G A O
Accountability * Integrity * Reliability

Comptroller General
of the United States

United States Government Accountability Office
Washington, DC 20548

Decision

Matter of: Ceradyne, Inc.

File: B-406479.2

Date: August 15, 2012

Michael A. Hordell, Esq., Michael R. Golden, Esq., Heather Kilgore Weiner, Esq., and Samuel W. Jack, Esq., Pepper Hamilton LLP, for the protester.
Shantay N. Clarke, Esq., Defense Logistics Agency, for the agency.
Jennifer D. Westfall-McGrail, Esq., and Edward Goldstein, Esq., Office of the General Counsel, GAO, participated in the preparation of the decision.

DIGEST

Protest is denied where protester fails to demonstrate that agency's statistical testing methodology unreasonably restricts competition.

DECISION

Ceradyne, Inc., of Costa Mesa, California, protests the terms of request for proposals (RFP) No. SPM1C1-11-R-0118, issued by the Defense Logistics Agency for lightweight advanced combat helmets. The protester contends that the agency's methodology for evaluating the results of required ballistic testing of the helmets is unreasonable and restrictive of competition.

We deny the protest.

The RFP, which was issued on November 28, 2011, contemplates the award of 3 indefinite-delivery/indefinite-quantity (IDIQ) contracts (for a 24-month base period) to the 3 lowest-priced offerors of technically acceptable proposals. The only factor to be considered in determining the technical acceptability is whether the offeror has furnished passing ballistic test reports--that is, test reports demonstrating that the offeror will be able to meet the ballistic requirements for first article testing. The closing date for receipt of proposals was July 19, 2012.[1]

[1] At the time Ceradyne filed its protest with our Office on May 9, 2012, the closing date was set for May 30; the agency subsequently amended the RFP to extend the closing date to July 19. RFP, amend. No. 22.

The ballistic test requirements are set out in Army Purchase Description 10-2 Revision A with Change 4, May 8, 2012, (AR/PD) which was incorporated in the RFP by amendment No. 19. The purchase description provides for the testing of each size of helmet (small, medium, large and x-large) in each of 4 environmental conditions: ambient temperature (68° F), hot (160° F), cold (-60° F), and seawater (tested at ambient temperature after submersion in the water). Testing is conducted by placing a helmet on a clay-filled headform in an "as worn" position and firing a 9 mm full metal jacket round nose bullet at the helmet. Only one size of headform is used for the testing, and the testing procedures specify required clearances between the helmets and the headform that vary depending on the size of the helmet.[2] The helmet is evaluated for resistance to penetration (RTP) and ballistic transient deformation (BTD), which is the imprint left in the clay as the helmet bulges in response to a shot. It is the latter test that is at issue in this protest.

Each helmet is divided into 5 sections--front, back, right, left, and crown--and during testing, each helmet is shot 5 times, once in each section. The order in which the shots are taken varies in accordance with a pre-established matrix. After each shot, the helmet is removed from the headform and the imprint in the clay is measured. The clay is then repaired and the helmet remounted on the headform. The measurements are compiled and used to calculate an upper tolerance limit (UTL) for each section--that is, separate UTLs are calculated for the front, back, crown,

[2] Internal Operating Procedure (IOP) PED-003 (Helmet Ballistic Testing Procedures), issued by the U.S. Army Aberdeen Test Center, defines the inner crown shell to top of crown clay stand-off distances as follows:

Helmet Size	Distance in mm
S	36.0 ± 0.5
M	29.0 ± 0.5
L	23.0 ± 0.5
XL	22.0 ± 0.5

IOP-PED 003, at 51, Agency Report (AR), Exh. 3. The IOP explains that setting the crown stand-off distances increases the repeatability of the testing and helps to ensure consistent mounting of the helmet on the headform.

left/right sides of the helmet.[3] The UTL calculation for each section includes all of the test results pertaining to that section--i.e., it includes the results for all sizes of helmet in all environmental conditions For the BTD testing, the statistical analysis methodology is 90% UTL with 90% Confidence. Id. at 74. The agency explains that this statistical methodology was implemented to provide the agency 90% confidence that at least 90% of the BTD measurements will be at or below the UTL requirement. Contracting Officer's Report at 12. The purchase description requires that the BTD not result in a 90% UTL with 90% confidence calculation in excess of 16 mm for shots made to the right side, left side, and crown, and that it not result in a 90% UTL with 90% confidence calculation in excess of 25.4 mm for shots to the front and back. AR/PD at 15. The purchase description further provides that for acceptance, all of the 90% UTL with 90% confidence calculations must comply with the foregoing requirements. Id. at 75.

The solicitation, as issued, gave offerors the alternative of submitting with their proposals the results of ballistics testing on a limited quantity of helmets in lieu of the full quantity required for first article testing (FAT). Under the limited quantity qualification testing, offerors were required to furnish the results of testing on one helmet in each size in each environmental condition (i.e.,16 shots to each of the 5 helmet sections, or 80 shots total). Under the full testing requirements for FAT, in contrast, testing on 3 helmets in each size in each environmental condition is required (i.e., 48 shots to each of the 5 helmet sections, or 240 shots total).[4]

Ceradyne maintains that the results of its limited quantity testing show that the BTDs resulting from shots to the crown of the small helmets are substantially smaller than the BTDs resulting from shots to the crowns of helmets of other sizes.[5] The protester attributes this to the fact that, as noted above, the testing procedures call for the small helmets to have greater clearance from the headform when positioned for ballistic testing. The protester contends that inclusion of the low data points resulting from testing of the small helmets in the same sample with the higher

---

[3] The AR/PD further provides that if the BTD measurements for the right and left sides form two distinct distributions, separate BTD UTL calculations will be conducted for each side. AR/PD at 15 and 75.

[4] In amendment No. 22, issued on June 4, 2012, the agency amended the RFP to require offerors performing their testing after the date of issuance of the amendment to conduct full BTD testing.

[5] The protester has furnished test results that support the foregoing assertion--that is, its results show BTDs of [deleted] for shots to the crown of small helmets, in contrast to results of [deleted] for shots to the crown of medium helmets, [deleted] for shots to the crown of large helmets, and [deleted] for shots to the crown of x-large helmets. Protester's Comments, Exh. 1-4.

data points resulting from testing of the medium, large, and x-large helmets results in a high standard deviation (i.e., [deleted]), and that the high standard deviation in turn results in a UTL exceeding the limit of 16 mm.[6] That is, the protester maintains that despite an average BTD well below the allowable UTL and no individual results exceeding the allowable UTL of 16 mm, its helmets will fail BTD testing for the crown location simply because the results for its small helmets are so much lower than the results for the larger helmets. In support of its position, the protester cites the following crown shot results from its recent qualification testing:

BTD Average			**BTD Standard Deviation**			**90/90 UTL**		
S only	M, L & XL	All sizes	S only	M, L & XL	All Sizes	S only	M, L, & XL	All Sizes
[deleted]	[deleted]	[deleted]	[deleted]	[deleted]	[deleted]	[deleted]	[deleted]	[deleted]

Protester's Comments, June 21, 2012, at 13. The protester argues that the agency should be required to use a different sized headform for the testing of each different size of helmet (so as to reduce variations in the results attributable to the differing stand-off distances), and/or the agency should separately calculate UTLs for each size of helmet so that the low data points for the small helmets do not inflate the standard deviation. Ceradyne also objects to the agency's randomized shot sequence, arguing that it increases the standard deviation among the data points, which makes it more difficult for an offeror to achieve the required UTLs.

We review testing requirements using the same standard applicable to any other challenge of a solicitation's evaluation procedures; the establishment of testing or qualification procedures or standards is a matter within the technical expertise of the procuring activity, and we will not object to the imposition of certain terms unless they are shown to be without a reasonable basis. RSL Elec. Ltd., B-404117.3 et al., Mar. 28, 2011, 2011 CPD ¶ 90 at 5. A protester's mere disagreement with the agency's judgment concerning its needs and how to accommodate them does not show that the agency's judgment is unreasonable. Id.

First, with regard to the protester's complaint that the agency should be required to use different sized headforms during testing, the agency reports that "varying-sized headforms are not available for testing within the foreseeable future."[7] Agency

[6] UTL is calculated using a statistical formula that takes into consideration, among other things, average result, number of shots, and standard deviation.

[7] The agency also cited as justification for its use of a single size headform that ballistic testing with a single headform was the industry standard and had been accredited by the National Institute of Justice. We are not persuaded by this argument given the absence of evidence that other entities performing testing are performing the same sort of statistical analysis of their results as the agency here.

(continued...)

Supplemental Response, July 6, 2012, at 4. The fact that headforms of varying sizes are not currently available clearly provides a reasonable basis for the agency not to have used them. Moreover, given that the helmets are considered critical safety items, we do not think that the agency should be required to delay the procurement in order to develop them.[8] See TLC Sys., B-277095, Sept. 2, 1997, 97-2 CPD ¶ 61 at 4.

Second, in response to the protester's argument that combining the results for shots to the crown of helmets of differing sizes inflates the standard deviation of the sample, which in turn raises the calculated UTL above the prescribed limit, the agency explains the basis for adoption of the given methodology as follows:

> Prior to the current testing methodology, which is based on a statistical analysis method, helmets were evaluated by individual helmet size on a pass/fail basis. Under the old method, vendors would produce a helmet design that was effective for certain sizes (e.g., medium, large, and extra-large) and risk failures in other sizes where the design was not as effective (e.g., small). [Citation omitted]. By assuming this risk, vendors were effectively "gaming the system" by providing helmet designs that were not the highest quality for all helmet sizes. [Citation omitted] With the implementation of the "test grouping" and the statistical analysis methodologies, vendors are required to show with 90% confidence that their helmet design will fall within the BTL [sic] upper tolerance limit threshold (≤ 16mm/≤ 25.4mm) at least 90% of the time for all helmet sizes. This statistical methodology better tests the overall helmet's design performance as opposed to testing a particular helmet size's performance. Given that the Agency is procuring all sizes from each awardee, this methodology enables the Agency to determine which vendor's helmet design provides the safest and highest quality for all helmet sizes and, therefore, better meets the Government's needs.

Contracting Officer's Report at 17-18.

(...continued)
(Put another way, the protester's complaint regarding disparities in the test results linked to the size of the headform arises only because the agency is evaluating the test results using a statistical analysis that considers helmets of all sizes as part of the same sample.)

[8] We would, however, encourage the agency to consider the development of multiple sizes of headform since it does appear that use of differing sized headforms would increase the reliability of the test data by eliminating variations in BTD size attributable to differences in stand-off distance.

It is clear from the foregoing explanation that the agency's goal in providing for the given statistical analysis methodology was to assure consistency in the quality of protection afforded by helmets of differing sizes. While it is true that the agency's methodology (of including results for all sizes of helmet in the same UTL calculation) can result in an unacceptable UTL calculation where there is a high standard deviation due to a high disparity among shot results, all of which may be passing on an individual basis (e.g., the protester, which has several BTD values for its small helmet that are significantly smaller than the BTD values for its other size helmets), we have no basis to conclude that such a methodology is flawed. Rather, the agency's methodology appears to be a reasonable means of ensuring that there is limited variability in testing between helmets of different sizes and thereby offer similar levels of protection between sizes.[9]

We also find that the limited qualification test results provided by the protester in support of its arguments do not support the protester's contention that calculating the UTL for the crown section of each size of helmet separately yields results below the 16 mm limit. That is, when the UTLs are calculated using the data points furnished by the protester as exhibits to its June 21 comments and the formula for calculation of a UTL explained by the protester in its August 2 response to our Office's inquiry, the result for each size of helmet other than small exceeds the 16 mm limit. Accordingly, the record here does not demonstrate that the protester suffered any prejudice as a result of the agency's failure to calculate a separate UTL for each size of helmet.

Finally, in response to the protester's complaint that the randomized test sequence increases the variability of the BTD results, the agency notes that prior testing of the helmets using a fixed shot sequence "allowed vendors to potentially 'game the system' by adding additional protection, different material, etc. in areas that they knew would be impacted later in the shot sequence." Memorandum from Department of the Army, Product Manager, Soldier Protective Equipment, May 30,

---

[9] We note that the protester does not offer independent support for its premise that the smaller BTD test results for the crown of its small helmet is attributable to the greater stand-off distance used during testing, as opposed to inherent variability in its product. As noted supra at note 2, each size helmet is tested using a different stand-off distance, yet the protester does not identify any BTD variability for any of its other helmet sizes. To the contrary, throughout its protest, the protester calculates its UTL for the M, L, and XL helmets combined, thereby singling out the small helmet. Furthermore, the protester's suggestion that a greater stand-off distance yields a lower BTD is not borne out by its own test results. The mean BTD of its L size helmet [deleted] is higher than the mean for its XL size helmet [deleted], notwithstanding the fact that the L size helmet testing requires a greater stand-off distance.

2012, at 2; AR, Exh. 19. In other words, while the randomized shot sequence may increase the variability of results, it was implemented to overcome a bigger problem, i.e., manufacturers designing the helmets to beat a particular test shot sequence rather than designing the helmet to offer the overall best protection. In our view, it was reasonable for the agency to choose to address what it perceived to be the larger problem. In addition, while the protester asserts that the randomized test sequence increases the variability of the test results, making it more difficult for an offeror to achieve passing results, we were not able to discern from the record any such variability attributable to the shot sequence. In this regard, there is no apparent correlation between position of the shot in the test sequence and size of the BTD in the qualification test results furnished to us by the protester (nor was the protester able to point one out to us when we asked it to explain how the test results cited in its comments supported its position).

The protest is denied.

Lynn H. Gibson
General Counsel