Sensor-Based Assessment of Basic Rifle Marksmanship

Sam O. Nagashima, Gregory K. W. K. Chung, Paul D. Espinosa
CRESST/UCLA
Los Angeles, CA
samn@ucla.edu, greg@ucla.edu, paulespi@ucla.edu

Chris Berka
Advanced Brain Monitoring, Inc.
Carlsbad, CA
chris@b-alert.com

ABSTRACT

Chung, Delacruz, de Vries, Bewley, and Baker (2006) describe five dimensions that underlie marksmanship performance—perceptual-motor, cognitive, affective, equipment, and environment. While prior research on the impact of cognitive (Harllee, 1916), affective (Tierney, Cartner, & Thompson, 1979; Sade, Bar-Eli, Bresler, & Tenenbaum, 1990; Chung, O'Neil, Delacruz, & Bewley, 2005), equipment, and environment (Osborne, Morey, & Smith, 1980) variables abound, studies on the perceptual-motor dimension are relatively incomplete. A possible reason for this lack of attention is the inherent difficulty in measuring the fine-motor movements involved in the practice of marksmanship. In training environments, a marksmanship instructor typically relies on two methods to evaluate shooter performance, visual observations of the shooter and inspection of shot distribution on the target, also known as shot placement analysis. Both methods, while commonplace, are subjective and dependent on a human observer, making it difficult to consistently assemble accurate and reliable information about a shooter's skill performance. This in turn leads to inefficiencies in diagnosis and feedback during the training process. A reliable and objective method is needed to help make proper diagnosis possible so that appropriate instruction is made available.

This paper presents work on sensor-based measures of rifle marksmanship skill performance. Sensors were developed to collect information on four skill areas, three related to breath control and one for trigger control. Data collected from experts (n = 9) and novices (n = 30) were used to fit and test a model of skill performance. Using sensing information alone, a two-class model differentiating between expert and novice skill performance achieved an accuracy of >90%. These preliminary results suggest that sensor-based skill assessment is a viable option as a reliable and objective measure to discriminate levels of skill performance in rifle marksmanship.

ABOUT THE AUTHORS

Sam O. Nagashima is a research associate at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST). His research activities revolve around assessment development and use. He holds a B.S. in Biochemistry & Cell Biology from the University of California, San Diego, and an M.S. in Education from the University of Pennsylvania. He is currently a doctoral student at the University of California, Los Angeles, where his study focuses on measurement, assessment, and research methodologies.

Gregory K. W. K. Chung is a senior researcher at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST). His current work at CRESST involves developing computer-based assessments to measure problem solving and content knowledge in military and engineering domains. He has experience in developing Web-based assessment tools for diagnostic and embedded assessment purposes using Bayesian networks, domain ontologies, and other advanced computational tools. He holds a Ph.D. in educational psychology from the University of California, Los Angeles, an M.S. in educational technology from Pepperdine University at Los Angeles, and a B.S. in electrical engineering from the University of Hawaii at Manoa.

Paul D. Espinosa is a software engineer at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) where his core responsibility is to build data collection software. He has managed various software projects where the design emphasis is on extensibility and usability. He earned a B.S. in Computer Science and Engineering at the University of California, Los Angeles.

Chris Berka, CEO and Co-Founder of Advanced Brain Monitoring, Inc., has over 25 years experience managing clinical research and developing and commercializing new technologies. She is co-inventor of 7 patented and 7 patent-pending technologies and principal investigator for grants awarded by National Institutes of Health, DARPA, ONR, and IARPA providing over $17 million of research funds to ABM. She received her B.A. with distinction in Psychology/Biology at Ohio State University and completed graduate studies in Neuroscience at the University of California, San Diego.

Sensor-Based Assessment of Basic Rifle Marksmanship

Sam O. Nagashima, Gregory K. W. K. Chung, Paul D. Espinosa
CRESST/UCLA
Los Angeles, CA
samn@ucla.edu, greg@ucla.edu, paulespi@ucla.edu

Chris Berka
Advanced Brain Monitoring, Inc.
Carlsbad, CA
chris@b-alert.com

INTRODUCTION

A common method of assessing marksmanship performance is the examination of shots on target (e.g., score, accuracy, tightness of shots) to make judgments about a shooter's skill level. Research on marksmanship performance such as correlations between simulators and live fire (Hagman, 1998; Schendel, Heller, Finley, & Hawley, 1985; Smith & Hagman, 2000), the impact of nutrition on performance (Tharion & Moore, 1993), and the role of anxiety on novice shooters (Chung, O'Neil, Delacruz, & Bewley, 2005) have all used shot placement as the primary measure of shooter skill.

While the ultimate goal in marksmanship training is to achieve accurate and consistent shot placement, evaluation of a shooter solely on this basis can conceal underlying differences in shooter skill. For example, individual shooters can exhibit identical shot placement while executing radically different procedures during the shooting process (e.g., breath control, trigger control, sight alignment). Consequently, within a training context, the use of shot placement alone is not sufficient for the identification and subsequent remediation of problematic aspects of performance, that is, performance inconsistent with doctrine (AERA, APA, NCME, 1999; Wiggins, 1998).

Rifle marksmanship coaches and instructors augment shot placement analysis with visual observations of the shooter to help with the diagnosis of skill gaps. Although visual observation of the shooter during the shooting process is a direct measure of marksmanship skill performance, the use of human observers is not ideal. Some aspects of performance, such as overall position quality, coarse movement of the muzzle, and the trigger break, are easily distinguished by a human evaluator; on the other hand, many steps leading up to the trigger break (e.g., aiming, trigger squeeze, control of respiration) are less perceptible, making direct visual observation and proper diagnosis difficult.

A good measure, or metric, must be objective, at the level of detail appropriate for decision making, precise, generalizable, sensitive, reliable, and most importantly, valid (ANSI, 1993). By these standards, both shot placement analysis and visual observation of the shooter may not be the most appropriate measures by which to diagnose shooter skill performance. A new method, one that directly captures and measures the subtle movements employed during the practice of marksmanship, is needed to support the assessment needs of coaches and instructors during training.

Sensors have the potential to serve as a reliable and unobtrusive surrogate in situations where human observations are impractical or unreliable (De Ketelaere, Bamelis, Kemps, Decuypere & De Baerdemaeker, 2004; Wide, Winquist, Bergsten, & Petriu, 1998). As a methodology for evaluating human performance, sensors have already been shown to be effective in the medical field differentiating levels of experience of arthroscopic surgeons (Chami, Ward, Phillips, & Sherman, 2008) and laparoscopic surgeons (Rosen, Solazzo, Hannaford, & Sinanan, 2001).

The goal of this study was to test whether sensor-based measures designed to assess key aspects of marksmanship skill are sensitive enough to differentiate between levels of rifle marksmanship skill performance. A complete sensing system was developed concentrating on two areas of marksmanship performance believed to be important in known distance precision shooting, breath control and trigger control. Each shot was evaluated, using expert criteria, to judge the quality of performance.

METHODS

Participants

Shots were collected from 39 participants, 30 novices and 9 experts. Novices ranged in age from 19 to 29 years (M = 22.20, SD = 2.57). Of the 30 novices, 23

(77%) were male, and 7 (23%) were female. Twelve (40%) reported having prior experience shooting a rifle. Of those reporting prior experience, 3 (25%) reported having shot a rifle within the last year, 3 (25%) within 3 to 5 years, and 6 (50%) reported firing a rifle over 5 years ago. None of the novices reported experience with competitive shooting.

The nine experts selected for this study were active-duty members in the same branch of the armed forces with a primary military occupation specialty (MOS) as marksmanship coaches. All were male and ranged in age from 21 to 25 years (M = 23.33, SD = 1.41). Coaching experience ranged from 1 to 24 months (M = 12.44, SD = 7.52). In addition to being rifle marksmanship coaches, five (56%) were also qualified as rifle marksmanship instructors.

While several participants in the novice sample had some familiarity with marksmanship, none had training consistent with marksmanship instruction as delivered in the armed forces. Accordingly, all the novices used in this study were regarded as novices.

Study Design

Holdout validation was used to assess the quality of skill classifications based on estimated model parameters (Kerlinger & Pedhazur, 1973). Participants were randomly assigned to two groups, model training and model testing. Individuals assigned to model training were used to estimate model parameters, while those in the model testing group were held back from the estimation procedure and later fitted to the data. Sample distribution of subjects across the two data files is presented in Table 1.

Table 1. Distribution of Subjects

	Data		
Status	**Training**	**Testing**	**Total**
Novice	15	15	30
Expert	5	4	9
Summary	20	19	39

Apparatus

Data were collected in an indoor controlled environment. An instrumented weapon was developed using off-the-shelf sensing components and a demilitarized M16/A2 housing a pneumatic recoil system designed to approximate the weight, noise, and action of a real weapon firing live rounds (LaserShot, 2008).

Four performance skill measures were collected using two sensors, a force-pressure sensor attached to the trigger and a respiration belt. Both sensors were wired to a processing unit that was designed in-house and contained circuitry to handle the sensor signals (e.g., signal conditioning, amplifiers) and a microcontroller. The microcontroller sampled the sensor signals at a rate of 128 samples/sec with 10-bit resolution. The sensor data (i.e., each sample of each sensor) were packaged into a data frame and sent via Bluetooth to a data integration laptop.

Shots were directed against a projection of a circular target equivalent to 20 inches at 200 yards. A camera identified shot placement on target by recognizing infrared laser strikes delivered by the rifle. A data integration program, Fusion 4000, was developed to display and collect shot performance data. Fully developed in-house, Fusion 4000 is a combination of earlier data collection applications intended to streamline the data collection process. Using this single program, set-up of the targeting system to collect shooting performance and interfacing with the microcontroller to collect sensor data were possible. Additionally, data collection within a single application enabled proper synchronization and logging of the data. For detailed information on the development of the sensor-based measures, targeting system, and signal processing procedure, see Espinosa, Nagashima, Chung, Parks, and Baker (2009).

Procedure

The novice sample selected for this study was intended to represent a typical trainee. With that in mind, all novices were provided basic instruction on shooting position, weapons handling, and proper sight alignment; however, instruction regarding specific details related to doctrine, such as breath control and trigger control, were withheld.

All participants were instructed to shoot in the kneeling position. They were given the option of choosing between low, medium, or high kneeling. Variations in the kneeling position were modeled by the instructor; in addition, illustrations depicting left- and right-handed variations on the kneeling position were provided.

Ten shots were collected and analyzed from each subject across two trials. No time constraints were imposed on the shooters and they were not provided feedback regarding shot placement until the end of each trial.

Measures

Four performance skill measures were evaluated for each shot, three related to breath control (*breath location*, *breath duration*, and *shot-percent breath*) and one for trigger control (*trigger duration*).

Breath location represents the location of the trigger break in the respiratory cycle. Values range from 0% to 100%, with 0% indicating a shot taken while fully exhaled, and 100% indicating that the shooter was fully inhaled during the trigger break. Doctrine dictates shots be taken during a natural respiratory pause, therefore a value near zero is desirable. Figure 1 illustrates a trigger break during a natural respiratory pause.

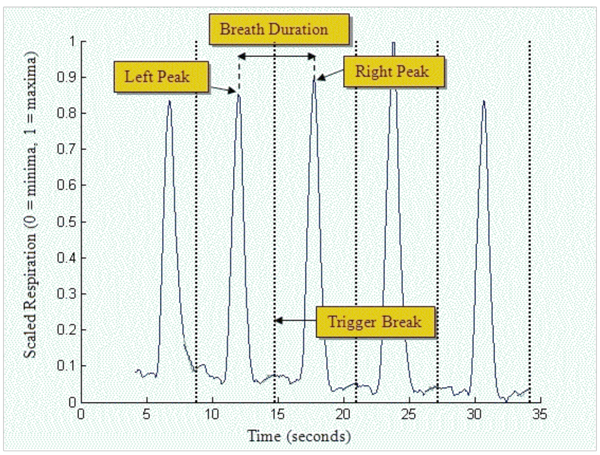

Figure 1. Breath Duration (Focused on Second Trigger Break)

Breath duration is a measure of the time, in seconds, between full inhales flanking the shot. This is illustrated in Figure 1 as the time between left peak and right peak. Larger values indicate longer periods of time between breaths (slower rate of respiration); conversely, smaller values indicate a shorter period of time between breaths (faster rate of respiration). While doctrine does not specify an exact value for preferred rate of respiration, extremely quick respiration (signified by a small value) is undesirable.

Shot-percent breath is used to approximate the location, in percent, of where the trigger break occurred relative to the full inhales spanning the trigger break. This calculation relies on breath duration in conjunction with the trigger break. For example, .50 indicates that a shot was fired in the middle of a respiratory cycle. There is no desired value for this particular variable; it is used to measure a potential difference between expert and novice performance.

Trigger duration was the only measure of trigger control and represents the amount of time, in seconds, pressure is exerted on the trigger prior to a shot being fired. Larger values indicate a greater amount of time taken to pull the trigger, conversely, smaller values indicate less time pulling the trigger. Extremely short trigger duration, indicative of jerking the trigger, is undesirable.

Analysis

A logistic regression model was developed to test the extent to which shots can be classified as originating from a novice or expert using the skill measures as predictors. The skill measures were used as four continuous predictor variables—*breath location, breath duration, trigger duration,* and *shot-breath location*. The logistic regression analysis was carried out using the binary logistic regression command in Statistical Package for the Social Sciences (SPSS,® 1999) version 16 in Windows 2000 environment.

The statistical significance of individual coefficients was tested using the Wald chi-square statistic, and the Hosmer-Lemeshow (H–L) test was used to assess the goodness-of-fit for the final model.

Several indices describing the predictive performance were calculated to assess predicted model classifications—sensitivity specificity, false positive, false negative, and c-statistic. Sensitivity, or true positive fraction, is the proportion of correctly classified experts. Specificity, or true negative fraction, represents the proportion of correctly classified novices. False positive is the proportion of cases misclassified as experts, while false negative is the proportion of cases misclassified as novices.

The c-statistic is a measure of discrimination. Ranging from 0.5 to 1, a value of 0.5 indicates that the model is no better than assigning observations randomly into outcome categories; A value of 1 indicates that the model assigns higher probabilities to all observations with the event outcome, compared with nonevent observations.

RESULTS

Descriptive Statistics

Mean and standard deviations for the skill variables are provided in Table 2. For all shots, *breath location* ranges from 0.00 to 91.80 (M = 35.34, SD = 23.86), *breath duration* ranges from 0.31 to 13.16 seconds (M = 3.45, SD = 2.35), *shot-percent breath* ranges from

0.01 to 1.00 (M = .54, SD = .25), and *trigger duration* ranges from 0.00 to 95.27 seconds (M = 4.32, SD = 8.06).

Table 2. Mean and Standard Deviation for Skill Measures

Variables	Status		
	Novice *M (SD)*	**Expert** *M (SD)*	**All** *M (SD)*
Breath control			
Breath location	42.05 (22.85)	12.97 (8.71)	35.34 (23.86)
Breath duration	2.54 (1.10)	6.46 (2.85)	3.45 (2.35)
Shot-percent breath	0.52 (0.27)	0.64 (0.17)	0.54 (0.25)
Trigger control			
Trigger duration	5.20 (8.93)	1.42 (2.29)	4.32 (8.06)

Note. n = 300 for novice group and n = 90 for expert cases.

Mean *breath location* was 42.1 (SD = 22.9) for novices, and 13.0 (SD = 8.7) for experts. When a shot was fired, novices, on average, were partially inhaled, while experts were nearly fully exhaled at trigger break. The mean *breath duration* for novices was only 2.5 seconds (SD = 1.1), and 6.5 seconds (SD = 2.8) for experts. These values indicate that the average respiratory cycle for novices around the trigger break lasts 2.5 seconds, whereas for experts, the respiratory cycle lasts 6.6 seconds. The mean *shot-percent breath* for novice was .52 (SD = .27) and .64 (SD = .17) for experts. The novice group mean of .52, or 52%, indicates that the average shot was fired midway between two full inhales, while the expert group has a mean of .64, or 64%, which indicates that experts take shots closer toward the end of a respiratory cycle. For the measure of trigger control, the mean *trigger duration* for novices was 5.2 sec (SD = 8.9), and 1.4 sec (SD = 2.3) for the expert group. Novices appear to take longer pulling the trigger, 5.2 seconds, compared to experts, 1.4 seconds.

Pearson correlations among the skill variables are reported in Table 3. The correlation of *breath location* and *breath duration* was significant, r (388) = -.488, p < .001, as was *breath location* and *shot-percent breath* at r (388) = -.216, p < .001, and *breath duration* and *shot-percent breath*, r (388) = .280, p < .001. *Trigger duration* did not correlate significantly with the other variables.

Table 3. Correlations for Measures of Skill Performance

Variables	1	2	3	4
1. Breath location	--			
2. Breath duration	-.488**	--		
3. Shot-percent breath	-.216**	.280**	--	
4. Trigger duration	.096	-0.049	0.005	--

Note. N = 390.
$^{**}p$ < 0.01 (two-tailed).

Given the significant correlation values between the variables related to breath control, tolerance values were calculated to assess the threat of collinearity. Tolerance values are reported in Table 4. Tolerance values range from .730 for *breath duration* to .990 for *trigger duration*. Based on the critical value tolerance < .2, the potential threat of collinearity is negligible (Menard, 1995).

Table 4. Collinearity Statistics for Independent Variables

Variable	Tolerance
Breath control	
Breath location	.750
Breath duration	.730
Shot-percent breath	.913
Trigger control	
Trigger duration	.990

Logistic Regression Model

We first estimated individual univariate logistic regression models for the variables in the training data to test the research hypothesis regarding the relationship between the likelihood of classification as expert based on the individual measures of skill performance. The outcome variable, *status*, was used to designate the shot classification as expert marksman (1 = yes, 0 = no). The four continuous predictor variables include the three variables for breath control (*breath location*, *breath duration*, *shot-percent breath*) and one measure for trigger control (*trigger duration*). Table 5 presents the results from the analysis for the univariate relationship between the skill measures and marksmanship status.

Table 5. Summary of Univariate Logistic Regression Results for Marksmanship Skill Variables Using Training Data (n = 190)

Variables	*B*	*SE*	Wald statistic	OR	95% CI
Breath control					
Breath location	-0.137	0.022	37.559	0.872**	.834 - .911
Breath duration	2.039	0.335	37.060	7.686**	3.986 - 14.820
Shot-percent breath	2.892	0.758	14.572	18.029**	4.084 - 79.584
Trigger control					
Trigger duration	-0.237	0.075	9.902	0.789*	.681 - .915

Note. OR = odds ratio, CI = confidence interval.
$^{*}p < .01.$ $^{**}p < .001.$

Considered individually, all four variables are significant (p < .01) predictors relative to the null model. Next, we estimated a multiple logistic regression model to investigate the simultaneous effects of all four skill measures on status. Given the significance of the four predictor variables in the univariate model, a four-predictor multiple logistic model was fitted to the data. Table 6 presents the results of multiple regression analysis.

Table 6. Summary of Multiple Logistic Regression Results for Marksman Skill Variables Using Training Data (n = 200)

Variables	*B*	*SE*	Wald statistic	OR	95% CI
Breath control					
Breath location	-0.148	0.052	7.946	0.862**	.778 - .956
Breath duration	2.111	0.491	18.502	8.256***	3.155 - 21.604
Shot-percent breath	2.241	3.752	0.357	9.398	.006 - 14691
Trigger control					
Trigger duration	-0.540	0.200	7.282	0.583**	.393 - .863
Constant	-6.381	2.604	6.003	0.002*	

Note. OR = odds ratio, CI = confidence interval.
$^{*}p < .05.$ $^{**}p < .01.$ $^{***}p < .0001.$

When all four predictors are considered jointly, the overall model significantly differentiates between expert and novice skill performance relative to the null model, χ^2 =188.18, df = 4, p < .001. The variables *breath location*, *breath duration*, and *trigger duration* are significant (p < .05). The variable *shot-percent breath*, while a significant predictor when used alone, was not a significant predictor when used concurrently with all four variables. The test of the intercept (i.e., constant in Table 6) suggests the intercept should be included in the model.

When interpreting the logistic regression results, an odds ratio greater than 1.0 implies a positive association between the skill measure and status, while an odds ratio less than 1.0 implies a negative association. Odds ratios close to 1.0 indicate that unit changes in that skill variable do not affect the odds of predicted status. The variable *breath location* with an odds ratio of .862 indicates that as *breath location* increases, the odds of expert skill diminish. Specifically, the odds of expert classification diminished by a factor of .137, for one unit increase in location, controlling for other variables in the model. Additionally, for *breath duration*, a one-second increase in *breath duration* results in an 8.26 times greater chance of expert classification. Lastly, the odds ratio of .583 for *trigger duration* signifies that for every one-second increase in *trigger duration*, the odds of expert classification decreases by a factor of .417. The variable *shot-percent breath* was not a significant predictor.

Overall Model Evaluation

The Hosmer–Lemeshow test of inferential goodness-of-fit yielded a $\chi^2(8)$ of .196 and is non-significant ($p >$ .05), suggesting that the model exhibits a considerable degree of fit to the data. In other words, the null hypothesis of a good model fit to data was tenable. The logistic model resulted in a c-statistic of .973, indicating that for 97.3% of all possible pairs of shots—one expert and the other novice—the model correctly assigned a higher probability to those who were expert.

For the model training data, 147 of 150 novice shots and 46 of 50 expert shots were accurately classified. Accordingly, the sensitivity, the ability to identify expert shots, was 92%, and the specificity, the power to identify novice shots, was 98% for the training data.

For the model testing data, 144 of 150 novice shots and 13 of 40 expert shots were accurately classified, resulting in a sensitivity of 96% and specificity of 67.5%. A 2 × 2 classification table showing observed versus predicted classifications, based on a cutoff value of .50, or 50%, can be found in Table 7 for the training data, and Table 8 for the testing data.

Table 7. Summary of Predicted Classification for Model Training Data

	Predicted		
Observed	**Novice**	**Expert**	**% Correct**
Novice	147	3	98.0
Expert	4	46	92.0
Overall % correct			96.5

Note. Cut value set at .50.

Table 8. Summary of Predicted Classification for Model Testing Data

	Predicted		
Observed	**Novice**	**Expert**	**% Correct**
Novice	144	6	96.0
Expert	13	27	67.5
Overall % correct			90.0

Note. Cut value set at .50.

As illustrated in Table 9, the classifying of experts was less accurate than novices. This observation is supported by the magnitude of sensitivity (67.5%) compared to that of specificity (96.0%). Both false positive and false negative rates were modest at 18.2% and 8.3% respectively. Given the distribution of expert and novice across the two data files, the default accuracy in classification by identifying all cases as novice (the most prominent classification) in the training data was 75% and 78.9% in the testing data. Compared with the overall percent correct classification in the training data (96.5%) and the testing data (90.0%), there was a 21.5% and 11.5% improvement, respectively.

Table 9. Classification Performance for Testing Data

Measure	Value	Definition
Sensitivity	0.675	Proportion of correctly classified events (expert).
Specificity	0.960	Proportion of correctly classified nonevents (novice).
False positive	0.182	Proportion of observations misclassified as expert over all of those classified as experts.
False negative	0.083	The proportion of observations misclassified as novices over all of those classified as novice.

DISCUSSION

In this study, our objective was to consider whether sensor-based skill measures have the discriminatory power to differentiate expert-novice skill performance in marksmanship. A key finding in our analysis is that a collection of sensor-based skill measures, specifically breath location, breath duration, and trigger duration, considered jointly, provides a reliable method for discriminating between expert-novice skill performance in rifle marksmanship. The resulting evidence from this study justifies the use of sensors as a valid and reliable tool to measure marksmanship skill performance.

While these results are promising, it must be noted that in evaluating the predicted probabilities, the training data exhibited an accuracy rate of 96.5% and the testing data only 90.0%. One possible explanation for this discrepancy is the variability in skill performance across experts. Whereas the 6 false positive cases in the testing data are distributed nearly evenly across five novice shooters, all 13 false negative cases come from only two experts; one expert shooter accounted for 9 (69.2%) of the false negative classifications, with the remaining 4 (30.8%) attributed to another. Given that in the testing data, all 4 false negative cases came from a single expert shooter, there is reason to believe that, across experts, there is a considerable amount of variability in skill performance. An extension to this study, one that examines skill performance differences across different experts is warranted. Moreover, further refinement of the expert group into subgroups (e.g., marksman, sharpshooter, expert) may lead to improved predictions and shed light on additional levels of skill performance.

Lastly, although we are confident in the results, for the time being, we remain cautious in extending the generalizability of the results to live-fire environments. The use of live rounds introduces a number of factors, such as heat and noise, which may impact the reliability of the sensors. Additional studies are needed to assess the validity and reliability of sensor-based assessment of skill performance in live-fire environments to support skill diagnosis.

ACKNOWLEDGEMENTS

We would like to thank the staff at Camp Pendleton WTBN. We would also like to thank the following people from UCLA/CRESST: Joanne Michiuye for her help with the preparation of this manuscript and with data collection, and Daniel Parks for hardware design and development.

REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (AERA, APA, NCME). (1999). *Standards for educational and psychological testing.* Washington, DC: American Educational Research Association.

American National Standards Institute (ANSI). (1993). *Guide to human performance measurement* (G-035-1992). Washington, DC: Author.

Chami, G., Ward, J. W., Phillips, R., & Sherman, K. P. (2008). Haptic feedback can provide an objective assessment of arthroscopic skills [Electronic version]. *Clinical Orthopaedics and Related Research, 466,* 963–968.

Chung, G. K. W. K., Delacruz, G. C., de Vries, L. F., Bewley, W. L., & Baker, E. L. (2006). New directions in rifle marksmanship research. *Military Psychology, 18,* 161–179.

Chung, G. K. W. K., O'Neil, H. F., Delacruz, G. C., & Bewley, W. L. (2005). The role of anxiety on novices' rifle marksmanship performance. *Educational Assessment, 10*(3), 257–275.

De Ketelaere, B., Bamelis, F., Kemps, B., Decuypere, E., & De Baerdemaeker, J. (2004). Non-destructive measurements of the egg quality. *World's Poultry Science Journal, 60,* 289–302.

Espinosa, P. D., Nagashima, S. O., Chung, G. K. W. K., Parks, D., & Baker, E. (2009). *Development of sensor-based measures of rifle marksmanship skill and performance* (CRESST Tech. Rep. No. 756). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Hagman, J. D. (1998). Using the Engagement Skills Trainer to predict rifle marksmanship performance. *Military Psychology, 10,* 215–224.

Harllee, W. C. (1916). *U.S. Marine Corps Score Book and Rifleman's Instructor*. Philadelphia: International Printing Company.

Kerlinger, F. N., & Pedhazur, E. J. (1973). *Multiple regression in behavioral research.* New York: Holt, Rinehart & Winston.

LaserShot. (2008). Military Skills Engagement Trainer [Military training simulator system]. Stafford, TX: Author.

Menard, S. (1995). *Applied logistic regression analysis.* Thousand Oaks, CA: Sage Publications.

Osborne, A. D., Morey, J. C., & Smith, S. (1980). *Adequacy of M16A1 rifle performance and its implications for marksmanship training* (Research Report 1265). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Rosen J., Solazzo, M., Hannaford, B., & Sinanan, M. (2001). Objective laparoscopic skills assessments of surgical residents using hidden Markov models based on haptic information and tool/tissue interactions. *Studies in Health Technology and Informatics, 81,* 417–23.

Sade, S., Bar-Eli, M., Bresler, S., & Tenenbaum, G. (1990). Anxiety, self-control and shooting performance. *Perceptual and Motor Skills, 71,* 3-6.

Schendel, J. D., Heller, F. H., Finley, D. L., & Hawley, J. K. (1985). Use of Weaponeer marksmanship trainer in predicting M16A1 rifle qualification performance. *Human Factors, 27,* 313–325.

Smith, M. D., & Hagman, J. D. (2000). *Predicting rifle and pistol marksmanship performance with the Laser Marksmanship Training System* (Tech. Rep. No. 1106). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Statistics Package for the Social Sciences (SPSS®). (1999). Version 16. Base Application Guide [Computer Software]. Chicago: Author.

Tharion, W. J., & Moore, R. J. (1993). *Effects of carbohydrate intake and load bearing exercise on rifle marksmanship performance* (T5-93). Natick, MA: U.S. Army Research Institute of Environmental Medicine.

Tierney, T. J., Cartner, J. A., & Thompson, T. J. (1979). *Basic rifle marksmanship test: trainee pretest and posttest attitudes* (Technical Paper 354). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Wide, P., Winquist, F., Bergsten, P., & Petriu, E. (1998). The human-based multisensor fusion method for artificial nose and tongue sensor data. *IEEE Transactions on Instrumentation and Measurement, 47,* 1072–1077.

Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance.* San Francisco: Jossey-Bass.