheses, questioning the link between these variables and coaching attendance. These eight variables are gender, previous education, nationality, relative study cost, self-efficacy, motivation and anxiety.

For variables such as gender, previous education and nationality, group membership is quite clear to the participant him- or herself. All of these variables are

---
[1] Gent University, Belgium.

frequently quoted in research concerning the coaching effects on admission tests [15], [14], [10]. The amount of previous participations could affect the test result due to test and item familiarity, while it might also serve as a self-selection variable into coaching because of the dissatisfaction with the earlier results [4], [14], [8]. As individuals who make a smaller estimation of their odds to succeed, are probably willing to put more effort to increase these chances, the first hypothesis is:

Hypothesis 1: Individuals who, based on their gender, previous education, nationality or amount of participations, expect a lower chance to succeed, will more quickly turn to coaching.

One can also expect that more wealthy participants can more easily afford the high cost of coaching programs [7], [1], [14]. Hence the following hypothesis:

Hypothesis 2: Individuals who have fewer problems to bear the study costs, will more quickly turn to coaching.

Self-efficacy and motivation can be considered in the same assumption. Participants who approach the test with more trust and motivation, might be feeling this way because of their confidence with their preparation [15], [5]. Moreover, it is likely that highly motivated participants don't spare any effort to prepare in the best possible way [4], [10]. This led to another hypothesis:

Hypothesis 3: Individuals with a high score on self-efficacy and motivation will more quickly turn to coaching.

As a last self-selection variable, the participants' anxiety was included. Where Ryan et al. [15] dealt with it as a dependent variable, one might also assume that anxiety is an expression of stress-tolerance. Participants might choose for coaching in the hope of having enough confidence during the test administration, which led to the following:

Hypothesis 4: Individuals with a high level of anxiety will more quickly turn to coaching.

Reviewing the literature of coaching effects on similar admission tests to the FMAT, such as the Scholastic Aptitude Test (SAT) [13], [4], [14], [8], or the Medical College Admission Test (MCAT) [12] in North-America and the Israeli Psychometric Test (PET) [1] shows that coaching programs mostly have a rather small effect on the eventual test result. Thus, a small coaching effect on the FMAT can be expected, which led to the fifth hypothesis:

Hypothesis 5: Coached participants get a higher result than uncoached participants.

In an explorative part of this dissertation, the effect of online coaching on the FMAT is examined. The effectiveness of such programs has rarely been covered in research, and due to the ever evolving technology, studies from as far as ten years ago cannot be used as comparison [16], [12]. More interesting it is to compare the effectiveness of online coaching with traditional coaching programs. In this paper, online coaching is considered to be the consulting of commercial and free websites and bulletin boards by the participants, in preparation of the FMAT and its subtests. This description is broad but can be justified by the explorative nature and the recent development of the subject. The sixth hypothesis in this paper is:

Hypothesis 6: Online coached participants get higher results than not online coached participants.

Several studies noted that participants who attended a coaching program, already spent more time on preparation activities and this in an intensive way [5], [11], [13]. Therefore, it is interesting to examine whether the combination of both coaching and online coaching leads to a complementary, substitution or synergistic effect. If both concepts are complementary, it means both forms of coaching are independent of each other. If the coaching forms are substitutes, it would mean that the contents of the coaching programs and the

online coaching programs overlap. A synergistic effect would imply that coaching attendance and online coaching strengthen each other when applied simultaneously. This translated into the following hypothesis:

Hypothesis 7: The effect of a combination of coaching and online coaching is larger than the sum of the separate effects.

2. Method

2.1. Sample

In 2008, 3214 individuals participated at the Flemish Medicine Admission Test. After this test, in February 2009, all participants received a questionnaire by e-mail. 40 participants gave an unreachable address. There were 965 respondents after the first mailing. The second mailing delivered 569 respondents. The third and last mailing added an extra 558 respondents. There were 1792 respondents in total (56.5% response rate). 1 respondent refused to participate and 3 respondents did not finish secondary school at the time of the test administration. 40 respondents only partly completed the questionnaire. Out of all participants at the FMAT, 55.2% provided useful data.

33.8% of the participants are male, 66.2% are female. In total, 21 nationalities can be distinguished. Most of the participants are Belgian (83.1%). The Dutch are strongly represented in the group with other nationalities, with 15.4%.

2.2. The Admission Test

The Flemish Medicine Admission Test is organized twice a year. In 2008, the first administration of the FMAT was on July 1st while the second administration took place in August 26th.

The FMAT consists of two equally important parts: "Knowledge of and insight into sciences" (KIW) and "Acquiring and processing information" (IVV). The first part, KIW, focuses on biology, physics, chemistry and mathematics with an expected grade of difficulty around the average of the educational requirements for third grade secondary school (age 17 to 18). Each scientific subject arises in an equal amount of questions. The questions in the KIW part are mostly exercises. No mathematical proof or theorems are asked. Questions can be based on graphical representations, which also make the ability to analyze these representations important. Due to the nature of the questions, no difficult calculations are required. If necessary, data, formulas and constants are given. Participants can consult example items on the website of the FMAT. 10 out 20 points on the KIW test is the minimum required result to pass the FMAT.

The second part, IVV, tests the participant's ability to acquire and process information and the ability to solve problems based on given information. Thus, this second leg investigates the learning ability of the candidate, as the learning ability predicts future study results. This part is divided into a case study, with a reading test and a doctor-patient conversation, and a reasoning test. The reading test has a scientific theme. Participants have to read a text silently and have to answer related questions in which their assimilative capacity is tested. The doctor-patient conversation has a medical theme. Participants have to analyze a meeting between a doctor or dentist and a patient. In the conversation, a health issue of the patient arises. The participants are asked to make a fitting analysis and processing of the doctor-patient communication during these encounters. The reasoning test is designed to measure the participant's ability to process information. All acquired information needs to be processed mentally, to become problem solving. The reasoning test

contains rigid time limits, as effective tackling of a problem is necessary in acute situations. Example items for all three subtests of the IVV test can be found on the website of the admission test. Example items for the reading test and the doctor-patient conversation can also be found in the information brochure. 10 out of 20 points on the IVV test is the minimum required result to pass the FMAT.

A result of 22 out of 40 points on the complete FMAT is the minimum requirement to pass. All questions are multiple choices; firstly, answers can be noted in a workbook, and afterwards, they can be copied on an answer sheet. To discourage guessing, correction for guessing is applied. Each question has one correct answer which leads to a gain of 1 point. When an incorrect answer is chosen, 1 divided by the amount of incorrect answers is subtracted. By not answering a question, a participant gains nor loses points. Participants get a written announcement declaring whether or not they have passed the FMAT. They can also consult their result online on the website of the admission test, using a password.

2.3. Questionnaire

The questionnaire was divided into three parts. The first part contained questions about the participant's situation before the administration of the FMAT. The first 17 questions are about personal matters such as previous education, family situation and future studies. Concerning previous education, participants were asked how long their secondary studies had lasted, what study they have followed, how their result was compared to their colleagues, how their week schedule had been configured regarding 6 subjects and which their best subject was. Concerning the family background, participants were asked about the diploma and profession of both their parents. Regarding their future

studies, the questions were whether the study costs were bearable, how many times they previously participated at the FMAT and whether they planned to specialize after studying medicine. The next 51 questions are about the participant's preparation activities. These questions are divided into three subtests which means 17 questions were asked 3 times: for the science test, the reasoning test and for the doctor-patient conversation. It was tried to map the complete preparation process: information sessions at all kind of locations, free or paid coaching programs, reading, studying, consulting friends and using the internet as a preparatory instrument. If a participant confirmed using one of the 15 different activities in his or her preparation, the participant was also asked to precise the amount of hours spent on this activity. The last two questions for each subtest were whether the participants were satisfied with their preparation activities and whether they performed these activities mainly before or after the first administration of the FMAT.

The second part investigated the test experiences of the participants. For each of the three subtests, 9 propositions were posed. The respondents could respond to each of these propositions using a 1 (*completely agree*) to 5 (*completely disagree*) point Likert scale. An example of such a proposition is "I was very motivated to perform well on these tests". The first three propositions measured Self-efficacy, proposition 4 to 6 measured Motivation, while proposition 7 to 9 measured Anxiety. After the 3 sets of 9 propositions, participants were asked about their final result and about the way the test result was communicated. Participants also had the possibility to suggest improvements for the admission test.

The third part contained questions about the current situation of the participants, after the admission test. This part was divided in two parts: one for those who currently study medicine and one for the

others. Medicine students were asked at which university they are currently studying. They were also asked to respond to 11 propositions about their choice of university, using a 1 (*totally not*) to 5 (*in a very strong way*) point Likert scale. An example of such a proposition is "I chose this university because their exam system suits me". Non-medicine students were asked about the degree of disappointment for not studying medicine. They were also asked about their current study or professional activity.

A final 16 propositions were stated about the admission test as a whole. Participants could respond using a 1 (*completely disagree*) to 5 (*completely agree*) point Likert scale. An example of such a proposition is "Doing well on intellectual tasks is very important to me".

2.4. Data

All the results of the participants for the FMAT subtests that were used in this dissertation were collected from the admission test committee. All other data used in this dissertation were collected or constructed from the questionnaire. The following part gives an overview of the manner in which the variables were obtained from the admission test or the questionnaire.

Coaching The respondents were asked whether they attended training sessions by an official tutor or tutoring company, for both the sciences or reasoning subtest, or the doctor-patient conversation. In this dissertation, coaching is regarded as a paying activity in which the participant is helped by a third party and in direct interaction, in his or her preparation for the admission test. By using this operationalization, 13.5% of the participants attended coaching programs for the science subtest, 9.1% for the reasoning subtest and 8.7% for the doctor-patient conversation. For each of the three subtests, participants were also asked

whether they took the majority of their preparation activities before or after the first admission test (July). If a participant attended training sessions by an official tutor or tutoring company and if this participant did the majority of the preparation activities before July, he or she is regarded as coached for the first administration of the corresponding FMAT subtest. In case that the participant attended training sessions by an official tutor or tutoring company and this participant did the majority of preparation activities after July, he or she is regarded as coached for the second administration of the corresponding FMAT subtest.

Online coaching The consulting of paying or free websites and bulletin boards during the preparation for the admission test, is regarded as online coaching. As most of the participants did use the internet at some point during their preparation, this definition would make the group of online coached participants very large. In order to obtain useful comparison groups, the following division was made: online coached participants visited paying websites, or both free websites and bulletin boards, or a combination of these three sources. By using this operationalization, 46.7% of the participants were coached online for the science subtest, 37.6% for the reasoning subtest and 34.9% for the doctor-patient conversation. For each subtest, the group of online coached participants was divided on the fact whether they did the majority of their preparation activities before or after the first administration of the FMAT (July).

Previous education When only considering the recognized courses in Flanders and The Netherlands, 19 different courses can be distinguished. As meaningful quantitative analysis cannot be conducted with this many levels, a cluster analysis, based on the participants' school schedule, was done. Only the subjects relevant to the admission test were included. The weekly amount of hours of

mathematics, physics, chemistry, biology, Latin and Greek were standardized before being included in the analysis. This resulted in three interpretable clusters. The first cluster (n=773) groups participants with many hours of mathematics, few hours of Latin and Greek and/or an average amount of sciences in their schedule. The second cluster (n=645) mainly consists of participants with Latin or Greek as main subjects, or participants with few science subjects in their schedule. The third cluster (n=239) contains participants with many science subjects in their curriculum.

Nationality The nationalities of participants were recoded twice in an attempt to have useful comparison groups. At first, Belgian, Dutch and other were distinguished. A second recoded group Dutch and other nationalities, resulting in two groups: Belgian (n=1456) and other (n=296).

Amount of participations As the amount of participations is unlimited, respondents could fill in a number between 1 and 24.

Relative study cost The financial power of the participants was questioned in an item that asked for the feasibility of study financing. The item was: "The cost my parents, guardian or me will spend on financing of my study is…" with a possible suffix ranging from "not a problem at all" (1) until "a big issue" (5).

Self-efficacy The confidence of participants in their test capacities was questioned in three items, with answers ranging from "totally disagree" (1) to "totally agree" (5), using a scale constructed by Bauer, Maertz, Dolen and Campion [3]. An example of such an item is: "I'm confident about my capacity to perform well on this kind of tests".

Motivation The motivation of participants to successfully conclude the test was questioned in three items, with answers ranging from "totally disagree" (1) to "totally agree" (5), using a scale constructed by Arvey, Strickland, Drauder

and Martin [2]. An example of such an item is: "I urged myself to do the utmost on this test".

Anxiety The experienced anxiety during the completion of the admission test was questioned in three items, with answers ranging from "totally disagree" (1) to "totally agree" (5), using a scale constructed by Arvey et al. [2]. An example of such an item is: "During the completion of the test, I often thought I wasn't doing well".

Test results These are the weighted results by the respondents on the science and reasoning subtest and the doctor-patient conversation. For the science subtest, results from after the deliberation are taken into account.

3. Results

Tables 1 to 3 (see tables at the end of this paper) present the sample sizes, means, standard deviations, and inter correlations of measures for the first and second administration of, respectively, the science test, the reasoning test and the doctor-patient conversation.

3.1. Self-Selection and Individual Differences

Hypotheses 1 to 4 relate to the question which individuals seek coaching whilst others prefer to prepare without. In order to answer this question, three logistic regressions were conducted, corresponding the three subtests. Each time, coaching attendance for one of the three subtests was the dependent variable and relative study cost, amount of participations, self-efficacy, motivation, anxiety, gender, nationality and previous education were the independent variables. The results of these analyses are presented in table 4.

In the upper part of table 4, the results of the analysis with coaching attendance for the science subtest as outcome are

presented. The variables relative study cost, amount of participations and motivation show significant relations with coaching attendance. Participants with a lower relative study cost, higher amount of participations and a higher motivation were more likely to attend coaching for the science test.

In the middle section of the table, the results of the analysis with coaching attendance for the reasoning test as dependent variable are shown. The relative study cost, amount of participations, motivation, anxiety, nationality and previous education are related with coaching attendance. Belgian participants were more likely to attend coaching, with 10.7% of the Belgian respondents being coached for the reasoning test compared to 2.1% of the other respondents. Participants from educational clusters 1 (9.1%) and 2 (9.9%) are more likely to be coached for the reasoning test than participants out of the third cluster (7.9%). Participants with a lower relative study cost, higher amount of participations, higher motivation and higher anxiety were more likely to attend coaching for the reasoning test.

The results of the analysis with coaching for the doctor-patient conversation as dependent variable can be found in the lower part of the table. The variables relative study cost, amount of participations, motivation, nationality and previous education show significant relations with coaching attendance. Belgian participants were more likely to attend coaching programs, with 10.4% of the Belgian respondents being coached for the doctor-patient conversation compared to 1.7% of the other respondents. Participants from the first (9.4%) and second cluster (9.3%) are more likely to attend coaching for the doctor-patient conversation than participants out of the third cluster (6.8%).

3.2. The Coaching Effect and Online Coaching

The fifth research question was whether coached participants get a higher result than uncoached participants. To answer this question, a hierarchical regression analysis was conducted for every subtest and both test administrations. The control variables were entered on the first step. Even though Self-efficacy and Gender do not seem to be self-selection variables in this study, they are included on the first step in order to have a good overview throughout all performed analyses. Coaching attendance was entered on the second step.

The sixth research question, whether online coached participants get higher results than participants who are not coached online, is addressed in the same analysis by entering online coaching next to coaching attendance on the second step.

The seventh research question was whether the combination of coaching and online coaching leads to a larger effect than the sum of the separate effects. In order to test this hypothesis, the interaction between coaching and online coaching was entered on the third step. If this variable proved to have a positive significant relation with the test result, one could conclude that coaching and online coaching strengthen each other's effect when combined. However, if the relation is negative, both coaching forms weaken each other's effect. If no relation is found, then both forms are complementary. Tables 5 to 7 show the results for each subtest and test date.

Table 5 presents the results when the outcome is the science subtest, first administration. The performance on this subtest is related to the relative study cost, self-efficacy, motivation, anxiety, gender and nationality. Men score about ½ standard deviation higher than women (Male $M=10.05$, $SD=3.43$; Female $M=8.73$, $SD=3.22$). Belgians get a result

that is around 1/5 standard deviation higher than other nationalities (Belgian M= 9.23, SD=3.35; Other M=8.64, SD=3.26). Higher test performance is associated with a lower relative study cost, higher self-efficacy, higher motivation and lower anxiety. Coaching attendance and online coaching for the science subtest did not explain any variance in the performance on this subtest, nor was there an interaction effect of both types of coaching.

The lower part of table 5 shows the results of the analysis with the result on the science subtest, second administration as dependent variable. A relation with the relative study cost, amount of participations, gender, nationality, self-efficacy, motivation, anxiety and previous education were found. Men score ¼ standard deviation higher than women (Male *M*= 8.07, *SD*= 3.90; Female *M*= 7.16, *SD*= 3.53). Belgians score around 2/3 standard deviation higher than other nationalities (Belgian *M*=7.88, *SD*=3.67; Other *M*=5.61, *SD*=3.10). Participants from the second cluster score significantly higher (1/4 SD) than the participants from the third cluster (1st cluster *M*=27.14, *SD*=8.80; 2nd cluster *M*=28.22, *SD*=8.69; 3rd cluster *M*=25.76, *SD*=8.57). Higher test results are associated with a lower relative study cost, more participations, higher self-efficacy, higher motivation and lower anxiety. Coaching attendance had a significant effect; coached students score about 1/2 standard deviation higher than uncoached students (Coached *M*=9.23, *SD*=3.67; Other *M*=7.33, *SD*=3.65). Online coaching also had a significant relation with the test result. Online coached participants scored less than ¼ standard deviation higher than those who did not prepare through the internet (Online coached *M*=7.90, *SD*=3.74; Other *M*=7.35, *SD*=3.66). The interaction between coaching attendance and online coaching did not have a significant effect.

The results of the analysis found in table 6, with performance on the first administration of the reasoning test as dependent variable, show a relation with the relative study cost, amount of participations, gender, nationality, self-efficacy, motivation and anxiety. Men score about 1/6 standard deviation higher than women (Male *M*= 28.23, *SD*= 8.76, Female *M*= 26.83, *SD*= 8,91). The result of Belgians is just less than ½ standard deviation higher than the result of other nationalities (Belgian *M*=27.70 *SD*=8.82; Other *M*=24.22, *SD*=8.80). Higher test results are associated with a lower relative study cost, more participations, higher self-efficacy, higher motivation and lower anxiety. Coaching did not have a significant relation with the test result. Online coaching did have a significant relation with the result: Online coached participants scored less than ¼ standard deviation higher than those who did not prepare through online coaching (Online coached *M*=28.88, *SD*=11.31; Other *M*=27.30, *SD*=8.65). The interaction of both forms also had a significant effect. The cell means are given in table 7.

The performance on the second administration of the reasoning subtest was related to the relative study cost, the amount of participations, nationality, motivation, anxiety and previous education. Belgians score more than ½ standard deviation higher than other nationalities (Belgian *M*=34.87, *SD*=9.52; Other *M*=27.08, *SD*=9.40). Participants from the first and second cluster score about 1/6 standard deviation higher than participants from the third cluster (1st cluster *M*=34.39, *SD*=9.50; 2nd cluster *M*=34.52, *SD*=9.65; 3rd cluster *M*=28.54, *SD*=9.37). Higher test results are associated with a lower relative study cost, more participations, higher motivation and lower anxiety. Coaching and online coaching had significant effects on the test performance; coached participants were found to score more than 2/3 standard deviation higher than uncoached participants (Coached *M*=40.31 *SD*=6.38;

Other *M*=33.00, *SD*=9.91). Comparing the group means, the result of online coached participants is over 1/2 standard deviation higher than that of other participants (Online coached *M*=37.63, *SD*=8.41; Other *M*=32.14, *SD*=9.95). The interaction between both forms of coaching also had a significant effect on the test result. The cell means are given in table 8.

The results for the analysis with the first administration of the doctor-client conversation as outcome can be found in table 9. A relation with the amount of participations, gender, nationality, self-efficacy, motivation and anxiety were found. Women were likely to get a higher score than men (Male *M*= 10.12, *SD*= 4.88, Female *M*= 11.02, *SD*= 4.50). Belgians get a result that is higher than that of other nationalities (Belgian *M*=10.85, *SD*=4.64; Other *M*=9.82, *SD*=4.61). Higher test results are associated with a higher amount of participations, higher self-efficacy, higher motivation and lower anxiety. There was no significant effect of coaching for the doctor-client conversation. Online coaching did have a significant relation with the result: Online coached participants scored less than 1/6 standard deviation higher than those who did not prepare through online coaching (Online coached *M*=11.64, *SD*=4.57; Other *M*=10.95, *SD*=4.58). The interaction of both coaching forms did not have an impact on the test performance.

The performance of the second administration of the doctor-client conversation was related with the relative study cost, the amount of participations, nationality, motivation and anxiety. Belgians had a higher group mean than other nationalities (Belgian *M*=13.01 *SD*=4.85; Other *M*=10.66, *SD*=5.08). Higher test results were found to be related with a lower relative study cost, more participations, higher motivation and lower anxiety. There was a significant effect of coaching and online coaching. Coaching attendance had a significant effect;

coached students score about 1/2 standard deviation higher than uncoached students (Coached *M*=15.00, *SD*=3.94; Other *M*=12.48, *SD*=4.95). Participants who were online coached for the doctor-patient conversation scored less than 1/3 standard deviation higher than the other participants (Online coached *M*=13.81, *SD*=4.56; Other *M*=12.29, *SD*=4.98). The interaction between both coaching forms had no significant relation with the test result.

4. Discussion

In order to separate coaching effects from the effects of self-selection into coaching, the individual differences that might be responsible for self-selection needed to be examined. Therefore, research questions one to four investigate the possible effects of certain individual differences on self-selection into coaching. The analyses indicated that participants were more likely to attend coaching, for all three subtests, when they had a low relative study cost, a high amount of previous participations and a high motivation. For both the reasoning test and the doctor-patient conversation, participants were more likely to attend coaching when they were Belgian and members of the first or second educational cluster, mainly leaving out foreign participants and their corresponding courses.

The fifth research question was whether coaching positively influenced the result of the FMAT. The analyses showed that coaching had a significant effect on the second administration of all three subtests. The sixth research question was whether online coaching influenced the result of the FMAT. Online coaching proved to be effective for the first and second administration of all three subtests, except for the first administration of the science subtest. The seventh research question was whether coaching and online coaching strengthened each other's effects when used in a combination. This was for none

of the subtests the case, though an opposite, thus substitution effect was found for the first and second administration of the reasoning test.

4.1. Self-Selection and Individual Differences

Over all three subtests, coaching was related to the relative study cost, the amount of participations and motivation. Participants were more likely to attend coaching when they had a lower relative study cost, higher amount of participations and higher motivation. These findings fully support the second hypothesis, stating that participants who have fewer problems to bear the study costs are more likely to attend a coaching program. It also partly supports the first hypothesis, which stated that individuals who expect a lower chance to pass the admission test will more easily turn to coaching, and the third hypothesis, as highly motivated participants invest more time and energy in their preparation. Nationality and previous education were related to coaching for the reasoning test and the doctor-patient conversation. Belgians and participants originating from the first or second cluster were more likely to attend coaching for these subtests. As the third educational cluster mainly contains participants from other nationalities, both findings contradict with the first hypothesis and support the earlier remark that foreigners are less likely to attend coaching due to the perceived distance with the coaching program. The analyses found no effect of gender and self-efficacy. The perception of most of the participants probably was that both genders had equal chances to pass the admission test, though 4 out of 6 of the analyses regarding the coaching effect refute this thought.

4.2. The Coaching Effect

Both coaching and online coaching were defined in a rather broad way, as the used data originated from a questionnaire that tried to capture the complete preparation of the participants. Nevertheless, some interesting results came up. For coaching attendance, significant effects on the test result were only found for the second administration of the test. These effects were found for the science test, the reasoning test as well as the doctor-patient conversation. For online coaching, effects were found for the first administration of the reasoning test and the doctor patient conversation, and the second administration of all three subtests.

There are a few possible explanations for these findings regarding coaching attendance. One might assume that, for any reason, coaching attendance gets more effective after the participants took the complete admission test a first time. Following this reasoning, it is possible that participants who turn to coaching after the first admission test benefit more from these effortful programs, as they can relate the offered information better to the actual test and test situation. In other words, participants might benefit more from coaching due to a practice effect. Another possible and maybe more plausible explanation would be that participants who got a dissatisfying result on the first administration turned from no test preparation or normal test preparation activities to more time consuming coaching programs for the second administration. This is supported by the fact that, for all three subtests, the scores on the first administration of the participants who attended coaching for the second administration are lower than those of non-coached participants on both administrations and coached participants for the first administration. In this case, attending a coaching program might have been a correct choice. It is assumable though that, as motivation self-selected participants into coaching for all three subtests, these participants also invest more time in other preparation activities.

The same explanation could be applied

for the effect of online coaching on the second administration of each of the subtests. Contrary to coaching attendance, there is an effect of online coaching on the first administration of the reasoning test and the doctor-patient conversation. These findings support the general assumption that online coaching is helpful when its users lack free time to prepare through other ways. Before July, most of the participants are confronted with exams for the sixth grade of secondary school or the first year of university. As a result, most of their time is devoted to studying for these exams. As online coaching is a form of asynchronous learning, it gives participants the opportunity to prepare for the admission test whenever they have time and motivation for it. However, during summer holidays more free time is available, which might be why traditional coaching catches up on online coaching, offering the benefit of interaction with a tutor and possibly some other participants.

4.3. Limitations and Future Research

Several study limitations should be discussed. All self-selection variables were derived from previous studies and based on the available data. However, it cannot be ruled out that other possible self-selection variables, which were not discussed in previous studies or not included in the questionnaire, should have been considered. For example, participants with a different degree goal (dentist, general practitioner, surgeon) might also differ in their test preparation. Or, participants from catholic and public schools might have a similar time table, though the content of the courses is possibly different. The only way to bypass this problem is by random assignment of examinees to coaching treatment groups and non-coaching control groups [13]. This study was conducted in a observational setting though, as opposed to an experimental setting where participants would be randomly assigned to the

coached or uncoached group. This limitation was partly addressed by the use of individual differences as control variables in the analyses. Setting up an experimental design would have been better in terms of effectiveness of the procedure, however, it would also produce logistical and ethical problems. The researchers would have to organize a decent coaching program for the coached group and deny the uncoached group access to this coaching program. As some participants of the coached group might be less motivated than participants of the uncoached group, it could prove to be a difficult task to mobilize the coached group for each coaching session. This observational study got around these difficulties. Few studies concerning coaching effects on academic admission tests used random assignment and, as Messick and Jungeblut [13] stated, those who did have problems maintaining realistic control conditions [1]. Next to randomization, the statistical technique of propensity scoring might be another solution to cover the problems of self-selection. By using propensity scoring, the way how participants have been assigned to treatment and control conditions is statistically modeled. Using matching, stratification or regression analysis, coached groups could be linked with uncoached groups while the group members have an equal possibility of belonging to the coached or the uncoached group [14], [7]. As the estimated propensity scores are based on a set of entered covariates, the problem of overlooked individual differences remains.

This dissertation used the data of a study into the test preparation of participants at the FMAT. As the questionnaire tries to map the complete preparation, the focus is much broader than on coaching alone. A specific questionnaire would have made it possible to replace insignificant variables by more experimental variables which were not considered in previous research.

For example, the question when participants attended coaching was not asked. Instead, the question was when most of the preparation activities of the participant took place. While the data show that this gives an accurate image of the timing of coaching, it cannot be ruled out that participants, who invested their time in many activities except coaching before July, turned to coaching after July due to an unsatisfying result on the first administration.

In the part dealing with test preparation activities of the self-report questionnaire, participants were asked whether they attended a coaching program using one question: "Did you attend coaching program with a private tutor?". This question covers the operationalization that was used in this dissertation. However, some participants possibly misinterpret this question, as the term "private tutor" may not be generally known. The following question, regarding the total amount of hours during which coaching was attended, was left open by the majority of the coached participants. This might also be caused by the unfamiliarity with the coaching concept, or by the half year gap between the second test administration and the questionnaire. As a result, the amount of coaching could not be included in the analyses. Whether participants attended coaching programs and to which degree was questioned, but further insight into the content of these coaching programs was not addressed. Again, this is because of the aim of the questionnaire, as being a tool to map the complete preparation of the participants.

One could make the same remarks regarding online coaching, though the lack of previous research regarding this concept makes the operationalization more open. To divide the participants in an online coached and a not online coached group, three variables were used. This resulted in good comparison groups for all three subtests. Again, the amount of hours spent on online coaching could not be included because of missing data. A possible solution could be replacing the open question by a 5-point Likert scale question in analogy to most of the other questions, where only few data were missing.

The supposed pre-test variables self-efficacy, motivation and anxiety were questioned in a post-test measure. While this is not methodologically sound, from a pragmatic point of view the variables did prove to be important in the process of self-selection into coaching or as control variables in the analyses. Furthermore, it would have been practically impossible to measure these variables before participants started their preparation, as the inscription procedure for the FMAT lasts until roughly one month before the date of the admission test.

A last criticism could be the use of a self-report questionnaire, completed by the mainly young participants [14], [5]. They might over- or underestimate the effort they put in their preparation or respond in an emotional rather than rational way due to their results. However, there were six months between the second test administration and the questionnaire, the replies were treated confidentially and test takers were not obligated to respond.

Other information may be obtained from these addresses: lars.vandekerckhove@ugent.be, larsvdk@gmail.com

References

1. Allalouf, A., Ben-Shakhar, G.: *The effect of coaching on the predictive validity of scholastic aptitude tests.* In: Journal of educational measurement (1998) Vol. 35, pp. 31-47.
2. Arvey, R.D., Strickland, W., Drauden G., Martin C.: *Motivational Components of Test Taking.* In: Personnel Psychology (1990) Vol. 43, pp. 695-716.

3. Bauer, T. L., Maertz, C. P., Dolen, M. R., Campion M. A.: *Longitudinal Assessment of Applicant Reactions to Employment Testing and Test Outcome Feedback.* In: Journal of Applied Psychology (1998) Vol. 83, pp. 892-903.

4. Becker, B. J.: *Coaching for the scholastic aptitude test: Further synthesis and appraisal.* In: Review of Educational Research (1990) Vol. 60, pp. 373-417.

5. Clause, C. S., Delbridge, K., Schmitt, N., Chan D., Jennings D.: *Test preparation activities and employment test performance.* In: Human Performance (2001) Vol. 14, pp. 149-167.

6. Connelly, B. S., Sackett, P. R., Waters S. D.: *Reducing bias through propensity scoring: A study of SAT coaching.* In: Proceedings of the 23rd Annual Conference of the Society for Industrial and Organizational Psychology, San Francisco. CA., 2008.

7. Federal Trade Commission: *Effects of coaching on standardized admission examinations: Revised statistical analyses of data gathered by Boston Regional Office of the Federal Trade Commission.* Federal Trade Commission, Boston Regional Office. 1979.

8. Hausknecht, J. P., Halpert, J. A., Di Paolo, N. T., Moriarty Gerrard, M. O.: *Retesting in selection: A meta-analysis of coaching and practice effects for tests of cognitive ability.* In: Journal of Applied Psychology (2007) Vol. 92, pp. 373-385.

9. Janssen, P.: *Vlaanderens toelatingsexamen arts en tandarts: Resultaten na negen jaar werking. (The Flemish Medical Entrance Exam: Results after nine years).* In: Tijdschrift voor geneeskunde (2006) Vol. 22, pp. 1569-1581.

10. Maurer, T. J., Solamon, J. M., Andrews, K. D., Troxtel, D. D.: *Interviewee coaching, preparation strategies, and response strategies in relation to performance in situational employment interviews: An extension of Maurer, Salomon, and Troxtel (1998).* In: Journal of Applied Psychology (2001) Vol. 86, pp. 709-717.

11. Maurer, T., Solamon, J., Troxtel, D.: *Relationship of coaching with performance in situational employment interviews.* In: Journal of Applied Psychology (1998) Vol. 83, pp. 128-136.

12. McGaghie, W. C., Downing, S. D., Kubilus R.: *What is the impact of commercial test preparation courses on medical examination performance?* In: Teaching and Learning in Medicine (2004) Vol. 16, pp. 202-211.

13. Messick, S., Jungeblut, A.: *Time and Method in Coaching for the SAT.* In: Psychological Bulletin (1981) Vol. 89, pp. 191-216.

14. Powers, D. E., Rock, D. A.: *Effects of coaching on SAT I: Reasoning test scores.* In: Journal of Educational Measurement (1999) Vol. 36, pp. 93-118.

15. Ryan, A. M., Ployheart, R. E., Greguras G. J., Schmit M. J.: *Test preparation programs in selection contexts: self-selection and program effectiveness.* In: Personnel Psychology (1998) Vol. 51, pp. 599-621.

16. Welsh, E. T., Wanberg, C. R., Brown K. G., Simmering, M. J.: *E-learning: emerging uses, emperical results and future directions.* In: International Journal of Training and Development (2003) Vol. 7, pp. 245-258.

Sample Sizes, Means, Standard Deviations, and Intercorrelations of Measures   Table 1
for the First Administration of the Science Test

	N	M	Max. Value	SD	1	2	3	4	5	6	7	8
1.Relative study cost	1744	1,98	5	1,03	-							
2.Participations	1747	1,90	8	1,05	-,02	-						
3.Self-Efficacy	1721	3,40	5	,80	-,04	-,17	-					
4.Motivation	1720	4,31	5	,72	-,05	,19	,10	-				
5.Anxiety	1719	2,74	5	,90	,06	,23	-,36	,09	-			
6.Coaching	1713	-,82	1	,57	-,09	,13	-,03	,09	,03	-		
7.Online Coaching	1713	-,31	1	,95	-,06	,12	,02	,17	-,01	,17	-	
8.Science Test Result	1539	9,16	20	3,35	-,10	-,08	,37	,16	-,37	,00	,07	-

Sample Sizes, Means, Standard Deviations, and Intercorrelations of Measures
for the Second Administration of the Science Test

	N	M	Max. Value	SD	1	2	3	4	5	6	7	8
1.Relative study cost	1744	1,98	5	1,03	-							
2.Participations	1747	1,90	8	1,05	-,02	-						
3.Self-Efficacy	1721	3,40	5	,80	-,04	-,17	-					
4.Motivation	1720	4,31	5	,72	-,05	,19	,10	-				
5.Anxiety	1719	2,74	5	,90	,06	,23	-,36	,09	-			
6.Coaching	1713	9,16	1	,43	-,10	,12	-,04	,07	,06	-		
7.Online Coaching	1713	-,90	1	,67	-,05	,08	-,09	,04	,13	,33	-	
8.Science Test Result	1061	-,75	20	3,68	-,10	,13	,19	,14	-,21	,13	,06	-

Note: Correlations above .06 are significant at p< .05 and above .10 at p< .001 for all variables.

Sample Sizes, Means, Standard Deviations, and Intercorrelations of Measures   Table 2
for the First Administration of the Reasoning Test

	N	M	Max. Value	SD	1	2	3	4	5	6	7	8
1.Relative study cost	1744	1,98	5	1,03	-							
2.Participations	1747	1,90	8	1,05	-,02	-						
3.Self-Efficacy	1717	3,63	5	,81	-,04	,07	-					
4.Motivation	1715	4,32	5	,70	-,05	,20	,23	-				
5.Anxiety	1715	2,36	5	,89	,08	-,01	-,45	-,06	-			
6.Coaching	1675	-,90	1	,44	-,10	,07	,03	,10	,01	-		
7.Online Coaching	1675	-,54	1	,84	-,06	,09	,06	,16	-,05	,15	-	
8. Test Result	1539	27,29	50	8,88	-,09	,17	,22	,15	-,28	,04	,18	-

Sample Sizes, Means, Standard Deviations, and Intercorrelations of Measures Table 2 (cont.)
for the Second Administration of the Reasoning Test

	N	M	Max. Value	SD	1	2	3	4	5	6	7	8
1.Relative study cost	1744	1,98	5	1,03	-							
2.Participations	1747	1,90	8	1,05	-,02	-						
3.Self-Efficacy	1717	3,63	5	,81	-,04	,07	-					
4.Motivation	1715	4,32	5	,70	-,05	,20	,23	-				
5.Anxiety	1715	2,36	5	,89	,08	-,01	-,45	-,06	-			
6.Coaching	1675	-,91	1	,41	-,10	,07	,03	,10	,01	-		
7.Online Coaching	1675	-,68	1	,73	-,06	,09	,06	,16	-,05	,15	-	
8. Test Result	1061	33,42	50	9,97	-,09	,17	,22	,15	-,28	,04	,18	-

Note: Correlations above .06 are significant at $p < .05$ and above .10 at $p < .001$ for all variables.

Sample Sizes, Means, Standard Deviations, and Intercorrelations of Measures Table 3
for the First Administration of the Doctor-Patient Conversation

	N	M	Max. Value	SD	1	2	3	4	5	6	7	8
1.Relative study cost	1744	1,98	5	1,03	-							
2.Participations	1747	1,90	8	1,05	-,02	-						
3.Self-Efficacy	1718	3,04	5	,84	,00	-,07	-					
4.Motivation	1718	4,20	5	,76	-,02	,17	,16	-				
5.Anxiety	1718	2,71	5	,86	,08	,08	-,31	,02	-			
6.Coaching	1650	-,91	1	,42	-,08	,07	,01	,06	,02	-		
7.Online Coaching	1650	-,57	1	,82	-,05	,11	,03	,18	,04	,15	-	
8. Test Result	1539	10,73	25	4,65	-,02	,07	,17	,14	-,17	-,01	,10	-

Sample Sizes, Means, Standard Deviations, and Intercorrelations of Measures
for the Second Administration of the Doctor-Patient Conversation

	N	M	Max. Value	SD	1	2	3	4	5	6	7	8
1.Relative study cost	1744	1,98	5	1,03	-							
2.Participations	1747	1,90	8	1,05	-,02	-						
3.Self-Efficacy	1718	3,04	5	,84	,00	-,07	-					
4.Motivation	1718	4,20	5	,76	-,02	,17	,16	-				
5.Anxiety	1718	2,71	5	,86	,08	,08	-,31	,02	-			
6.Coaching	1650	-,91	1	,41	-,10	,11	-,03	,04	,01	-		
7.Online Coaching	1650	-,70	1	,71	-,03	,09	-,10	,06	,05	,23	-	
8. Test Result	1061	12,58	25	4,97	-,06	,17	,08	,13	-,13	,13	,13	-

Note: Correlations above .06 are significant at $p < .05$ and above .10 at $p < .001$ for all variables.

Logistic Regression of Coaching Attendance for the Science Test   Table 4
on Self-Selection Variables

	Parameter Estimate	Standard Error	Wald Chi-square	P	Exp(B)
Relative study cost	-,413	,083	24,631	,000	0,662
Participations	,335	,064	27,094	,000	1,398
Self-effectiveness	-,038	,101	00,681	,409	0,920
Motivation	,458	,125	13,380	,000	1,582
Anxiety	,057	,092	00,385	,535	1,059
Gender	-,113	,086	01,725	,189	0,893
Nationality	,260	,154	02,859	,091	1,297
Educational cluster 1	-,107	,122	00,764	,382	0,899
Educational cluster 2	-,102	,127	00,646	,421	0,903

Logistic Regression of Coaching Attendance for the Reasoning Test
on Self-Selection Variables

	Parameter Estimate	Standard Error	Wald Chi-square	P	Exp(B)
Relative study cost	0-,578	,107	29,297	,000	0,561
Participations	0 ,205	,075	07,474	,006	1,228
Self-effectiveness	0 ,012	,123	00,010	,920	1,012
Motivation	0 ,491	,154	10,129	,001	1,634
Anxiety	0 ,221	,107	04,262	,039	1,247
Gender	0-,135	,100	01,819	,177	0,874
Nationality	1,136	,261	18,865	,000	3,113
Educational cluster 1	0-,306	,142	04,676	,031	0,736
Educational cluster 2	0-,307	,143	04,587	,032	0,736

Logistic Regression of Coaching Attendance for the Doctor-Patient Conversation
on Self-Selection Variables

	Parameter Estimate	Standard Error	Wald Chi-square	P	Exp(B)
Relative study cost	0-,497	,106	22,084	,000	0,608
Participations	0,261	,074	12,425	,006	1,299
Self-effectiveness	0,041	,113	00,131	,718	1,042
Motivation	0,331	,134	06,122	,013	1,392
Anxiety	0,115	,111	01,074	,300	1,122
Gender	0-,031	,099	00,099	,753	0,969
Nationality	1,378	,322	18,339	,000	3,966
Educational cluster 1	0-,245	,144	02,919	,088	0,782
Educational cluster 2	0-,324	,147	04,846	,028	0,723

Regression of Science Test Result on Self-Selection Variables,   Table 5
Coaching Attendance and Online Coaching, and their Interaction

First Administration

	Parameter Estimate	Standard Error	t	p	R² Increment
Step 1					
Relative study cost	-,147	,076	-1,989	,047	
Participations	-,016	,077	-,184	,854	
Gender	,277	,087	3,200	,001	
Nationality	,398	,153	2,585	,010	
Educational cluster 1	,177	,127	1,402	,161	
Educational cluster 2	,043	,133	,300	,764	
Self-efficacy	1,012	,108	9,395	,000	
Motivation	,682	,112	6,206	,000	
Anxiety	-1,026	,097	-10,658	,000	.25
Step 2					
Coaching	,018	,131	,043	,965	
Online Coaching	,014	,081	,901	,368	.00
Step 3					
Coaching x Online coaching	,075	,132	-,566	,571	.00

Second Administration

	Parameter Estimate	Standard Error	t	p	R² Increment
Step 1					
Relative study cost	-,223	,103	-2,610	,009	
Participations	,379	,111	3,367	,001	
Gender	,266	,121	2,368	,018	
Nationality	1,100	,183	6,108	,000	
Educational cluster 1	,431	,167	2,456	,014	
Educational cluster 2	-,179	,177	-1,087	,277	
Self-efficacy	,754	,146	4,895	,000	
Motivation	,647	,164	4,243	,000	
Anxiety	-,780	,132	-5,759	,000	.18
Step 2					
Coaching	,573	,222	2,589	,010	
Online Coaching	,302	,142	2,026	,043	.01
Step 3					
Coaching x Online coaching	,018	,225	0,79	,937	.00

Regression of Reasoning Test Result on Self-Selection Variables,   Table 6
Coaching Attendance and Online Coaching, and their Interaction

First Administration					
	Parameter Estimate	Standard Error	t	p	R² Increment
Step 1					
Relative study cost	-,378	,216	-2,130	,033	
Participations	1,135	,211	5,392	,000	
Gender	,488	,237	2,056	,040	
Nationality	1,580	,429	3,822	,000	
Educational cluster 1	,352	,358	,763	,445	
Educational cluster 2	,479	,374	1,177	,239	
Self-efficacy	1,009	,315	3,153	,002	
Motivation	,776	,331	3,192	,001	
Anxiety	-2,157	,279	-8,104	,000	.15
Step 2					
Coaching	,322	,470	,714	,475	
Online Coaching	,591	,251	5,606	,000	.02
Step 3					
Coaching x Online coaching	,965	,465	-2,073	,038	.00

Second Administration					
	Parameter Estimate	Standard Error	t	p	R² Increment
Step 1					
Relative study cost	-,519	,271	-2,636	,009	
Participations	2,589	,294	8,290	,000	
Gender	,262	,309	,865	,387	
Nationality	2,151	,481	5,164	,000	
Educational cluster 1	,526	,442	1,963	,050	
Educational cluster 2	,465	,465	,632	,527	
Self-efficacy	-,035	,408	-,679	,497	
Motivation	1,370	,453	3,365	,001	
Anxiety	-1,912	,344	-5,719	,000	.21
Step 2					
Coaching	2,248	,555	3,832	,000	
Online Coaching	,990	,324	7,397	,000	.06
Step 3					
Coaching x Online coaching	-1,742	,551	-3,162	,002	.01

Coaching x Online Coaching Interaction: Table 7
Cell Means for the First Administration of the Reasoning Test

	Not Online coached	Online Coached
Non-coached	26,36	30,26
Coached	29,38	28,40

Coaching x Online Coaching interaction: Table 8
Cell Means for the Second Administration of the Reasoning Test

	Not Online coached	Online Coached
Non-coached	31,81	37,09
Coached	40,01	40,56

Regression of Doctor-Patient Conversation Result on Self-Selection Variables,   Table 9
Coaching Attendance and Online Coaching, and their Interaction

First Administration					
	Parameter Estimate	Standard Error	t	p	R² Increment
Step 1					
Relative study cost	-,024	,119	-,244	,807	
Participations	,210	,115	1,993	,047	
Gender	-,398	,130	-3,027	,003	
Nationality	,686	,241	2,716	,007	
Educational cluster 1	-,040	,198	-,089	,929	
Educational cluster 2	,013	,207	,135	,893	
Self-efficacy	,541	,155	3,667	,000	
Motivation	,626	,162	4,253	,000	
Anxiety	-,961	,152	-6,230	,000	.08
Step 2					
Coaching	-,301	,270	-1,191	,234	
Online Coaching	,618	,143	2,805	,005	.01
Step 3					
Coaching x Online coaching	,252	,269	,937	,349	.00

Second Administration

	Parameter Estimate	Standard Error	t	p	R² Increment
Step 1					
Relative study cost	-,254	,151	-2,195	,028	
Participations	,582	,162	3,375	,001	
Gender	-,216	,172	-1,258	,209	
Nationality	,921	,273	3,701	,000	
Educational cluster 1	,209	,244	1,118	,264	
Educational cluster 2	,050	,258	,050	,960	
Self-efficacy	,511	,197	1,925	,055	
Motivation	,540	,227	2,680	,007	
Anxiety	-,664	,193	-3,509	,000	.09
Step 2					
Coaching	,867	,311	2,738	,006	
Online Coaching	,337	,187	3,580	,000	.02
Step 3					
Coaching x Online coaching	-,412	,309	-1,333	,183	.00