



Making ethics teaching more effective with a three step model

Hans Teke¹ 

Accepted: 12 October 2020 / Published online: 22 October 2020
© The Author(s) 2020

Abstract

In this study, the impacts of two different “methods” for teaching ethics as part of the religious education in the Swedish upper secondary school were compared by means of a non-randomized controlled trial in two parts, involving 542 students. The question was which “method” had the greatest capacity to generate long-term ethical awareness in the students. The intervention condition consisted of students whose teachers were instructed to teach according to the Three Step Model, a teaching method influenced by research concerning how moral autonomy and ethical awareness could be increased by means of instruction and training. The control condition consisted of students whose teachers were instructed to teach basically as usual but with some added guidelines. During the trial, all students were given a pre-test before the ethics section had started and a post-test 10–12 weeks after it was finished. When quantified and summarized, the results showed an advantage of the intervention condition in measure B (development of demonstrable knowledge) but an advantage of the control condition in measure A (self-assessed ethical awareness); however, the advantage of the intervention condition was clearer and stronger. Even though the intervention students did not *experience* a stronger development, they appeared to have *learned* significantly more, not least in terms of procedural knowledge in ethical problem solving. The tentative conclusion is therefore that the Three Step Model is a more effective method for increasing ethical awareness, at least if one defines ethical awareness and measures it the way it was done in this study.

Keywords Ethics · Ethics teaching · Didactics · Religious education · Upper secondary school · Impact study · Long-term effects

✉ Hans Teke
hans.teke@uvtet.lu.se

¹ Department of Educational Sciences, Joint Faculties of Humanities and Theology, Lund University, Lund, Sweden

About the study

It might be the dream of every person who teaches ethics that the teaching will make an actual difference in the minds of the students, months or even years afterwards; either that they think differently about ethical issues, that they get a wider perspective on right and wrong or even better: that they will make better choices in the face of a moral problem. However, it is not that easy to find out how such long-term effects could be achieved. One way of doing it would be to compare at least two different ways of teaching ethics in this regard and see which one is the most effective. This was done in a large-scale quantitative study, a non-randomized controlled trial, in which an intervention was made in the ethics section of the religious education (a mandatory but non-confessional subject) in the Swedish upper secondary school (covering the ages 16–19). The aim of the study was to find out what the most effective way would be to increase the long-term ethical awareness in the students, whether it was by using the Three Step Model or teaching in a more traditional way. We will come back to the results of this comparison.

Previous research

Regarding previous research about the effects of ethics teaching (here defined as a separate curricular event exclusively devoted to ethics or moral philosophy), most of it has *not* been conducted at the secondary school level. There are, however, a few such studies deserving to be mentioned. In a so-called natural experiment among secondary school students in Belgium, it was concluded that students who had attended a Roman Catholic religion course scored higher in the Defining Issues Test (a psychometric measure of moral development) than those who had attended a course in non-confessional ethics. It was also concluded, though, that the students who considered themselves non-religious scored higher than those who considered themselves Christians (Mortier 1995, p. 11 f.) In a non-randomized controlled trial at an American high school, it was concluded that an ordinary introductory ethics class and an economic ethics class respectively were more effective for promoting “comprehensive moral maturity”, compared to a role-model ethics class and a control group without ethics teaching respectively (DeHaan et al. 1997, p. 5 f). Moreover, in a trial among American high school students, it was concluded that a profound introduction in economy and ethics called “Ethical Foundations” increased the degree of *theoretical* knowledge about ethics in the students who attended it, compared to the students who did not attend it, whereas the ethical *attitudes* were not at all affected by the introduction (Niederjohn et al. 2009, p. 78).

The overwhelming majority of the existing studies on the effects of ethics teaching, however, have concerned ethics instruction at the university level, and this material is so rich and diversified that it is difficult to overview. It was therefore very helpful to learn that in 2009, two meta-analyses were published, regarding science and business ethics respectively. Both of them aimed to improve the teaching practice in their respective areas by identifying the characteristics of the instructions that generated the largest effect sizes. What both analyses indicated was that the best results were achieved when the instruction had a *cognitive* orientation, i.e. aimed at giving the

students strategies for ethical problem solving or reasoning – that is, without the dependence on normative theories (Antes et al. 2009, p. 380 f.; Waples et al. 2009, p. 139 f.) One example of this, relatively successful, approach to instruction was an experiment among social work students in Canada, in which it was concluded that a profound workshop in ethical decision making had a significant impact on the participants' abilities to make a well-informed decision, compared to those who did not receive the same workshop (Gawthrop and Uhlemann 1992, p. 39 f.)

A Swedish researcher who has developed moral education of this problem-solving kind is Kavathatzopoulos. Following the cognitive developmental tradition of Piaget and Kohlberg, he has addressed the skill for moral autonomy, i.e. rational and independent moral problem solving, and implemented a program for stimulating the development of it in a lasting and measurable way. In a number of studies, he has shown that it is possible to make people change the ways in which they approach a moral problem, by just giving them a workshop with instructions and training. They then go from solving the problem *heteronomously*, i.e. with reference to a moral rule or authority, to solving it more *autonomously*, i.e. with arguments from the concrete situation (Kavathatzopoulos 1993, p. 384.) In a word, they become better at solving ethical problems deliberately by themselves, which does not necessary mean that they become “better” or more altruistic people, but rather that they become less prone to make blind, hasty or one-eyed decisions, which we easily do when strong emotions are involved. (Kavathatzopoulos and Rigas 2006, p. 55 f.)

In his later research, he has focused on training people in applying the autonomous skill to real and personal, normally work related, ethical problems – as he claims that doing this is necessary for the autonomy to “spill over” to real life conduct – with promising results in follow-up tests given months and even years later. The increased degree of moral autonomy appears to have remained and in addition, participants have reported that the training has increased their ethical awareness, i.e. made them more perceptive to the ethical problems in the first place (Kavathatzopoulos 2004, p. 285; Kavathatzopoulos 2012, p. 389 f.) When the Three Step Model was developed, it was influenced by Kavathatzopoulos' research, showing how desirable long-term effects could be achieved by means of a relatively small intervention (the autonomy training), which could be easily transferred to the ethics teaching in upper secondary school.

“Regular” ethics teaching VS. the three step model

As a preparation for this study, 18 upper secondary school teachers (a majority of which later became informants) were met with and asked to describe how they normally structured the ethics section of the religious education. On average, the section was reported to take about five or six sessions to complete. The most common introduction was to give examples of, and initiate a discussion about, everyday ethical problems without immediately connecting those to normative theories. But once the theories had been explained, which was normally done in session two or three, they were used as the most important guides regarding what is right and wrong: the students were taught how to argue for a solution to an ethical problem on the basis of a normative theory. This was typically done by letting the students discuss a number of problems in small groups and later in the whole class (during a few sessions), often

in ways that gave them the opportunity to take turns being a utilitarian, a duty ethicist etc. When the examination task was given, it typically had the character of an essay in which the students were asked to discuss an important ethical problem (such as capital punishment or euthanasia) from different normative points of view and formulate a personal opinion about it. This was done either at home or in class, and the students were normally allowed to have the normative theories in front of them when doing it (only one teacher gave the students a more traditional written test). Even though not necessarily representative of a larger population, these reports can together give us an idea of what “regular” ethics teaching is normally about.

The Three Step Model, which was the challenger in this study, is intended to be a new and more effective formula for ethics teaching, not least regarding long-term results. It is a method rather than a content, based on the conviction that before it is relevant to introduce the students to normative theories, one has to help them develop their *personal* abilities to deal with (emotionally charged) real-life problems – otherwise there will be no link in their minds between ethical theory and moral practice. Here, the development of moral autonomy is a key component, as it means becoming more independent and rational in the face of a moral problem. The method is meant to take 6–9 lessons to implement, is made up by 3 distinct steps and 5 different exercise types (with several examples of each type); all of them worked out beforehand and gradually preparing the students for the examination test, also worked out beforehand, which the teacher finally gives them.

- Step 1 is that the students learn how to recognize an ethical problem. This means that they become able to derive an ethical problem from a hypothetical situation where no such problem has yet been formulated, and make explicit why they derived this particular problem (exercise type 1). The point of this step is that it helps them assigning an ethical problem with words and thereby becoming more aware of it; this is a preparation for the autonomy training in Step 2.
- Step 2 is that the students learn how to solve an ethical problem with arguments from the concrete situation. This means that they become able, in the face of an ethical problem, to recognize the alternative ways of solving it and, with arguments from the concrete situation, justify the decision that they chose. In other words: they learn how to solve the problem autonomously instead of just referring to a rule or authority saying what is *always* right or wrong in a similar situation. Doing this, they first work with hypothetical problems (exercise type 2A) and then with real problems from their own lives which they exemplify (exercise type 2B, the most important of them all). The point of this step is that it takes their ethical awareness yet a bit further, as it means applying a conscious problem-solving process to situations in which they normally just act by instinct or habit.
- Step 3 is that the students learn how to relate a chosen solution to a normative theory; as a suggestion classical utilitarianism and Kant’s duty ethics respectively. This means that they, after having been introduced to these theories, become able to tell how a particular solution to an ethical problem would most likely be judged by a representative of the respective theory. (But importantly, the solutions in question have not been *derived* from the use of normative theories but from an autonomous problem-solving process.) When doing this, they

first work with the solutions to hypothetical problems (exercise type 3A), then with the solutions to their real and personal problems from step 2 (exercise type 3B). Thereby they learn how to, philosophically, test a given solution to an ethical problem and see if it holds for a more universal way of determining right or wrong, which can be helpful in situations when they find themselves unable to really determine the right way of action.

The examination test, finally, is made up by a hypothetical but realistic situation that the students are exposed to in a written test, where they should 1: recognize the ethical problem; 2: solve it with arguments from the concrete situation and 3: relate the solution to utilitarianism and Kant's duty ethics respectively. In other words, all the three steps in the model are repeated once again, to make it more likely that what they have learnt during the ethics section will be embedded in long-term memory.

Besides providing the students with basic knowledge about normative ethics, the purpose of the Three Step Model is to increase their long-term ethical awareness. In the specific sense that the term is used here, it means that they become more attentive to situations in which a moral decision has to be made (by someone). A desirable implication of this is that they also become more attentive to their own moral behavior and thus, hopefully, more inclined to make more well-reasoned decisions. It is an ethical awareness of a kind implying that one sees ethical problems in situations where one did not see them before, perhaps because one did not look for them. One could say that this is a necessary but insufficient condition for being able and ready to solve ethical problems autonomously, but it is here hypothesized that exercising moral autonomy will increase ethical awareness as well.

Assignment of participants to conditions

The question that this study set out to answer was whether the Three Step Model was a more effective way of increasing ethical awareness than more "regular" forms of ethics teaching. In order to answer this, a non-randomized controlled trial in two parts (sub-studies) was conducted in which some teachers, along with their respective students, were assigned to the intervention and others to the control condition. The assignment was merely based on which teachers agreed to participate in the trial and what categories of students were needed at that point in the process. The striving was to obtain a balance between the two teaching conditions, by having the same proportion (or disproportion) in both of them between males and females, and between students at vocational programs (not aspiring for higher studies) and students at higher education preparatory programs. Since a randomization would have required a different planning of the study from the beginning, this was probably the best way to minimize the risk of a systematic skewed distribution.

Guidelines for the control condition

It was also important to take some steps for preventing the design of the study from favoring the method it set out to test. One of these was to let regular ethics teaching be

represented not merely by “habitual” ways of teaching, but by the use of a *basic teaching guide* which allowed the teachers in the control condition to teach roughly the way they were used to, but with some added guidelines that would make their teaching comparable to the Three Step Model. These guidelines were, in short: (1) give the students at least six lessons of ethics teaching, (2) explain that the purpose is to make them better at dealing with moral problems in real life, (3) focus the discussion on realistic and not too dramatic problems, (4) tell them to justify their solutions with arguments, and (5) explain utilitarianism and Kantianism to them. Thereby it was hoped that there would not be any outcome differences only because some teachers had received helpful instructions and some had not.

The pre- and post-test

Another step was to ensure that the effectiveness of the two “methods” was measured by a relatively neutral standard: an assessment of the potential for increasing ethical awareness irrespective of which “method” the teacher had used. The assessment was divided into a pre-test for the students to complete before the teacher had introduced the ethics section, and a post-test for them to complete 10–12 weeks after the examination.

The purpose of the *pre-test* was to find out which knowledge they had in advance. They were therefore initially asked to tell whether they had studied ethics before (yes or no). After that they were asked to (1) tell why a problem having to do with matching clothes is not an ethical problem, (2) give an example of an ethical problem, (3) suggest a solution the problem and justify it properly, (4) say something about the essence of utilitarianism and (5) Kantianism.

The purpose of the *post-test* was twofold. First, to see how the students’ *self-assessed* ethical awareness had increased after the ethics teaching (part A). This was done by asking them to consider to what degree they felt that they (1–2) had become better at discovering ethical problems and their possible solutions; (3) reflected more upon their behavior towards others; (4) had become better at giving arguments for what they considered right or wrong and (5) had become more interested in ethical issues. Second, to see how much their *demonstrable* knowledge about ethics had increased since the pre-test (part B). This was done by repeating the pre-test questions to see how the quality of the answers had improved Figs. 1, 2 and 3.

Thus, measure A of students’ development was the score in the first part of the post-test (self-assessment), while measure B was the increase in points from pre-test to the second part of the post-test (demonstrable knowledge). In both measures, there were some items that were considered more relevant than others, such as the exemplification of an ethical problem and the justification of a solution to it in measure B; therefore,

<i>Pre-test</i>	<i>Post-test</i>
Measured what demonstrable knowledge about ethics the students had in advance.	Measured (A) how the students’ self-assessed ethical awareness had increased as a result of the teaching and (B) how their demonstrable knowledge about ethics had increased since the pre-test.

Fig. 1 The two parts of the assessment

<i>Sub-study 1</i>	Seven teachers followed the Three Step Model (intervention condition) and eight followed the basic teaching guide (control condition).
<i>Sub-study 2</i>	Three teachers who followed the basic teaching guide in Sub-study 1 (control condition) also agreed to follow the Three Step Model in a few additional study groups (intervention condition)

Fig. 2 The two parts of the study

these items were assessed separately to make the assessment sharper. The hypothesis was that the intervention condition would score significantly higher than the control condition in both measures A and B.

The two sub-studies

In Sub-study 1, a cohort of students whose teachers had followed the Three Step Model (intervention group) was compared to a cohort whose teachers had followed the basic teaching guide but otherwise taught as usual (control group). The primary purpose was to find out if there would be any significant outcome differences between the conditions, indicating that one “method” was stronger than the other. The secondary purpose was to find out if there would be any significant outcome differences between male and female students, and between vocational and higher education preparatory programs, as there were previous research indicating that females and higher education preparatory students, who normally also have a higher grade point average (The Swedish National Agency for Education 2017, p.3) should have an advantage (Altmeyer et al. 2011, p 14.)

There was, however, a limitation in Sub-study 1 as the teachers in the two conditions were not the same, which entailed a risk that there would be outcome differences only because one of the conditions may have had more skillful or charismatic teachers. To balance this risk, it was decided that a smaller, complementary study would be conducted in which the teachers would be the same. So in Sub-study 2, three of the teachers who had used the basic teaching guide (control condition) in Sub study 1, also agreed to use the Three Step Model (intervention condition) in some additional classes. The results from these classes were then compared to the teachers’ results in Sub-study 1, in order to find out if the same teachers using different “methods” would lead to any differences in the results. Answering this could help clarify whether or not the “methods” the teachers used were really the key variable.

Supervising of teachers

All teachers who had agreed to participate were informed about the purpose of the study (to find out what the most effective way of teaching ethics would be) and told that

Part of hypothesis	The intervention group in Sub-study 1 would score significantly higher in measure A	The intervention group in Sub-study 1 would score significantly higher in measure B	The intervention group in Sub-study 2 would score significantly higher in measure A	The intervention group in Sub-study 2 would score significantly higher in measure B
Confirmed	No	Yes	No	Yes

Fig. 3 Different parts of the hypothesis and whether it was confirmed or not

there were other teachers who were instructed to teach differently, but they were not told whether they were in the intervention- or control condition. Neither were they told anything about the content of the pre- and post-tests. The teachers in the control group, who just had to adhere to the basic teaching guide, had only one formal supervising occasion. The teachers in the intervention group, who had to follow the Three Step Model, were supervised twice (one time before and one in the middle of the teaching period) to make sure that they had not misunderstood anything in the very detailed instruction.

Structure of trial

In both sub-studies, the trials were conducted according the same structure. Just before the respective teacher introduced the ethics section, the classes were visited, informed about the study and their potential part in it. They were told that their participation was voluntarily and that they were guaranteed anonymity. After giving written consent, they had 20 min to fill out the pre-tests with paper and pencil. In case they had questions about the content, they were free to ask. After that, the teacher was left alone with the students for the rest of the teaching period, during which s/he was expected to follow the instructions as closely as possible. 10–12 weeks after the examination (a period when the teacher was instructed not to repeat any ethics), the students were visited again, unexpectedly to them, and given the post-tests to fill out. When this was done, a personal conversation was held with the teacher, during which s/he shared a number of details about the teaching, such as how many sessions of ethics the students had been given and how long these had been. S/he was also asked to share information whether there were any students who had been absent to such an extent during the teaching period (more than 50%) that they could not be counted as reliable participants in the study. If so, these students were classified as non-answers – just as those who were reported to have studied philosophy (and thereby ethics from a different source) during the same semester and, of course, those who had not filled out *both* the pre- and the post-tests.

Statistical analysis

When the students' answers in the pre- and post-tests had been quantified in cooperation with a supervisor, they were objects for statistical analysis. Initially we ran an independent samples t-test on the data, but since the data were not normally distributed (which in a t-test could give a false impression of statistical significance), we were advised by a statistician to also run a Mann-Whitney U-test (a non-parametric test in which the median score is calculated instead of the mean). This we did, just to discover that the results in terms of statistical significance were the same as when we ran the independent samples t-test. Thus, the statistician advised us to stick to the initial t-test analysis (as this would be easier to present and explain in context like this), and do some additional analysis of the variance, the results of which is presented below, together with some of the most important information gathered in the conversations with the teachers.

Results, sub-study 1

Participants In all, 456 students participated in Sub-study 1 (when 346 non-answers had been excluded). Of them, 244 were in the intervention condition and 212 in the control condition.

Intervention condition In the intervention condition, 41% of the students came from vocational programs, 59% came from higher education preparatory programs, 56% were males and 44% were females.

Control condition In the control condition, 42% came from vocational programs and 58% came from higher educational programs, 54% were males and 46% were females.

Previous ethics studies In the intervention condition, 10% of the students reported that they had studied ethics before, while in the control condition 17% reported that they had studied ethics before.

Teaching hours On average, the teachers in the intervention condition reported that they had used 510 min for ethics teaching while the teachers in the control condition reported that they had used 556 min (Tables 1, 2, 3 and 4).

The control group scored significantly higher in post-test part A in total ($P < .032$) but there were no significant differences between the groups in post-test part A items 1–2 and 3. This means that the part of the hypothesis saying that the students in the intervention condition would score significantly higher in measure A (self-assessment) was *not* confirmed.

Even though both groups developed from pre-test to post-test part B, this development was significantly stronger in the intervention group than in the control group, both in total ($P < .000$) and in item 2–3 ($P < .000$). This means that the part of the hypothesis saying that the students in the intervention condition would score significantly higher in measure B (development of demonstrable knowledge) was confirmed.

In the additional comparison between male and female students, the female students scored significantly higher in both measures A and B, and in the comparison between vocational and higher education preparatory programs, the higher education preparatory programs scored higher in both measures A and B. Moreover, an analysis of the variance (ANOVA) showed that among both male and female students, and among both students

Table 1 Differences between the two conditions (in which the teachers followed the Three Step Model and the basic teaching guide, respectively) in Sub-study 1 regarding their mean scores in the self-assessment, post-test part A in total, and the items that were assessed separately

Part of self- assessment	Intervention condition (n = 244)	Control condition (n = 212)	P value
Post-test part A in total (measure A)	Mean = 10.40 (Std. dev. = 3.17)	Mean = 11.08 (Std. dev. = 3.17)	$P < .032$
Post-test part A, item 1–2	Mean = 4.49 (Std. dev. = 1.48)	Mean = 4.37 (Std. dev. = 1.59)	$P < .412$
Post-test part A, item 3	Mean = 2.05 (Std. dev. = .92)	Mean = 2.19 (Std. dev. = .94)	$P < .122$

Table 2 Differences between the two conditions (in which the teachers followed the Three Step Model and the basic teaching guide, respectively) in Sub-study 1 regarding mean development from pre-test to post-test part B in total, and the item that was assessed separately

Development	Intervention condition (n = 244)	Control condition (n = 212)	P value
From pre-test to post-test part B in total (measure B)	Mean = 3.86 (Std. dev. = 2.48)	Mean = 1.60 (Std. dev. = 2.33)	$P < .000$
From pre-test to post-test part B, item 2–3	Mean = 2.20 (Std. dev. = 1.87)	Mean = .63 (Std. dev. = 1.82)	$P < .000$

from vocational and higher education preparatory programs, the control students scored significantly higher in measure A, and the intervention students scored significantly higher in measure B. This discrepancy was, in other words, a general tendency in Sub-study 1.

Results, sub-study 2

Participants In all, 123 students participated in the complementary Sub-study 2 (when 89 non-answers had been excluded). Of them, 68 were in the intervention condition and 55 in the control condition.

Intervention condition In the intervention condition, 69% of the students came from vocational programs and 31% from higher education preparatory programs, 47% were males and 53% were females.

Control condition In the control condition, 58% of the students were from vocational programs and 42% were from higher education preparatory programs, 42% were males and 58% were females.

Previous ethics studies In the intervention condition 19% of the students reported that they had studied ethics before, whereas in the control condition 13% reported that they had studied ethics before.

Table 3 Differences between the two conditions (in which the teachers followed the Three Step Model and the basic teaching guide, respectively) in Sub-study 2 regarding their mean scores in different parts of the self-assessment, post-test part A in total, and the items that were assessed separately

Part of self-assessment	Intervention condition (n = 68)	Control condition (n = 55)	P value
Post-test part A in total (measure A)	Mean = 10.47 (Std. dev. = 2.98)	Mean = 10.80 (Std. dev. = 3.62)	$P < .581$
Post-test part A, item 1–2	Mean = 4.45 (Std. dev. = 1.36)	Mean = 4.49 (Std. dev. = 1.66)	$P < .898$
Post-test part A, item 3	Mean = 1.13 (Std. dev. = .89)	Mean = 2.20 (Std. dev. = .98)	$P < .692$

Table 4 Differences between the two conditions (in which the teachers followed the Three Step Model and the basic teaching guide, respectively) in Sub-study 2 regarding mean development from pre-test to post-test part B in total, and the item that was assessed separately

Development	Intervention condition (n = 68)	Control condition (n = 55)	P value
From pre-test to post-test part B in total (measure B)	Mean = 3.32 (Std. dev. = 2.94)	Mean = 1.72 (Std. dev. = 2.26)	$P < .001$
From pre-test to post-test part B, item 2–3	Mean = 1.66 (Std. dev. = 2.06)	Mean = .94 (Std. dev. = 1.77)	$P < .044$

Teaching hours On average, the teachers in the intervention condition reported that they had used 521 min for ethics teaching, whereas the teachers in the control condition reported that they had used 540 min.

There were no significant differences between the intervention and the control groups in either the post-test part A in total or the items that were assessed separately. This means that the part of the hypothesis saying that students in the intervention condition would score significantly higher in measure A was *not* confirmed.

Even though both groups developed from pre- to post-test part B, the development was significantly stronger in the intervention group, both in total ($P < .001$) and in item 2–3 ($P < .044$). This means that the part of the hypothesis saying that the students in the intervention condition would score significantly higher in measure B was confirmed.

If we then look individually at the three teachers, who all taught in both the intervention and the control group, we see that in all three cases the intervention students scored higher in measure B and the control students scored higher in measure A, even though the differences were not in all cases significant. In other words then, this discrepancy was a general tendency in Sub-study 2 as well.

Discussion

What we have seen in this study is that the intervention group, whose teachers followed the Three Step Model, scored higher in measure B (development of demonstrable knowledge) while the control group, whose teachers just followed the basic teaching guide, scored higher in measure A (self-assessed ethical awareness). This was a tendency which was not affected by whether the students were males or females, or vocational or higher preparatory; it was not even affected much by who the teacher was. So what could be the reason for this discrepancy, indicating that the intervention group actually *learned* more, while the control group *experienced* a stronger development?

If we look at how the two groups were composed regarding male and female, and vocational and higher education preparatory students, we find no obvious clues to this, as the proportions (or disproportions) were roughly the same in both sub-studies. Neither do we find any obvious clues if we look at how many students in the two teaching conditions reported they had studied ethics before: in Sub-study 1 there was a larger percentage in the control than in the intervention condition who reported this, while the proportions in Sub-study 2 were reversed. The fact that the teachers in the control group reported having used more time for ethics teaching than the teachers in the intervention group in both sub-studies,

can perhaps give a clue to the advantage of the control condition in measure A, but it does not help us understand the advantage of the intervention condition in measure B. Of course, there are factors which have not been controlled for in this study (such as the participating students' general academic abilities), but due to what we know it is very likely that the results, at least to some degree, were due to the different "methods" for ethics teaching that were used. Or more precisely: the ways in which the teachers were *instructed* to teach, since one cannot be sure that all teachers followed the instructions to the letter.

So how can the discrepancy be explained? Could it be that there is no positive correlation between the two aspects of development which were measured by the post-test? Or is there even a negative one: the more you learn in terms of demonstrable knowledge (measure B), the less you tend to experience that you have developed (measure A)? This explanation, implying that learning makes you more realistic about yourself, is quite unlikely for the following reason. In the comparisons between males and females, and between vocational and higher education preparatory students, which was done in Sub-study 1 as a complement to the comparison between teaching conditions, there was no such discrepancy. Females and higher education preparatory students scored significantly higher than their counterparts in *both* measures (just as expected) which they would not have done if there had been this negative correlation. Instead, this indicates that there was rather something specific, having to do with the comparison between the two teaching conditions, that caused the discrepancy.

If we look more specifically at the advantage of the intervention group in measure B (which in both sub-studies was significant and confirmed by the item that was assessed separately), this could be explained by the fact that the Three Step Model was structured so that the students *had* to learn what an ethical problem was, how it could be solved properly etc.; otherwise they would not have been able to complete all the exercises or pass the examination test. If this explanation is true, it lends support to the research telling us that the best results of ethics teaching are achieved if one directly exercises the students' cognitive skills for ethical problem-solving/reasoning (Antes et al. 2009, p. 380 f.; Waples et al. 2009, p. 133 f.). Even more directly, it lends support to the research indicating that training in autonomous problem solving can make a long-term difference with regard to how an individual approaches a moral problem (Kavathatzopoulos 2004, p. 285; Kavathatzopoulos 2012, p. 389 f.).

Another explanation to the advantage of the intervention group in measure B could be the recurring element of *repetition* that was built into the method. It was repeated over and over again what an ethical problem was and how it could be solved properly. The repetition was carried out by the students themselves when they were doing the exercises, in which they had to retrieve the memories they had previously encoded. This may have generated the same kind of long-term effects we know that testing and re-testing may do (Dunlosky et al. 2013, p. 29 f.) especially in combination with the examination test at the end, in which all the steps were repeated again. On the basis of the reports we have seen about "regular" ethics teaching, it is unlikely that the teachers in the control group used repetition in any way similar to this.

Then, if we look more specifically at the advantage of the control group in measure A, which was only significant in Sub-study 1 and *not* confirmed by the items that were assessed separately, this could be explained by a possible *downside* with teaching that builds on repetition and many exercises, namely that it makes the students lose interest. Through a couple of comments from the teachers in the intervention group, it was revealed that some students felt the exercises were too many. Perhaps they did not remember the ethics section

as particularly stimulating or emotionally engaging, because it was more *demanding* than they were used to (especially in subjects like religious education).

Likewise, it could be that it is difficult for a teacher to, suddenly, start following a very detailed set of guidelines without losing some of the *vitality* that comes with a more “spontaneous” way of presenting the material; a predicament that the teachers in the intervention group may have found themselves in. This too could have affected the students’ impressions of the teaching. To some extent, the Three Step Model leaves room for the teachers to do things as they prefer, but first and foremost it requires them to follow the structure and make the students do all the exercise types. This may have posed an obstacle to creativity for some teachers, especially if they had not given themselves time to integrate the method with their own ways of thinking, so that they could follow it without becoming “slaves” under it or doing things mechanically. Thus, a lack of vitality could be the reason as well, *if* the lower score in the self-assessment was caused by the students in the intervention group finding the teaching less emotionally engaging.

If this is true, it is a setback as one of the goals with the Three Step Model was to make the students *more* engaged by the ethics teaching than they had been before (not least because it involved discussion about real and personal ethical problems). Since the hypothesis that the intervention group would score significantly higher than the control group was only confirmed by measure B, it has to be admitted that the Three Step Model did not turn out quite as effective as had been anticipated. However: the tendency that the intervention group scored higher in measure B was clearer and stronger than the tendency that the control group achieved better according to measure A. So in this sense the intervention students developed more.

But does that really imply an increase in ethical awareness? One could argue like the following: “There were indeed some things that the intervention students appeared to have learned better, such as giving examples of ethical problems and their potential solutions. However, it would be far-fetched to call this an increase in ethical awareness, as being able to *exemplify* an ethical problem does not automatically imply the ability to actually *recognize* one. If measure B would have assessed the ability to actually recognize an ethical problem, it would have been different. But since it did not, we cannot draw the conclusion that intervention students developed more in terms of ethical awareness.” Then, the response would be the following: “It is true that the ability to exemplify an ethical problem does not automatically imply the ability to recognize one, and this may well be a limitation with the assessment. But the *likelihood* that you will recognize an ethical problem should be greater if you are able to exemplify one than if you are not, because if you have a good picture of what an ethical problem is, you will know better what to look for. And this is a development we saw much more of in the intervention group.”

Therefore the tentative conclusion, nevertheless, will be that the Three Step Model is more effective than regular ethics teaching when it comes to increasing ethical awareness – at least if one defines ethical awareness and measures it in the two complementary ways that was done in this study. Future studies, however, may show how the method can be developed. One interesting idea would be to make the students keep a journal of the ethical problems they encounter in everyday life, as a preparation for the part of the ethics teaching which in all likelihood is most important: when they discuss their real and personal ethical problems. By means of this journal, it hopefully will be more recent and emotionally charged ethical problems that the autonomous problem-solving skill is applied upon. Thereby the ability is more likely to be embedded. This would be an intervention with a good possibility to increase the long-term effects of the Three Step Model.

Approval by ethics board

This research project was examined by the ethics board in Lund and approved on January 27, 2016.

Funding Open access funding provided by Lund University.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Altmyer, Don et al., 2011. Student ethical awareness and business program matriculation: evidence from the U.S., *Business Ethics Education & Accreditation* 3 (1), p. 41.
- Antes, Alison et al., 2009. A meta-analysis of ethics instruction effectiveness in the sciences. *Ethics and Behavior*, 19 (5), p. 389.
- DeHaan, Robert et al., 1997. Promoting Ethical Reasoning, Affect and Behavior among High School Students: an Evaluation of Three Teaching Strategies. *Journal of Moral Education*, 26 (1), p 5 f.
- Dunlosky, John et al., 2013. Improving students' learning with effective learning techniques: Promising directions from cognitive and Educational Psychology. *Psychological Science in the Public Interest*, p. 29 f.
- Gawthrop, John C., and Uhlemann, Max R., 1992. Effects of the Problem-Solving Approach in Ethics Training. *Professional Psychology, Research and Practice* 23 (1), p. 39 f.
- Kavathatzopoulos, Iordanis. 1993. Development of a Cognitive Skill in Solving Business Ethics Problems: The Effect of Instruction. *Journal of Business Ethics* 12 (5), p. 380.
- Kavathatzopoulos, Iordanis, 2004. Making ethical decisions in professional life, printed in Montgomery et al. (eds.): *How professionals make decisions*, p. 285. Mahwah: Lawrence Erlbaum Associates, Inc.
- Kavathatzopoulos, Iordanis, 2012. "Assessing and Acquiring Ethical Leadership Competence", printed in G.P. Prastacos et al. (eds.), *Leadership through the Classics*, p. 392. Springer-Verlag Berlin Heidelberg.
- Kavathatzopoulos, Iordanis, and Rigas, Giorgios. 2006. A Measurement Model for Ethical Competence in Business. *Journal of Business Ethics Education* 3, p. 51.
- Mortier, Freddy, 1995. Separate Values Education and Moral Development in Flemish Secondary Schools. *Journal of Moral Education* n4, p. 11 f.
- Niederjohn, M. Scott et al., 2009. Teaching ethics to high school students: Virtue meets economics. *Social Education* 72 (2), p 78.
- The Swedish National Agency for Education. 2017. *Grades and results in upper secondary school 2017*, p. 3. <http://www.skolverket.se>
- Waples, Ethan P. et al., 2009. A meta-analytic investigation of business ethics instruction. *Journal of Business Ethics* 87, p. 133 f.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.