

## 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  
National Center for Research on Evaluation, Standards, and Student Testing (CRESST)  
University of California, Los Angeles  
Los Angeles, CA  
[gdelacruz@cse.ucla.edu](mailto:gdelacruz@cse.ucla.edu), [greg@ucla.edu](mailto:greg@ucla.edu), [bewley@cse.ucla.edu](mailto:bewley@cse.ucla.edu)

### 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 gradually become 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) as a tool for microgenetic analysis to characterize learners in the cognitive phase using Marine Corps 2<sup>nd</sup> Lieutenants going through entry-level marksmanship training. The capability to characterize learners may direct the level of instruction or practice she is 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 a Marine's 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 has the potential to identify four types of learners in the cognitive phase: *growing, declining, stable, and erratic*.

### ABOUT THE AUTHORS

**Girlie C. Delacruz** is a Research Associate at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Her current work at CRESST involves measuring human learning, developing Internet-based assessment tools for diagnostic and embedded assessment purposes, evaluating technology-based learning environments, and developing problem-solving assessments for computer-based assessments. Ms. Delacruz earned her B.A. in Psychology with a minor in Applied Developmental Psychology from the University of California at Los Angeles (UCLA), and is pursuing her Ph.D. in Psychological Studies in Education at UCLA.

**Gregory K.W.K. Chung** is a Senior Research Associate at CRESST. His current work at CRESST involves developing problem-solving assessments for computer-based assessments, evaluating technology-based learning environments, conducting research on the measurement of team processes using networked simulation, and developing Internet-based assessment tools for diagnostic and embedded assessment purposes. Dr. Chung earned a Ph.D. in Educational Psychology from the University of California, at Los Angeles, an M.S. degree in Educational Technology from Pepperdine University at Los Angeles, and a B.S. degree in Electrical Engineering from the University of Hawaii at Manoa.

**William L. Bewley** is an Assistant Director at CRESST. His experience combines research on human learning and performance, instructional systems design, education and training product development, software development, and management. His current work is concerned with research on applications of advanced technology to human performance assessment and evaluation models for e-learning. Dr. Bewley earned a Ph.D. and M.S. in Psychology and Computer Science from the University of Wisconsin and a B.A. in Psychology from Stanford University.

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

Girly C. Delacruz, Gregory K.W.K. Chung, and William L. Bewley  
National Center for Research on Evaluation, Standards, and Student Testing (CRESST)  
University of California, Los Angeles  
Los Angeles, CA  
[gdelacruz@cse.ucla.edu](mailto:gdelacruz@cse.ucla.edu), [greg@ucla.edu](mailto:greg@ucla.edu), [bewley@cse.ucla.edu](mailto:bewley@cse.ucla.edu)

## 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 by providing information of their trainees in a low-cost, feasible manner.

Using Fitts and Posner’s (1967) model of skill acquisition, we assumed our trainees were in the cognitive phase of rifle marksmanship because our sample consisted entirely of entry-level lieutenants who have not had prior experience with the M-16. 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 elements 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.

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 et al., 2002). 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 et al., 1999).

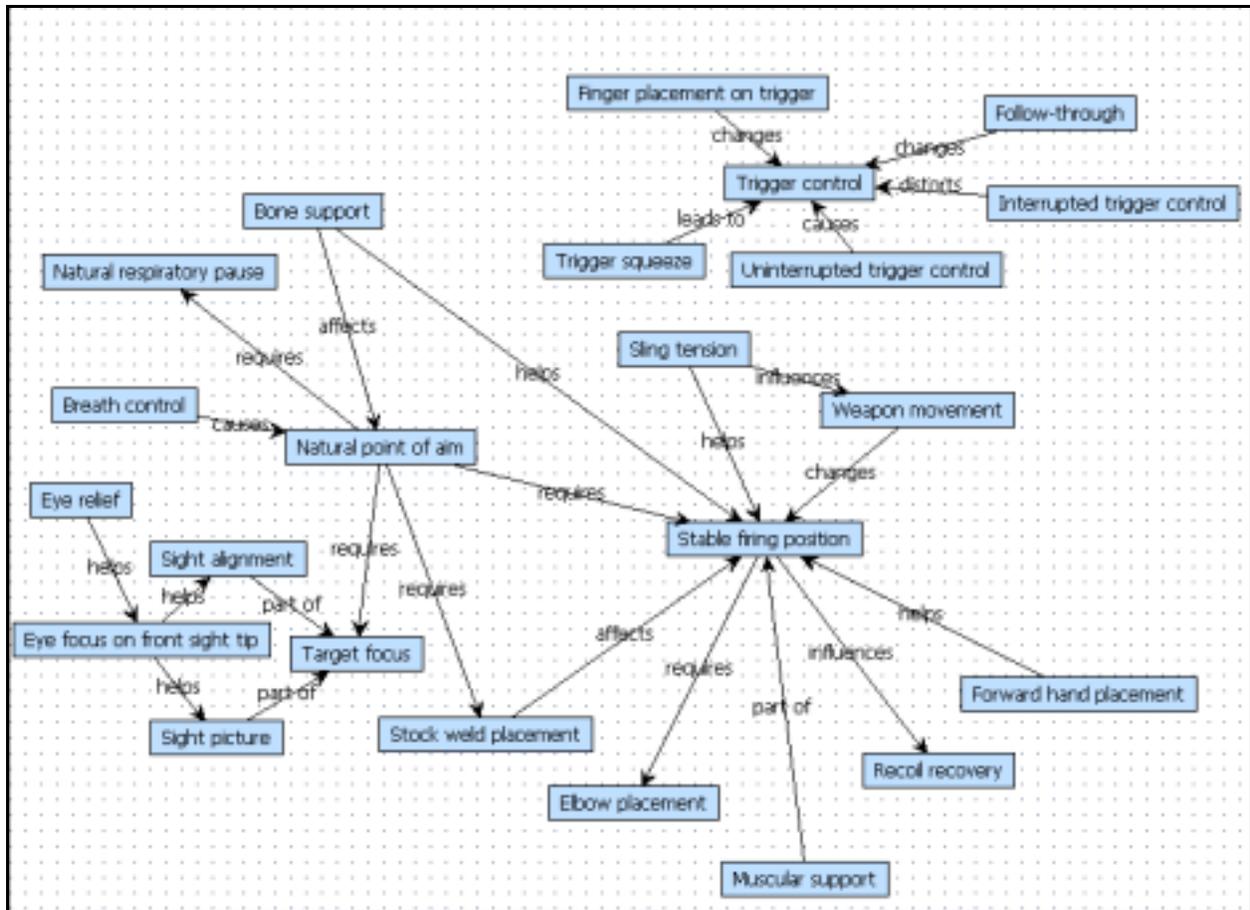


Figure 1. Example of a knowledge map

### 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 on the test being the indication of learning. However, by using the pre- and posttests as measures, one can only infer that learning has occurred because Posttests provide information about the products of learning, not the process (Lavelli et al., in press).

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, B., 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.

These 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, 1992; 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 tend to increase rapidly (Anderson, 1992; 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, 1992; 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 potentially useful information 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, two times during live-fire practice, and after qualification. We could then qualitatively analyze knowledge maps at various stages of the learning process providing snapshots of their knowledge representation over time.

Likewise, because research has shown that the HPKMT can detect changes in learning (Chung et al., 2001; Osmundson et al., 1999; Schacter et al, 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 should increase after classroom instruction and perhaps either stabilizes, or the rate of growth would either decrease or increase in smaller increments.

## **METHOD**

### **Participants**

Fifty-three 2<sup>nd</sup> 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 examined the differences in outcomes of the HPKMT by scoring the knowledge maps holistically against an expert map 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 applying quality ratings to each individual proposition. We modified Osmundson et al.'s (1999) coding scheme to reflect the causal relationships in the marksmanship domain. The scoring rubric is 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 of the domain applied the quality ratings to all of the individual propositions.

**Table 1.** Proposition Scoring Rubric

Score	Definition	Example
0	Illogical or nonsensical	<ul style="list-style-type: none"> <li>rifle butt placement affects rifle butt placement</li> </ul>
1	Superficial: Reflects superficial knowledge of marksmanship (categorical links are used, i.e. <i>part of</i> )	<ul style="list-style-type: none"> <li>aiming process is part of fundamentals of marksmanship</li> <li>forward hand placement is part of 7 common factors</li> </ul>
2	Limited: Indicates a cause-effect relationship exists	<ul style="list-style-type: none"> <li>forward elbow placements affects controlled muscular tension</li> </ul>
3	Principled: Explains how the concepts are related	<ul style="list-style-type: none"> <li>breath control requires natural respiratory pause</li> <li>stable firing position requires 3 elements of a good shooting position</li> </ul>

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 across time and remained stable the other days.
Declining	Weighted scores for these individuals decreased at least twice across 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 day 2-4, 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. The difference between the mean scores from Day 1 to Day 2-Day 5 were significant,  $p < .05$ . The differences between the other days were not significant.

**Table 3.** Means and Standard Deviations for Expert Content Score

	Expert content score	
	<i>M</i>	<i>SD</i>
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

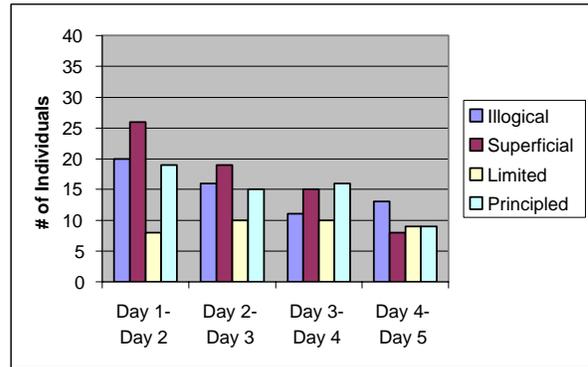
**Proposition Quality Ratings**

Table 4 shows the means and standard deviations of each proposition rating by occasion. Changes were minimal across days and any differences across time were not significant.

**Table 4.** Means and Standard Deviations for Proposition Scores

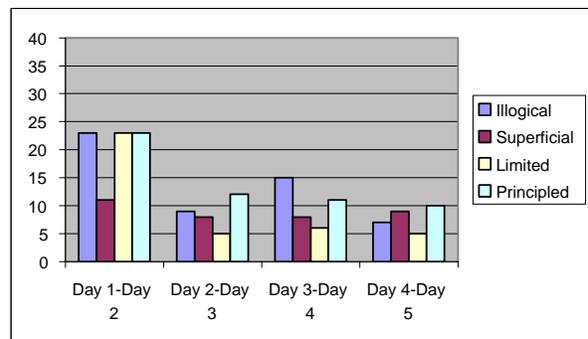
		Quality of proposition scores			
		0	1	2	3
Day 1	<i>M</i>	5.48	9.62	1.62	7.46
	<i>SD</i>	3.63	5.92	1.96	4.74
Day 2	<i>M</i>	6.12	9.33	1.76	7.35
	<i>SD</i>	4.59	6.26	2.16	4.60
Day 3	<i>M</i>	5.58	9.00	1.48	7.06
	<i>SD</i>	3.81	5.96	1.78	5.08
Day 4	<i>M</i>	5.26	9.38	1.68	6.04
	<i>SD</i>	4.27	6.56	2.02	4.98
Day 5	<i>M</i>	5.15	8.61	1.67	8.43
	<i>SD</i>	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 each day. The number of individuals whose scores for illogical, superficial, and principled knowledge increased was greatest on Day 2.



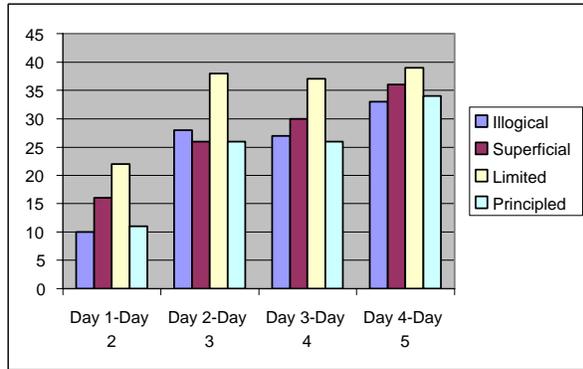
**Figure 2.** Number of Individuals Whose Proposition Scores Increased Daily.

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- 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

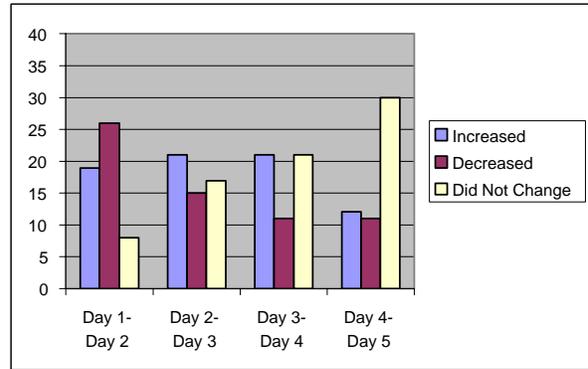
**Weighted Scores and Learning Characteristics**

Our second set of analyses examined the change of the weighted scores. Table 5 shows the means and standard deviations of the weighted scores. Again, any minimal differences between average weighted score were not significant.

**Table 5.** Means and standard deviations for weighted scores

Weighted scores		
	<i>M</i>	<i>SD</i>
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

Results of the number of individuals whose weighted scores increased, decreased, or did not change are displayed in Figure 5. The number of individuals whose weighted scores increased decreased on the last day existing maps was modified (Day 5). The number of individuals whose weighted scores decreased continues to decline each day. The number of individuals whose weighted scores remain stable increased over time.



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

Using the coding scheme in Table 2, we classified the trainees. 38 individuals who completed a map each day were used. The breakdown of the classifications is found in Table 6.

**Table 6.** Breakdown of Classifications

	Growing	Declining	Stable	Erratic
# 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 to 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 demonstrate

that there is potential for HPKMT to provide evidence of characterizing trainees in the cognitive phase.

We examined the maps using three ratings: expert content scores, quality of individual propositions, and weighted scores. Because Fitts and Posner (1967) suggested that learners in the cognitive phase should benefit from instruction, we expected that all three ratings would increase post-classroom instruction. We did find that the trainees' mean scores were significantly higher on Day 2 using the expert rating. However, our qualitative ratings (quality of individual propositions and weighted scores) did not have the same increase. In fact, although the results were not significant, several ratings decreased on Day 2 from Day 1. For example, the mean weighted score decreased by 2 points. A little less than half of our sample decreased in limited proposition scores and principled proposition scores. This result was very unusual in comparison to similar analyses that we have done with other samples. We typically find that all three ratings increase post-classroom instruction.

We then examined the knowledge maps to uncover an explanation for these results. As it turns out, many trainees changed their links from Day 1 to Day 2. Maps on Day 1 were filled with links such as "requires", "helps" and "uses" which 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 happened because marksmanship instruction is often presented on a surface level primarily because it is directed toward novice learners. It is with experience and more instruction that one begins to understand that the concepts are related. Therefore, trainees who might have used a lot of "causal" links on Day 1, but changed to more superficial links on Day 2 to reflect what and how they were taught would have had the decrease in limited proposition scores, principled proposition scores and weighted mean scores. Evidence supporting this explanation is found with the expert maps. They tended to have more superficial connections such as "part of" and "type of" explains why expert content ratings increased on Day 2 whereas quality ratings decreased.

To begin classifying groups, we used their weighted scores to classify the trainees *growing*, *declining*, *stable*, or *erratic*. Out of the 38 individuals we classified, 7 were classified as *growing*, 2 were classified as *declining*, 14 were classified as *stable*, and 14 were classified as *erratic*. The characteristics of the cognitive phase suggest that those who would be

classified as *growing* and *stable* might shoot well while those classified as *declining* and *erratic* might shoot poorly. As expected, 5 out of the 7 who were classified as *growing* shot well, qualifying as sharpshooters and experts. 2 out of the 3 who were classified as *declining* qualified as a marksmen.

We were surprised to find that 7 out of the 14 of those who we classified as *stable* qualified as marksmen, 2 as sharpshooter, and 1 as expert. One potential explanation could be that those who were classified as *stable* made no changes to their knowledge maps. However, close examination of their maps revealed that only 1 individual did not make changes.

We then examined each individual's quality of proposition scores to see which scores were higher: superficial or principled. Conceivably, those with more superficial propositions would shoot worse than those with more principled propositions. Interestingly, for the group who qualified as marksmen, 4 out of the 7 had more superficial propositions than principled, and the 2 who had more principled propositions qualified as high marksmen. For the group who qualified as sharpshooter or expert, 4 out of the 7 had more principled propositions and 2 individuals had equal scores for both. This suggests that although the 14 trainees' knowledge stabilized across instruction, there were really two distinct groups within the category. One group had a more superficial understanding of rifle marksmanship, whereas the other group had a deeper, more principled understanding of the domain which was reflected in their shooting scores.

Again, contrary to expectations, only 4 out of the 10 that were classified as *erratic* qualified as marksmen. Unfortunately, we did not find two distinct groups as we did with the *stable* category. While 3 out of the 4 marksmen had more superficial propositions, only half of those who qualified as sharpshooter or expert had more principled propositions. This finding could be a result of our small sample, or perhaps it just might be more difficult to make assumptions about individuals whose knowledge map scores follow no trend.

By classifying the learner into one of the four groups an instructor would have some insight into whether or not an individual was not benefiting from the instruction and practice. We shared 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 analysis has the potential to assist in classifying a marksmanship

trainee. To strengthen our results, we are planning on using the HPKMT to conduct similar microgenetic analyses on the sample of individuals who went through sustainment-level rifle marksmanship.

### ACKNOWLEDGEMENTS

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.

### 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: 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* (pp.269-293). Cambridge, U.K.: Cambridge University Press.
- 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).
- 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 Psychologist*, 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* (pp.3-18). New York: John Wiley & Sons, Inc.