A Metric for Military Ethics Instruction

Recent ethical failures, especially among senior military leaders, have focused attention on how the US military addresses ethics and professionalism in both training and education. There has been a renewed emphasis on ethics education and training in all of the military PME institutions over the past several months. In other words, ethics education in the military has become a “growth industry”.

However, the metrics used to assess ethics training and education are often flawed or inconsistent. Robinson, Lee, and Carrack (2008) concluded “We likewise have no solid data comparing the different techniques used in different systems to confirm which are the more effective. Inevitably, our conclusions are somewhat impressionistic and anecdotal” (p. 199). There has been little, if any, research done to confirm the impact of military ethics instruction. The purpose of this essay is to evaluate the metrics used to assess military ethics instruction and make recommendations on how they can be modified or improved.

There are several factors that make it difficult to assess ethics training and education. As Wolfendale (2008) noted “… the different methods used to teach ethics revealed an unresolved tension between the twin aims of simply getting military personnel to behave correctly, and trying to also make them more sophisticated moral thinkers and people of good moral character” (p. 165). This tension makes it difficult to assess outcomes. When participants self-report data on ethical learning outcomes there are potential validity problems. Participants may exaggerate responses or be embarrassed to report ethical failings which lead to social desirability bias.

Other professions including law, medicine, and business all have established codes of ethics and conduct ethical training on a regular basis. Over the past several years, instruments have been introduced, and tested, that actually measure moral reasoning and judgment. These include both Kohlberg’s Moral Judgment Interview (MJI) and Rest’s Defining Issues Test (DIT). Turner (2008) capitalized on this idea, and developed an instrument to measure moral reasoning in the Army profession for USMA cadets at West Point.

This paper will discuss the assessment process, some difficulties with the current process used to assess ethics instruction, findings from the 2015 GAO Report on ethics, assessment instruments from other disciplines, and make recommendations to help military PME institutions better determine the effectiveness of their ethics instruction.

As ethics instruction continues to develop and grow, the US military must develop valid metrics and put systems in place to measure the quality and performance of this instruction.

Dr. Tom Gibbons

Dr. Tom Gibbons, a retired Army Colonel, is currently the Associate Professor of Professional Military and Graduate Education Effectiveness working for the Associate Provost at the US Naval War College. Dr. Gibbons served on active duty in the US Army for 29 years.

His teaching interests include leadership and ethics. His research interests include college honor codes, influences on ethical behavior, and assessment in higher education.
He holds a Bachelor of Science degree from the U.S. Military Academy, a Master’s degree in Engineering Administration from The George Washington University, a Master of Arts in National Security and Strategic Studies from the U.S. Naval War College, and a Doctorate in Education Leadership from Johnson & Wales University. He received the 2007 Clifton J. Boyle Dissertation of the Year Award from Johnson & Wales University for his dissertation “Relationship of College Honor Codes and Core Values to Unethical Behavior in the Military Workplace.” In 2011, he was recognized by Johnson & Wales University as an honoree of the Alan Shawn Feinstein Graduate School for his support to their doctoral program.
A Metric for Military Ethics Instruction

Fort Leavenworth Ethics Symposium

2016

Submitted by

Thomas J. Gibbons

U.S. Naval War College

Author’s Note
The views expressed in this essay are those of the author and do not necessarily reflect the official policy or position of the Department of the Navy, the Department of Defense or the U.S. Government.
Introduction

Recent ethical failures, especially among senior military leaders have focused attention on how the US military addresses ethics and professionalism in both training and education. There has been a renewed emphasis on ethics education and training in all of the professional military education (PME) institutions over the past few years.

Throughout the Department of Defense there have been many initiatives implemented to highlight professional military ethics and the military as a profession. SECDEF Hagel created a new position on his staff, the Senior Advisor for Military Professionalism (SAMP), and appointed RADM Margaret “Peg” Klein to the position as his new ethics advisor.\(^1\) The latest version of the Officer Professional Military Education Policy (OPMEP), which provides guidelines for officer PME, was released in May 2015 and includes specific joint learning areas on ethics, leadership, and the profession of arms. As a result, many of the Professional Military Education (PME) institutions have re-vamped their curricula. The resources devoted to military ethics have also increased significantly, making ethics a growth industry in the PME institutions.

However, the metrics used by military PME institutions to assess the effectiveness of ethics training and education are often flawed or inconsistent. Robinson, Lee, and Carrack (2008) concluded, “We likewise have no solid data comparing the different techniques used in different systems to confirm which are the more effective. Inevitably, our conclusions are somewhat impressionistic and anecdotal.”\(^2\) They went on to explain, “Moreover, even if ethics training and development appear on the surface to be relatively successful, there is considerable room for improvement.”\(^3\) There has been little, if any, research done to confirm the effectiveness of military ethics instruction.
As ethics instruction continues to develop, the PME institutions must develop valid metrics and put systems in place to measure and assess the quality and performance of this instruction. This paper will discuss the assessment process, some difficulties with the current process used to assess ethics instruction, findings from the 2015 GAO Report on ethics, assessment instruments from other disciplines, and make recommendations to help military PME institutions better determine the effectiveness of their ethics instruction.

**Assessment Fundamentals**

Assessment is a continuous process designed to improve student learning. Walvoord (2010) identified three fundamental steps in any assessment model. They are:

1. Goals/outcomes – What is the end result of the assessment?
2. Information – What information or data will indicate how students are achieving the desired outcomes and goals?
3. Action – What action should be taken and what can be done with the information collected to improve student learning?

Each of these steps is critical to making the overall process work. Often, the outcomes are identified and information is collected to track the outcomes but there is no action taken. In other words, the assessment process is not completed because nothing is done to improve student learning. Additionally, the information collected and metrics used should provide an accurate picture of the goals or outcomes.

The OPMEP provides general guidance and policies for joint education from the CJCS and the Joint Staff. “Outcome of the joint learning process is a descriptive process to be undertaken by each JPME accredited college or school.” Each PME institution independently develops their own outcomes and internally manages the assessment process to track those outcomes.
**Ethics Training/Education Assessment**

There are several factors that make it difficult to assess ethics training and education. As Wolfendale (2008) expressed “… the different methods used to teach ethics revealed an unresolved tension between the twin aims of simply getting military personnel to behave correctly, and trying to also make them more sophisticated moral thinkers and people of good moral character.” Are the PME institutions trying to make their students more ethical or to improve their ethical reasoning skills? This tension makes it difficult to assess the learning outcomes.

Currently, most PME institutions administer student surveys to solicit feedback about the ethics courses, faculty, and instructional materials used. Some also receive comments about the ethics instruction from recent graduates on their alumni surveys. One faculty member acknowledged “What these surveys do not measure is the level of ethical knowledge of our students.” Additionally, most schools conduct after-action reviews or “hot washes” with faculty moderators to solicit their feedback. Much of this information is anecdotal and subjective. Administrators and course directors subsequently review the information and take necessary actions to complete the assessment process.

However, when students or graduates self-report data on ethical learning outcomes there are potential validity problems. Ethics for many of us, especially military professionals, can be personal and some may perceive themselves as better than they really are. This is human nature. Participants may exaggerate responses or be embarrassed to report ethical failings which may lead to social desirability bias. Gall, Gall, and Borg (2003) define social desirability as “the tendency to present oneself in a
favorable light." Data collected on surveys where participants self-report is subject to this bias and may not be accurate.

Wong and Gerras (2015) released a report “Lying to Ourselves: Dishonesty in the Army Profession. “This study found that Army officers, after repeated exposure to the overwhelming demands and the associated need to put their honor on the line to verify compliance, have become ethically numb.” The tendency to stretch the truth may bleed into the PME institutions. A quantitative metric may be more appropriate and easier to measure.

**GAO Report**

The Government Accounting Office (GAO) released a report GAO-15-711 on military ethics, in September 2015, entitled *Military Personnel Additional Steps Are Needed to Strengthen DoD’s Oversight of Ethics and Professionalism Issues*. The report highlighted the DoD’s failure to establish value-based ethics programs or “… programs that focus on upholding a set of ethical principles in order to achieve high standards of conduct.” In the past, much of the DoD ethics has been compliance-based where strict enforcement of the rules and regulations is important. The report also emphasized that “… DoD is unable to determine whether its ethics and professionalism initiatives are achieving their intended effect because it has not yet developed metrics to measure the department’s progress in addressing ethics and professionalism issues.” The GAO concluded that metrics are essential to measure the impact of military ethics instruction.

The Senior Advisor for Military Professionalism (SAMP) leads the DoD effort concerning ethics and military professionalism. “The purpose of the SAMP office is to coordinate and ensure the integration of the department’s ongoing efforts to improve
professionalism, and to make recommendations to senior DoD Leadership that complement and enhance such efforts." SAMP initially considered using misconduct as a metric to assess ethics and professionalism. However, the nuances related to each different reporting body would make this task almost prohibitive. Besides the difficulties associated with reporting, using misconduct as a metric for ethics training/education effectiveness is unreasonable. There are quantitative instruments and associated metrics currently available to assess moral judgment and reasoning in an academic environment.

**The Defining Issues Test and Moral Judgment Interview**

The assessment of ethics education/training has been a difficult problem for many professions. Turner (2010) suggested “… one of the main problems in studying the effects of professional ethics education is that measures have lagged behind the expansion of the moral development theory and conceptions of ethics education.” Over the past several years, however, instruments have been introduced, and tested, that actually measure moral reasoning and judgment. These include both Kohlberg’s Moral Judgment Interview (MJI) and Rest’s Defining Issues Test (DIT).

Lawrence Kohlberg is the patriarch of moral development theory and laid the groundwork for those that followed. Baily (2010) pointed out “Lawrence Kohlberg’s cognitive-developmental theory of ethical judgment has been the reigning paradigm in ethics-related research for the past half century.” Kohlberg’s Cognitive-Moral Development theory instituted six stages and three levels of moral development that individuals progress through. Kohlberg focused on why individuals made the decisions they did. The lowest level is Pre-conventional where self-interest is the primary motivator. The next level is the Conventional level where law and order and obeying the
rules are primary motivators. This is the level where most adults are. At the highest level (Post-conventional), general moral principles are primary motivators. At this level, individuals use their moral principles to make decisions. Few individuals consistently function at the Post-conventional level.\textsuperscript{15}

Kohlberg and his associates developed the Moral Judgment Interview (MJI) to measure an individual’s level of moral development. The MJI consists of the participants reading a series of three ethical dilemmas and then being interviewed to determine their capacity for moral reasoning. Participants are scored based upon the content of moral reasoning in their responses.\textsuperscript{16} However, there are some difficulties with the MJI. As Elm & Weber (1994) confirmed “Coding of the subject’s responses has been problematic for Kohlberg and his associates over the years.”\textsuperscript{17} Additionally, subjects need to actually “formulate a moral response” to the question which may be difficult for some based upon their language skills. Furthermore, the MJI is time intensive because it involves face-to-face oral interviews conducted by highly trained facilitators and researchers.

James Rest developed the Defining Issues Test (DIT) as an alternative to the MJI\textsuperscript{18}. “The DIT is the most extensively validated and most widely used measure of moral judgment.”\textsuperscript{19} The DIT is based upon the stages of Kohlberg’s moral development model. Subjects read six different ethical dilemmas and then answer a series of multiple-choice questions using paper and pencil. Subjects then rank their responses based upon how each would influence their decision-making. Each subject receives a P-score based upon the percentage of upper level reasoning used in their decision making.\textsuperscript{20} The DIT is also more direct because it allows subjects to select from a list of responses rather than
formulate a response to an interviewee’s open-ended question. Scoring for the DIT is quick and objective. Rest updated the DIT in 1998 and released the Defining Issues Test-2 (DIT-20) which is shorter with more current ethical dilemmas. Baily (2011) observed “The Defining Issues Test of ethical judgment is a widely accepted instrument with a long track record: about 500 researchers use the DIT every year and have done so at a steady pace for the last 15 years.” Bebeau (2002) noted that the DIT has been used successfully in academic environments by many professions including; business, medicine, law, nursing, and dentistry.

Based upon research conducted at the University of Minnesota with his colleagues, Rest (1983) also introduced the Four Component Model (FCM) of Morality. According to the FCM, the four components that must be developed to function morally include moral sensitivity, moral reasoning, moral motivation, and moral character. Rest postulated that all four components should be developed for an individual to reach moral maturity. They should have the sensitivity to recognize moral issues, judgment to make the ethical choices, motivation to take action based upon the choice, and character to stand by the decisions made.

Both the MJI and the DIT are useful instruments to provide valuable information on subjects’ moral reasoning. As Elm and Weber (1994) concluded “A critical advantage of both the MJI and the DIT is one that is not shared with numerous methodologies used in business ethics research, that is, well-established reliability and construct validity of the instruments.” However, as Turner (2008) emphasized “Current measures of moral judgment development (e.g., Defining Issues Test (DIT), Moral Judgment Interview (MJI)) measure general moral reasoning and do not target reasoning specific to a
Each profession is different so it may be advantageous to have instruments designed with that particular profession in mind.

**ALERT**

As Woodward (2007) cautioned “Bebeau and Thoma also concluded that each discipline (profession) should work toward the development of profession-specific assessments directed at more accurately measuring ethical development.”

The Dental Ethical Reasoning and Judgment Test (DERJT) introduced by Bebeau (2002) for the dental profession is a prime example. Turner (2008) capitalized on this idea, and developed an instrument to measure moral reasoning in the Army profession for USMA cadets at West Point. He wrote:

> Following the view that profession-specific measure of ethical reasoning and judgment provide more useful information, this study describes the development and testing of the Army Leader Ethical Reasoning Test (ALERT), which targets the Army profession and measures the moral reasoning and judgment of USMA cadets at the intermediate concepts level.

Turner developed five military-specific vignettes that addressed ethical decision-making. He solicited feedback about them from military and education professionals to verify their content validity. Next he tested the scenarios with both USMA and ROTC cadets at different stages in their cadet careers. Based on the pilot study, he modified the scenarios and had another group of military professionals assess each scenario and design a scoring system. Turner subsequently administered both ALERT and DIT-2 to USMA and ROTC cadets. There was a positive correlation between the results of ALERT and DIT-2. Overall, Turner’s findings support the validity and reliability of ALERT to measure moral reasoning and judgment. An instrument similar to ALERT, modified to accommodate students attending PME institutions, could be used to provide feedback and
metrics to the institutions on their ethics curricula. This instrument may also be valuable to provide information to students concerning their moral reasoning skills.

**Counterarguments**

Moral reasoning, however, is only one component of Rest’s Four Component Model. Rest and his colleagues determined that all four components (moral sensitivity, moral reasoning, moral motivation, and moral character) should be developed to reach moral maturity. Instruments like the MJI, DIT or DIT-2, and ALERT only provide feedback on an individual’s moral reasoning skills. As Turner (2008) concluded, “It is unwise to assume that post-conventional moral reasoning ensures moral behavior.”

Likewise, Stromberg, Wakin, and Callahan (1982) postulated “There is no empirical data establishing that the person who masters the objectives set forth for the teaching of ethics will, in fact, act in a morally responsible way in his profession.” In other words, exceptional moral reasoning skills do not necessarily guarantee moral behavior. The DIT provides a quantitative measure of moral reasoning skills. However, there are many other factors that actually influence moral behavior.

The DIT or any other instrument that provides a measure of moral reasoning skills is not the Holy Grail to influence moral behavior. It simply provides a quantitative measure of one’s post-conventional moral reasoning. As my colleague, Martin Cook (2013) related, “… situational factors play a large and perhaps decisive role in shaping behavior.” Arthur Caplan (1980) summarizes it best, “He (Rest) notes that the DIT is intended solely as a measure of moral judgment, not of moral worth, or of likely moral conduct.”
Another concern is that the curricula for the PME institutions normally only encompasses one academic year. The pre- and post-tests for subjects in the DIT and ALERT studies were done over a period of academic years. For example, ALERT subjects from both the freshman and senior class were evaluated. Other studies were also conducted over a period of years. Attempting to do a longitudinal study with a pre- and post-test during the course of only one academic year may not be sufficient time to elicit growth in moral reasoning skills. However, when conducted at the end of the academic year, the results may be useful to both graduates and curricula developers.

**Recommendations**

The instruments discussed have been validated and provide useful information about students’ moral reasoning skills. Moreover, they have been used by professional schools at a number of institutions over the years to provide valuable feedback and data about their ethics curricula.

The author makes the following recommendations for PME institutions:

1. **Use the DIT-2 on a limited basis to supplement the existing survey instruments used to collect information on the ethics curricula.** Administer the DIT-2 to a select group of students on a limited basis at the outset of the academic year and then again just prior to graduation to determine if there was a change in moral reasoning skills.

2. **Develop an ALERT-like instrument for use with intermediate and senior-level military students.** The instrument would consist of military-related ethical dilemmas and unique scoring specifically based on the profession of arms. Tailored for the military professional, this instrument would provide useful data especially for military leaders.
3. Offer the DIT-2 to interested students, faculty, and staff to provide an indication of their moral reasoning skills based upon Rest’s model. This information may be useful to military professionals interested in leader development. As Hartwell (2004) observed, “Moral reasoning is conceived as a cognitive process on a continuum along which individuals progress when they are goaded by the realization that their present moral reasoning is inadequate to provide them with rational and acceptable moral decisions.”

**Conclusion**

Ethics and professionalism are “hot button” topics in the military today. Considerable resources have been devoted to increasing the scope of ethics education and training in all of the PME institutions. Yet the instruments currently used to assess ethics education and training, are subject to social desirability bias and do not provide valid metrics.

Some professions use the DIT or DIT-2 to assess moral reasoning skills in their ethics curricula. “For educators wanting to develop an assessment programme to accompany a curriculum in professional ethics, the DIT is certainly a starting point.” The DIT or DIT-2 could provide a valid measure of students’ moral reasoning skills. Developing a military-specific instrument like ALERT would provide an even more accurate assessment of students’ moral reasoning skills. Used in conjunction with the current qualitative assessments, these instruments could provide curriculum developers and administrators with better information to improve student learning. As ethics instruction continues to develop, the US military PME institutions must develop valid
metrics and put systems into place to measure the quality and performance of this instruction.
Endnotes


3. Ibid., 199.


5. Chairman of the Joint Chiefs of Staff Instruction, *Officer Professional Military Education Policy* (29 May 2015), CJCSI 1800.01, E2.


11. Ibid., 31.

12. Ibid., 18.


18. Information concerning the DIT was summarized from James R. Rest, Development in Judging Moral Issues (Minneapolis: University of Minnesota Press, 1979), 75-105.


20. Information concerning the P-score was summarized from Rest, Development, 75-105, and Elm and Weber, “Measuring Moral Judgment,” 341-355.


27. The fact that moral reasoning skills are only one component of Rest’s Four Component Model is highlighted in Bebeau, “The Defining Issues Test and the Four Component Model,”; Steven Hartwell, “Moral Growth or Moral Angst? A Clinical Approach,” Clinical Law Review 11, (Fall 2004): 115-147; and Rest, “Morality.”


Bibliography


Chairman of the Joint Chiefs of Staff Instruction, CJCSI 1800.01, *Officer Professional Military Education Policy*, (May 29, 2015).


