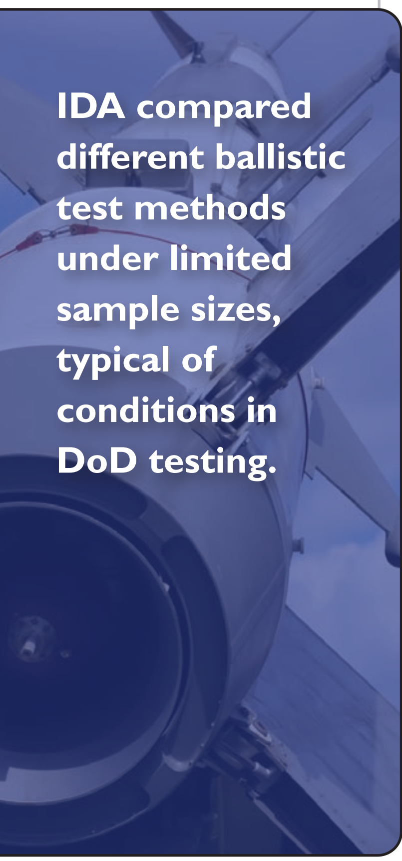
Ballistic Resistance Testing Techniques

Thomas Johnson, Laura Freeman, Janice Hester, Jonathan Bell

The Problem

The Department of Defense conducts ballistic resistance testing to estimate the probability that a projectile will perforate the armor of a system under test. Ballistic resistance testing routinely employs sensitivity experiment techniques in which sequential test designs are used to estimate armor performance. Statistical procedures used to estimate the ballistic resistance of armor in the DoD have remained relatively unchanged for decades. New test design methods can lead to improved testing efficiency, which is critical for test and evaluation in the current fiscal climate.



In reviewing sequential methods used in DoD and comparing them to those using Monte Carlo simulation, we found that newer test design and analysis techniques provide significant improvements over current methods. Newer methods can reduce test sizes, reduce bias in estimates, and support the estimation of the full probability of perforation curve instead of only a single metric.

Armor Testing

Various fields use sensitivity experiments to characterize the probability of a binary outcome as a function of a stimulus or stress. DoD ballistic characterization tests employ sensitivity experiments to characterize the probability of perforating armor or another material as a function of a projectile's velocity. One such example is shown in Figure 1.

Ballistic characterization tests, which are essential to understanding the vulnerability and lethality of military equipment, are conducted on systems ranging from body armor to vehicle and aircraft armor. In recent years, ballistic characterization tests were conducted for the Enhanced Small Arms Protection Inserts (body armor plates), the Enhanced Combat Helmet, the new floor paneling of the CH-47F, cockpit armor for the KC-46, and armor paneling of the Joint Light Tactical Vehicle, to name a few.

Ballistic testing is destructive, and DoD ballistic testing can be expensive in terms of both test and material costs. Sample sizes are generally limited, and ballistic characterization tests are almost always limited to fewer than 20 shots.

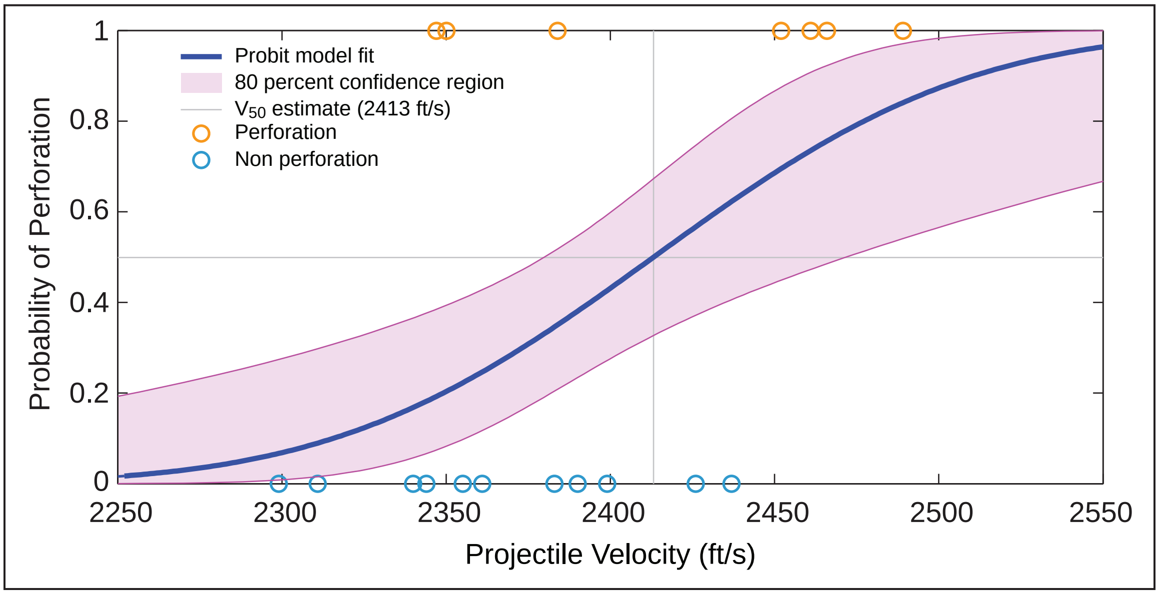


Figure 1. Example Data from a Ballistic Limit Test

Ballistic characterization testing frequently focuses on a specific percentile of interest and must therefore provide sufficient data to accurately estimate that percentile. The most common percentile of interest is the velocity at which the projectile has a 50 percent probability of perforating the armor—called the ballistic limit or V_{50} of the armor for the particular projectile (V_{50} Ballistic Test for Armor 1997). Historically, V_{50} was sufficient to characterize armor. It can also be estimated more precisely and with fewer shots than other percentiles. With modern armors, however, V_{50} might not be sufficient to characterize an armor; users might be more interested in the velocity at which the probability of perforation is 10 percent, or even lower. Estimating V_{10} is another reason to explore the efficiency of test techniques based on maximum likelihood estimation (MLE).

IDA compared different ballistic test methods under limited sample sizes, typical of conditions in DoD testing. Three simulation studies compared test methods in terms of their efficiency and accuracy at estimating V_{50} and V_{10}.

To estimate V_{50} and V_{10}, we used the probit maximum likelihood estimator (Probit-MLE). In a probit model, the response of an armor target to a ballistic projectile can be characterized as perforation or non-perforation. Let $y_i=1$ or 0 denote the binary outcome of the ith shot, perforation or non-perforation, where $i=1,2,3,\ldots,N$ are the first, second, third, and final shot, respectively. Let $F(x_i)$ denote the probability that $y_i=1$ for the velocity of the ith shot. The location-scale probit model, which we used in our article to characterize the ballistic resistance of armor, is $F(x,\mu,\sigma)=\Phi((x-\mu)/\sigma)$, where Φ is the standard normal cumulative distribution function. We define x_P to be the Pth quantile of the distribution, where $F(x_P)=P$. In this formulation, μ, the estimator of V_{50}, is estimated using maximum likelihood estimation, and is referred to as the Probit-MLE estimator (Collins 2012).

Numerous aspects of a ballistic resistance test can affect the quality and consistency of the results: the laboratory setup, location of projectile impact, obliquity angle, temperature, and projectile type, among others. We focused on statistical aspects, namely, the sequential methods.

Sequential Methods

A sequential method dictates the velocity setting of each projectile fired in a test. We investigated seven sequential methods, chosen based on their prevalence in military armor testing, ease of implementation, and overall effectiveness at estimating V_{50} and V_{10}. The sequential methods compared are:

- Up and Down Method (UD) (Dixon and Mood 1948)
- Langlie Method (LM) (Langlie 1962)
- Delayed Robbins Monroe Method (DRM) (Hodges and Lehmann 1956)
- Wu's three-phased optimal design approach (3Pod) (Wu and Tian, Three-Phase Sequential Design for Sensitivity Experiments 2013)
- Neyer's Method (NM) (Neyer 1994)
- Robbins Monroe Joseph Method (RMJ) (Wu and Tian, Three-Phase Optimal Design of Sensitivity Experiments 2014)
- K-in-a-row (KR) (Gezmu 1996).

To illustrate their utility, Figure 2 shows notional tests for selected methods.

Simulation Comparison Study

We compared sequential methods, estimators, and stopping criteria using Monte Carlo simulation. We used a Probit model to represent the true relationship between probability of perforation and projectile velocity. We considered two sets of true parameters that are reflective of the combat helmet example shown in Figure 1: (1) μ_T = 2,400 ft/s, σ_T=75 ft/s, and (2) μ_T=2,400 ft/s, σ_T= 150 ft/s.

A simulated test is carried out in a manner similar to a physical one except that no projectiles are fired, and the outcome of whether the projectile perforated the armor is determined using a random Bernoulli draw from the probability of perforation estimated from the true model. For example, if a given simulated shot (x) is fired at 2,300 ft/s, according to the first set of true parameters, the probability that that projectile perforates the armor is $\Phi((x-\mu_T)/\sigma_T) = \Phi((2{,}300-2{,}400)/75)=0.09$, where Φ is the cumulative distribution function from the normal distribution. Then, a random Bernoulli number is generated that has a 9 percent chance of being a perforation. To instill more realism into the simulation, we include a velocity set point error. For each calculated velocity, we add a random error drawn from a uniform distribution between plus or minus 10 ft/s.

The simulation employs a full factorial experiment to compare the different test designs in terms of their ability to estimate V_{50} and V_{10}. Table 1 shows the variables considered in the simulation experiment. Ideally, we hope to find a method that results

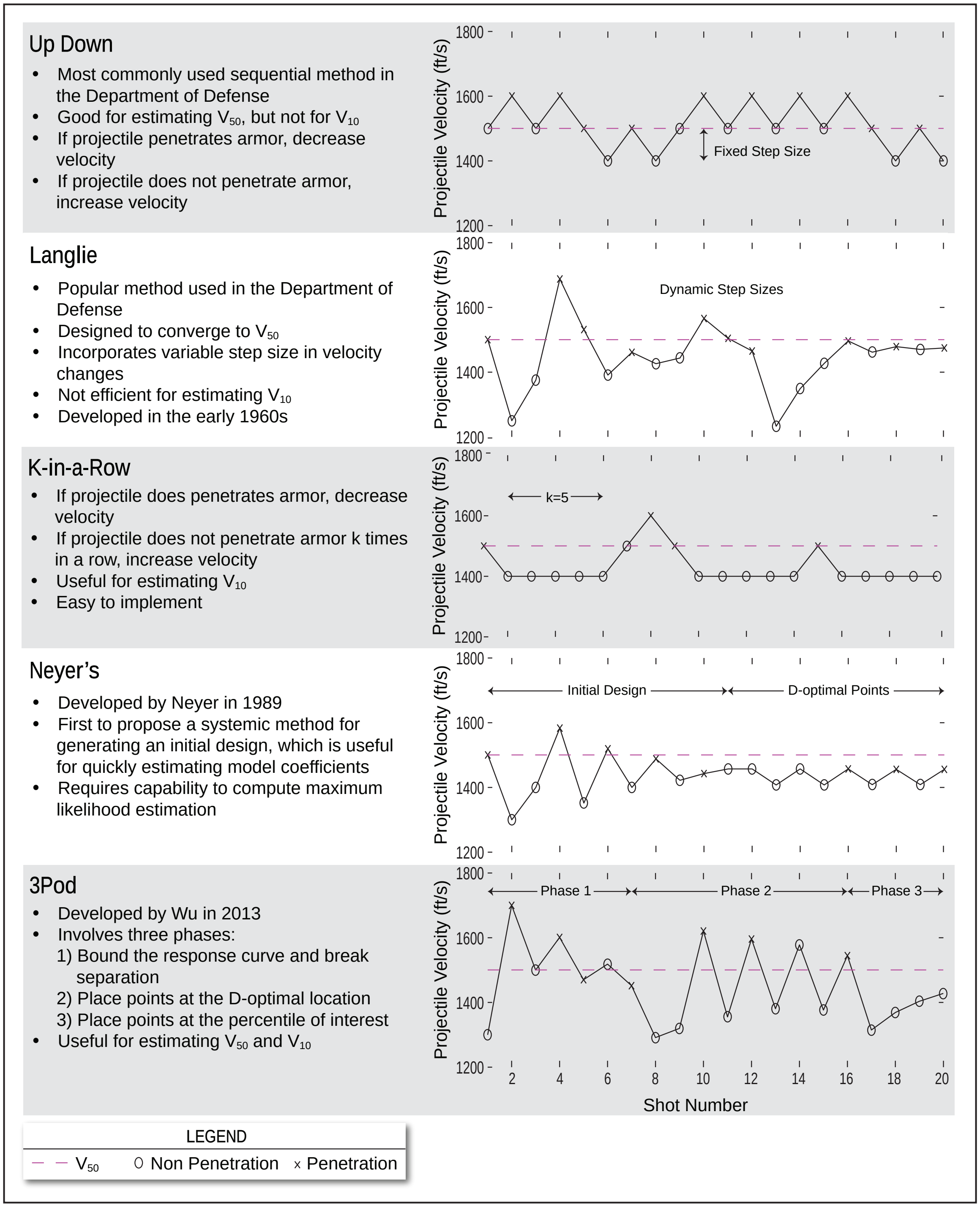


Figure 2. Example Tests for Sequential Methods

in improved estimates and is robust to poor starting estimates. Note that we intentionally consider cases where the mean and variance of the data are incorrect to represent test cases where there is poor understanding of the armors true performance. The response variables are the median and

Table 1. Factors and Levels

Sequential Method	Up and Down (UD), Langlie (LM), Delayed Robbins Monroe (DRM), Three Pod (3Pod), Neyer's (NM), Robbins Monroe Joseph (RMJ), K-in-a-row (KR)
Sample Size	N=20, N=40
σT	75 ft/s, 150 ft/s
σG/σT	1/3,1/2,2,3
µG	µT-2σT, µT, µT+2σT

interquartile range of the V_{50} and V_{10} bias.

The full factorial experiments consist of 336 trials. One thousand simulations are executed per trial. A simulation is representative of a single live fire test, consisting of either 20 or 40 sequentially fired projectiles. After each simulation, the V_{50} bias is calculated as the difference between the assumed "true" V_{50} and the V_{50} estimated from the simulation. V_{10} is calculated similarly. The median and interquartile range of the V_{50} and V_{10} bias are the response variables for each factorial trial.

Results

Figure 3 shows the median and interquartile range of the V_{50} and V_{10} errors from the 1,000 simulation runs. The figure illustrates that RMJ and DRM reduce the V_{10} median bias more than the other sequential methods. 3Pod is the next best performing, followed by KR, LM, NM, and finally UD. The advantage of reduced V_{10} bias by RMJ comes at the expense of V_{50} bias. Figure 3 also shows that the V_{10} median error is bias interaction between KR and σ_G when σ_G is equal to 2 in one direction for all methods except DRM and RMJ. This result occurs because the other sequential methods place runs closer to V_{50}, thereby biasing the V_{10} estimate closer to V_{50}. This result is magnified for LM, NM, and UD, since these sequential methods place runs closer to V_{50} by design.

Figure 4 shows the results of effect screening, which is a more robust way of understanding the results from Figure 3. Effect screening is an efficient way to summarize and compare the results of highly dimensional factorial experiments. The effects show the impacts of the factors, and interactions between factors, on the response variables. The effects are calculated by regressing each simulation outcome (median V_{50} and V_{10} error) on the factors of the factorial experiment. Coefficient estimates are shown for all main effects and two-factor interactions. The intercept of the regression model is the grand mean of the response variable, shown in the bottom left of the effects plot. The coefficient of a particular level of a factor describes the difference between the grand mean and the average response at that level.

Figure 4a shows that the effect that had the largest detrimental effect on the V_{10} error was the interaction between KR and σ_G when σ_G is equal to 2. In that case, KR is unable to converge

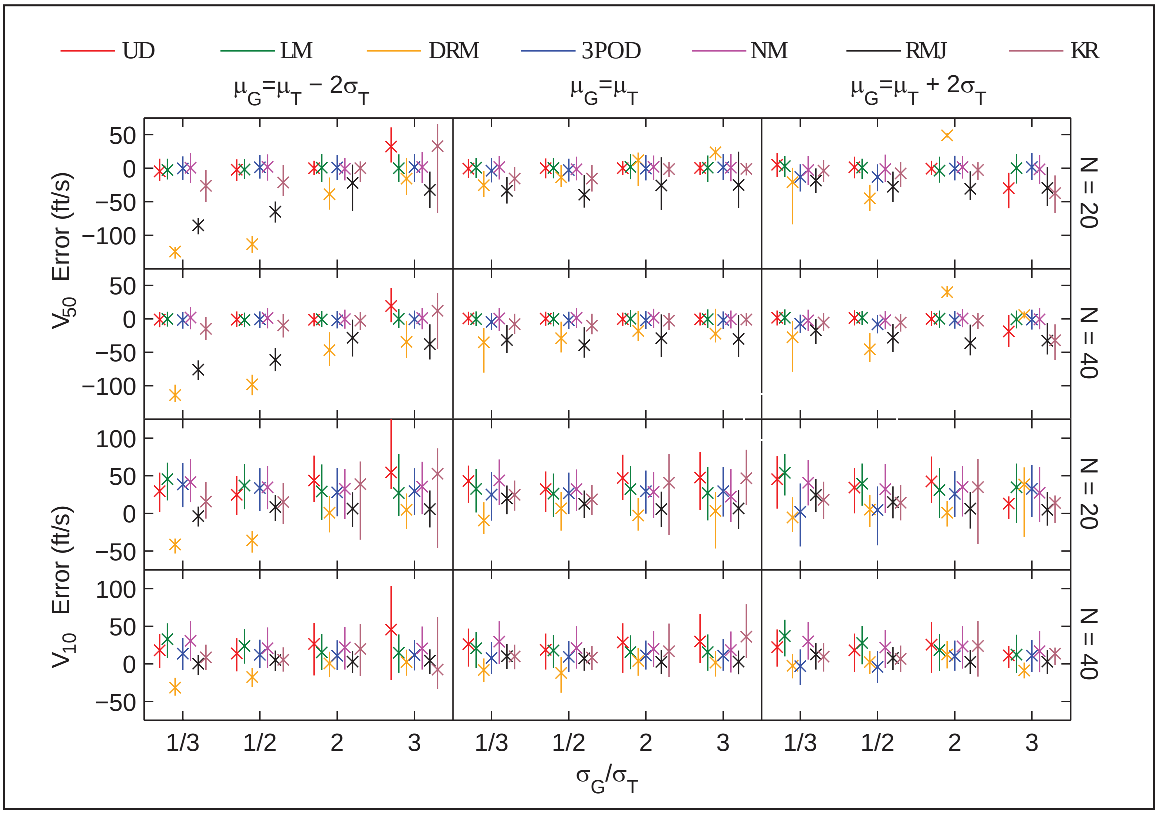


Figure 3. Simulation Outputs

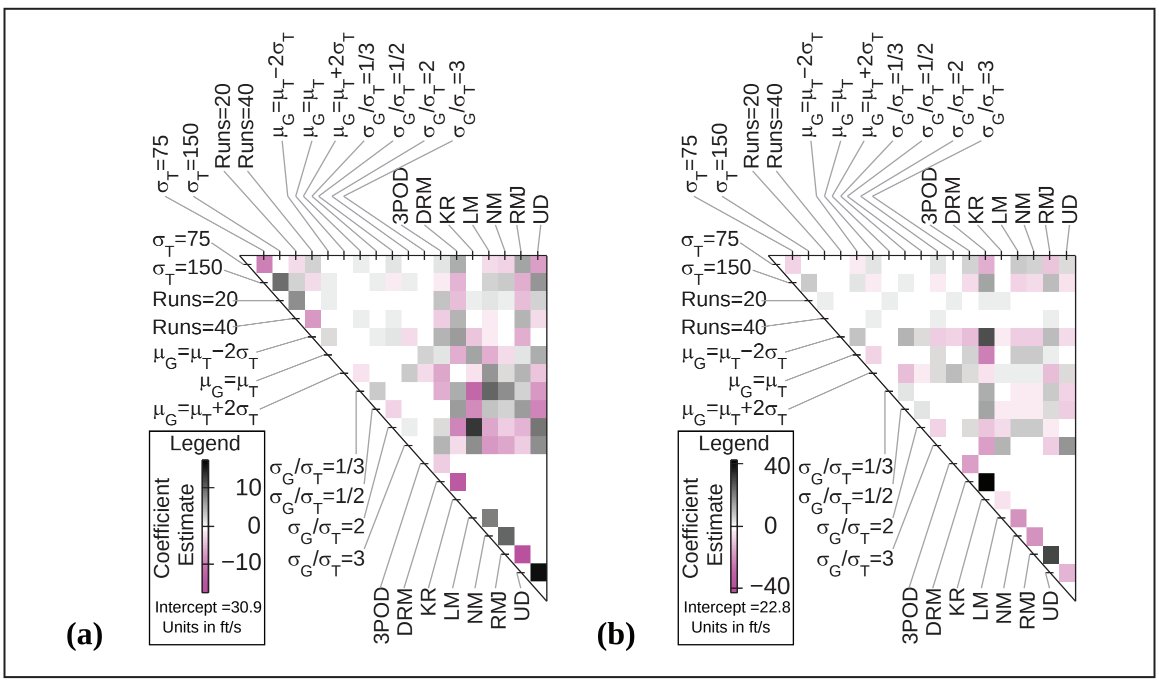


Figure 4. Screening Results for (a) Magnitude of Median V_{10} Error (b) Magnitude of Median V_{50} Error

to V_{10} because of its large step size. UD and KR seem to suffer estimation inaccuracies from interactions, more than the other sequential methods. Meanwhile, 3Pod and NM appear to be most robust.

Figure 4b shows that the three best sequential methods for reducing V_{50} bias are LM, NM, and 3Pod. RMJ and DRM yield the worst V_{50} bias. This result is not surprising because DRM and RMJ forgo initial designs and do not place points near V_{50}. Meanwhile, 3Pod and NM employ initial designs and D-optimal selection criteria that balance the design space.

In general, we found that the methods compared in this study perform commensurate with the goal of the test design. The top three sequential methods that reduce V_{10} bias are, in descending order, RMJ, DRM, and 3Pod. However, 3Pod is more robust to incorrectly specified values of μ_G and σ_G / σ_T than DRM. We also noted that DRM performs erratically for tests with greater than 20 samples because its step size becomes smaller than the velocity set point error. UD, LM, 3Pod, and NM resulted in the lowest bias on V_{50}. The 3Pod method appears to be the most robust method of estimating multiple quantiles.

Conclusions

The DoD uses sensitivity experiments to assess the ballistic resistance of various types of armor. We have shown that employing more recent sensitivity test design methods such as 3Pod and Neyer's Method can lead to improved testing efficiency, increased accuracy, and supports estimation of the entire response curve. Use of these new methods requires that the test community perform real-time statistical analysis of the data during test to select sequential test shots. We have also demonstrated the advantage of using maximum likelihood estimation and generalized linear models in the analysis and execution of ballistic limit testing. Maximum likelihood estimation techniques permit generation of the full perforation response curve, providing more information for the same test resources.

Dr. Johnson *is a Research Staff Member in IDA's Operational Evaluation Division. He holds a Doctor of Philosophy in aerospace engineering from Old Dominion University.*

Dr. Freeman *is an Assistant Director in IDA's Operational Evaluation Division. She holds a Doctor of Philosophy in statistics from Virginia Polytechnic Institute and State University.*

Dr. Hester *is an Adjunct Research Staff Member in IDA's Operational Evaluation Division. She holds a Doctor of Philosophy in physics from Princeton University.*

Dr. Bell *is a Research Staff Member in IDA's Operational Evaluation Division. He holds a Doctor of Philosophy in materials science and engineering from the University of Illinois at Urbana-Champaign.*

The original article was published in the *Institute of Electrical and Electronics Engineers (IEEE) Access*, December 2014.

"A Comparison of Ballistic Resistance Testing Techniques in the Department of Defense"

http://dx.doi.org/10.1109/ACCESS.2014.2377633

References

Collins, J. C. *Quantile response: Practical Sensitivity Testing*. Technical Report, Adelphi, MD: US Army Research Lab, 2012.

Dixon, W. J., and A.M. Mood. "A Method for Obtaining and Analyzing Sensitivity Data." (Journal of American Statistisical Association) 43, no. 241 (March 1948): 109-126.

Gezmu, M. *The Geometric Up-and-Down Design for Allocating Dosage Levels.* Ph.D. Dissertation, Department of Mathematics and Statistics, Washington D.C.: American University, 1996.

Hodges, J.L., and E.L. Lehmann. "Two Approximations to the Robbins-Monroe Process." *Proc. Third Berkeley Symp. 1 (1956).*

Langlie, H.J. *A Reliability Test Method for 'One-Shot' Items*. Ford Motor Company, 1962.

Neyer, B. "A D-Optimality-Based Sensitivity Test." *Technometrics* 36, no. 1 (February 1994): 61-70.

V_{50} Ballistic Test for Armor. Department of Defense Test Method Standard, 1997.

Wu, C.F.J., and Y. Tian. "Three-Phase Optimal Design of Sensitivity Experiments." *Journal of Statistical Planning and Inference* 149 (June 2014): 1-15.

Wu, C.F.J., and Y. Tian. "Three-Phase Sequential Design for Sensitivity Experiments." (Technical report, Georgia Institute of Technology) 2013.