EXPLOITING TECHNOLOGY TO HARNESS THE POWER OF PEERS

Karen Lochbaum
Knowledge Analysis Technologies
Boulder, CO

Joseph Psotka
Army Research Institute
Alexandria, VA

Lynn Streeter
Knowledge Analysis Technologies
Boulder, CO

ABSTRACT

Over the past year, several hundred Army officers have discussed tactical and military management scenarios using an electronic collaborative learning environment. The scenarios were either Think Like a Commander scenarios developed at Ft. Leavenworth or Tacit Knowledge of Military Leadership scenarios developed jointly by the Army Research Institute and Yale University. A standard threaded discussion platform was modified to allow participants to automatically find semantically similar notes and relevant reference material in an online library using Latent Semantic Analysis, a machine learning algorithm for text understanding.

Greater equality of participation has long been known to be an outcome of electronic discussion groups with anonymity enhancing this effect. These effects were replicated here as well in discussions involving officers of the same rank. However, we also found that electronic discussions produced higher quality initial and final solutions to complex military scenarios than comparable face-to-face discussions. The quality of the responses was graded by military experts and by an automated grading program developed by Knowledge Analysis Technologies. The automatic grading system exhibited reliability as high as the military experts.

Two forces appear to drive the superior performance and greater learning evidenced in the online environment: (1) peer pressure knowing that others will read and comment on one s solution produces more thoughtful and complete responses even under conditions of anonymity; (2) learning from peers reading, reacting and commenting on each other s notes produces a superior final solution over face-to-face discussion. Lower ranking officers (i.e., Lieutenants and Captains) demonstrated greater learning from the online discussion than did Lt. Colonels, whose initial responses were nearly optimal at the outset.

We discuss these results as well as planned enhancements to the online environment to create an even more effective e-learning environment.

ABOUT THE AUTHORS

Karen Lochbaum is a Senior Scientist at Knowledge Analysis Technologies and PI of the work described in this paper. She received her doctorate in Computer Science from Harvard University specializing in computational linguistics. For five years she was a Member of Technical Staff at U S WEST Advanced Technologies, where she did exploratory software R&D, including work on Latent Semantic Analysis applications that led to patents. She has published many technical articles in areas relevant to this research.

Joseph Psotka is a research psychologist at the Army Research Institute where he is a Senior Scientist in the Training Laboratory, in Basic Research, and in Organization and Personnel Research. He was also a Program Director for the Applications of Advanced Technologies program at the NSF, and has served on numerous boards and organizations. He has published widely and edited several books and journals. His research currently focuses on the application of cognitive science and automated technologies to learning and instruction, with a special emphasis on their role in developing leadership knowledge, skills, values, and attitudes.

Lynn Streeter is a Senior Scientist with Knowledge Analysis Technologies working officer leadership and training applications, field trials of Latent Semantic Analysis (LSA)-based essay scoring for Army officers, and other LSA measurement applications. She holds a Ph.D. from Columbia University, and has advanced Computer Science training and experience, as well as formal executive and business training. She was a Research Director at Bell Labs and Bellcore, directing human computer interaction work, AI and linguistics research, and software engineering research. She participated in the invention and early applications of LSA. Thereafter she was a General Manager of an R&D division of U S WEST Advanced Technologies.

EXPLOITING TECHNOLOGY TO HARNESS THE POWER OF PEERS

Karen Lochbaum
Knowledge Analysis Technologies
Boulder, CO

Joseph Psotka
Army Research Institute
Alexandria, VA

Lynn Streeter
Knowledge Analysis Technologies
Boulder, CO

INTRODUCTION

Increasingly frequent and diverse coalition contingency missions in sensitive geopolitical situations call for broader and renewable strategic leadership skills in officers. The knowledge behind such skills is largely tacit, based on voluminous experience in a wide variety of possible situations, and demands up-to-date understanding of the myriad political, cultural, economic, and operational factors involved in crisis situations. An emerging technology for improving such leadership skills can be provided by discussion and sharing of knowledge and experiences in a computer-supported cooperative learning environment. One such example is *Knowledge Post* developed in joint partnership with the Army Research Institute (ARI). In *Knowledge Post* officers are presented with vignettes describing situations that exemplify difficult leadership problems. The officers debate ways to deal with the situations, recount similar experiences, introduce and consider relevant facts and data, and voice and hear a broad range of opinions. An important aspect of *Knowledge Post*, as well as other threaded discussion groups, is that participants can interact without necessarily being in the same place or being engaged at the same time on portable, wireless, or desktop internet connections. Our data, collected with over 200 Army officers using the system in the past year, demonstrates that officers become more actively and creatively involved; they more openly contribute experiences and ideas in this form of interchange than in traditional face-to-face discussions and briefings. We show that their responses are also more thorough and thoughtful, and are rated by military experts to be of higher quality than those collected in face-to-face discussions.

Knowledge Post

Knowledge Post is a semantically enhanced collaborative discussion environment. Its main features are as follows:

1. Users read scenarios or discussion prompts stored as notes within the discussion forum.
2. Users write notes in response to those scenarios or prompts.
3. Users can read and respond to the notes of other users.
4. Users can search for notes that are semantically similar to others in the current discussion area or the special library.
5. Users can have their own responses evaluated by the *Intelligent Essay Assessor* (Foltz et al., 1999).

Features 1-3 are commonly available in other threaded discussion groups. Features 4 and 5 however are unique to *Knowledge Post* and are built using Latent Semantic Analysis (LSA). LSA is a machine-learning technology for simulating human meaning of words and text passages. It is both a model of human knowledge representation and a method for extracting and representing the meaning of words mathematically (for greater detail, see Landauer and Dumais, 1997; Landauer et al., 1998). LSA induces word and passage meanings by mathematically analyzing a large corpus of relevant text. In the resulting LSA space, every word and every passage is represented as a point in a high-dimensional "semantic space." This space defines the degree of estimated semantic similarity between any two words or passages. Simulations of many linguistic, psycholinguistic, and human learning phenomena, as well as several other educational and personnel applications, show that LSA reflects very accurately corresponding similarities of meaning as judged or used by humans (Landauer and Dumais, 1997; Foltz et al., 1999).

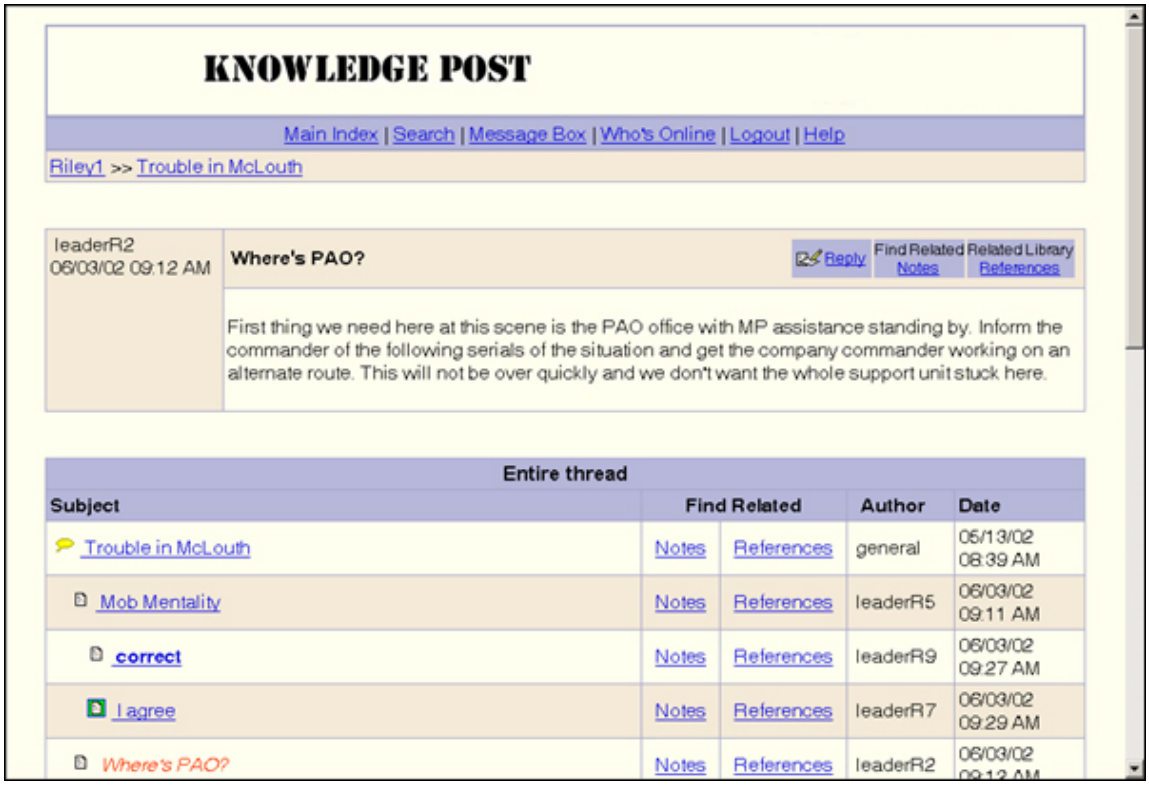

Figure 1. The note Where s PAO?

The *Knowledge Post* research and development effort has involved collecting open-ended responses to Think like a Commander (TLAC) scenarios developed by the Army Research Institute in Ft. Leavenworth and to scenarios of Tacit Knowledge of Military Leadership (TKML) developed by Yale University and the Army Research Institute. These are discussed in more detail in the next sections.

Several screenshots of *Knowledge Post* are shown in Figures 1, 2, and 3 with Figures 2 and 3 highlighting some of the semantic search additions that it includes. Figure 1 shows an example of a note entitled Where s PAO? written by leaderR2, a captain. This note is addressing the TLAC scenario Trouble in McLouth described in the next sections. The user can reply to the note as well as find semantically related notes and references. The hierarchical thread of which this note is a part is displayed below it.

Figure 2 shows the result of finding semantically related notes to the Where s PAO? posting. The five notes most similar in content were returned. Similarity scores are given to the right of the note titles. In general, similarities of 50 or above are very close in meaning. As shown in the figure, when a user rolls their mouse over a note, the most semantically salient sentence from that note is displayed. These two features of finding related notes and displaying summary sentences aid users in navigating large quantities of notes to find those that are of most interest to them.

Figure 3 shows the result of asking the system to assess the entirety of one s contributions. The assessment includes a) an overall rating (Needs Improvement, Good, Excellent) based on the assessment model derived from military experts scores of previous answers, and b) the two themes that have the lowest computed similarity to the user s combined contributions. The assessment feature is described in more detail below.

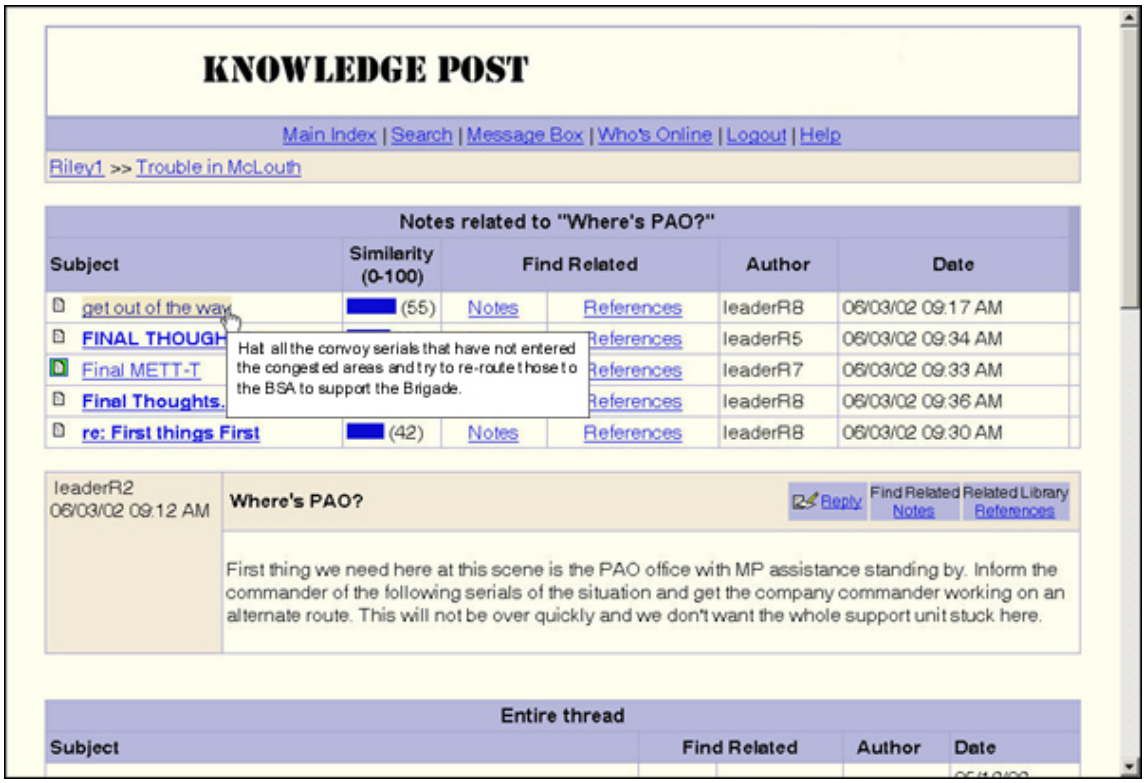

Figure 2. Notes semantically related to Where s PAO?

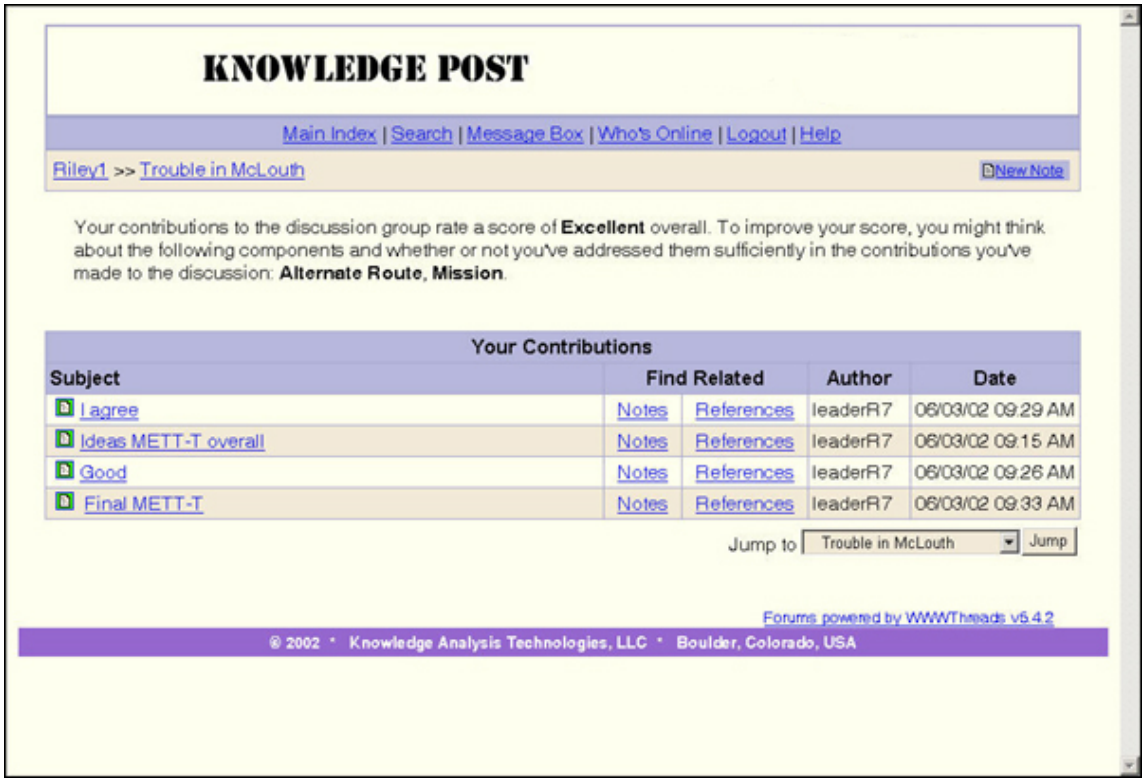

Figure 3. Assessment results for leaderR7.

Tacit Knowledge of Military Leadership (TKML)

Yale University, in collaboration with the Army Research Institute, has developed a technique for assessing tacit knowledge of military leadership (TKML) (see Hedlund et al., 1998; Sternberg et al., 1993). The method is based on a carefully developed set of representative scenarios of challenging interpersonal leadership situations that are commonly encountered by Army officers, along with sets of alternative actions that a leader might take. Interviews with experienced officers originally suggested the scenarios and alternatives. After initial construction of a large number of scenarios, a subset was selected and improved through a process of editing, expert judgments, trial uses, and empirical evaluations. In previous TKML administrations the leader or trainee first reads one of the scenarios, then rates on a nine-point scale the appropriateness of each of six to ten or more alternative actions that are described. This is not a multiple-choice test in the sense that only one answer is correct. The alternative actions are all acceptable or unacceptable to various degrees, the mix varying from scenario to scenario, and the right rating of each alternative defined by expert consensus. The test was first validated as a survey over three levels of leadership: platoon, company, and battalion command**.** As expected, more experienced leaders agreed on the most effective courses of action to a much greater extent than less experienced officers, reflecting tacit leadership knowledge acquired on the job.

In past research, the TKML scale proved to be an accurate measure of leadership effectiveness. The usual scoring of the instrument consists of determining the degree to which an individual agrees with the norms of the most experienced leaders using a 9-point Likert scale. These scores are significantly, although not highly, correlated with leadership ratings given by superiors, peers, and subordinates. Details of these earlier studies and findings can be found in Hedlund et al., 1998.

An example of one of the platoon-level scenarios is:

> **You are a new platoon leader who takes charge of your platoon when they return from a lengthy combat deployment. All members of the platoon are war veterans, but you did not serve in the conflict. In addition, you failed to graduate from Ranger School. You are concerned about building credibility with your soldiers. What should you do?**

One of the important perceived limitations of the original TKML technique is that the test-taker rates alternative actions that were generated by someone else. It is at least as important, if not more, for a leader to be able to think of the possible actions for a situation by him- or herself. We thus asked participants to use *Knowledge Post* to generate free-form written responses describing what they thought should be done in the situation and why. Such a measure would have been impractical until very recently because of the large amount of expert labor and time required to evaluate the answers experienced leaders or well-trained judges would have to read and score each essay. This would be especially prohibitive for use in an on-line tutorial environment where instant feedback is desirable. We believed, however, that appropriate analysis and automatic scoring of such constructed responses for both content and quality could be provided by our LSA-based intelligent assessment technology which only requires a sample of officer responses to be graded. In fact, we have begun to use the automated assessment method for TLAC scenarios. (See Figure 3 and the next section.)

We installed four platoon-level TKML scenarios on *Knowledge Post* to ensure that all levels of officers would have some relevant experience. Each group of same rank officers (Lieutenants, Captains, Majors, and Lt. Colonels) discussed two of the four scenarios. Answers were recorded in the synchronous 30-minute online *Knowledge Post* forum and LSA calculations were performed as well. Each response was later graded for the quality of response by two military experts. Figures 4a and 4b show the average quality of response within rank via paper and pencil and in Knowledge Post both after the initial response (labeled First Knowledge Post in the figure) and at the end of the discussion. The results show that officers using *Knowledge Post* contributed higher quality responses than those in face-to-face discussions. This effect was somewhat larger for Lieutenants than the higher-ranking officers.

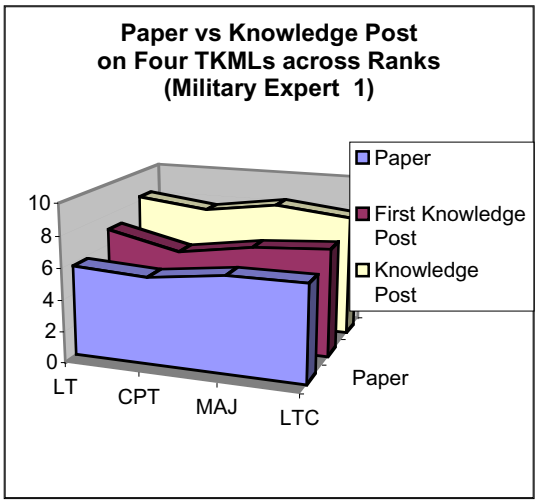

Figure 4a. Comparisons of Paper and Pencil to Knowledge Post by Officer Level for the TKML scenarios for Military Expert 1 s grades (1-9 with 9 being the highest grade).

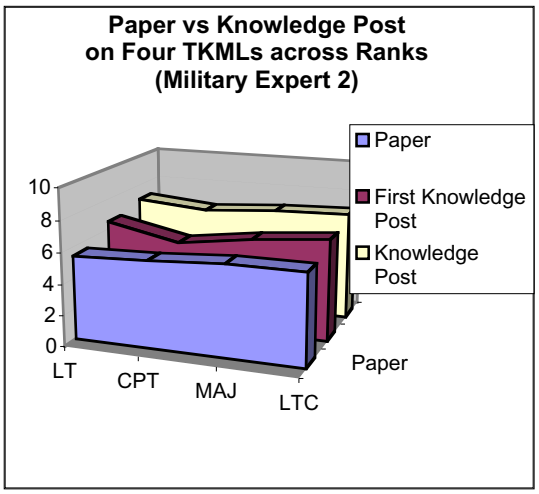

Figure 4b. Comparisons of Paper and Pencil to Knowledge Post by Officer Level for the TKML scenarios for Military Expert 2 s grades (1-9 with 9 being the highest grade).

Think Like a Commander (TLAC)

The Think Like a Commander scenarios were developed by the Army Research Institute at Ft. Leavenworth to teach tactical and strategic skills. The scenarios are presented as multi-media PowerPoint presentations. In the scenario we discuss here, Trouble in McLouth , a large group of refugees is climbing over and onto a serial of Bradleys and tankers en route to a refueling station. Another serial is approximately 10 minutes behind the first. The news media are present observing the conflict between the Army personnel and the refugees. A verbal description of the scenario is given below:

Trouble in McLouth

Narrator: Your BDE is moving north from AA Hawk to assigned sector. Advance party from FSB arrived at FAA Rose, 6 KM north-northeast of McLouth, to set up Task Force refueling and Class I distribution. The rest of the FSB, plus the TCF, is en route to FAA Rose.

Zoom in on McLouth.

Narrator: The refugee camp southwest of McLouth was designed for 2k -- but has swollen to over 8k refugees in past 2 days. Atchison camp abandoned. A 4-hour old report indicated over half of those refugees leaving camp, complaining Centralian Gov't intended to keep them there. Refugees are departing heading South of Kansas River -- no matter what. This report has not reached your brigade headquarters.
The MP Plat Ldr reports that the serial containing fuel tankers was stopped by a large number of civilians at the TCP. You were north of the problem site. Anticipating a problem, you decided to return there. On arrival, you find the lead FSB serial stopped by 1000 + refugees and receive OIC report.

Lt. N's Report: Lt. N's 6 POL tankers and 2Bradleys are stopped because of an accident and are swarmed by civilian refugees, blocking way and climbing over vehicles. Lt. N. can't proceed or turn around without risking injuries to civilians. Lt N. also can't find a civilian leader and tells you that journalists are interviewing civilians and want to interview you. He tells you two more serials are inbound, due to arrive in 10-15 minutes.

Narrator Continues: You are also informed that TF 4-81 Scouts report over 1000 civilians blocking Hwy 92 4 KM west of McLouth. Centralian militia have arrived on site and report to you that many more refugees are congregating in McLouth and in intersection 1k West of McLouth. This ends the situational updates.

After the presentation, officers are asked to formulate a response and discuss the scenario either in face-to-face discussion groups or using *Knowledge Post.* The discussion groups were again comprised of same rank officers ranging from Lieutenants to Lt. Colonels. About half of the 72 officers discussed the scenario face-to-face and wrote their responses on pencil and paper, while the other half discussed the scenarios using

Knowledge Post. The electronic discussion group entered an initial response and then a final response at the conclusion of the online synchronous discussion.

All responses were randomized, and the rank of the authors removed before they were blind graded by two military leadership experts as to the quality of the proposed solution. The results are shown in Figure 5a and 5b. There was a highly significant effect of media — those who used the electronic discussion group contributed higher quality initial responses (shown as First Knowledge Post in the figure below) than those in the pencil and paper group. In addition, the lower ranking officers learned more using the electronic discussion group than the face-to-face participants.

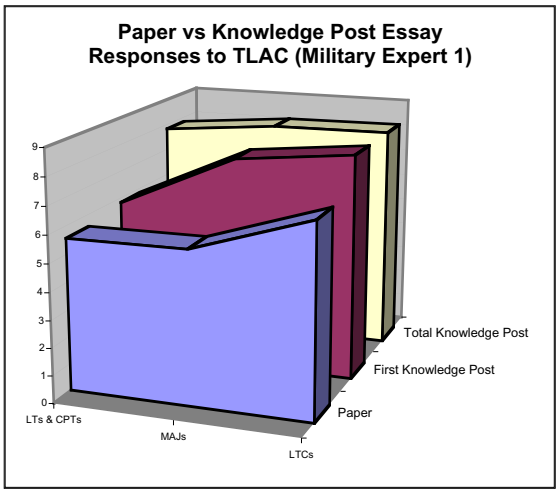

Figure 5a. Comparisons of Paper and Pencil to Knowledge Post by Officer Level for one TLAC scenario for Military Expert 1 on a 1 to 9 scale.

The results above show that *Knowledge Post* resulted in superior responses overall compared with the paper-based face-to-face discussion. Although senior officers had a slight superiority in the paper-based version, all officer groups improved through the LSA-mediated *Knowledge Post* discussion. Even the first response in the online format was superior to the final responses in the face to face discussion; this may have been facilitated by the fact that each officer knew his or her responses in the *Knowledge Post* environment would be seen by a peer, a fellow officer, and not just a civilian researcher, or an anonymous retired General.

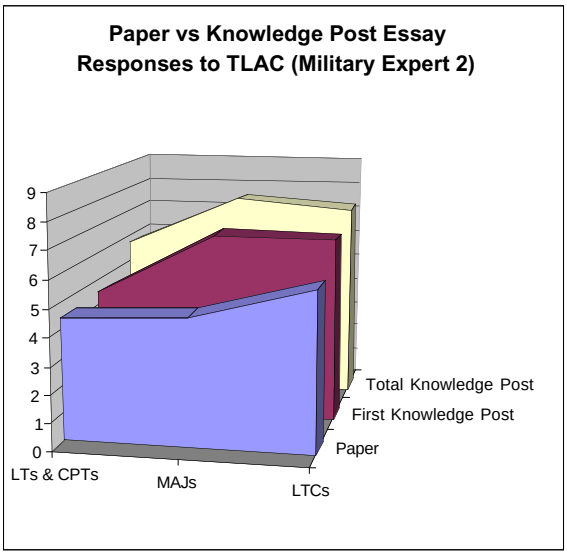

Figure 5b. Comparisons of Paper and Pencil to Knowledge Post by Officer Level for one TLAC scenario for Military Expert 2 on a 1 to 9 scale.

These results show the efficacy of the *Knowledge Post* system for leadership training. One of the goals of our research is not only to facilitate such learning, but also to provide automatic feedback to the officers on their progress as they are using the system. Using the *Intelligent Essay Assessor* (IEA) (Foltz et al., 1999), we built a system for automatically scoring the officers contributions based on the grades assigned by the military experts. The correlation between the grades assigned by the system and those assigned by the military experts was .87. The correlation between the two experts ratings was .67. The IEA thus agrees with the human experts better than the two humans agree with each other.

Figure 3 shows the initial integration of the IEA into *Knowledge Post.* The assessment returned by the system in this case is:

> **Your contributions to the discussion group rate a score of *Excellent* overall. To improve your score, you might think about the following components and whether or not you ve addressed them sufficiently in the contributions you ve made to the discussion: *Alternate Route*, *Media*.**

Officers thus receive an overall evaluation of their combined contributions in the form of a score of Needs Improvement, Good, or Excellent. This range of scores is a compression of the 9 point scale assigned by the military experts and learned by the system. The officers also receive some feedback on the themes they

might consider to improve their scores. These themes are represented by text descriptions generated by military experts. The officer s contributions are semantically compared against each of the theme descriptions and those that are least similar are returned.

CONCLUSION

We have shown that an enhanced threaded discussion group, in this case *Knowledge Post,* fosters higher quality discussion and more learning, particularly for junior officers, than similar face-to-face discussions. These findings have important implications for structuring distance learning courses. One of the major deficiencies in many distance learning courses is diminished human-to-human interaction. Recent data from Forrester research indicates that 70% of distance learning students abandon their online courses and perhaps one reason for this is the lack of discussion with peers and instructors. *Knowledge Post* can fill this void and even increase the quality of interaction among peers.

Currently, we do not know how much of the enhanced learning is due to the rich semantic features that have been incorporated into the *Knowledge Post* environment and how much of the improvement would be found in any vanilla-flavored threaded discussion environment. We suspect both have positive effects. One of the oldest results from the work on electronic chat groups pioneered by Hiltz and Turoff (1978) is the greater equality of participation in the electronic medium over face-to-face groups. Anonymity certainly contributes to the increased participation nobody knows you re a dog on the internet. In post discussion feedback we have collected on *Knowledge Post*, anonymity is always mentioned as a positive feature, producing richer, more honest communication.

Thus, several factors conspire to produce better discussion and learning in the electronic medium. One is a natural desire of humans to communicate with each other. Another, is the parallel nature of the discussion members of an electronic discussion can contribute simultaneously, thereby making more effective use of the time available. This is not possible in face-to-face discussions. The simultaneity of input coupled with greater equality of participation, result in a richer set of ideas generated by a greater number of people. In face-to-face discussions only a few people contribute the bulk of the remarks in small groups the most vocal two people do over 60% of the talking (Stephan & Mishler, 1952).

The instructional task is also an important factor. We have found that compelling scenarios capture the minds of the participants. The scenarios that we have used present either a moral, managerial or tactical/strategic military dilemma. These vignettes are relevant to each officer s daily life and their military leadership. They are intellectually challenging and are designed to elicit a broad range of responses, depending on the officer s experience and level. Thus, there are novice and expert points of view which can be codified, and usually one or more participant articulates the best solutions for each discussion group. It is noteworthy that companycommand.com, a website popular with Army officers uses the critical incident or scenario as the preferred discussion mode.

The technological advance that we have brought to electronic threaded discussions is turbo-powering them with Latent Semantic Analysis. This provides a natural way to materially improve collaborative learning and decision making. LSA tools can add automatic interactive monitoring, assessment, and feedback based on the content of free-form verbal contributions and do so in real-time. In addition, LSA can be automatically trained on existing bodies of text relevant to the discussion, thus giving LSA the intelligence to interpret and interrelate participants comments. Its evaluation of the content of verbal dialog takes only a few seconds and is robust over broad topical domains.

LSA-based technology can also provide continuous and cumulative assessment of the progress of individual participants and whole groups, and do so without the often dampening effect of close human supervision. LSA-based systems have already been implemented to provide tutorial feedback, to connect related comments to each other, and to provide instant access to large libraries of history, doctrine, and other relevant information. Instantaneous assessment (grading of discussion notes) and feedback was added to *Knowledge Post* quite recently. We are only now beginning to understand how to fully exploit its usefulness, but our initial attempts have energized these discussions and met with encouraging approval by the officers whose comments have been critiqued online by the impartial automated assessor. We are actively exploring the implications of this powerful new capability.

ACKNOWLEDGEMENTS

This research was funded by the Army Research Institute under contract DASW01-00-C-3051.

REFERENCES

Foltz, P. W., Laham, D., Landauer T. K. (1999). The Intelligent Essay Assessor: Applications to Educational Technology. *Interactive Multimedia Education Journal*. Available online at: (http://imej.wfu.edu/articles/1999/2/04/index.asp).

Hedlund, J., Horvath, J. A., Forsythe, G. B., Snook, S., Williams, W. M., Bullis, R. C., Dennis, M., & Sternberg, R. J. (1998). *Tacit knowledge for military leadership: Evidence of construct validity.* Alexandria, Virginia: U.S. Army Research Institute for the Behavioral and Social Sciences.

Hiltz, S. R. & Turoff, M. (1978). The Network Nation: Human Communication via Computer. Reading, Massachusetts: Addison Wesley Advanced Book Program.

Landauer, T. K. & Psotka, J. (in press). Education applications of Latent Semantic Analysis: Introduction. *Interactive Learning Environments.*

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211-240.

Landauer, T. K., Foltz, P. & Laham, D. (1998). An introduction to Latent Semantic Analysis, *Discourse Processes*. 25, 259-284.

Stephen, F. F., & Mishler, E. G. (1952). The distribution of participation in small groups: An exponential approximation. *American Sociological Review,* 17, 598-608.

Sternberg, R. J., & Wagner, R. K., & Okagaki, L. (1993). Practical intelligence: The nature and role of tacit knowledge in work and at school. In H. Reese & J. Puckett (Eds.), *Advances in lifespan development* (pp. 205-227). Hillsdale, NJ: Erlbaum.