Characterizing Trainees in the Cognitive Phase using the
Human Performance Knowledge Mapping Tool (HPKMT)
and Microgenetic Analysis

CSE Technical Report 699

Girlie C. Delacruz, Gregory K. W. K. Chung,
and William L. Bewley
UCLA/CRESST

September 2006

National Center for Research on Evaluation,
Standards, and Student Testing (CRESST)
Center for the Study of Evaluation (CSE)
Graduate School of Education & Information Studies
University of California, Los Angeles
GSE&IS Bldg., Box 951522
Los Angeles, CA 90095-1522
(310) 206-1532

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305B960002, as administered by the Office of Educational Research and Improvement, U.S. Department of Education, and the Office of Naval Research Award Number N00014-02-1-0179, as administered by the Office of Naval Research. The findings and opinions expressed in this report do not reflect the positions or policies of the National Institute on Student Achievement, Curriculum, and Assessment, the Office of Educational Research and Improvement, the U.S. Department of Education, or the Office of Naval Research.

CHARACTERIZING TRAINEES IN THE COGNITIVE PHASE USING THE HUMAN PERFORMANCE KNOWLEDGE MAPPING TOOL (HPKMT) AND MICROGENETIC ANALYSIS

Girlie C. Delacruz, Gregory K.W.K. Chung, and William L. Bewley
CRESST/UCLA

Abstract

Models of skill acquisition suggest that learners go through three phrases: (1) cognitive phase—when instruction is most effective, errors are frequent, and performance is inconsistent; (2) associative phase—when the learner begins to integrate the parts of the process or domain as a whole, and errors are gradually eliminated; and (3) autonomous phase—when the process becomes more automatic and less moderated by cognition, and there is less interference from outside distracters. In this paper, we will examine the use of the CRESST Human Performance Knowledge Mapping Tool (HPKMT) to characterize learners in the cognitive phase using Marine Corps 2nd Lieutenants going through entry-level marksmanship training. The capability to characterize learners may direct the level of instruction or practice they are given. HPKMT is designed to measure a learner's knowledge of a domain. Learners express their understanding of a domain by graphically depicting the relations among concepts. Further, the microgenetic analysis methodology provides a finer picture of the learning process by using repeated observations throughout the period of change giving detailed analysis of how and when change occurs. By measuring Marines' knowledge of marksmanship during classroom training, dry-fire practice, live-fire practice, and after qualification, we will have observations of their performance on the HPKMT at key stages of their learning. Our results suggest that the HPKMT can identify four types of learners in the cognitive phase: (1) growing, (2) declining, (3) stable, and (4) inconsistent. The HPKMT is also sensitive to instruction such that the mean difference between expert content scores on the knowledge maps are significantly different from those expert content scores on subsequent days.

Introduction

"Every Marine is a Rifleman" is a Marine credo, reflecting the importance of marksmanship training in the Corps. Annually, thousands of Marines are in marksmanship training either for the first time in boot camp or The Basic School, or at various rifle ranges such as Stone Bay or Quantico for requisite annual qualification.

Marksmanship scores play an important role in determining promotions, pay scales, and candidacies to become an officer.

Teaching marksmanship to the vast numbers of trainees can be a difficult feat for any instructor. Instruction often occurs in a gymnasium or on outdoor bleachers with poor acoustic quality. It is fairly difficult to target individual trainees in these conditions. Marines who get inadequate training may have difficulty qualifying, which is costly in terms of remediation and ammunition.

Currently, instructors use no assessment of knowledge and rely on human observation only during live-fire practice. A multiple-choice exam could be given at the end of classroom instruction, but it is often difficult and laborious to repeatedly test trainees with multiple-choice exams throughout the training process. Our research team is examining the use of the Human Performance Knowledge Mapping Tool (HPKMT) as a tool for microgenetic analysis to assist instructors with information on the characteristics of their trainees in a low-cost, feasible manner.

Our sample consisted entirely of entry-level lieutenants who did not have prior experience with the M-16. Using Fitts and Posner's (1967) model of skill acquisition, we assumed the trainees were in the cognitive phase of rifle marksmanship because they were new to the domain. We characterized four categories of trainees with characteristics typically found in the cognitive phase: *growing, declining, stable,* and *erratic.*

This paper will first provide an overview of the HPKMT and how it has been used and validated as an assessment tool. We will then discuss how researchers have used the microgenetic analysis methodology to investigate the process of learning and describe the various tools that have been employed. We then present Fitts and Posner's (1967) model of skill acquisition and its implications for instruction and training. Finally, we will illustrate how we applied all three in a marksmanship context.

Human Performance Knowledge Mapping Tool

The HPKMT is designed to measure an individual's understanding of any given domain. Individuals are asked to graphically depict what they know in a knowledge map. A network of nodes and links represents their knowledge. Nodes represent the concepts in a domain and links represent how the concepts are interrelated (see Figure 1). In the context of marksmanship, examples of concepts are *breath control, trigger control,* and *eye relief*. Examples of links are *requires, uses, affects,* and *part of*. These interrelationships (concept-link-concept) are called propositions.

Maps can be scored holistically by comparing students' maps against a referent map (typically constructed by an expert or an instructor). Points are given to matching propositions. Likewise, propositions can be scored individually for quality ratings independent of a referent map (Osmundson, Chung, Herl, & Klein, 1999).

Research has demonstrated that knowledge maps are a valid measure of assessing what someone knows in various contexts (Chung & Baker, 1997; Chung, Harmon, & Baker, 2001; Osmundson et al., 1999; Ruiz-Primo & Shavelson, 1996) and that outcome scores on knowledge maps relate to external measures such as essays (Herl, 1995; Herl et al., 1996; Klein, Chung, Osmundson, Herl, & O'Neil, 2002).

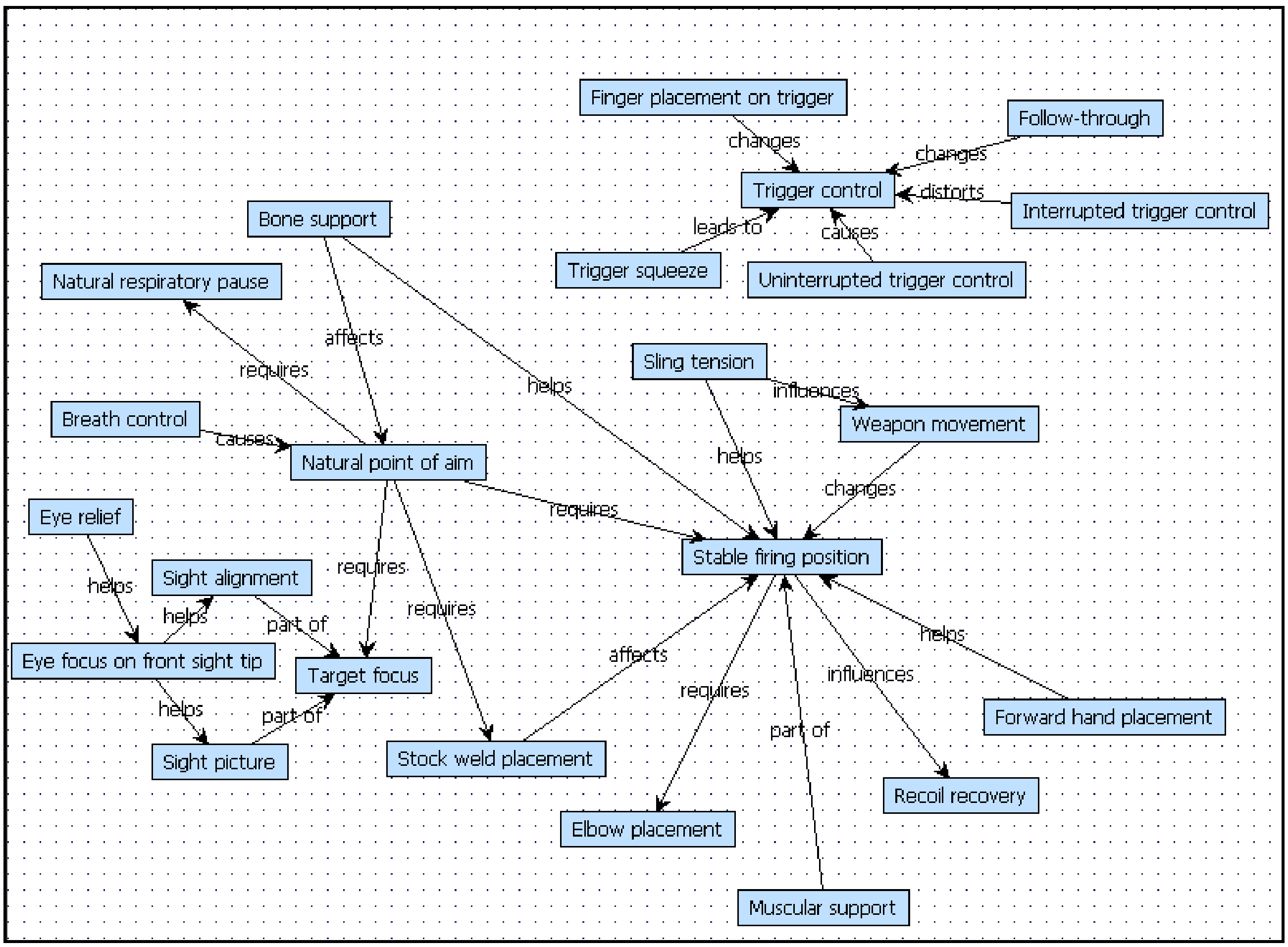

Figure 1. Example of a knowledge map.

Microgenetic Analysis Methodology

The measurement of learning is typically accomplished using pre- and post-task outcomes. A learner is given a pretest, goes through a treatment (or no treatment at all), and is given a posttest, with improvement being an indication of learning. However, by using the pre- and posttests as measures, one can only infer that learning has occurred. Posttests provide information about the products of learning, not the process (Lavelli, Pantoja, Hsu, Messinger, & Fogel, 2001).

Insight into the process of learning can prove to be valuable to instructors and researchers. Microgenetic analysis enables researchers to get a sense of the process of learning using repeated measures in a short span of time (Gelman, Romo, & Francis, 2002; Osmundson et al., 1999; Rittle-Johnson, Siegler & Alibali, 2001; Siegler & Crowley, 1991). This methodology is characterized by dense observations during the period of change. Researchers using this methodology employ various tools for observation. Siegler and Crowley (1991) have used human observation for strategy discovery, Gelman et al. (2002) have used journal entries to track science learning, and Osmundson et al. (1999) used an earlier version of the HPKMT to assess science learning in an elementary school classroom.

The repeated observations during the period of conceptual change enable an instructor or researcher to see how one's representation of knowledge develops. For example, the students' journal entries in Gelman et al.'s (2002) studies reflected increasing scientific understanding. Because this methodology enables one to explore the process of learning, it may assist in identifying where a learner is in the process. Understanding where a learner is could be useful for determining the level of instruction that is needed (Fitts & Posner, 1967).

Learning Phases

Research has suggested that the process of skill acquisition may be defined by phases (Anderson, 1982; Fitts & Posner, 1967). Fitts and Posner (1967) proposed that learners go through three phases during skill acquisition: cognitive, associative, and autonomous. The cognitive phase can be viewed as the stage of learning where the learner is new to the domain and is in the beginning stage of the learning process. This phase is often characterized by gross errors and inconsistencies because the learner is testing out various strategies. Fitts and Posner have also suggested that it is in this phase that a learner is most sensitive to instruction. The learner may benefit from explicit instruction on which perceptual cues and response characteristics are important to attend to. During this period, with practice and instruction, a learner's skills also tend to increase rapidly (Anderson, 1982; Fitts and Posner, 1967; Wrisberg, 2001).

The associative phase is characterized by an intermediate stage of the learning process. The learner has grown more efficient in his use of strategies, gross errors are gradually eliminated, skills are refined, and attention is paid to relevant cues. Upon entering the autonomous phase, the learner can be categorized as having successfully acquired the skill. In this phase, performance becomes efficient and automatic, requiring less cognitive processing.

Most learners remain in the cognitive phase because it takes many hours of quality instruction and practice to reach proficiency (Anderson, 1982; Wrisberg, 2001). Learners in domains such as marksmanship often have no previous experience. In the cognitive phase, many learners are faced with the task of mapping novel data and information onto existing knowledge structures. This may be the cause of numerous errors and inconsistencies. With frequent and quality practice and instruction, learners are better able to recognize the relevant information and errors start to decrease (Wrisberg, 2001).

Marksmanship Context

It is very likely that reaching proficiency is difficult without practice and instruction. The Marines going through marksmanship training, both at the entry level and the sustainment level, all have varying degrees of ability and experience with the rifle. Some entry-level lieutenants are prior enlisted while others have recently graduated from college. Those in the latter group may never have handled a rifle in the past, whereas the former might consist of expert marksmen. Individuals going through sustainment-level qualification vary in expertise as well. Some trainees are cooks or engineers who only handle the weapon once a year during annual qualification, while others are those in the infantry unit who shoot the M-16 on a daily basis.

Because of the various proficiencies of trainees, it is unfeasible for an instructor to adapt instruction to each level of expertise. This is a problem that the primary marksmanship instructors face with training. An instructor has little insight into the individual ability of the trainee until live-fire practice commences. Therefore, any information regarding how the trainee is grasping knowledge of rifle marksmanship would be useful for the instructors. For example, if we could characterize a trainee as declining or erratic, then more classroom instruction, additional remediation, or individual training might be prescribed. Likewise, if a trainee is characterized as stable or growing, less emphasis might be placed on classroom instruction; rather, more emphasis would be placed on hands-on experience with the rifle.

To conduct a microgenetic analysis of the learning process using the HPKMT, we administered the HPKMT at strategic times during the training process: before and after classroom instruction, three times during live-fire practice, and after qualification. We could then qualitatively analyze knowledge maps at various stages of the learning process. This methodology should give us a better sense of the stability or inconsistency of map scores over time in comparison to the expert maps, as well as the quality rating of propositions over time.

Likewise, because research has shown that the HPKMT can detect changes in learning (Chung et al., 2001; Osmundson et al., 1999; Schacter, Herl, Chung, Dennis, & O'Neil, 1999), we expect that maps administered repeatedly should reflect the changes in learning of marksmanship fundamentals as the entry-level lieutenants gain instruction and experience. We will investigate whether the growth is related to the amount of instruction and practice received, such that map scores and proposition ratings increase after classroom instruction and then perhaps either stabilize, or the rate of growth either decreases or increases in smaller increments.

Method

Participants

Fifty-three 2nd Lieutenants going through entry-level marksmanship training at a Marine base participated in the study.

Design

Data were collected from three groups. The first and second groups constructed their maps from a predetermined set of links. The third group constructed their maps from the same predetermined set of links but had the option of typing in their own links. The HPKMT was administered five times.

Coding Scheme

Based on a similar study by Osmundson et al. (1999) we chose to examine the differences in outcomes of the HPKMT by scoring the knowledge maps holistically and scoring their individual propositions. All knowledge maps received three scores based on three different ratings: (a) expert content score, (b) quality of individual propositions, and (c) weighted score.

Our first rating was the expert content score. We used an automated scoring algorithm to score the knowledge maps against the expert maps. Scores were generated by holistically comparing knowledge maps against eight expert maps. A knowledge map received a point for each matching proposition.

We then began our qualitative analysis by rating the quality of each individual proposition. We modified Osmundson et al.'s (1999) coding scheme to reflect the causal relationships in the marksmanship domain. The scoring rubric can be found in Table 1. Each proposition was judged to be nonsensical, superficial, limited, or principled, and received a corresponding score of 0 to 3. One rater who is knowledgeable in the domain applied the quality ratings to all of the individual propositions.

Table 1

Proposition Scoring Rubric

Score	Definition	Example
0	Illogical or nonsensical	rifle butt placement affects rifle butt placement
1	Superficial: Reflects superficial knowledge of marksmanship (categorical links are used, i.e., "part of")	aiming process is part of fundamentals of marksmanship
		forward hand placement is part of 7 common factors
2	Limited: Indicates a cause-effect relationship exists	forward elbow placements affects controlled muscular tension
3	Principled: Explains how the concepts are related	breath control requires natural respiratory pause
		stable firing position requires 3 elements of a good shooting position

Our last group of analyses examined the change of the sum of an individual's proposition scores. The score is weighted, such that propositions that received a higher quality rating (e.g., 2 or 3) received more weight in the formula. The weighted score used this formula:

$$\text{weighted score} = \sum 0(0)+1(1)+2(2)+3(3)$$

We used the change in the weighted scores of each individual's knowledge maps to characterize the four types of trainees using the coding scheme found in Table 2. Two coders achieved 90% reliability in their classifications.

Table 2

Trainee Characterization Coding Scheme

Classification	Explanation
Growing	Weighted scores for individuals in this category have increased at least twice over time and remained stable the other days.
Declining	Weighted scores for these individuals decreased at least twice over time and remained stable the other days.
Stable	Weighted scores remained the same for the majority of the latter days. If scores did not change from Days 2 to 4, but changed on Day 5, individual would not be considered stable.
Erratic	Weighted scores for these individuals exhibit no trends.

Results

Expert Content Ratings

The means and standard deviations of the knowledge map scores with respect to the expert maps are found in Table 3.

Table 3

Means and Standard Deviations for Expert Content Score

	Expert content score	
	M	SD
Day 1	14.24	10.63
Day 2	19.27	13.36
Day 3	21.24	14.19
Day 4	21.73	13.90
Day 5	20.88	13.87

The difference between the mean scores from Day 1 to Day 2-Day 6 were significant, $p < .05$. The differences between the other days were not significant.

Proposition Quality Ratings

Table 4 shows the means and standard deviations of each proposition rating (0-3) by occasion. The differences across time were not significant.

Table 4

Means and Standard Deviations for Proposition Scores

		Quality of proposition scores			
		0	1	2	3
Day 1	*M*	5.48	9.62	1.62	7.46
	SD	3.63	5.92	1.96	4.74
Day 2	*M*	6.12	9.33	1.76	7.35
	SD	4.59	6.26	2.16	4.60
Day 3	*M*	5.58	9.00	1.48	7.06
	SD	3.81	5.96	1.78	5.08
Day 4	*M*	5.26	9.38	1.68	6.04
	SD	4.27	6.56	2.02	4.98
Day 5	*M*	5.15	8.61	1.67	8.43
	SD	3.43	6.53	1.76	5.26

We also examined the number of individuals whose proposition scores increased, decreased, or remained stable over time. Figure 2 shows the number of individuals whose proposition scores increased daily. The number of individuals whose scores for both superficial and principled knowledge increased was greatest on Day 2, which was post-classroom-instruction.

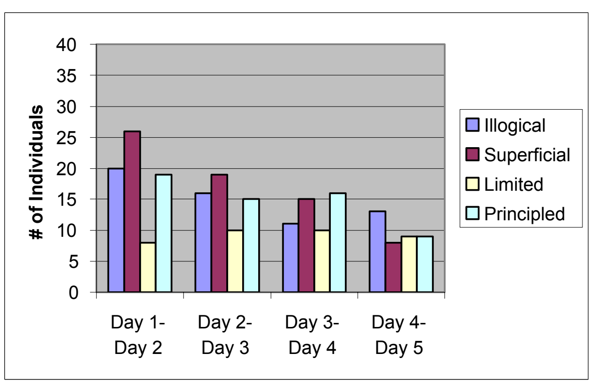

Figure 2. Number of individuals whose proposition scores increased daily.

Figure 3 shows the number of individuals whose proposition scores decreased from day to day. Overall, the number of individuals whose proposition scores decreased is highest on Day 2.

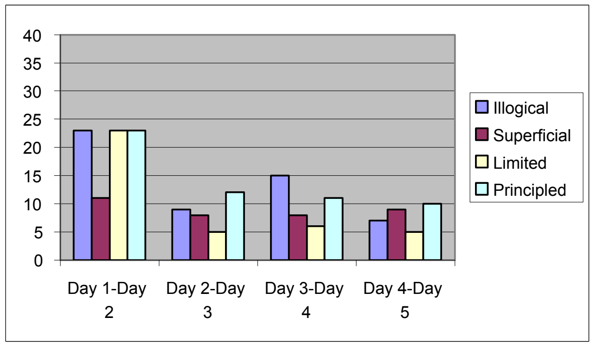

Figure 3. Number of individuals whose proposition scores decreased daily.

Figure 4 shows the number of individuals whose proposition scores did not change over time. The number of individuals whose scores did not change from Day 1 to Day 2 was lower than the number of individuals whose proposition scores did not change on other days.

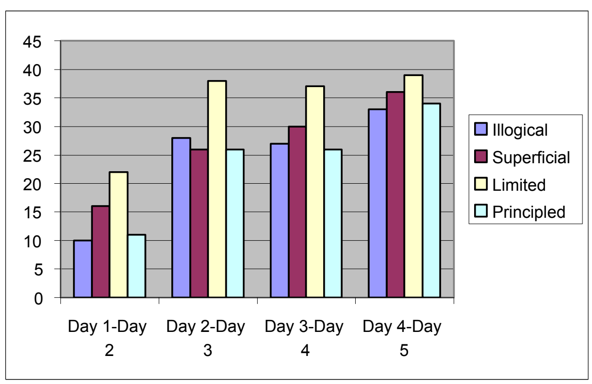

Figure 4. Number of Individuals Whose Proposition Scores Did Not Change Daily.

Weighted Scores and Learning Characteristics

Our second set of analyses examined the change of the weighted scores. Table 5 shows the means and standard deviations of the weighted scores. The average weighted score remained pretty constant across days with no significant change.

Table 5

Means and standard deviations for weighted scores

	Weighted Scores	
	M	SD
Day 1	32.43	13.63
Day 2	30.49	14.86
Day 3	32.17	15.53
Day 4	30.85	16.15
Day 5	31.09	17.55
Day 6	31.75	12.80

Results of the number of individuals whose weighted scores increased, decreased, or did not change are displayed in Figure 5. In general, the number of individuals whose weighted scores increased decreases on the last day existing maps were modified (Day 5). The number of individuals whose weighted scores decreased declines until the last day. The number of individuals whose weighted scores remained stable increases until the last day.

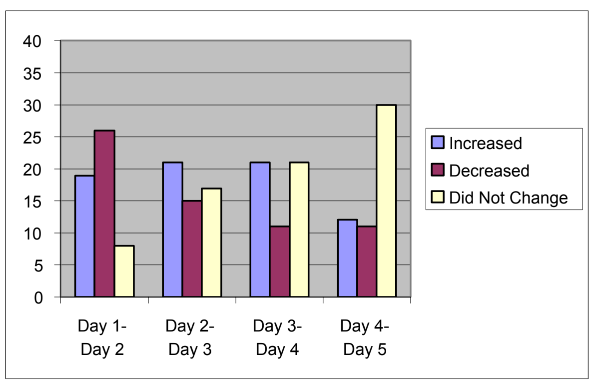

Figure 5. Number of Individuals Whose Weighted Scores Changed Daily.

To begin classification of learning characteristics, we classified the change trends in the weighted map scores using the coding scheme in Table 2. Thirty-eight individuals who completed a map each day were used. The breakdown of the classifications are found in Table 6.

Table 6

Breakdown of classified learning characteristics

	Growing	Declining	Stable	Erratic
Number of Individuals	7	3	14	14

We then looked to see how our classifications related to the trainees' shooting scores. The breakdown is found in Table 7.

Table 7

Comparison to shooting scores

Classification	Marksman (0-189)	Sharpshooter (190-209)	Expert (220-250)
Growing	2	1	4
Declining	2	0	1
Stable	7	2	5
Erratic	4	2	6

Conclusion

The purpose of this study was to examine the use of the HPKMT as a tool for microgenetic analysis to provide additional information for an instructor about their trainees. We used the microgenetic analysis approach so we could qualitatively examine the learner's knowledge of marksmanship fundamentals at key stages of the learning process. Our results suggest that there is potential for the tool to provide evidence of characterizing individuals in the cognitive phase.

We chose to examine the maps by giving them three ratings: expert content scores, quality of propositions, and weighted scores. Fitts and Posner (1967) suggested that learners in the cognitive phase should benefit from instruction. As expected, using the expert rating, we found that the trainees' mean scores were significantly higher on Day 2 which was post-classroom instruction. However, when we did our qualitative ratings (quality of propositions and weighted scores) map scores did not demonstrate the same spike; in fact some of the ratings decreased from Day 1 to Day 2. This result is very unusual in comparison to similar analyses that we have done with other marksmanship studies. We usually find that all three ratings increase post-classroom instruction.

We examined the knowledge maps to find an explanation for the difference in ratings. It was discovered that many trainees changed their links from Day 1 to Day 2 such that maps on Day 1 were filled with links such as "requires," "helps," and "uses". Propositions that used these links are given a score of 3. On Day 2, many links were changed to "part of" and "type of" which are given a score of 1. We think this might be due to the fact that our participants were all novices who had no prior rifle marksmanship training or instruction. Marksmanship instruction often presents the information in a hierarchical fashion as opposed to making explicit causal connections between the concepts, partly because the latter is directed to more novice learners, whereas with experience and more instruction, one begins to understand that the concepts are related. Therefore, trainees might have changed their "causal" propositions to that which reflect what and how they were taught. This would explain why quality ratings would have been lower post-classroom instruction.

To begin classifying groups, we wanted to see how many individuals could be classified as growing, decreasing, stable, or erratic. Out of the 38 individuals we looked at, 24 demonstrated a detectable trend in their learning characteristics, while 14 were erratic in their scores. As expected, 5 out of the 7 who were classified as growing shot well, qualifying as sharpshooters and experts. Two out of the 3 who were classified as declining qualified as a marksman. We were surprised to find that 7 out of the 14 of those who we classified as stable qualified as marksmen, 2 as sharpshooters, and 1 as expert. A potential explanation why map scores did not vary was that no changes were made to the maps. However, close examination of the maps of those classified as stable revealed that changes were made. It could be that this category would need to be broken down further, perhaps reflecting those who have not grasped the knowledge and are continuously not grasping the knowledge (who would potentially qualify as marksmen), versus those who grasped the knowledge fairly well from the beginning and continued to do so across time (who would potentially qualify as sharpshooter and expert). We found a similar pattern for those classified as erratic which leads us to believe that a similar breakdown should be done for this category.

This information can be useful for an instructor who wanted to distinguish those individuals who were clearly benefiting from the instruction and practice from those who were having difficulty grasping the concepts. By classifying the learner into one of the four groups an instructor would have insight into whether or not an individual was benefiting from the instruction and practice. Maps scored against experts reflected classroom instruction. However, we are very tentative in our classifications because our *N* was small, and we were only able to relate it to one qualification score. To strengthen our results, we recommend repeating the study with trainees going through sustainment-level rifle marksmanship annual qualification so that we can have a larger sample with varied levels of experience. We did share the results with the instructors but we cautioned any interpretation of success due to our reservations about the small sample size and limited external outcome measures.

Nevertheless, our study has demonstrated that using the HPKMT as a tool for microgenetic analyis has the potential to help classify a marksmanship trainee. We are planning on using the HPKMT to conduct similar microgenetic analyses on the sample of individuals who went through sustainment-level rifle marksmanship.

References

Anderson, J. R. (1992). Acquisition of cognitive skill. *Psychological Review, 89*, 369-406.

Chung, G. K. W. K., & Baker, E. L. (1997). *Year 1 technology studies: Implications for technology in assessment* (CSE Tech. Rep. No. 459). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Chung, G. K. W. K., Harmon, T. C., & Baker, E. L. (2001). The impact of a simulation-based learning design project on student learning. *IEEE Transactions on Education, 44, 390-398*.

Fitts, P. M., & Posner, M. I. (1967). *Human Performance*. Belmont, CA: Brooks/Cole Publishing Company.

Gelman, R., Romo, L., & Francis, W. S. (2002). Notebooks as windows on learning: The case of a science-into-ESL program. In N. Gronott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning.* Cambridge, U.K.: Cambridge University Press.

Gutheil, G., Vera, A., & Keil, F. C. (1998). Do houseflies think? Patterns of induction and biological beliefs in development. *Cognition, 66,* 33-49.

Hatano, G., & Inagaki, K. (1994). Young children's naive theory of biology. *Cognition, 50,* 171-188.

Herl, H. E. (1995). *Construct validation of an approach to modeling cognitive structures of experts' and novices' U.S. history knowledge.* Unpublished doctoral dissertation, University of California, Los Angeles.

Herl, H. E., O'Neil, H. F. Jr., Chung, G. K. W. K., Bianchi, C., Wang, S.-L., Mayer, R. E., Lee, C-Y., Choi, A., Suen T., & Tu, A. (1999). *Final report for validation of problem-solving measures* (CSE Tech. Rep. No. 501). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Klein, D. C. D., Chung, G. K. W. K., Osmundson, E., Herl, H. E., & O'Neil, H. F., Jr. (1999). *The validity of knowledge mapping as a measure of elementary students' scientific understanding* (CSE Tech. Rep. No. 557). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Lavelli, M., Pantoja, A. P. F., Hsu, H.-C., Messinger, D., & Fogel, A. (2001). Using microgenetic designs to study change processes. In D. G. Teti. (Ed.), *Handbook of Research Methods in Developmental Psychology*. Baltimore, MD: Blackwell Publishers.

Osmundson, E., Chung, G. K. W. K., Herl, H. E., & Klein, D. C. D. (1999). *Concept mapping in the classroom: A tool for examining the development of students' conceptual understandings* (CSE Tech. Rep. No. 507). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

McGuigan, D. (1953). *A comparison of the whole and part methods of marksmanship training*. Fort Knox, KY: Army Armor Center.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93,* 346-362.

Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching, 33,* 569-600.

Schacter, J., Herl, H. E., Chung, G. K. W. K., Dennis, R. A., & O'Neil, H. F., Jr. (1999). Computer-based performance assessments: A solution to the narrow measurement and reporting of problem-solving. *Computers in Human Behavior, 15*, 403-418.

Siegler, R. S., & Crowley, K. (1991). The microgenetic method. A direct means for studying cognitive development. *American Psychologies, 46,* 606-620.

Wrisberg, C. A. (2001). Level of performance skill: From beginners to experts. In R.N. Singer, H. A. Hausenblas, & C. M. Janelle (Eds)., *Handbook of sport psychology*. New York: John Wiley & Sons, Inc.