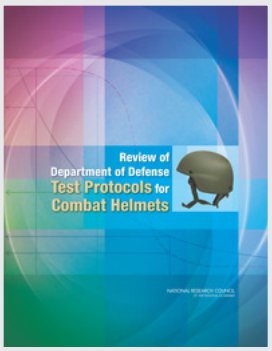
The National Academies of SCIENCES ENGINEERING MEDICINE
THE NATIONAL ACADEMIES PRESS

This PDF is available at http://nap.edu/18621

SHARE



Review of Department of Defense Test Protocols for Combat Helmets

DETAILS

158 pages | 8.5 x 11 | HARDBACK
ISBN 978-0-309-38744-6 | DOI 10.17226/18621

GET THIS BOOK

FIND RELATED TITLES

CONTRIBUTORS

Committee on Review of Test Protocols Used by the DoD to Test Combat Helmets; Board on Army Science and Technology; Division on Engineering and Physical Sciences; National Research Council

Visit the National Academies Press at NAP.edu and login or register to get:

- Access to free PDF downloads of thousands of scientific reports
- 10% off the price of print titles
- Email or social media notifications of new titles related to your interests
- Special offers and discounts

Distribution, posting, or copying of this PDF is strictly prohibited without written permission of the National Academies Press. (Request Permission) Unless otherwise indicated, all materials in this PDF are copyrighted by the National Academy of Sciences.

Copyright © National Academy of Sciences. All rights reserved.

9

Characterization Tests for the Advanced Combat Helmet and Future Helmets

9.0 SUMMARY

The statement of task to the committee includes the following: "Evaluate the scope of characterization testing relative to the benefit of the information obtained." The term "characterization" is broad and is used in different ways in different contexts. However, the Office of the Director, Operational Test and Evaluation (DOT&E) provided additional information to elaborate on this task. Most of the issues raised by DOT&E that are relevant to this portion of the statement of task are addressed in this chapter. The committee also describes additional characterization tests that are needed. Some of these are intended for future helmet designs. A number of these additional tests have been discussed in earlier chapters and are repeated here because they can be viewed as being related to characterization studies. These include the following: evaluating helmet performance across a broader range of, and more realistic, threats; assessing the effect of aging; understanding the relationship between helmet offsets and helmet protection; and conducting gauge repeatability and reproducibility (R&R) studies to understand the different sources of variation in the test process and possibly providing opportunities to reduce some of the variation. This chapter also includes a discussion of current V_{50} testing and an alternative methodology as well as a discussion of industrial practices in characterizing process capability.

9.1 INTRODUCTION

The committee's task to "evaluate the scope of characterization testing relative to the benefit of the information obtained" was added after the committee had started its deliberations, apparently in response to issues raised in the Department of Defense (DoD) Inspector General Report (DoD IG, 2013).

Chris Moosmann from DOT&E provided additional information on the task during a presentation to the committee on March 21-22, 2013. He said:

- ACH (Advanced Combat Helmet) characterization was not done prior to release of the helmet test protocol;
- DOT&E and PEO (Program Executive Office) Soldier have committed to characterize ACH helmets;
- DOT&E indicated that the ECH (Enhanced Combat Helmet) would also be characterized;
- DOT&E will use the results of characterization to determine whether any changes to current protocol standards are appropriate; and
- DOT&E/program offices will consider characterization of new future designs during developmental testing to assess any need for protocol changes.[1]

Mr. Moosmann's presentation noted that the following questions will be addressed as part of the above characterization testing:

1. What is the lower confidence limit (90% confidence) on P(nP) as measured with n shots?
2. What percent of the population (90% confidence) meets the backface deformation (BFD) requirement by location?
3. Do shot location, helmet size, environment, and shot sequence affect P(nP) or BFD?
4. What effect do shot location, helmet size, and shot sequence have on the slope of the ballistic characterization curve?
5. What are the V_0 and V_{50} velocities associated with the fragment simulating projectiles (FSPs) and right circular cylinders (RCCs) currently used during helmet testing?

[1]Chris Moosmann, Live Fire Test & Evaluation, DOT&E, "DOT&E Issues Update," presentation to the committee on March 21, 2013.

Copyright National Academy of Sciences. All rights reserved.

Review of Department of Defense Test Protocols for Combat Helmets

6. What BFDs are associated with FSPs/RCCs currently used during helmet testing?
7. How do helmets perform against foreign threats?[2] (slide 5)

The presentation requested that "the committee review and comment on the scope of characterization testing relative to the benefit of the information obtained and the resources required to do so." In particular,

I. Are there additional questions that should be addressed (threats, conditions, etc.)?
II. Should characterization address issues such as durability and aging ("shelf life")?
III. Should there be a common (minimum) set of questions all characterization efforts should address and what should those include?[3] (slide 6)

The rest of this chapter is aimed at identifying the relevant aspects of characterization, addressing the questions posed by DOT&E, and providing a general discussion of industrial practices involved in studying process capability.

9.2 CHARACTERIZATION OF THE ADVANCED COMBAT HELMET USING EXISTING TEST DATA

For the ACH, existing test data from first article testing (FAT), lot acceptance testing (LAT), and other sources can be used to answer most of the questions posed above by DOT&E. In fact, Question 1 was the subject of Recommendation 6-3 in Chapter 6. It notes that upper confidence bounds (UCBs) should be computed and reported based on the observed number of penetrations in FAT. In addition to characterizing the actual penetration probability, the UCBs can be used to monitor how the penetration levels vary over time and among manufacturers. The same kinds of analyses should also be done with LAT data to monitor a manufacturer's performance over time.

A similar recommendation was made in relation to Question 2 in Chapter 7. Recommendation 7-3 states that the BFD measurements (from FAT) should be analyzed to determine the margins (number of standard deviations between the mean BFD and its threshold) and tracked over time to assess changes. Since the BFD thresholds lack scientific basis, it is better to track changes in the margins or examine the exceedance probabilities at multiple thresholds. It is straightforward to compute the point estimates and associated confidence intervals (or upper bounds) for the exceedance probabilities. Again, similar analyses should be done with LAT data to track a manufacturer over time. Recommendation 7-4 suggests replacing the current ad hoc threshold for BFD (at different locations) using data-based limits obtained from historical BFD test data. Developing such limits can be viewed as a characterization study.

Similarly, existing data for ACH can be used to answer Questions 3 and 4 above. The suite of resistance-to-penetration (RTP)/BFD tests for FAT (see Table 4-1) consists of a designed "full factorial experiment" with three factors: helmet size (Small, Medium, Large, Extra Large), conditioning environment (ambient, hot, and cold temperatures, and seawater), and shot location (front, back, left, right, crown). While the procurement decision rules are based on aggregated data, the full data provide the necessary information to characterize differences among helmet size, shot location, and environment, as specified in Questions 3 and 4 above. In fact, Chapter 5 (Section 5.3) reports some answers to these questions from the committee's analyses of FAT and LAT data that were made available to it. Moreover, the "clustering" analysis already being done by DOT&E and the Institute for Defense Analysis is aimed at characterizing exactly these differences.[4,5]

The current goal of the clustering analysis is to do preliminary tests to see if the data can be pooled across the different factors (environment, locations, etc.), and the committee has noted in Chapter 7 that such preliminary tests are not to be recommended. However, the analyses to estimate the differences among the factors and to monitor them over time (Questions 3 and 4 above) are certainly important and should be continued.

V_{50} testing, raised in Question 5, is discussed in Section 9.4 in this chapter. Regarding Question 6, the committee does not know if data from fragment simulating projectiles (FSPs) and right circular cylinders (RCCs) are stored from past FAT studies for ACH. If they are, Question 6 can also be readily answered.

The issue of testing helmets against other threats has been discussed extensively in the report. The committee will return to this point in Section 9.3.

ACH test data can also be used to characterize many other aspects of helmet performance. For example, FAT and LAT data can be compared over time to find trends and patterns associated with the production process for an individual manufacturer. Data can also be compared across manufacturers to detect possible differences across manufacturers. Further, data from the drop-tests can be used to track performance of manufacturers over time in terms of blunt-force trauma.

9.3 EXPANDED CHARACTERIZATION REQUIRING ADDITIONAL DATA

DOT&E also asked if there were additional topics that should be part of its characterization studies. The committee describes selected topics here. This class of characterization

[2]Ibid.
[3]Ibid.

[4]Janice Hester, Research Staff Member, Institute for Defense Analysis, "DOT&E Helmet Test Protocols Overview: Statistical Considerations and Concerns," presentation to the committee on January 25, 2013.

[5]Laura Freeman, Research Staff Member, Institute for Defense Analysis, "Protocol Analyses and Statistical Issues Related to Testing Methodologies," presentation to the committee on March 21, 2013.

Copyright National Academy of Sciences. All rights reserved.

Review of Department of Defense Test Protocols for Combat Helmets

studies is intended to explore the properties of the helmet beyond the current DOT&E protocol. Several of these suggestions are of a longer-term nature and intended for the ECH and newer generations of helmets rather than the ACH.

- *Evaluate helmet performance for a variety of different threats.* As noted in Chapter 3, the primary focus of DOT&E's (and the Army's) test protocols is gunfire threats. Recommendations 3-1, 3-2, and 3-5 emphasize the importance of expanding the test profile to cover emerging threats as well as more realistic blunt-impact threats. For example, improvised explosive devices (IEDs) have dramatically different distributions of fragment sizes and velocities compared to those from artillery. Recommendation 3-3 asks DoD to reassess helmet requirements for current and potential future fragment threats, especially those energized by blast. Such a reassessment would include examining redundancy in the current profile of threats, such as the 2-grain versus 4-grain, and may lead to elimination of some tests. Resources can then be redistributed to cover a wider range of realistic ballistic threats, including larger mass artillery fragments, bullets other than 9-mm, and IED fragments. A comprehensive examination of threat profiles would involve considerable additional resources and consist of much more than characterization studies. Nevertheless, the committee believes that this is a very important direction for future efforts by DoD.
- *Evaluate the sources of variation in the test process.* As noted in Chapter 4, there are many sources of variation in the test process and test measurements. Recommendation 4-2 recommends that the DoD conduct formal gauge R&R studies to understand the different sources of variation (test methodology, helmets, use of clay, headforms, etc.) and use the results to improve the test process. The committee judges that this should be a high priority, given the high costs of testing and the benefits to be gained from such an R&R study.
- *Evaluate helmet performance at selected areas of the helmet not currently tested*. The test protocols do not assess the helmet in some regions, such as edges and around the ear covering. While it may be reasonable to exclude them in the formal test process, it is still of importance to understand the range of protection afforded at these helmet locations. Potential differences in manufacturing choices could be better understood and might lead to improvements in overall design.
- *Evaluate performance for different helmet pad configurations*. Current testing procedures test the five locations with padding directly in the line of fire of the shot (crown, front, and back) or in a gap between pads (left and right). Anecdotal evidence suggests that many soldiers change the padding locations or remove some of the pads from their helmets in the field. Understanding the differences between testing results and what would be experienced by the soldier would help quantify relevance of the testing. One option for such a characterization study would be to obtain samples of common pad configurations in the field and perform the standard RTP and BFD testing. This would allow better connection of results to soldier experience and may suggest additional recommendations or requirements for soldiers.
- *Evaluate the relationship between helmet offsets and helmet protection*. With the availability of 5 headform sizes, it should be straightforward to characterize differences in BFD by location as a function of helmet offset. It is widely assumed that increased offset provides improved protection through reduced BFD magnitude. (However, Figure 5-3 in Chapter 5 shows that this may not be the case.) Quantifying this improvement, if it exists, could lead to changes to helmet assignment or a reassessment of the tradeoffs between functionality and protection.
- *Evaluate the aging characteristics of the helmets to determine if there is any meaningful degradation of the protection performance of the helmets over time.* An approach to this testing might be to store some of the helmets from a given lot and perform a test similar to FAT testing on helmets of different ages. For example, if helmets were generally thought to be used for 2 years before they were replaced, then a testing regimen could be established that tests helmets at ages 0, 6, 12, 18, 24, and 30 months to determine if there are changes in protection performance. An alternative would be to develop an accelerated testing program in which the helmets are exposed to stressful environmental or to use conditions that would simulate accelerated aging. This testing would provide reassurances that the helmets are not degrading over time.

Program and oversight personnel can identify other potentially important characterization tests that would provide additional information about a helmet's protective capabilities. DOT&E's charge to the committee specifically asked for an evaluation of "the scope of characterization testing relative to the information obtained." The committee does not have the necessary information or the expertise to do a cost-benefit analysis. On the other hand, the Department of Defense has the relevant expertise and information as to which information is important for soldier safety in the battlefield. DoD is better equipped to make the decision on which tests should be done, how to fund them, and whether funds should be redistributed from current test resources for important characterization tests.

Copyright National Academy of Sciences. All rights reserved.

Review of Department of Defense Test Protocols for Combat Helmets

Chris Moosmann's presentation to the committee[6] listed some possible studies that are being planned to characterize the ACH (from different vendors) and compare its performance with the lightweight ACH. If the ACH will no longer be procured (only current manufacturers who have passed FAT will produce them), then it is not wise to invest considerable additional resources to characterize the ACH. New tests and characterization studies should focus on new helmet designs.

When DoD adopts new helmets with changes to the design (such as lighter weight and added mobility), it will be necessary to reevaluate the test protocols. For example, it may not be possible for manufacturers to produce lighter helmets at current levels of penetration.

Recommendation 9-1. When combat helmets with new designs are introduced, the Department of Defense should conduct appropriate characterization studies and cost-benefit analyses to evaluate the design changes before making decisions. It is not advisable to automatically apply the same standard (such as the 90/90 rule or others) when these tests could potentially be across different protective equipment (body armor, helmets, etc.), different numbers of tests (e.g., 96 tests for the enhanced combat helmet, 240 tests for the advanced combat helmet), or over time.

9.4 V_{50} TESTING

Description

V_{50} refers to the "the velocity at which complete penetration and incomplete penetration are equally likely to occur" (DoD, 1997, p. 3). That is, V_{50} is the median of the velocity-penetration distribution or curve. (This is analogous to dose-response studies that arise in pharmaceutical studies.) This theoretical quantity is currently estimated from a series of ballistic tests using the methodology of Military Standard (MIL-STD) 662F (DoD, 1997).

V_{50} testing is an important component of the overall DOT&E protocol. The estimated value of V_{50} is used informally to track and compare helmet performance. The nature of the test suite and the subsequent data analysis are quite different from the RTP and BFD protocols. For these reasons, the committee considers V_{50} testing to be a part of characterization.

Table 4-1 (in Chapter 4) shows the test matrix and requirements for V_{50} testing under DOT&E's FAT protocol. It is performed for 2-, 4-, 16-, 17-, and 64-grain threats as well as a small arms threat (if required). The Army's lightweight ACH Purchase Description (which also specifies MIL-STD-622) further requires that helmets achieve a minimum V_{50} for each of the fragmentation threats (U.S. Army, 2012).

[6]Chris Moosmann, Live Fire Test & Evaluation, DOT&E, "DOT&E Issues Update," presentation to the committee on March 21, 2013.

The V_{50} testing procedure under MIL-STD 662F is as follows:

- A first round is shot with a striking velocity that is approximately 75 to l00 feet per second (ft/s) above the minimum V_{50} required per specification. (Previous V_{50} testing on comparable helmets could also provide a good starting velocity.)
- If the first round results in a complete penetration, the velocity of the second round is decreased by 50 to 100 ft/s from the velocity of the first round. If it results in no or partial penetration, the velocity is increased by 50 to 100 ft/s.
- In subsequent shots, the velocity is increased or decreased, as applicable, until one partial and one complete penetration is obtained.
- After obtaining at least one partial and one complete penetration, the velocity is increased or decreased in increments of 50 ft/s. Firing is continued until sufficient partial and complete penetrations are obtained to estimate V_{50} by taking the average of the velocities corresponding to an equal number of the highest partial and the lowest complete penetration, as specified in the contract (DoD, 1997, p. 10).[7] Typically 8-14 shots are used.

The committee notes that the protocol allows multiple shots per helmet, but it does not explicitly specify a maximum number of shots or shots per helmet: "If a valid V_{50} cannot be obtained with a single finished shell, the V_{50} will continue on an additional finished shell(s)" (IOP PED 003, Paragraph 5.2.1.1).

Finding 9-1. The current V_{50} testing protocol does not clearly specify the maximum number of shots per helmet.

During the committee's discussions with representatives of PEO Soldier[8] (Lozano, 2013) and DOT&E, the following reasons were given for collecting V_{50}-related data:

- It is a commonly understood metric that characterizes the performance of the helmet, both in the United States and in member countries of the North Atlantic Treaty Organization.
- It is easier to estimate than potentially more relevant velocity quantities such as V_0 or V_{10}.

[7]This estimation methodology is similar to the NATO Standardization Agreement (STANAG) 2920, Ballistic Test Method for Personal Armour Materials and Combat Clothing, promulgated 31 July 2003. STANAG 2920 requires an even number of at least six shots, half of which perforate and half of which do not, and all of which are have velocities of within 40 meters per second. Then the V_{50} is estimated as the mean velocity of the shots meeting these conditions (NSA, 2003).

[8]Frank J. Lozano, Product Manager, Soldier Protective Equipment, "Setting the Specifications for Ballistic Helmets," presentation to the committee on April 25, 2013.

Copyright National Academy of Sciences. All rights reserved.

Review of Department of Defense Test Protocols for Combat Helmets

- It can be useful for comparing helmet performance between manufacturers and over time.
- PEO Soldier uses V_{50} time series data as a leading indicator of manufacturer process degradation.

V_{50} values are used informally. More structured analyses could be done to compare V_{50} estimates among manufacturers, over time, and among environments. Another potential characterization analysis would be to investigate the relationship between V_{50} and fragment grain size.

Additional V_{50} Testing and Characterization Analyses

The current goal of V_{50} testing is to estimate a single point (the median) on the velocity-penetration curve. In the committee's view, it would be beneficial to expand V_{50} testing so that the whole curve can be estimated with reasonable precision, without expending a lot more additional resources in terms of number of shots.

This expanded testing would involve taking multiple shots at different (selected) velocities and fitting a parametric curve to the velocity-penetration response data. Typical choices for the curve are logistic or normal distributions, leading to logit and probit curves, respectively. This approach allows for estimation of any quantile of the velocity-penetration distribution, not just the median. One can also compute the standard error associated with the estimated quantile. There is extensive literature on the design and analysis of such studies (Ruburg, 1995; Prentice, 1976).

The curves are typically described by two parameters for location and shape. The shape parameter provides an indication of the spread in the velocity-penetration distribution. It measures how consistent the penetration velocity is from helmet to helmet or among shot locations within a helmet. Changes in a production process, for example, could either increase or decrease the variability of penetration velocities. Certain environments might not affect V_{50} but could increase the standard deviation and, thereby, degrade a helmet's protective capability.

Recommendation 9-2. The Department of Defense should consider alternative approaches to its current methodology for estimating V_{50}. One alternative is to estimate the entire velocity-penetration distribution by varying the shot velocities over a prescribed range. Given the limited test resources (number of shots), the estimation methodology has to be based on fitting parametric curves. The approach also allows computation of standard errors associated with V_{50} and other quantiles of interest.

9.5 COMPARISON WITH INDUSTRIAL PRACTICES

So far, this chapter has focused on specific issues on characterization related to helmet testing. This section provides a more general discussion of industrial best practices.

Understanding the ability of a current product to conform to production requirements is a common aspect of industrial practice and product improvement and is often called *capability analysis* (Bothe, 1997; Pyzdek and Keller, 2003). It encompasses characterization of process stability as well as margin on performance relative to product requirements (Hoerl and Snee, 2012). It is applicable to understanding product conformance internal to a company and for external suppliers, customers, and users. Typically, formal product requirements such as acceptable failure rates and specification limits are based on understanding customer needs. In the helmet procurement process, this would likely be based on data collected during developmental testing. Developing a stronger connection to what is possible, given current helmet manufacturing capability, would allow the opportunity to leverage this into improved helmets for the soldier. Using legacy measures to define the standard a helmet is required to meet for FAT and LAT represents a lost opportunity and potentially an important sacrifice in helmet protection.

Recommendation 9-3. To be consistent with the goal of continuous improvement, developmental testing results from helmet design should be used to allow better calibration of current helmet capability and to help define more meaningful thresholds for helmet protection.

A key difference in DoD's approach used in the procurement process for helmets from the more common practice of industry is the focus on performance specifications instead of design specifications. In much of industry, and indeed for some military procurement processes involving complex products and systems, when a product is being developed, design specifications for material, structure, and assembly are the basis for assessing its adequacy. In other words, the manufacturing process is closely monitored and checked to make sure that the product matches the details for what is required. This provides a direct and easily measurable means of checking new products as they are completed.

On the other hand, the current DoD helmet procurement process allows manufacturers to build the helmet with any design specifications, and the sole test of the adequacy of the helmet is through performance tests during FAT and LAT testing. An advantage of this approach is that it allows the manufacturers the flexibility to change the process and update their production methods as technology evolves. However, it has the disadvantage of placing all of the burden for evaluation at the end of the production process through rigorous and expensive testing.

A potentially beneficial alternative—one that would encourage improved process monitoring while still allowing manufacturers flexibility to improve their product as new technologies are developed—would be to combine the design and performance specification approaches. Manufacturers could develop their own design specifications, which would then be tracked with reports given to the DOT&E.

Copyright National Academy of Sciences. All rights reserved.

Review of Department of Defense Test Protocols for Combat Helmets

This information would then be used to complement the performance-based testing currently used, particularly at the LAT testing stage. This additional information would allow DOT&E to have better understanding of the stability of the process, while having the reassurances of the performance-based testing.

Once the design specification requirements have been determined by the manufacturer, then the capability of the currently available product can be quantified using one of the common process capability metrics (Montgomery, 2012). In the absence of formally specified requirements, matching or surpassing current production capability is a common alternative for capability analysis methods. Characterizing product performance is an established practice in industry and is used to quantify current performance as well as establish a baseline from which target future improvements can be assessed.

The standard approach to monitoring stability of production is through control charts based on manufacturing characteristics (Hoerl and Snee, 2012), that allow for continuous supervision and monitoring of standards as products are being produced. Supervision and monitoring involve active management and watching real-time results to see if there is a problem. Current FAT and LAT testing is based on a paradigm of inspection, in which during post-production the products are evaluated to assess conformance. Standard practice in industry has evolved away from primarily using inspection to a model in which monitoring is a key aspect of ensuring ongoing product quality. Monitoring has the advantages of ensuring that a production process operates at its full potential, reducing waste, and detecting changes in performance quickly.

Recommendation 9-4. Manufacturers should be required to provide some documentation of ongoing process monitoring of the helmet production as a beneficial enhancement to the lot acceptance testing protocol.

9.6 CONCLUDING REMARKS

It is for DoD to choose the appropriate characterization tests and analyses that should be done, based on its assessment of the benefits, in terms of improving the understanding of helmet protective properties and improving those capabilities, relative to the costs and resources they require. A number of the proposed characterization studies can be done using data that are collected as part of the FAT and LAT test process. Others will require different types of testing and the investment of additional resources.

Recommendation 9-5. For new generations of helmets, the scope of characterization studies should be broader than what is currently being done. They should include many of the activities described in Section 9.3.

9.7 REFERENCES

Bothe, D.R. 1997. Measuring Process Capability: Techniques and Calculations for Quality and Manufacturing Engineers. McGraw-Hill, New York, N.Y.

DoD (Department of Defense). 1997. Department of Defense Test Method Standard: V_{50} Ballistic Test for Armor. MIL-STD-662F. U.S. Army Research Laboratory, Aberdeen Proving Ground, Md.

DoD IG (Department of Defense Inspector General). 2013. Advanced Combat Helmet Technical Assessment. DODIG-2013-079. Department of Defense, Washington, D.C.

Hoerl, R.W., and R.D. Snee. 2012. Statistical Thinking: Improving Business Performance. Wiley, Hoboken, N.J.

Lozano, F., Product 9 Manager, Soldier Protective Equipment, U.S. Army. 2013. V_{50} Ballistic Limit Testing. Information paper. June 18, 2013. U.S. Army, Fort Belvoir, Va.

Montgomery, D.C. 2012. Introduction to Statistical Quality Control. Wiley, Hoboken, N.J.

NSA (NATO Standardization Agency). 2003. Ballistic Test Method for Personal Armour Materials and Combat Clothing. NSA/0723-PPS-2920. STANAG 2920 PPS–Edition 2. NATO Standardization Agency, Brussels, Belgium.

Prentice, R.L. 1976. Generalization of the probit and logit methods for dose response curves. Biometrika 32:761-768.

Pyzdek, T., and P.A. Keller. 2003. Quality Engineering Handbook. CRC Press, Boca Raton, Fla.

Ruburg, S.J. 1995. Dose response studies I. Some design considerations. Journal of Biopharmaceutical Statistics 5(1):1-14.

U.S. Army. 2012. Advanced Combat Helmet (ACH) Purchase Description, Rev A with Change 4. AR/PD 10-02. Soldier Equipment, Program Executive Office—Soldier, Fort Belvoir, Va.

Copyright National Academy of Sciences. All rights reserved.