

Comparative study of the military academy moral education model

Yi-Ming Yu¹ 

Published online: 29 August 2015
© Springer International Publishing AG 2015

Abstract The objective of the present study is to compare the effect of three moral education models: the bag-of-virtues, value-clarification, and virtue-ethics models. Students from military colleges in Taiwan were used as research samples, and a questionnaire survey with counterbalanced designs was conducted. The research results indicate that the value-clarification model exhibited lower educational effectiveness than the other two models did, but the virtue-ethics model increased student moral behavior, and the bag-of-virtues model facilitated the construction of student moral values. Additionally, the students rated the bag-of-virtues model as the most educationally effective among the three models.

Keywords Professional military ethics · CIPP (context, input, process, and product evaluation) model · Bag of virtues · Value clarifications · Virtue ethics

Moral education models and virtue

Moral education can be divided into three educational models: the bag-of-virtues, value-clarification, and virtue-ethics models (Lee 2006). The educational beliefs underlying the bag-of-virtues model are primarily derived from Socrates's perspective that "knowledge is virtue." Socrates argued that the innate conscience and goodness of people lead them to act morally. Therefore, the principal function of morals is to help people differentiate between virtue and vice. The bag-of-virtues teaching model mainly involves a one-way teaching method in which authorities determine moral content, and teachers explain the content and introduce moral theories to students. Students thus learn passively under this approach; they seldom engage in interactive teaching and

✉ Yi-Ming Yu
yuimingtw@gmail.com

¹ Department of Political Science, R.O.C. National Defense University, No. 70 Sec. 2, Zhongyang N. Rd., Taipei 11258 Taiwan, Republic of China

discussion and are not involved in relevant classroom activities. The bag-of-virtues model represents the traditional approach to classroom teaching (Doyle 1997).

Scholars who advocate for the application of the virtue-clarification model have asserted that societal tolerance and the coexistence of diverse societal values have changed the identification and learning modes of students. These scholars have insisted that moral education must originate from each person's values and that morality must be shaped through consideration of personal values. Similarly, these scholars have asserted that students must contemplate moral topics to improve their moral judgment and establish their moral philosophies. Therefore, under the virtue-clarification model, moral education primarily guides students in developing appropriate moral behavior by encouraging them to individually consider their moral values (Carr 1995).

According to the virtue ethics advocated by Macintyre (1981), virtue is a quality that people acquire through practice, and proficiency in this quality guides people toward morality. This morality derives from practiced values rather than from external purpose. Therefore, the third model of moral education focuses on participation and action, emphasizing the concept of education through experience. Under virtue ethics, moral education is based on practical experience, whereas the bag-of-virtues and value-clarification models limit moral education to cognitive learning (Raths et al. 1978). However, whether moral cognition and knowledge can be transformed into moral behavior remains unclear. Scholars of virtue ethics of thought (e.g., Lickona 2009) have contended that the purpose of education is to enhance moral character rather than moral knowledge, recognition, or judgment, and moral judgment is not equivalent to moral practice. According to virtue ethics, education is grounded in daily practice because successful moral education must inspire learners to act. Kohlberg (1981) asserted that moral learning begins from observing the behavior of others. When students observe different types of behavior and their consequences (e.g., someone being rewarded for helping and someone being punished for inappropriate conduct), they can recognize the moral values of which society approves and disapproves as well as learn to imitate positive behavior.

Previous studies on moral education models have primarily assessed education from the perspective of behavioral outcome without comprehensively evaluating the three models of moral education (e.g., Yu 2013), thus forming a research gap. This study compared the three moral education models from a macro perspective to determine the variation among these models according to environment and the amount of educational resources available.

Evaluation and models of moral education

Traditional educational assessments primarily focus on reforming curricula and investigating drawbacks in course design through evaluations without using evaluation results as references for future curriculum revisions. Although course evaluations involve comparing the merits of various courses, most of them focus on intellectual or technical courses, and few are suitable for affective courses (Stenhouse 1975). Affective courses primarily involve changes in the attitudes and values of learners, which are difficult to confirm. Particularly in open and democratic societies, student values and attitudes may originate from families, schools, classmates, and media. Many students live in open societies; therefore, their values and attitudes are frequently

influenced by various social channels. Thus, accurately measuring teaching effects and identifying the utility of affective courses is difficult in open societies. Courses on moral education involve distinguishing right from wrong, and evaluations of these moral values are inevitably biased because of the influence of factors such as social expectations. Course evaluation models can be divided into four types.

A. Goal-Attainment Model

The goal-attainment model is regarded as the original course evaluation model and is used to evaluate whether the effects of courses are consistent with their purposes; in other words, this model is a goal-oriented evaluation model (Tyler 2013). When employed to evaluate moral education, this model is predominantly used to determine whether learner behavioral changes are consistent with the goals of courses.

B. Countenance Model

Stake (1967) contended that student learning is complex, and learning outcomes do not completely originate from courses. Therefore, the goal-attainment model that assesses the effects of courses only according to student learning outcomes cannot be used to objectively evaluate courses. The countenance model incorporates all social channels that may influence the attitudes, behavior, and values of students.

C. Introspection Model

Under the introspection model, courses are not independent but are instead implemented according to mutual interaction and influence between courses and scenarios. In other words, courses are merely tools for guiding student analysis, and the purpose of teaching course material is not to provide correct answers for students. Therefore, Stenhouse (1975) stated that the mission of course evaluators is not to provide programs or seek correct programs as the goal-attainment model does but instead to manage the development and evaluation of courses for educational research.

D. Context, Input, Process, and the Product Evaluation Model

In contrast with the context, input, process, and product evaluation (CIPP) model, the goal-attainment and introspection models are primarily designed to evaluate courses with different content. However, the two models have difficulty evaluating courses that include retention- and value-based content. Furthermore, the countenance model cannot objectively compare interschool differences and propose solutions to teaching problems because it disregards the variation among the teaching scenarios at different schools. After summarizing the aforementioned shortcomings, Stufflebeam et al. (1971) constructed the evaluation framework of the CIPP model. They claimed that this evaluation framework compares the teaching effects of different types of courses and evaluates the entire learning process, increasing the objectiveness of evaluations.

Stenhouse (1975) stated that different teaching content should be evaluated using different models and that educational programs focusing on information-based content must be taught through instruction. The goal-attainment model is more suitable than other evaluation models for evaluating courses in which students must demonstrate learning outcomes through retention. If learning content involves value-based norms, then teachers generally teach through initiation. Thus, the introspection model is more suitable than other evaluation models for

evaluating this type of course. Comparing the aforementioned evaluation models indicates that evaluation methods should vary according to teaching method. In addition, the goal-attainment, countenance, and introspection evaluation models are based on various teaching methods. The goal-attainment model is suitable for evaluating courses comprising retention-based content, the introspection model is applied in courses comprising content on value guidance, and the CIPP model is used to compare course processes, from input to final results, in a generalized manner. In addition, the CIPP model enables the effects of different course types to be compared to facilitate analysis of the differences in the effects of input at the early, middle, and late stages of education. Moreover, although the countenance model provides a comprehensive framework for observation, it focuses on teaching scenarios. Schools' teaching scenarios crucially affect their courses; understanding these scenarios is beneficial to developing appropriate and feasible course reforms. However, interschool standards of course comparison and references for comparison are lacking in the countenance model, reducing the explanatory efficiency of evaluations under this model. This study investigated the educational effects of different moral education models by applying the CIPP model.

CIPP evaluation model and content

Lee Chi-Ming (1999) designed operational indicators, namely CIPP, to evaluate moral education courses in Taiwan according to the evaluation framework of the CIPP model. Context factors were regarded as environmental variables for indicators (i.e., fundamental characteristics and the resultant contextual scenarios of schools). The remaining input, process, and product indicators were considered the principal evaluation items for moral education courses. Input indicators refer to the external factors of moral education courses, including students, teachers, and resources (e.g., the proportion of dysfunctional families among students, teaching professionalism, and moral education course funding). Process indicators evaluate the design and implementation of moral education courses, including formal, informal, and hidden curricula. Regarding evaluation of the formal curricula of moral educational courses, process indicators include the goals, textbooks, supplementary material, pedagogy, weekly scheduled course content, and promptness of courses. The process indicators for informal curricula include the proportion of training and consulting activity hours among total class hours, the rate of participation in intramural and extracurricular service clubs, autonomous groups, or volunteer programs among students, and the rate of promoting and executing training and consulting activities. The process indicators for hidden curricula involve the school moral atmosphere and rate of participation in intramural and extracurricular service clubs and autonomous groups or volunteer programs among teachers. Product indicators evaluate the outcomes of moral education courses, emphasizing overall student performance, positive behavior, deviant behavior, and the satisfaction of various parties with the courses. Therefore, the evaluation items comprise overall student performance, positive and deviant behavior, and satisfaction. The product indicators for overall student performance consist of student moral performance, conduct performance, and dropout ratio. The product indicators for positive student behavior include the ratio of being rewarded for moral achievements and number of students selected as

moral role models inside and outside schools. The product indicators for student deviant behavior involve the ratios of being punished for immoral actions, skipping class, cheating on exams, campus violence, and immoral conduct such as smoking. In addition, the product indicators for satisfaction include the satisfaction of students, teachers, parents, society, and educational supervisors with moral educational courses.

Other evaluation models apply only to single types of teaching content and therefore fail to provide standards of comparison for courses that include more than one type of teaching content. Furthermore, although the countenance model is useful for clarifying the unique environments of schools, this study aimed to compare different types of moral education, and the advantages of this model for analyzing schools with identical educational scenarios are not applicable in this study. Therefore, this study adopted the CIPP model and referred to the indicators constructed by Lee (1999) for subsequent analysis.

Research methods

Three types of moral education courses were compared in this study (bag-of-virtues, virtue-ethics, and value-clarification courses). Undergraduate students at the Fu Hsing Kang College of the National Defense University were the research participants. This college has adopted diverse moral educational strategies; its military ethics (ME) courses have incorporated the bag-of-virtues teaching model, its military qualities (MQ) courses have employed the principle of value clarification, and its senior discipline (SD) courses have incorporated the educational logic of virtue ethics (Wortel and Bosch 2011; Yu 2014). Three types of courses were adopted as the basis for evaluating the three educational models in this study.

By using Lee's (1999) evaluation model for moral education as the analytical framework, this study compared the three educational models from the perspectives of input, process, and product. The context indicators in the CIPP model were eliminated because this study involved comparing the application of the three educational models at the same school, during the same period, and with the same students and teachers. In addition, Lee's (1999) evaluation model was developed primarily for elementary and junior high schools, whereas the three models of moral education were applied to a military college in this study. Therefore, relevant indicators were simplified in this study according to the military college conditions.

A. Data Collection

Various strategies were adopted to collect different types of data. The data in the input category were objective physical data such as teacher–student ratio, class hours, and educational cost, which were primarily acquired from educational administrative units. The process and product categories included psychological data mainly for evaluating the influence of teaching and course content on students. Data were collected using questionnaires.

B. Indicator Setting

1. Input Indicators

a. Average Class Hours: This indicator was applied to compare the weekly class hours of the three types of courses. Generally, additional class hours generate superior educational outcomes. Average class hours denote the weekly class hours for each student.

b. Teacher–Student Ratio: This indicator was the ratio of the number of military instructors and general teachers to that of students. A higher proportion of students to teachers indicates that students receive less attention during classes; by contrast, a lower proportion indicates that teachers pay more attention to individual student learning, thereby enhancing teaching effectiveness.

c. Unit Cost of Education: The cost of the three course types was the average cost of each course, the product obtained by multiplying the total hours of the three course types by the hourly pay of the military instructor or teacher of each course, and each student's educational cost per semester.

2. Process Indicators

a. Preparation: The indicator of preparation before class evaluated the degree to which each educational model emphasized courses. Students learn more effectively when courses are prepared. Samples were extracted, and teachers, military instructors, and fourth-year students were interviewed to investigate whether they were fully prepared before classes and the duration of their class preparation.

b. Teacher Conscientiousness: The degree to which moral education courses are used for educational purposes was analyzed in this study. Lee's (1999) evaluation standard for this is whether course hours are used for teaching or for other purposes. Moral education is categorized as a hidden curriculum; without fixed class hours, the problem of using course hours for other purposes does not apply. Therefore, this study evaluated teacher conscientiousness by determining whether teachers frequently attended to student learning and the amount of time they spent teaching students.

c. Student Concentration: This indicator referred to student level of concentration during class. Whether teachers are diligent and courses appeal to students can be determined according to the performance of students during class. Therefore, student concentration during class was used to evaluate teacher conscientiousness.

3. Product Indicators

The product category referred to the results of teaching or influence of moral education on students. Product indicators included student reception to teaching methods (reception), student satisfaction with course content (satisfaction), and the influence of courses on students (influence).

a. Reception: This indicator examined the variation among the three models according to self-report evaluations by students and evaluations of students by others.

b. Satisfaction: This indicator examined course content and professionalism to determine the satisfaction of students with the courses.

c. Influence: This indicator referred to the effects of courses on students. According to the Points of Implementation for Strengthening the Education of

Military Virtues at the Military Academies of the National Army (Political Warfare Bureau at the Ministry of National Defense 2009), the objective of moral education in military academies is to enhance military virtues, strengthen central beliefs, establish the foundation for education, promote favorable military ethos, and reinforce the fundamental beliefs of the national army (nation, responsibility, and honor) among officers and soldiers. Thus, the degree of course influence on students was measured on the basis of student behavior and values in this study. Although the satisfaction of students with courses can be used to evaluate their reception to courses, it does not entirely determine the results of course evaluation. Therefore, changes in student behavior and values that were observed by others were used as the basis for determining the influence of courses on students in this study.

C. Research Tools

Among the courses to which the three moral education models were applied in this study, the ME courses are compulsory for first-year students, whereas the other two courses are taught at different times in the academic year for students of every grade.

1. Questionnaire Design

The questionnaire of this study can be divided into self- and other-report evaluation questionnaires:

a. Self-Report Questionnaire

A 5-point Likert scale was adopted in this questionnaire and consisted of scores ranging from 5 (*completely agree*) to 1 (*completely disagree*). Each subscale consisted of five items. The process and product categories contained six subscales. The process category involved the three subscales of preparation, teacher conscientiousness, and student concentration for self-evaluation to investigate the process and conditions of each moral education model. The product category consisted of the three subscales of satisfaction, reception, and influence. After answering the six subscale items in sequence for one educational model, the respondents then responded to the items of another educational model. After completing the items for the three moral educational models, the respondents then answered the items regarding demographic data. A counterbalanced design was adopted in order to avoid possible biases caused by different response sequences for the three moral education models. Three questionnaires comprising distinct response sequences were designed for the moral education models to balance the possible biases. Questionnaire A comprised the sequence ME, SD, and MQ; Questionnaire B comprised the sequence MQ, ME, and SD; and Questionnaire C comprised the sequence SD, MQ, and ME.

The preparation indicator in the process category referred to teacher level of preparation before class. Preparing additional or new material before class and frequently updating supplementary course material indicated superior preparation. One of the five items for this indicator was “the teacher prepares supplementary material for every class.” Teacher conscientiousness was indicated by the level of effort and attention paid by teachers regarding classroom instruction. Teachers who repeatedly explained the topics that students had difficulty comprehending and adjusted teaching

methods according to the needs of students demonstrated high conscientiousness. One of the five items for this indicator was “I think the teacher frequently attends to student learning needs.” The indicator of student concentration represented student attention level during class. The subscale for this indicator consisted of five items, one of which was “I think my classmates concentrate in class.” This subscale may be substantially influenced by social desirability bias because few people admit their lack of concentration in class when asked this question. Therefore, this subscale was not designed in a self-report form; instead, student concentration in class was assessed according to the observations of other students under a moral education model.

The indicator of reception in the product category referred to student level of reception to teaching methods. A high level of reception indicated that students were highly interested in the course content. This indicator included five items, one of which was “I can accept the teacher’s teaching method.” The indicator of satisfaction referred to students’ level of satisfaction with the course content taught by teachers, including whether they agreed that teachers had professional knowledge and whether they comprehended the course content. One of the items was “I am satisfied with the teacher’s answers to the questions asked in class.” High scores on this subscale indicated increased student satisfaction with course content. The indicator of influence referred to whether courses achieved the expected educational goals (e.g., whether the courses benefitted the future military careers of students or their adaptation to life in military academies). One of the five items on this subscale was “the course is helpful for daily social interaction”.

b. Other-Report Questionnaire

The other-report questionnaire was employed to evaluate the degree to which moral education influenced students. Behavioral changes induced by education are generally difficult to observe, and the behavioral observations of others are more objective than self-observations. Therefore, an other-report questionnaire was adopted to examine these changes. The questionnaire focused on changes in student behavior and values. The items for observed behavior were the same as those for influence in the self-report questionnaire, and both sets of items measured the influence of various moral education models on military academy students by examining whether the courses assisted students with their life adaptation, daily social interaction, and future military careers. Five items for behavioral observation were rated on a 5-point Likert scale. To avoid subjective and biased answers to these items, the other-report form was adopted for a separate measurement to compare the two dimensions (behavior and values) and assess variation between the observations made from various perspectives. The other dimension for measuring influence in the other-report questionnaire was changes in student values. Measuring only student behavioral changes cannot objectively determine educational effectiveness because courses have different educational goals; some moral education models emphasize student behavioral changes, whereas others focus on student value cultivation. Consequently, in this study, value changes were measured using the self-report and other-report questionnaires. One of the items on the other-report questionnaire regarding value changes was “the courses are beneficial for strengthening younger student identification with soldiers.” Because the questionnaire included few items, respondent fatigue did not affect the reliability and validity of the

questionnaire. Therefore, a counterbalanced design with different response sequences was unnecessary.

The questionnaire was designed according to the content of the CIPP model. The Delphi method with three repetitions was employed to increase the reliability and validity of the scale. The first two implementations of the Delphi method involved groups of three to five officers (currently management cadres of the student military academy) who graduated from the military college, and the last implementation of this method was conducted by a group of three psychologists and education experts. Adjustments to the scale were made following the three implementations of the Delphi method. To confirm the reliability and validity of the questionnaire, a pretest was conducted before the formal test. The reliability and validity of the various subscales in the formal questionnaire was between Cronbach's α 0.79–0.90 and 0.70–0.82, respectively.

2. Questionnaire Distribution

The self-report questionnaire was answered by second-year students ($n=100$), and the other-report questionnaire was completed by first-year students ($n=100$). Second-year students were selected to answer the self-report questionnaire primarily. Therefore, second-year students were selected as our respondents. The questionnaire was distributed 10 min before the MQ course during the fifth and sixth academic sessions on September 19, 2014, and 95 valid samples were returned. Surveyors explained the procedure before the respondents answered the questionnaire anonymously and voluntarily. In addition, a gift was presented to each volunteer respondent to encourage candid questionnaire responses.

Third-year students were selected as the respondents of the other-report questionnaire because they were most familiar with the attitudes and behavior of the second-year students. The time, date, and procedure for answering this questionnaire were the same as those for answering the self-report questionnaire. Furthermore, the other-report questionnaire was distributed where the MQ course was taught.

Research result

A total of 74 male (77.9 %) and 21 female (22.1 %) second-year students answered the self-report questionnaire, of which 95 valid samples were returned, generating a response rate of 95 % (100 questionnaires were distributed). Among these respondents, who were aged 19 years on average, 22 (23.2 %) were from the Department of Political Science, 35 (36.8 %) were from the Department of Psychological and Social Work, 26 (27.4 %) were from the Department of Applied Arts, and 12 (12.6 %) were from the Department of Journalism. Furthermore, 77 male (79.4 %) and 20 female (20.6 %) third-year students answered the other-report questionnaire, of which 97 valid samples were returned, generating a response rate of 97 % (100 questionnaires were distributed). Among these respondents, who were aged 20 years on average, 28 (28.9 %) were from the Department of Political Science, 34 (35.1 %) were from the Department of Psychological and Social Work, 18 (18.6 %) were from the Department of Applied Arts, and 17 (17.5 %) were from the Department of Journalism (Table 1).

Table 1 Basic participant demographic data

	Self-evaluation questionnaire	Preceptor-evaluation questionnaire
Gender		
Male (%)	74(77.9 %)	77(79.4 %)
Female (%)	21(22.1 %)	20(20.6 %)
Department		
Political Science (%)	22(23.2 %)	28(28.9 %)
Psychological and Social Work (%)	35(36.8 %)	34(35.1 %)
Applied Arts (%)	26(27.4 %)	18(18.6 %)
Journalism (%)	12(21.6 %)	17(17.5 %)
Age (<i>M</i>)	19	20
Effective samples (recovery %)	95(95 %)	97(97 %)

The three courses were taught by different instructors. The ME courses were credit courses and thus taught by instructors with qualifications in ethics. The MQ courses were taught by officers because the instructors for these courses were required to possess substantial practical military experience. Thus, officers with at least 3 years of military service taught the MQ courses. The SD courses were taught by senior students, who taught junior students the school regulations regarding daily activities. Although the SD courses were not official credit courses, senior students remained responsible for guiding the daily activities of junior students because all students lived together in the school (Yu 2014). All three types of courses were compulsory.

A. Input Category

Among the three course types, the ME courses were the most costly (approximately NT\$62.37 per student per semester; Table 4), followed by MQ (approximately NT\$6.91 per student per semester) and SD courses (NT\$0). However, the SD courses were most costly regarding time and human resources at a ratio of approximately 2:1 (i.e., two seniors in charge of one junior) and approximately 6 course hours every week, followed by ME and MQ courses, both consisting of 2 course hours every week. The ME courses yielded a teacher–student ratio of approximately 1:18, and the MQ courses yielded a ratio of 1 military instructor to 110 students. Therefore, although the SD courses had the highest cost regarding human resources and time, they were incorporated into the unique management systems of the military academies without generating excessive cost regarding human resources and time. In other words, the SD courses were integrated for moral teaching, rather than practice, in the military academies. Moral education that is integrated into daily life represents the educational model of virtue ethics.

B. Process Category

According to the analysis of variance (ANOVA) in the process category (Table 2), preparation ($F_{(2282)}=39.24, p<.01$), teacher conscientiousness ($F_{(2282)}=10.05, p<.01$),

and student concentration ($F_{(2282)}=3.48$, $p<.01$) reached the level of significance. A posteriori comparison conducted using a least significant difference (LSD) test indicated that the ME courses exhibited the highest level of preparation before class (20.65), followed by the MQ (19.28) and SD courses (15.62). Regarding teacher conscientiousness, an a posteriori LSD test revealed that the ME courses (20.46) significantly differed from the MQ courses (18.74) and SD courses (17.74), indicating that teacher conscientiousness in the ME courses was higher than that in the MQ and SD courses. Regarding student concentration, an a posteriori LSD test showed that the ME courses (17.82) significantly differed from the MQ (16.26) and SD courses (16.48), indicating that students concentrated more in the ME courses than they did in the other two types of courses.

C. Product Category

According to the ANOVA in the product category (Table 2), reception ($F_{(2282)}=22.23$, $p<.01$), satisfaction ($F_{(2282)}=24.52$, $p<.01$), and influence ($F_{(2282)}=41.96$, $p<.05$) reached the level of significance. Regarding reception, an a posteriori LSD test indicated that the ME courses (20.59) achieved the highest level of reception among students regarding teaching methods, followed by the MQ (19.02) and SD courses (16.85). Regarding satisfaction, an a posteriori LSD test indicated that the ME courses (20.57) exhibited the highest satisfaction with teaching content among students, followed by the MQ (19.20) and SD courses (16.76). However, the level of influence revealed a score ranking (SD > MQ > ME) different from the aforementioned score ranking. An a posteriori LSD test showed that the ME courses (13.22) significantly differed from the MQ (18.13) and SD courses (18.15), indicating that the ME courses influenced students than the MQ and SD courses did.

The results of the other-report questionnaire answered by third-year students regarding the influence of the courses on the behavior and values of the second-year students are presented as follows. According to the ANOVA (Table 3), the influence of the courses on the behavior ($F_{(2288)}=9.04$, $p<.01$) and values ($F_{(2288)}=4.58$, $p<.01$) of the second-year students that were evaluated by the third-year students reached the level of significance. An a posteriori LSD test indicated that the influence of the three course types on student behavior and values reached the level of significance. The SD courses (17.7) most

Table 2 A posteriori comparison of the process and product categories

		ME(1)		SD(2)		MQ(3)		$F_{(2282)}$	Posteriori comparison
		M	SD	M	SD	M	SD		
Process	Preparation	20.65	3.98	19.28	4.00	15.62	4.16	39.24**	1>2>3
	Teacher conscientiousness	20.46	3.76	18.74	3.93	17.74	3.93	12.05**	1>2, 1>3
	Students' concentration	17.82	4.56	16.26	4.05	16.48	3.95	3.84**	1>2, 1>3
Product	Reception	20.59	3.69	19.02	3.08	16.85	4.13	22.23**	1>2>3
	Satisfaction	20.57	3.45	19.20	3.81	16.76	4.12	24.52**	1>2>3
	Influence	13.22	2.97	18.13	4.94	18.15	4.64	41.96*	2>1, 3>1

* $P<.05$, ** $P<.01$

influenced student behavior, followed by the ME (16.46) and MQ courses (15.36). An a posteriori comparison indicated that the ME courses most influenced student values, followed by the SD (16.47) and MQ courses (16.22). Furthermore, the ME courses significantly differed from the MQ and SD courses, indicating that the ME courses influenced students more than the MQ and SD courses did. The results regarding the influence of courses on student behavior were similar according to the self-report and other-report questionnaires; the SD courses were shown to influence student behavior more than the other two courses did according to the results of both questionnaires. Furthermore, the ME courses were shown to influence student values more than the other two courses did.

Overall (Table 4), although the ME courses had the highest cost and required the shortest time, teachers were most prepared and conscientious during these courses. Furthermore, students enthusiastically accepted the teaching methods adopted in the ME courses and were most satisfied with the content of these courses. However, the ME courses did not substantially benefit students in adapting to a military lifestyle or understanding the characteristics of the military, which are crucial for developing military values and identification. In addition, the MQ courses had the lowest teacher-student ratio, ranked second in required class hours and funds, and had the second largest overall input. In the process category, teacher preclass preparation, teacher in-class conscientiousness, student in-class concentration, and the final educational outcomes in the MQ courses ranked second. In the SD courses, educational input, process, and product differed from those of the other courses. In the input category, the SD courses yielded the highest teacher-student ratio and required the most class hours but had the least funding. Furthermore, the second-year students reported that the teachers of the SD courses were less prepared and conscientious than those of the other courses. Therefore, the students were least focused, least receptive to teaching methods, and least satisfied with course content during the SD courses. However, the SD courses received higher scores on the self-report and other-report questionnaires than the other courses did.

Discussion

The bag-of-virtues and virtue-ethics models achieved the anticipated educational objective

Socrates believed that knowledge is the foundation for moral judgements. Without knowledge, reasonable moral judgements cannot be formed. Socrates' perspective is

Table 3 A posteriori comparison of the influence of courses on second-year students evaluated by third-year students

	ME(1)		SD(2)		MQ(3)		$F_{(2288)}$	Posteriori comparison
	M	SD	M	SD	M	SD		
Behavior	16.46	3.60	15.36	4.06	17.70	3.60	9.04**	3>1>2
Value	17.87	3.74	16.22	4.40	16.47	4.09	4.58**	1>2, 1>3

** $P < .01$

Table 4 Overall comparison of the three course types

	Input evaluation			Process evaluation				Product evaluation		
	Teacher-student ratio	Class hours	Cost	Preparation	Teacher conscientiousness	Students' concentration	Reception	Satisfaction	Influence	
ME	1:18	2(h)	62.73(NT\$)	20.65	20.46	17.82	20.59	20.57	13.22	
MQ	2:1	60(h)	0(NT\$)	16.25	17.74	16.48	16.85	16.75	18.15	
SD	1:110	2(h)	6.91(NT\$)	19.28	18.74	16.62	19.02	19.20	18.13	

adopted in the bag-of-virtues model, where the importance of moral knowledge in moral education is emphasized. Moral knowledge primarily refers to moral theory in ethics. Thus, the bag-of-virtues model achieves the objective of moral education through the incubation of moral knowledge such as moral theory. The value-clarification model is centered on the belief that values are the foundation of all moral behavior and that values cannot be developed through education. Under this model, values are instead shaped through self-contemplation (Doyle 1997). Thus, moral education must guide individuals to contemplate various value-related problems, thereby shaping individual values. When personal values have been established, moral behavior follows (Carr 1995). In summary, according to the value-clarification model, to facilitate the development of personal values, the aim of moral education should be to guide self-contemplation rather than instill moral knowledge. The virtue-ethics model separates knowledge and behavior; in other words, knowing does not entail acting under this model. The virtue-ethics model contends that moral education is different from other types of education because moral education should center on behavior rather than knowledge and that verbal instruction should be replaced by practical learning.

The different approaches to moral learning among the aforementioned three models indicate that these models were designed for different educational purposes. The objective of the bag-of-virtues model is to instill moral knowledge. By contrast, the value-clarification model aims to guide self-contemplation, thereby shaping personal values. Finally, the virtue-ethics model promotes moral behavior through behavioral imitation. The results in Tables 2 and 3 indicate that the influence of the virtue-ethics model on moral behavior was significantly higher than that of the two other models in this study, indicating that the virtue-ethics model promotes moral behavior. As shown in Table 2, the participant satisfaction regarding the bag-of-virtues model and acceptance of the teaching methods it prescribes was significantly higher than that of the value-clarification and virtue-ethics models, confirming that the moral knowledge instilled through the bag-of-virtues model was superior to that instilled through the other two models. The influence of the value-clarification model on the development of values among military college students was lower than the influence that was anticipated (Table 3), a result that could be attributable to the fact that the officials in this study were not educational professionals and could not employ teaching techniques to guide student analysis. In addition, each class had more than 100 students; therefore, discussions were ineffective, and the influence of the teaching model was not as strong as the influence that was expected.

Furthermore, ethnic Chinese students are accustomed to being instilled teaching approaches, where students are not in the habit of participating in course discussions. The traditional authoritative image of ethnic Chinese teachers has also resulted in students perceiving teachers as the sole source of knowledge; thus, students lack the ability to question teachers' perspectives. Furthermore, Taiwan's current testing system emphasizes the ability to memorize and lacks tests on critical thinking and reasoning abilities. A long-term unidirectional learning approach has left the students incapable of adapting to the value-clarifications model.

The bag-of-virtues model received the most positive responses

Under the bag-of-virtues model, teachers actively instill knowledge in students, and students passively absorb knowledge. Teachers explain theories and examples, thereby

enabling students to understand the importance of various values. Courses designed according to the bag-of-virtues model do not emphasize the cultivation of logic. Nevertheless, to teach students effectively, teachers must attend to student learning needs during courses. In this study, the teachers of the courses that were designed according to the bag-of-virtues model were highly qualified, causing the students to respond positively to these courses. Table 2 indicates that the application of the bag-of-virtues model was superior to that of the other two models regarding teacher preparedness (course preparation), teacher attentiveness to student learning needs (teacher conscientiousness), and the appeal of the course content (student concentration). This suggests that the traditional teaching model, which involves instilling knowledge in students, exerts a greater influence on moral education than other models do.

The virtue-ethics model increases moral behavior

Tables 2 and 3 show that the effect of the virtue-ethics model on moral behavior is significantly greater than that of the two other models. The teaching methods used in the virtue-ethics model differ from those applied in the other two models, which primarily involve verbal instruction; the virtue-ethics model is centered on teaching by example, which can be used to demonstrate moral behavior for students. Kohlberg (1981) asserted that social norms, culture, and systems provide examples of moral behavior. Positive behavioral imitation is encouraged tangibly or intangibly (e.g., becoming a role model or high ranking cadre intern). Conversely, negative behavioral imitation results in social punishment or rejection. In military colleges, senior student behavior is often mimicked by junior students. This type of imitation is often mandatory at these colleges, enabling students to learn military values.

Table 2 shows that, among all teaching styles, the participants were least satisfied with the senior student disciplinary methods. In other words, although the participants were highly unsatisfied with the senior student teaching approach, they asserted that this approach exerted the greatest influence on their moral behavior. According to Kohlberg (1981), military students learn military culture through senior student formal instruction and by imitating senior student behavior. In this study, immediate punishment was administered to students who disobeyed rules; however, these students received tangible or intangible encouragement. Teaching by example, which primarily involves observational learning, exerts a substantially greater influence on moral behavior than the other two models, which are centered on verbal instructions, do. These results are consistent with those of previous research (Yu 2013).

Conclusion

Courses in military ethics aid students in establishing military values

In the input category, the ME courses obtained the most funding but had the fewest course hours. Teachers of the ME courses received professional training and demonstrated effective teaching skills as well as professional knowledge; therefore, students expressed high satisfaction and reception regarding the ME courses. However, students did not regard the ME courses as benefitting their adaptation to military academy life or

their future military careers. This result was likely explained by the fact that the ME courses primarily focused on ethical knowledge and theories. However, in the other-report evaluation, third-year students considered the ME courses beneficial for establishing military values. In summation, the ME courses received the most funding, exhibited the highest teacher conscientiousness, and obtained higher student reception regarding teaching methods and course content than that obtained by the other types of courses. However, the ME courses did not substantially influence the development of military behavior. This result was likely attributable to the fact that the teachers of the ME courses specialized in ethics and lacked substantial military knowledge, causing the courses to be overly theoretical and inapplicable to the student military environment. This argument is supported by the other-report evaluation, in which third-year students reported that the course content was beneficial to the establishment of military values. In other words, the ME courses focused on increasing student professional knowledge, resulting in their inability inspire student behavioral changes.

Senior discipline strengthens the adaptability of military academy students

Compared with other types of educational input, the economic cost of the SD courses was the lowest; however, the SD courses substantially influenced the life adaptation of military academy students. A detailed investigation into the overall influence of the courses indicated that the SD courses yielded a higher teacher–student ratio and required more class hours than the other two course types did. In other words, military academy students are influenced by the education provided by their senior schoolmates at all times. Moreover, under the seniority system in military academies, seniors have the authority to educate juniors; thus, juniors must constantly accept discipline from seniors outside of classes, which enables them to adapt to the intense pressure of military life in a short period. Because the seniors responsible for educating juniors in this study did not receive professional educational training, their teaching methods were mostly based on imitating the teaching methods of the seniors who had previously instructed them. In this type of master–apprentice educational model, teaching by example is more crucial than verbal instruction. Specifically, in military academies, immersive learning and behavior facilitate skill acquisition because practical experience is cultivated through juniors’ imitation of seniors’ behavior. Although gaining practical experience does not require substantial economic cost, it necessitates long-term educational cultivation on the basis of environment. According to the data gathered from the respondents, the SD courses had the lowest student reception among the three educational models. Additionally, the respondents generally considered the teaching quality of the SD courses to be the lowest. Typically, low educational quality and unfavorable reception among students cause unfavorable educational outcomes. However, the self-report and other-report evaluations indicated that discipline from seniors can assist juniors with life adaptation, suggesting that, even though the juniors did not identify with the seniors’ disciplinary methods, they agreed that the SD courses facilitated the development of life skills at military academies (Fig. 1).

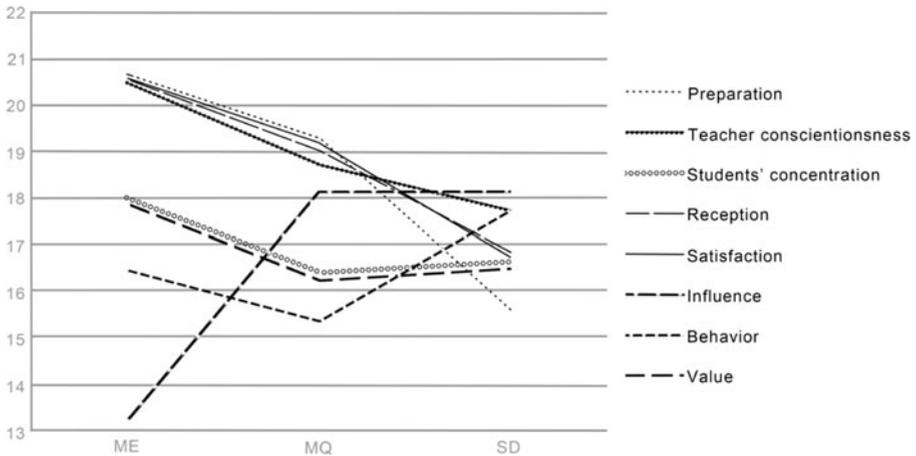


Fig. 1 Cadets on ME, MQ, SD course evaluation

Military qualities courses achieved moderate educational effects

The input of various resources for the MQ courses was similar to the input for the ME courses. However, the satisfaction of students with the teaching quality in the MQ courses was significantly lower than that in the ME courses; this was probably because the MQ courses were mostly taught by officers without professional MQ training. In addition to lacking relevant professional training, the officers that taught the MQ courses did not have classroom teaching experience; hence, their teaching skills, professional knowledge, and course influence were less satisfactory than those of the ME course instructors. Furthermore, the MQ courses included many students. More than 100 students in a grade congregated in a large classroom for these courses, and military instructors could present only lectures. Consequently, students could not express their opinions because of the lack of teacher–student interaction. Students could merely passively learn and were not encouraged to think independently; thus, the courses were unlikely to achieve the goal of shaping student values. Although the MQ course scores in the product category were lower than those of the ME courses, the abundant military experience of the MQ military instructors compensated for this deficiency. Therefore, although interactive teaching is difficult to apply in MQ courses because of the high number of students in these courses, the sharing of military experience by military instructors may still appeal to students. Overall, the MQ courses obtained moderate scores in the categories of input, process, and product.

The present study overcame many limitations of previous studies. A within-subjects design was employed to compare differences in the influence of three types of courses on student morality, thereby preventing interference caused by the participants' different backgrounds. Nevertheless, some study limitations should be noted. For example, the instructors in this study had different qualifications. The instructors of the three courses were teachers with qualified accreditations, military instructors with practical military experience, and senior students in the military college. The professionalism of these three types of instructors varied. The influence of the instructor professionalism on the research results was not properly controlled. In other words, the educational

effect of the three courses may have been influenced by the instructor professionalism rather than by the differences in the teaching models. In addition, the variables may have affected the teaching results to different degrees; thus, different effect weights should be used for these variables in the future to correctly calculate the effect of the courses. Furthermore, different amounts of funding and time were invested in the different models, possibly influencing the teaching effects. Although the present study had the previously mentioned limitations, the use of the CIPP model allowed for an effective evaluation of the influence of the teaching models on educational outcomes, and these results can be used as a reference for improving future course instruction. Nevertheless, future studies should aim to eliminate the aforementioned study limitations.

References

- Carr, W. 1995. *For education: Towards critical educational inquiry*. New York: McGraw-Hill International.
- Doyle, D.P. 1997. Education and character: A conservative view. *Phi Delta Kappan* 78(6): 440–443.
- Kohlberg, L. A. (Ed.). 1981. *The psychology of moral development: Essays on moral development (Vol. 2.)*. New York: Harper and Row.
- Lee, C. 1999. The construction for the evaluation indicator of moral education in elementary and junior high schools. *Bulletin of Civic and Moral Education* 8: 197–230.
- Lee, F. 2006. Transformation of the character, value, and ethics education in the United States. *Modern Education Forum* 14: 179–192.
- Lickona, T. 2009. *Educating for character: How our schools can teach respect and responsibility*. New York: Random House LLC.
- MacIntyre, A. 1981. *After virtue: A study in moral theory*. London: University of Notre Dame.
- Political Warfare Bureau at the Ministry of National Defense. 2009. Points of Implementation for Strengthening the Education of Military Virtues at the Military Academies of the National Army.
- Raths, L. E., M. Harmin, S. B. Simon. 1978. Values and teaching. OH Merrill Columbus.
- Stake, R. 1967. The countenance of educational evaluation. *Teachers College Record* 68(7): 523–540.
- Stenhouse, L. 1975. *An introduction to curriculum research and development*, vol. 46. London: Heinemann.
- Stufflebeam, D.L., W.J. Foley, W.J. Gephart, E.G. Guba, R.L. Hammond, H.O. Merriman, and M.M. Provus. 1971. *Educational evaluation & decision making*. Itasca: Peacock.
- Tyler, R. W. 2013. *Basic principles of curriculum and instruction*. University of Chicago press.
- Wortel, E., and J. Bosch. 2011. Strengthening moral competence: A ‘Train the trainer’ course on military ethics. *Journal of Military Ethics* 10(1): 17–35.
- Yu, Y.M. 2013. Between an example and a precept, which has greater importance? A comparison of the channels of socialization in military ethics. *Ethics & Behavior* 23(5): 341–359.
- Yu, Y.M. 2014. A multi-channel approach in military ethics courses: A case study of ethics education in Taiwan’s military academies. *Military Ethics* 13(4): 350–362.