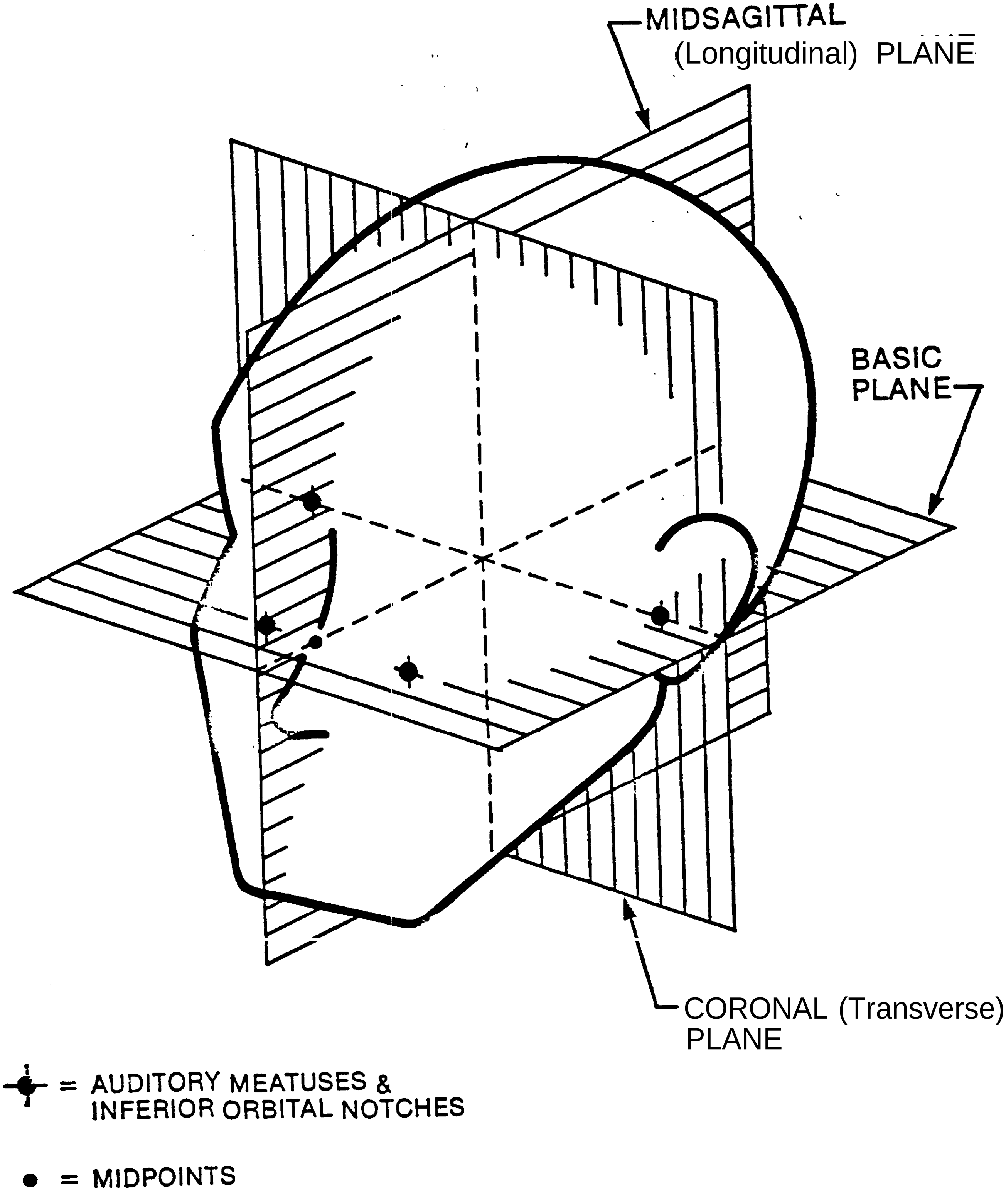


Figure '1. Anatomical Planes

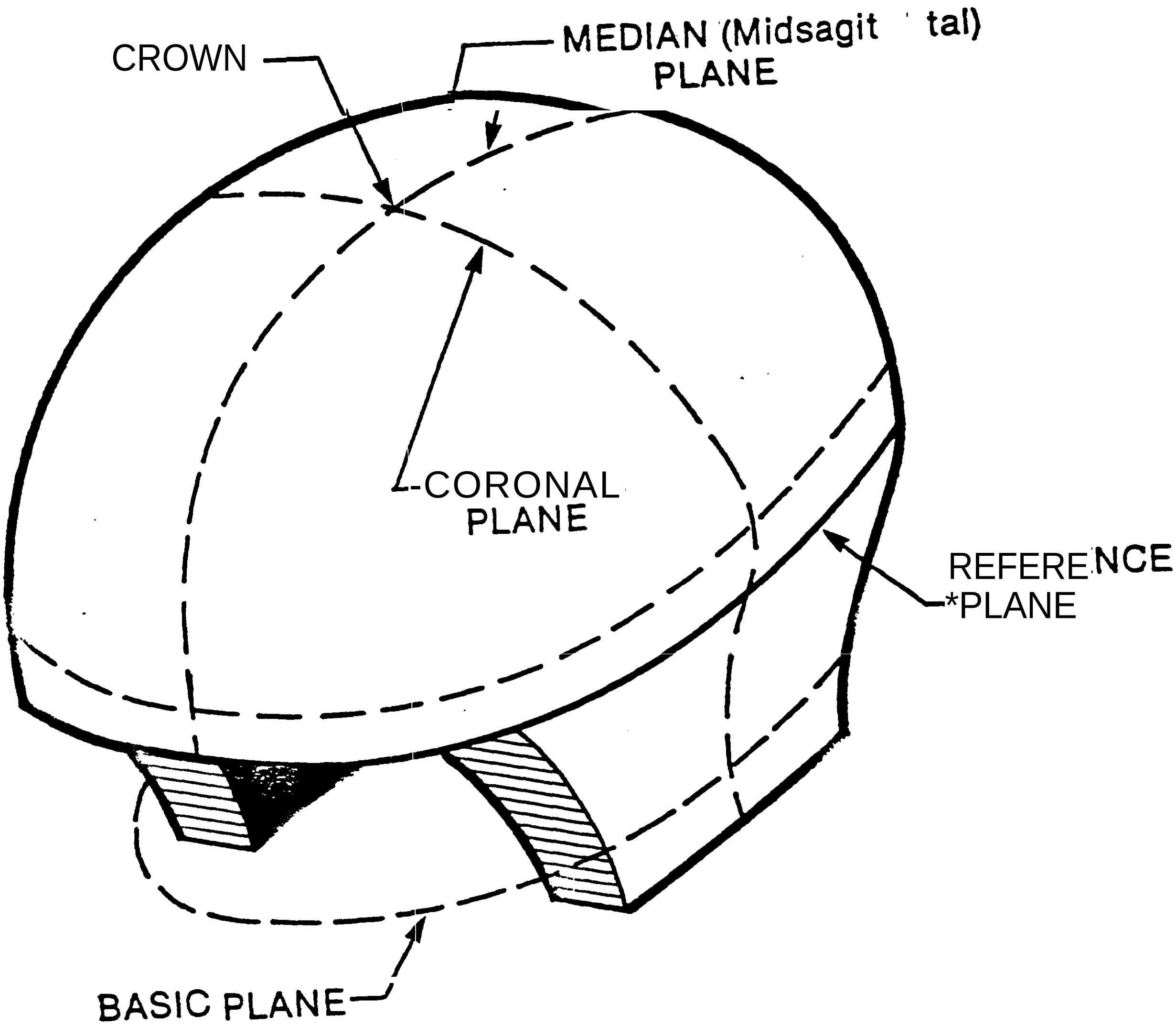
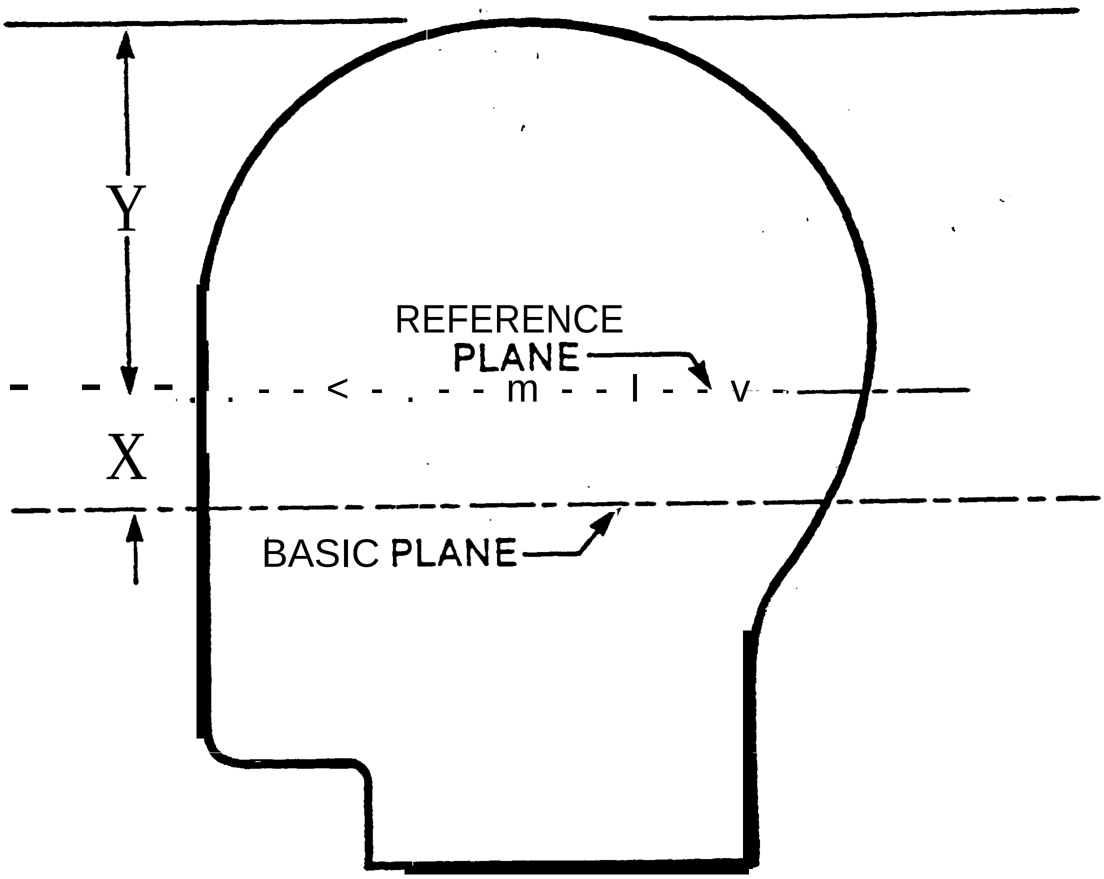


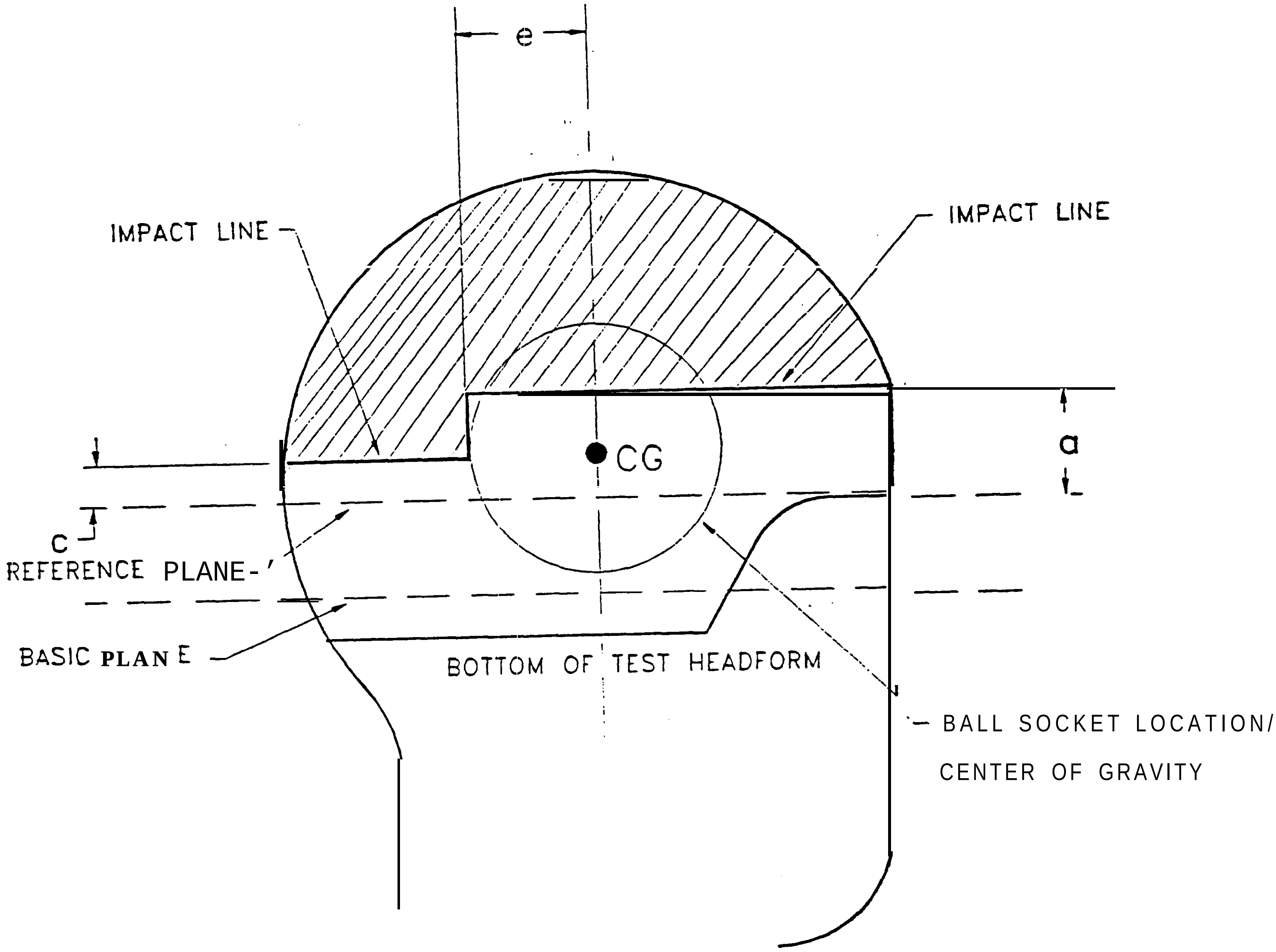
Figure 2. ISO Headform-Basic, Reference, and Median Planes



HEADFORM	SIZE	X	Y
A	500	24	90
E	540	26	96
J	570	27.5	102.5
M	600	29	107
O	620	30	110

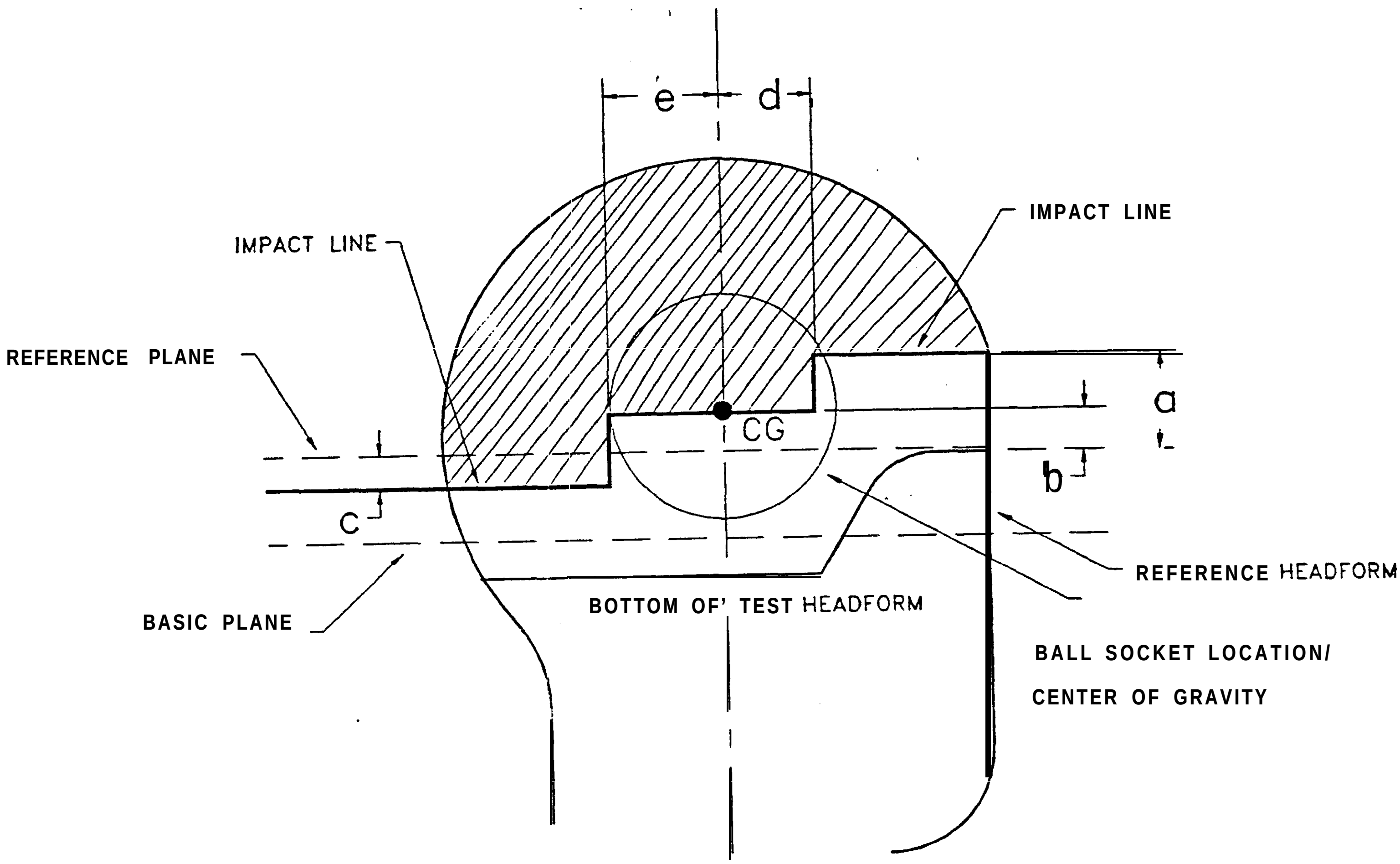
DIMENSIONS IN MILLIMETERS

Figure 3. Location of Reference Plane



HEADFORM	DIMENSIONS mm(in)		
	a	c	e
ISO A	38 (1.49)	27 (1.06)	49 (1.93)
ISO E	39 (1.54)	27 (1.06)	52 (2.05)
ISO J	41 (1.61)	27 (1.06)	54 (2.13)
ISO M	41 (1.61)	27 (1.06)	55 (2.16)
ISO O	42 (1.65)	27 (1.06)	56 (2.20)

Figure 4. Location of Test Lines for Helmets Intended for Persons Five (5) Years of Age and Older.



HEADFORM	DIMENSIONS mm (in)				
	a	b	c	d	e
ISO A	30 (1.18)	12.7 (0.50)	15 (0.59)	25 (0.98)	30 (1.18)
ISO E	32 (1.26)	12.7 (0.50)	16 (0.63)	27 (1.06)	32 (1.26)

Figure 5. Location of Test Lines for Helmets Intended for Persons Ages 1 and Older

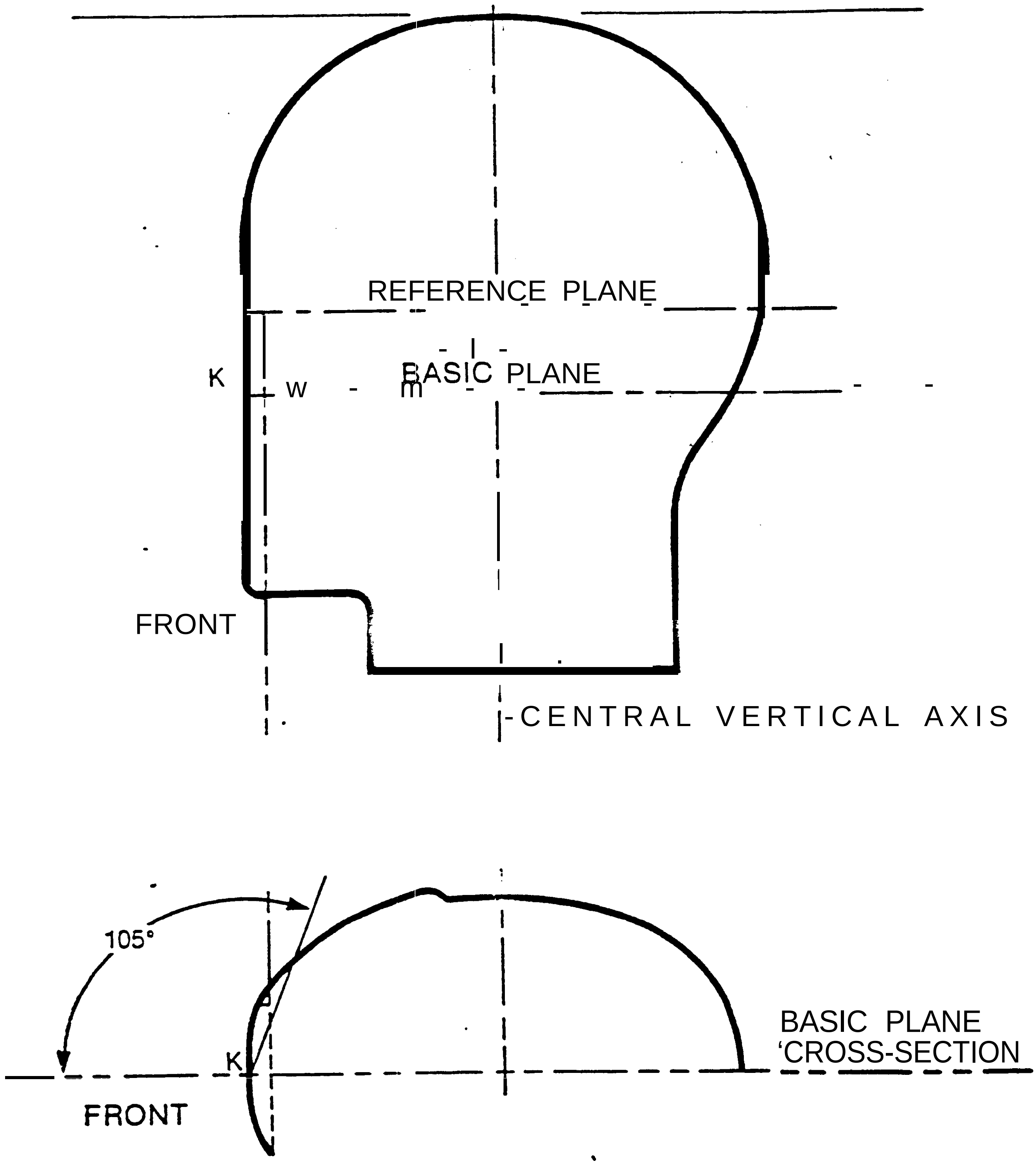


Figure 6. Field of Vision

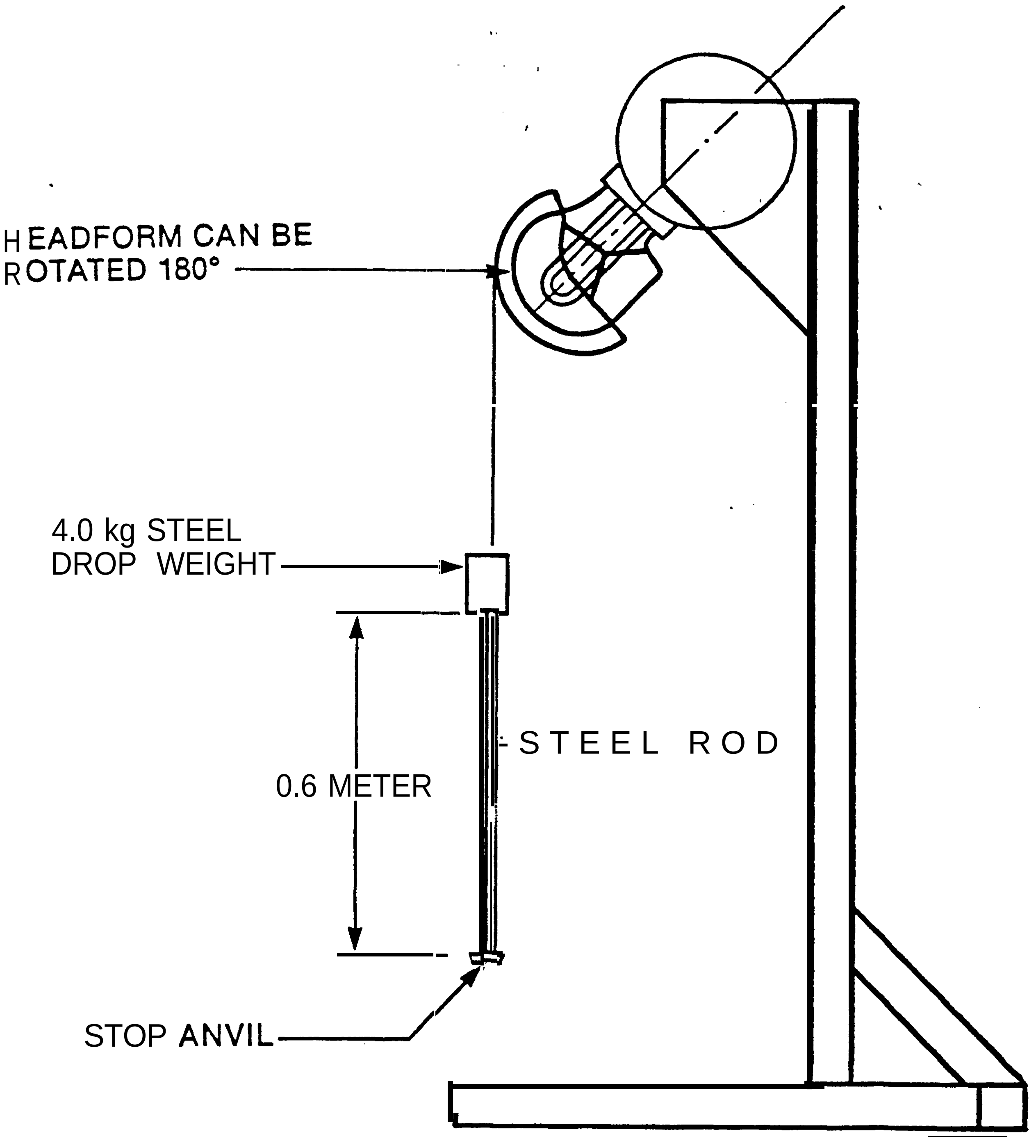


Figure 7. Typical Test Apparatus for Positional Stability Test

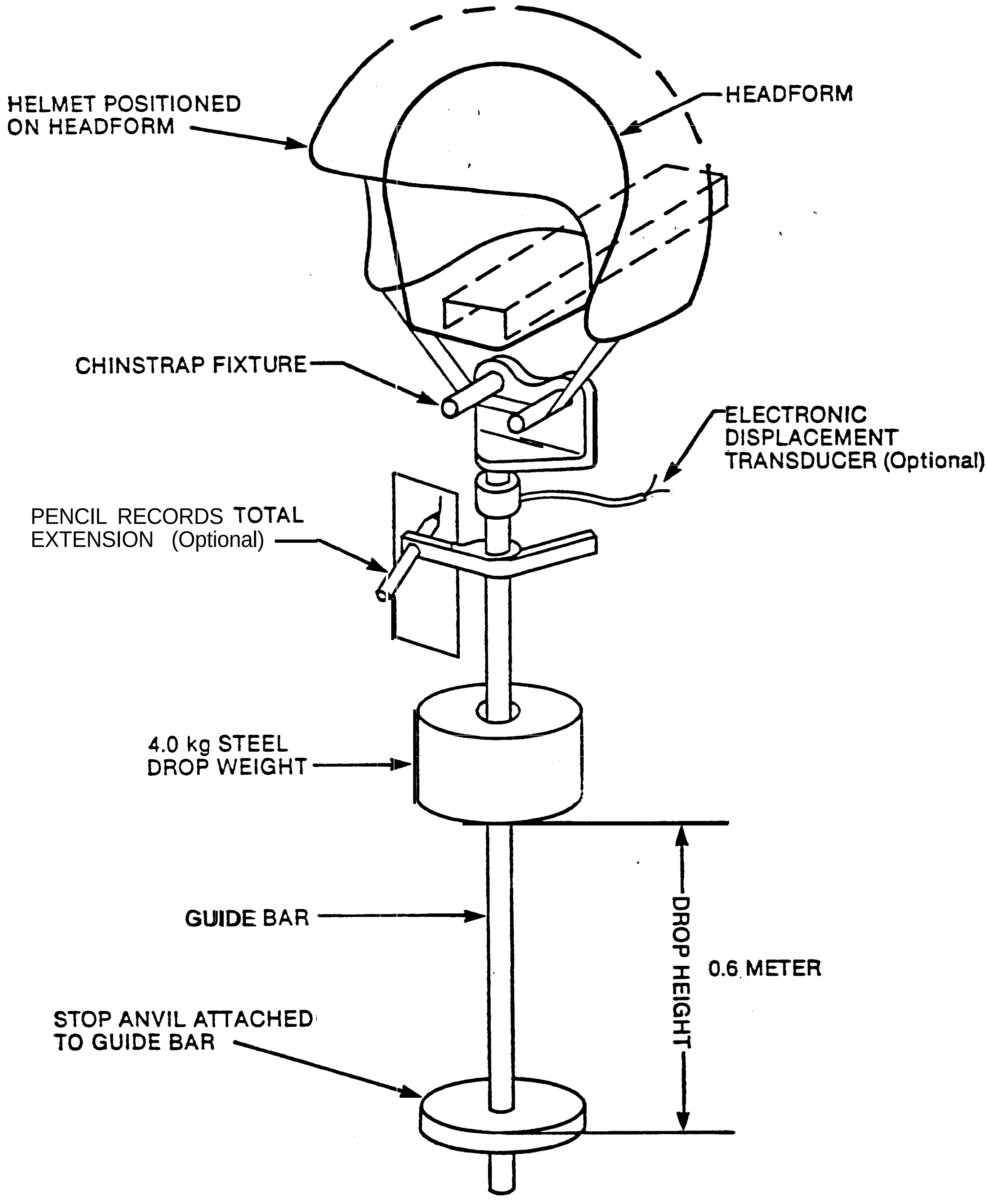


Figure 8. Apparatus for Test of Retention System Strength and Extention

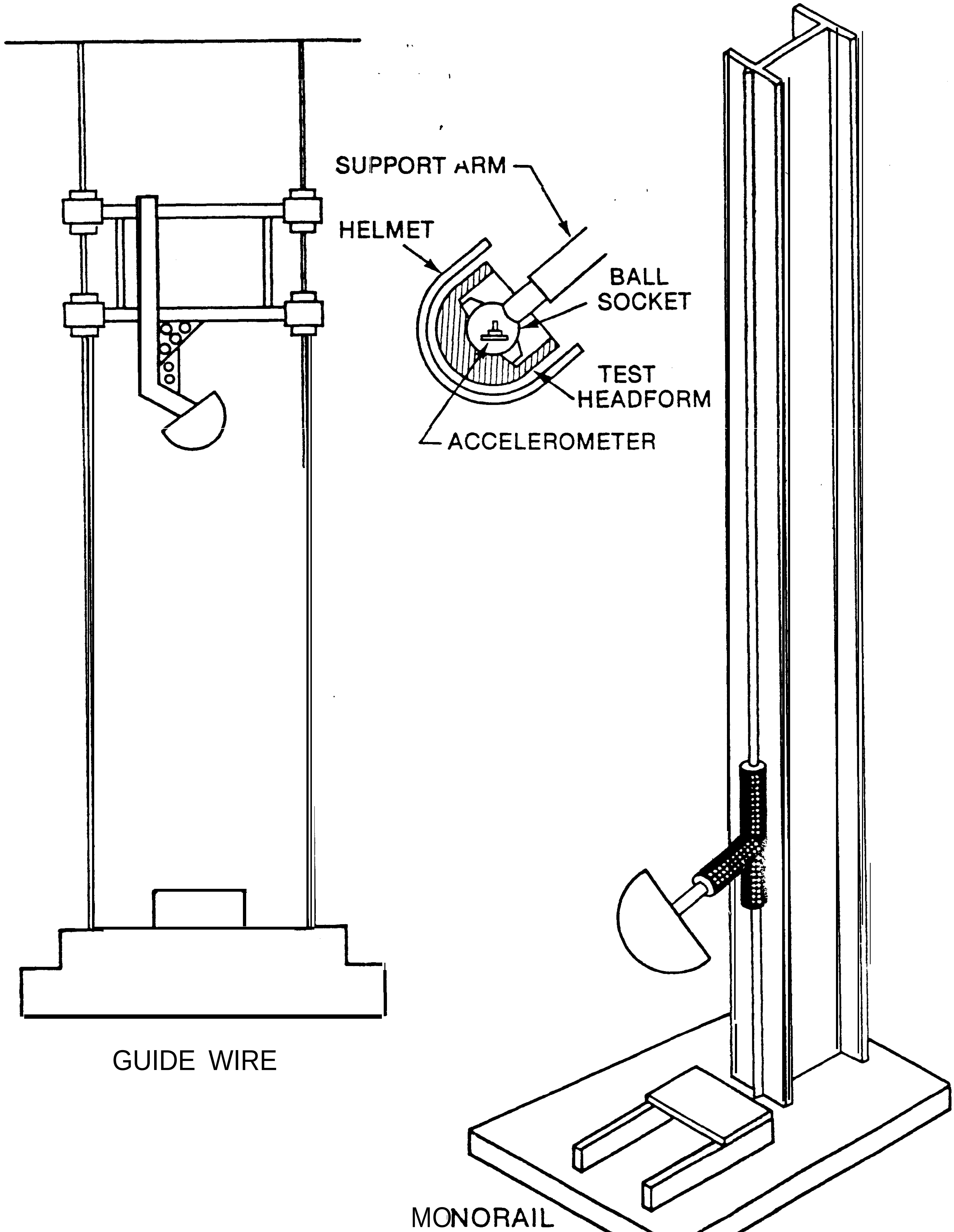


Figure 9. Impact Test Apparatus

Overhead View of Ball-Arm as Installed on Impact Test Apparatus

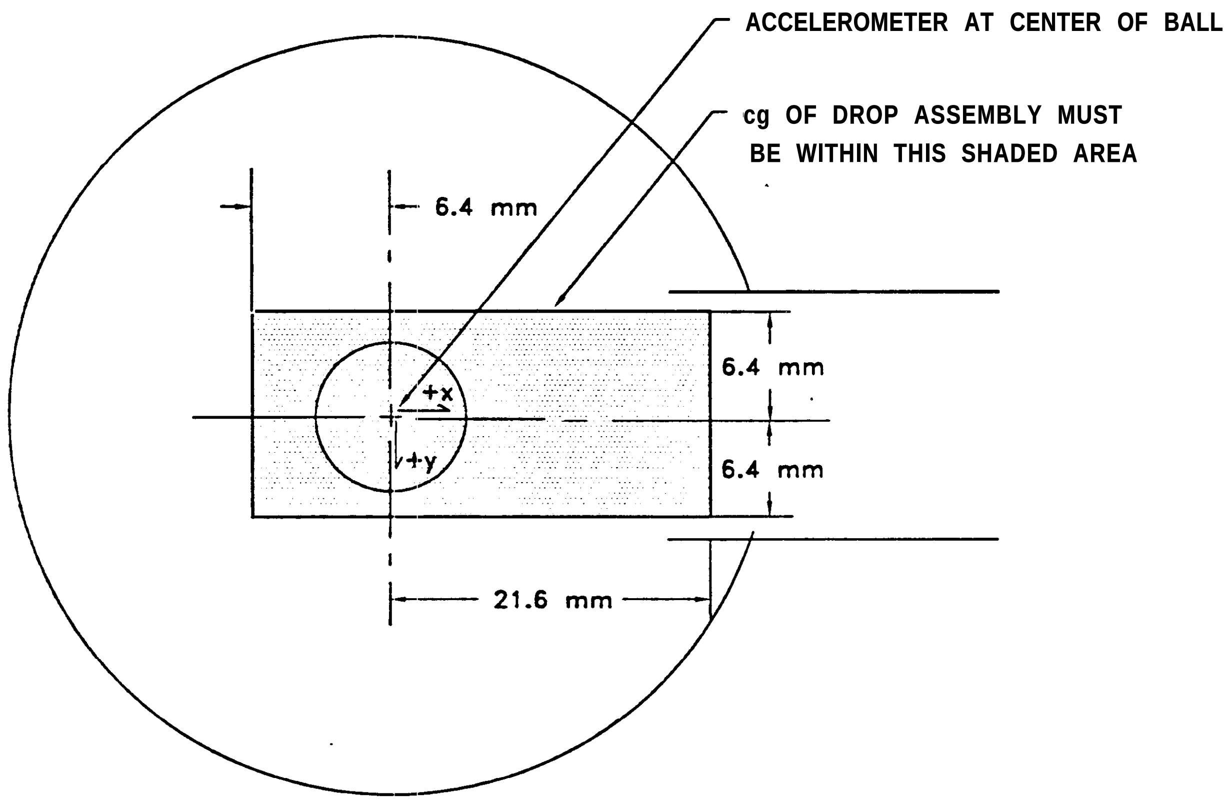


Figure 10. Center of Gravity for Drop Assembly

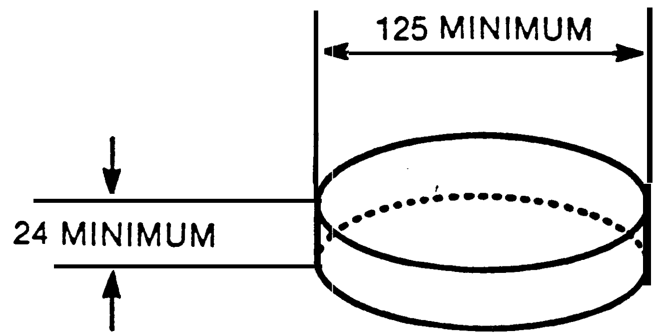


Figure 11. Flat Anvil

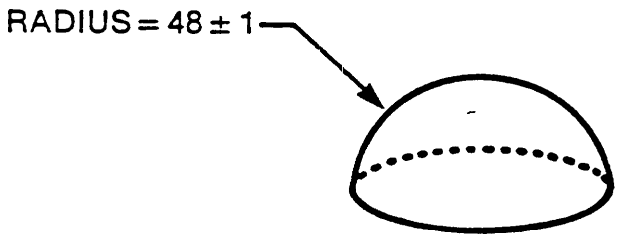


Figure 12. Hemispherical Anvil

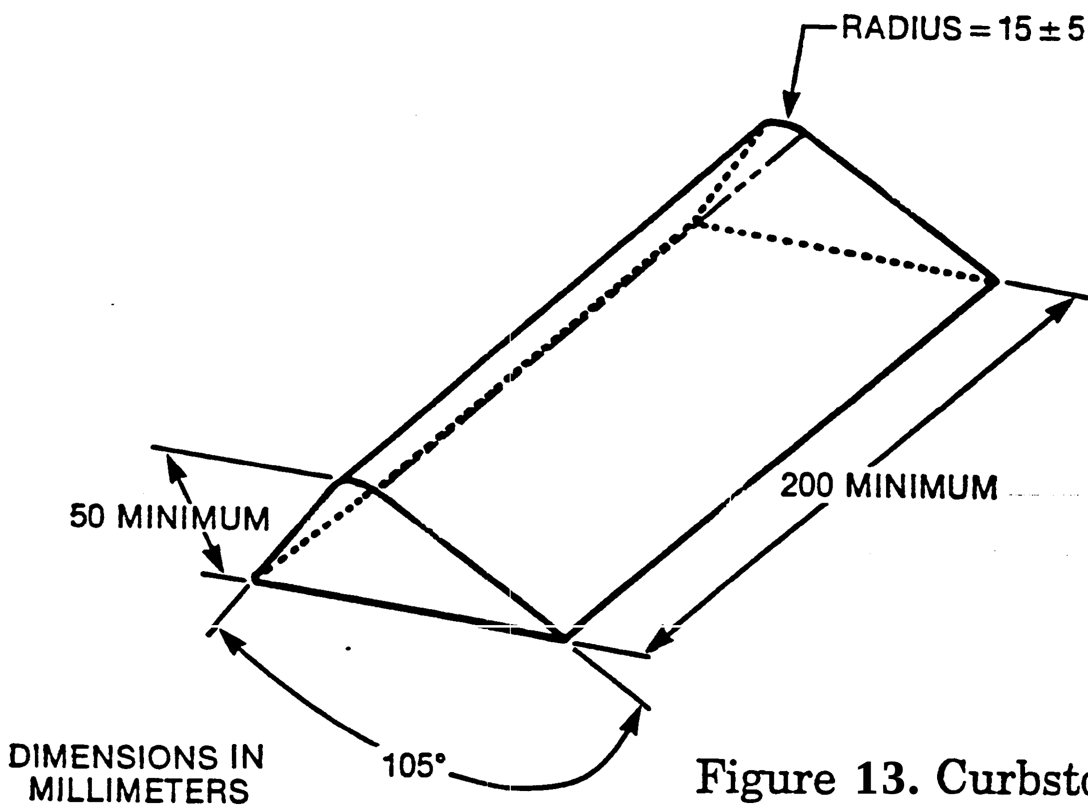


Figure 13. Curbstone Anvil

ATTACHMENT 2 - OTHER COMMENTS AND GENERAL ISSUES

Comment: Mr. Sabatano, President, London Bridge BMX Association [14], recommended that bike helmets be constructed so as to accommodate more serious accidents that might result from a child bicycle racing and jumping vs. merely riding on a path.

Response: While no helmet can protect against every conceivable impact, staff believes the available evidence supports that helmets designed to meet the CPSC standard will be very effective in protecting against serious injury within a wide range of common bicycle riding conditions. This would include many of the impact conditions that could occur during racing or jumping. Further, a standard for all bicycle helmets has to balance the benefits of more protective helmets against the additional cost, weight, bulk, and discomfort that more protection may impose. Such undesirable qualities may discourage many users from wearing helmets designed to protect against very severe impacts, which could more than cancel the effects of the additional protective qualities. Thus, the force with which the helmets are impacted in the standard's performance test has not been increased.

Comment: Randy Swart, Director of the Bicycle Helmet Safety Institute [16] suggested that the following items be considered as future revisions to the CPSC standard as progress in head protection research continues:

1. A test that requires the retention system to be easily adjusted for good fit
2. A test for protection against rotational injury.
3. A test to limit localized loads or "point loading"
4. A test for damage to the helmet by hair oil or other common consumer preparations.
5. A test of the retention system after impact to simulate field conditions
6. A test to ensure that visors and mirrors are shatter resistant and easily peel-off in the event of a crash.

Response: Staff agrees that it is important for the Commission to periodically review research related to improvements in head protection to determine if revisions should be considered for the CPSC bicycle helmet standard.

Comment: Dr. Richard Snyder, President, George Snively Research Foundation [19] referenced two studies relating helmet fit with head size and shape. The first study was conducted by Dr. Bruce Bradtmiller of the Anthropometry Research Project, Inc. Dr. Bradtmiller also responded to the proposed rule [20]. Dr. Bradtmiller concluded that for proper child helmet sizing, head breadth and length variables were more accurate guides than using age or head circumference. Dr. Bradtmiller urges caution in basing the CPSC rules for children's helmets on the draft IS0 DIS 6220-1983 standard for test headforms. Their study shows variation in the ratio of head length to head breadth. This ratio was found to be the prime determinant for helmet fit. The IS0 standard, however, maintains a constant head breadth/length ratio.

A second study also concluded that head circumference was not always a good indicator for helmet fit.

Response: IS0 headforms are the established norm for headgear testing in the U.S., Canada, Europe, and Australia. No other system of headforms is currently available that can be shown to provide more realistic results in terms of preventing injuries. ES recommends that the IS0 headform specification be maintained in the CPSC regulation. However, ES recommends that the staff stay current on developments of new test procedures and equipment that could lead to improvements in general helmet fit and in improvements that make it easier to fit and adjust helmets, especially for children.

Comment: NSKC [22] and CFA [23] recognized that the scope of the CPSC standard must be for bicycle helmets, but requested the Commission to move forward in investigating the issues related to multi-activity helmets. NSKC also urged the CPSC to work with community-based organizations to develop a comprehensive educational campaign regarding the importance of wearing a federally-approved bicycle helmet when participating in non-motorized activities other than bicycling.

Comment: Mr. Frank Sabatano, President London Bridge BMX Association [14] recommended that bicycle helmets should serve as multi-purpose protective devices for various sports such as bicycle riding, bicycle racing, skateboarding, and in-line skating.

Response: The Commission intends to monitor developments relevant to The multi-activity issue. ESHF (Tab G) concludes that wheeled recreational activities such as traditional rollerskating and in-line skating are typically conducted on the same surfaces as bicycling and can generate speeds similar to bicycling. Therefore, it is reasonable to assume that helmets that meet the requirements in the CPSC bike helmet standard will also provide head protection for roller/in-line skating and perhaps some other recreational activities.

However, as discussed in the December 6, 1995 Federal Register notice on the proposed rule, the Commission does not have sufficient data on the benefits and costs of additional features directed at injuries incurred other than bicycling to make the statutory findings that would be needed to issue a requirement for such features under either the CPSA or FHSA. Also, procedures in addition to those required by the Bicycle Helmet Safety Act would have to be followed. The Commission does not want to delay establishment of a mandatory bicycle helmet standard in order to pursue rulemaking for other types of helmets. Accordingly, this proposed regulation only addresses bicycle helmets.

As part its decision making in setting priorities for future activities, staff recommends that the Commission examine what actions it could take to encourage the use of bicycle helmets in activities that present head injury risks similar to those in bicycling.

Issue:

In his recommendations to the Commission, Duke University researcher Barry Myers M.D., Ph.D., suggested that a test for penetration resistance be considered for the final standard. He reasons that such a test would require helmets to have hard outer shells that would provide helmet users with improved protection. Dr. Myers contends that a hard shell will reduce the risk of penetration type traumas. He further contends that a hard shell will lessen friction between the helmet and the impact surface and that this has two benefits. First, it would reduce the total change in velocity (ΔV) of the head during impact. Second, by reducing the frictional constraints on the head during impact, it would reduce the risk of neck injury.

In support of hard-shell helmets, Dr. Myers references the latest Harborview[6] study, which reported a "consistent suggestion that hard-shell helmets are more protective against head and brain injuries than non-hard-shell helmets." Dr. Myers acknowledges that the differences measured were not statistically significant. However, he believes that a larger study, containing a sufficient number of severe brain injuries, might show this correlation with statistical significance.

In discussing protection against neck injury, Dr. Myers notes that automotive accidents cause serious neck injuries in about 15 to 25 percent of the persons who have serious head injuries, suggesting that neck injury is common among the most severely brain injured. However, since there were so few cases with severe brain injuries in Harborview's analysis of bicycling incidents, the significance of neck injury, and its mitigation by hard shell helmets among the severely brain injured, cannot be determined from the Harborview study.

Although Dr. Myers suggests a penetration test in order to require that bike helmets have a hard shell, he states that a detailed study of the 'most severe injuries is warranted. He also recommends that, before a requirement that all helmets have a hard shell is adopted, there should be an evaluation of whether this would reduce the number of riders who would wear bicycle helmets.

Response:

Protection Against Pen&ration

Currently available information does not show a need to address the hazard of penetration-type head impacts to bicyclists. One study[7] suggests that the majority of helmets involved in bicycle accidents suffer impacts on flat, hard surfaces (asphalt, cement, etc.) and that penetration type impacts are rare.

Protection Against Neck Injury

Injury Data

Bicycle-related injury data show a low incidence of serious neck injuries. In 1996, there were 566,400 bicycle-related injuries treated in U.S. hospital emergency rooms, based on data from CPSC's National Electronic Injury Surveillance System (NEISS). Of these, about 6,630 (1 percent) involved the neck. Of the neck injuries, about 4,520 (68 percent) involved strains/sprains, 1,155 (17 percent) involved contusions and abrasions, 275 (4 percent) involved lacerations, 240 (4 percent) involved fractures, and 440 (7 percent) involved other diagnoses. These numbers show that neck fractures accounted for about 0.04 percent of the total number of emergency room treated bicycle-related injuries in 1996. Detailed information was not available to analyze whether the use of a helmet or type of helmet had an effect on risk of neck injury.

---

[6] Thompson, Diane C., MS,: Rivara, Frederick P, MD, MPH; and Thompson, Robert S., MD. "Effectiveness of Bicycle Safety Helmets in Preventing Head Injuries." Journal of the American Medical Association. 276 (December 1996): 1968-1973.

[7] Dean Fisher and Terry Stern, "Helmets Work!," Bell Sports, Inc., AAAM/IRCOBI Conference, Lyon, France (September 1994)

The Harborview study also reported a low incidence of neck injury. Their report showed that 2.7% of the cases (including both helmeted and non-helmeted cases) suffered neck injury, ranging from sprain to nerve-cord injuries. There was no correlation between neck injury and helmet use or helmet type.

Dr. Myers cites that automotive accidents cause serious neck injuries in about 15 to 25 percent of the persons who have serious head injuries. However, this statistic may not be relevant to the issue of friction between the shell and the impact surface, since the neck injuries in automotive accidents are not necessarily caused by friction between the head and an impacting surface.

Sliding Resistance

Dr. Myers' advocacy of hard-shell helmets to reduce friction would seem to argue for a test to evaluate friction resistance of a helmet against typical impact surfaces, rather than for a penetration resistance test.

One study on this issue was done by Voigt Hodgson, Ph.D. at Wayne State University.* In this study, test helmets were secured to a modified Hybrid III dummy, and skid-type impacts were done on concrete at various angles from 30 to 60 degrees. Hodgson found that both hard-shell and micro-shell (or thin-shell) helmets tended to slide rather than "hang-up" on impact with concrete. (Thin-shell helmets are the type most commonly sold in the current market). No-shell helmets showed a larger tendency to hang-up on impacts with concrete. One of the conclusions of the study was that any helmet similar to those tested in the study (hard-, thin-, or no-shell) will protect the brain and neck much better than wearing no helmet.

Protection Against Head and Brain Injury

Harborview reports that there was a consistent trend indicating that hard-shell helmets provided better protection against head and brain injury than non-hard-shell helmets. However, in order for the results to be statistically significant, the number of people in the study would have had to be 11 times greater.

Recommendation

The following considerations are relevant to any possible requirement for hard-shell bicycle helmets:

(1) Studies of bicycle helmets damaged in accidents suggest that penetration-type helmet impacts are rare occurrences. In addition, bicycle-related injury data suggest a low incidence of serious neck injuries. For the small portion of incidents that involve serious neck injury or penetration-type hazards, available information is insufficient to estimate the degree of improved protective performance hard-shell helmets may offer over non-hard-shell helmets.

(2) Non-hard-shell bicycle helmets are effective in preventing serious head and brain injuries. There are no known studies that report a statistically significant finding that hard-shell helmets offer better protection than non-hard-shell helmets.

(3) A standard applying to all bicycle helmets has to balance the protective benefit that might be provided by a hard shell against the

---

[8] Voigt R. Hodgson, Ph.D., "Skid Tests on a Select Group of Bicycle Helmets to Determine Their Head-Neck Protective Characteristics," Department of Neurosurgery, Wayne State University, Detroit, MI (March 8, 1991)

additional cost, weight, bulk, and discomfort caused by such a requirement. Such undesirable qualities may discourage some users from wearing helmets, which could more than cancel the effects of any additional protective qualities. This is an especially important consideration given the popularity of non-hard-shell bicycle helmets.

Given these considerations, CPSC staff does not support for the final rule a penetration test or any other test that would require all bike helmets to have a hard shell. Available information is insufficient to support the inclusion of such a requirement. However, should future research provide evidence of the benefits of hard-shell helmets, and that the benefits would outweigh the potential negative impact on consumer use, the Commission should consider revising the mandatory standard.

ATTACHMENT 3

Engineering Sciences Recommendations on the Specification of the Impact Test Rig and Other Impact Testing Procedures

Scott Heh, Bicycle Helmet Project Manager, ESME

I. Background

On December 6, 1995, the Commission published in the Federal Register a proposed CPSC bicycle helmet standard. In the proposed standard, the CPSC specified the monorail type of test rig for bicycle helmet impact testing. Currently, U.S. voluntary bicycle helmet standards allow the use of either monorail or guidewire types of test rigs. The CPSC specified the monorail type to avoid the possibility that different results would be obtained with the two types of test rigs.

In their comments responding to the proposed rule, several helmet manufacturers and the Snell Memorial Foundation disagreed with the specification of the monorail test rig in the proposed CPSC standard. The respondents stated that guidewire-type rigs are more commonly used in the industry. Snell stated that there is no demonstrated improvement associated with the monorail rig in testing reliability and capability. Most respondents suggested that the CPSC standard allow the use of either guidewire or monorail rigs.

To respond to this issue, CPSC-ES initiated a seven-laboratory comparison test program. The main purpose of the study was to determine if there are statistically significant mean differences in test results when using monorail and guidewire test rigs under standardized testing conditions. The statistical analysis of the test results is at Tab F of the briefing package.

Seven laboratories participated in the test program, including the CPSC lab. Five of the laboratories tested on both monorail and guidewire rigs. Two laboratories only tested on monorail rigs. Three different helmet models were used. Each helmet was impacted twice, once at the rear of the helmet and once near the crown. Tests were conducted on the flat and curbstone anvils, and all testing was performed with ambient-conditioned helmets. This experiment provided for the analysis of the effect of the following variables: rig type, anvil, helmet model, laboratory, anvil impact sequence, and impact location.

II. Summary of the Interlaboratory Results and EH Recommendations

<u>Effect of Test Rig:</u>
When the data were summed across the laboratories having both types of test rigs, the type of test rig did not have an appreciable effect on test results in almost all examinations. However, the type of rig did have an effect when the Model I helmets struck the curbstone anvil on the second impact. In these conditions, the monorail rig yielded a significantly greater mean logarithm of peak-g than did the guidewire rig. This occurred at both the rear and crown impact sites.

EH reported that since it is possible to have the two test rigs yield significantly different results under specific testing conditions, it seems advisable to specify the test rig in the test procedure or make some modification to the test procedure to ensure that the two test rigs give similar results.

Other EH Recommendations

(1) In the instrument systems check procedure, include provisions for accuracy as well as precision.

(2) Modify the test proceciure in such a manner as to provide guidance to the tester in the selection of impact sites and order of use of anvil types.

(3) Test an appropriate sample of helmet model specimens under each test condition (instead of just one helmet model specimen) to take statistical variability into account.

Recommendation (1) pertains to ensuring that the data obtained in the instrument systems check procedure are sufficiently similar in different laboratories. Recommendation (2) deals with eliminating potential tester bias in the selection of impact locations, anvil types, and order of use on a helmet model specimen. Recommendation (3) ensures that more reliable results are obtained than those obtained by testing just one specimen under each test condition in CPSC compliance testing.

III. ES Discussion

Rig Type

The statistical analysis of the interlaboratory results showed that in almost all examinations of test variable combinations, the choice of test rig did not have an appreciable effect on test results. However, on the Model I helmets, and only when the second impact was on the curbstone anvil, the monorail showed a significantly higher mean logarithm for peak-g summed across laboratories having both types of test rigs. For reasons completely unrelated to these test results (see staff discussion in § 1203.13 on use of the curbstone anvil), a curbstone impact in combination with another impact on a single test helmet is no longer in the final standard being recommended by the staff. Since the interlaboratory data (summed across labs using both types of rigs) show no significant differences between guidewire and monorail rigs under test conditions within those defined in the draft final standard, the standard should allow either type of rig to be used for impact testing.

Over the last 15-20 years, voluntary standards in the U.S. have allowed the use of either monorail or guidewire types of test rigs. Both types of test rigs have been used extensively in both independent test laboratories and manufacturer's in-house test facilities. The Snell Memorial Foundation, one of the established helmet test organizations in the U.S., uses guidewire rigs to test conformance to their standards. The staff has no evidence to conclude that the allowance of both types of test rigs in voluntary standards has resulted in a compromise of safety for bicycle helmet users.

For the reasons discussed above, the technical staff recommends that both types of rigs are suitable for impact attenuation testing, and that the CPSC standard specify that either a monorail or a guidewire test rig may be used.

Systems Check Procedure

ES staff recommends that the following precision and accuracy procedure be added into the regulation so that laboratories can verify that their test equipment is recording accurately. The procedure requires that a spherical impact missile of a specified dimension be dropped with a certain impact velocity onto a Modular Elastomer Programmer (MEP). An MEP is a cylindrical pad of a polyurethane rubber that is used as a consistent impact medium for the systems check procedure. Pre-test and post-test impacts on an MEP to verify system recording is a standard practice of bicycle helmet test labs.

(1) Instrument system check (precision *and accuracy).* The impact-attenuation test instrumentation shall be checked before and after each series of tests (at least at the beginning and end of each test day) by dropping a spherical impactor onto an elastomeric test medium (MEP). The spherical impactor shall be a 1.46 mm (5.75 in) diameter aluminum sphere that is mounted on the ball-arm connector of the support assembly, The total mass of the spherical impactor and support assembly shall be 5.0 ±0.1 kg (11.0 ± 0.22 lb). The MEP shall be 152 mm (6 inches) in diameter and 25 mm (1 inch) thick, and shall have a durometer of 60 ± 2 Shore A. The MEP shall be affixed to the top surface of a flat 6.35 mm (¼ inch) thick aluminum plate. The geometric center of the MEP shall be aligned with the center vertical axis of the accelerometer (see § 1203.17(a)(2)). The impactor shall be dropped onto the MEP at an impact velocity of 5.44 m/s ± 2%. (Typically, this requires a minimum drop height of 1.50 meters (4.9 ft) plus a height adjustment to account for friction losses.) Six impacts, at intervals of 75 ± 15 seconds, shall be performed at the beginning and end of the test series (at a minimum at the beginning and end of each test day). The first three of six impacts shall be considered warm-up drops, and their impact values shall be discarded from the series. The second three impacts shall be recorded. All recorded impacts shall fall within the range of 380-g to 425-g. In addition, the difference between the high and low values of the three recorded impacts shall not be greater than 20-g.

The range of 380-g to 425-g represents an allowable tolerance of about 10%. The interlaboratory testing showed this tolerance to be attainable between laboratories. However,, test experience shows that even greater precision can be obtained for the systems check procedure within a given laboratory. The test data from the interlaboratory study shows that a target range of 380-g to 425-g and a precision range of 20-g can be achieved by bicycle helmet test labs in the U.S. and Canada.

<u>The effect of impact site, anvil type and anvil impact sequence</u>

Because the impact site, anvil type, and order of impacts can influence test results, the regulation must explicitly state that the test personnel will test helmets to the most severe conditions allowed by the standard. Since these conditions may vary depending on the design of the helmet, the test personnel must have flexibility in choosing how the helmet should be tested.

ES recommends that the following statement by added to section 1203.17(b)(2): "Impact sites, order of anvil use (flat and hemispherical), and curbstone anvil orientation shall be chosen by the test personnel in a manner that provides the most severe test for the helmet. Rivets and other mechanical fasteners, vents, and any other helmet feature within the test region are valid test sites."

In addition, the following statement should be added to Section 1203.12 (d) Impact attenuation *criteria:*

"(1) *General.* A helmet fails the impact attenuation performance test of this standard if a failure can be induced under any combination of impact site, anvil type, anvil impact order, or conditioning environment permissible under the standard, either with attachments or without attachments, or combinations of attachments, that are provided with the helmet. Thus, the Commission will test for a 'worst case" combination of test parameters. What constitutes a worst case may vary, depending on the particular helmet involved."

A hypothetical example of testing for a worst case condition might be the case of a helmet that comes with a detachable visor and has a larger than normal front air vent. Since the vent is larger than those ordinarily found on bicycle helmets, the Commission may choose to test the helmet with the curbstone anvil aligned within the vent to create a wedge or splitting action upon impact. If the helmet's visor were to interfere with achieving a clean impact of the anvil against the helmet shell, the Commission would test the helmet with the visor removed. This is just one example of how a particular helmet design may influence how the Commission will test for compliance to the standard.

Test specimen sample size

The purpose of the standard is to define the test procedures and set the mandatory performance criteria for bicycle helmets marketed in the U.S. For the helmet manufacturers, the issue of sample size must be addressed in the reasonable testing program that is required by the rule. The rule provides flexibility for each manufacturer to establish a testing program that best fits its production process. The rule calls for eight helmet samples to test to the provisions of the standard. As a matter of enforcement policy, the Commission may elect to test additional samples. As an example, if the Commission testing shows a "marginal" pass or fail for a particular helmet, the Commission may elect to collect one or more additional samples for retesting in order to verify the initial test results. Such actions will be considered on a case by case basis.

TAB D

United States
CONSUMER PRODUCT SAFETY COMMISSION
Washington, D.C. 20207

MEMORANDUM

DATE: OCT 06 1997

TO	:	File
THROUGH	:	Andrew G. Stadnik, Associate Executive Director for Engineering Sciences
FROM	:	Scott Heh, ESME, Bicycle Helmet Project Manager, 504-0494 ext. 1308
SUBJECT	:	Discussion of Special Provisions for Helmets for Children Ages 1-5, Test Headform Mass and Peak-g

I. BACKGROUND

One of the provisions of The Children's Bicycle Helmet Safety Act of 1994 was for the Commission to include in the final CPSC standard provisions that address the risk of injury to children. This does not require that children's helmets be subject to requirements that differ from those for adults' helmets; it requires only that the final standard be appropriate for children's helmets.

The issue of whether special standard provisions for young children's helmets are needed has been debated for several years by head protection experts. Voluntary standards organizations such as the ASTM and the Canadian Standards Association (CSA) have worked on developing standards specifically for helmets for children under the age of five years. CSA is the only North American standard to complete special provisions for young children's helmets. In examining how young children's helmets might be tested differently from helmets for older persons, fhere are three main items that are generally considered: (1) requiring an increased area of head coverage, (2) specifying a smaller mass for the test headform, and (3) requiring a lower allowable acceleration limit ("peak-g") for the impact test.

The Commission first proposed a safety standard for bicycle helmets on August **15, 1994.** In that proposal, the only special provision for helmets for children under five years was an increased area of head coverage.

On December 6, 1995, however, the Commission proposed special provisions for headform mass, peak-g limit, and head coverage for bicycle helmets for children under five years. The special children's provisions were based on the on-going work of voluntary standards organizations and proposals at that time in the technical literature. A comparison of the CPSC proposed test parameters for helmets for children under five years and for older persons is shown below.

	Under 5	5 and older
Mass of test headform	3.9 kg	5.0 kg
Peak-g limit	250-g	300-g
Head Coverage	more coverage at rear and sides of head	

The proposal for increased head coverage is relatively uncontroversial, continues to be recommended by staff, and is not discussed further in this memorandum. The headform mass and peak-g requirements have undergone extensive reassessment by the staff and are discussed in detail below.

II. DISCUSSION

A young child's skull has different mechanical properties than the skull of an older child and adult. These differences are especially evident for children under the age of five years. Their skulls have a lower degree of calcification, making them more flexible than adult skulls. During an impact to the head, the increased skull flexibility results in a greater transfer of kinetic energy from the impact site to the brain tissue. Besides the different mechanical properties, the mass of a young child's head is also different from that of a more mature person's head. Studies show that the head mass of children under the age of five years ranges from approximately 2.8 to 3.9 kg. This mass is lower than the 5-kg test headform mass specified in current U.S. bicycle helmet standards.

Proponents of special provisions for young children's helmets believe that these helmets should be tested under different test parameters than helmets intended for older persons. The current test parameters are based primarily on adult head injury tolerance and on a headform mass that is approximately that of an adult head. Supporters of special provisions contend that these adult test parameters result in a helmet with a liner that is too stiff to optimally protect a young child's head. By using a headform weight that better represents a young child's head (e.g., 3.9 kg), and reducing the allowable peak-g, helmets would need to be designed with a lower density ("less stiff") liner to further lessen the impact transmitted to the head.

The comments received by the Commission in response to the December 6, 1995 proposed standard illustrate the complexity of the issues concerning special provisions for children's helmets. A few respondents to the proposed rule (8,16)[1] supported the lower mass and lower peak-g provisions, believing that they will lead to an improvement in head protection for small children. One of these respondents., however, urged the Commission to consider the most recent research on this subject before including the special provisions in a final standard. One respondent (12) favored a reduced headform mass provision, but did not recommend a reduced peak-g provision, stating that it could result in a helmet with a lower margin of safety.

Several respondents (3,4,6,9,10,13,15,18,19,27,28,29,30) questioned if it is advisable to move forward with the provisions of a reduced-mass headform and a lower limit for peak acceleration. Some respondents suggested that special children's provisions should not be adopted since studies show that children's helmets as they exist today are protective.

Studies by researchers at the Harborview Injury Prevention and Research Center have shown that bicycle helmets that meet existing standards are effective in protecting against serious head and brain injuries.[2] [3] One of the items analyzed in the most recent Harborview study[3] was whether the protective effects of bicycle helmets vary by the age of the user. For four age groups of riders, they estimated the protective effect of helmets against three levels of injury listed in order of increasing severity: (1) head injury, (2) brain injury, and (3) severe brain injury. Due to a small number of helmeted case subjects that suffered brain injury and severe brain injury, Harborview researchers could not estimate the protective effect of helmets against these injuries fo:r the under six-year-old age group. However, one of Harborview's overall conclusions was that helmets

---

[1]The numbers in parentheses refer to the comment number assigned to the respondent by the CPSC Office of the Secretary. All numbers have the prefix "CC96-1-"

[2]Thompson, Robert S., MD; Rivara, Frederick P, MD, MPH; and Thompson, Diane C., MS "A Case Control Study of the Effectiveness of Bicycle Safety Helmets." The New England Journal of Medicine, 320 (May 1989): 1361-1367.

[3]Thompson, Diane C., MS; Rivara, Frederick P, MD, MPH; and Thompson, Robert S., MD. "Effectiveness of Bicycle Safety Helmets in Preventing Head Injuries." Journal of the American Medical Association, 276 (December 1996): 1968-1973.

are effective for all bicyclists, regardless of age, and that there is no evidence that children younger than six years need a different type of helmet.

The Commission requested technical views on this issue from Barry Myers, M.D., Ph.D., Associate Professor, Department of Biomedical Engineering, Duke University. In his report[4], Dr. Myers explains that modification to the standard should be considered only if it can be shown to improve performance. Improvements may be shown by epidemiological or biomechanical evidence. However, considering the degree of head injury protection provided by current helmets, incremental improvement would be difficult to detect, even with a large epidemiological study.

From a biomechanical perspective, it is important to assess how changes in test headform mass and peak-g criteria would affect helmet design and protective capability. This can be done by examining how a helmet functions to protect the head in an impact.

The helmet has a crushable liner typically made of expanded polystyrene foam. If the liner is crushed as the head presses against the inside of the helmet during impact, the liner allows the head to stop over a longer distance and time than would otherwise be the case. This reduces the impact energy that is transmitted to the head, thereby reducing the risk of injury.

The degree to which the liner resists being crushed affects the helmet's protective qualities. For a given impact, a helmet liner that is too soft will "bottom out," thereby losing its protective ability to allow relative movement between the head and the object being impacted. Conversely, a liner that is too hard will not allow sufficient crushing to adequately protect the head.

Effect on Helmet Design

A simple way to examine the effect of changing mass and peak-g is to model the helmet as a spring and apply the one-dimensional spring-mass impact formulas shown below. This approach is discussed by both Dr. Myers and by Mr. Jim Sundahl, Senior Engineer with Bell Sports, in his response to the proposed rule **(12)**.

---

[4]Myers, Barry, M.D., Ph.D. "An Evaluation of A Helmet Standard for Children." Report to the U.S. Consumer Product Safety Commission (July 1997)

$$a_{peak}=V_o\sqrt{\frac{k}{m}} \qquad (1)$$

$$x_{peak}=V_o\sqrt{\frac{m}{k}} \qquad (2)$$

where:

a_{peak} = peak acceleration (peak-g)
V_o = impact velocity
k = liner stiffness
m = headform mass
x_{peak} = required stopping distance (liner thickness)

Effect of changing mass and peak-g

If the value for m is reduced in Equation (1), the value for k must be reduced to achieve the same peak-g at the same impact velocity. This means that if a helmet meeting the standard's criteria with a 5-kg headform did not meet the peak-g requirement using a lighter headform, the helmet liner would need to be made softer so more crushing of the liner could occur. If the value for a_{peak} is reduced in Equation (1), and the other variables are held constant, the value :Eor k again must be reduced. Likewise, this means that a helmet that could not comply with a reduced peak-g criterion also would need a softer liner to allow more crushing. Equation (2) shows that with a decreased liner stiffness, a greater percentage of the helmet's available crush distance will be used during impact.

The biomechanical analysis shows that, for impact conditions that do not result in complete compression of the helmet's liner, it is possible to lessen the impact energy transmitted to the head (and reduce the risk of injury) by reducing the stiffness of the liner. However, as the impact energy increases, a helmet with a softer liner will bottom out (crush beyond its protective capacity) under less severe conditions than a helmet with a more rigid liner of the same thickness. To compensate, the softer helmet would have to be made thicker to prevent bottoming out. However,, there is a limit to how thick a helmet can be before it is no longer practical or appealing to the user. Therefore, the goal of helmet design is to optimize liner density and thickness to protect against the widest range of impact conditions and still have a product that people will use.

Effect on Protective Performance

The biomechanical analysis suggests that reducing the liner stiffness could have both a positive and a negative influence on the protection provided by helmets under existing criteria. Therefore, it is necessary to also examine available epidemiological data that relate to this issue. Decreasing the liner stiffness would benefit those who experience injuries with minimal or no liner deformation. However, a decrease in liner stiffness could increase the number of head injuries that occur during more severe impacts that cause the helmet liner to bottom out.

To learn the effect on level of protection offered by softer helmet liners for children under 5, two questions would need to be answered:

1. Are children suffering head injuries with minimal or no liner deformation of current helmets?

2. Are children suffering head injuries with a bottomed-out liner?

Unfortunately, currently available information is limited and does not answer either of these questions. Therefore, it is uncertain whether young children would benefit from special provisions for headform mass and peak-g.

The only known study to examine the relationship between helmet damage and head injury was completed in 1996 by the Snell Memorial Foundation and the Harborview Injury Prevention and Research Center.[5] Of those bicycle helmets collected from individuals (of various ages) who went to a hospital, 40% of the helmets had no deformation, 14% had significant damage in which the helmet was approaching a bottomed-out condition, and 7% of the helmets had catastrophic damage. The data were not presented specifically for the under-5 age group or any other specific age group. The study showed that there was a risk of head and brain injury even with no or minimal helmet damage. The risk of injury increased moderately as the severity of helmet damage increased, until catastrophic damage was reached. As expected, the risk of head and brain injury jumped dramatically when a helmet was damaged catastrophically. This study suggests that if helmets for all ages were designed with softer liners, there is a potential to both improve the protection for lower-severity impacts and increase the risk of injury at the higher-severity impacts.

Since the risk of injury rises dramatically with catastrophic helmet damage, and current helmets are effective in reducing the risk of head and brain injuries, the staff does not

---

[5] Rivara, Frederick P., MD, MPH, Thompson, Diane C., MS, Thompson, Robert S., MD "Circumstances and Severity of Bicycle Injuries." Snell Memorial Foundation/Harborview Injury Prevention and Research Center (1996).

support a change to require softer helmet liners for bicyclists of all ages. The available data are insufficient to determine that such a change would increase overall protection. When focussing on the age range of under five years, currently available information is even more sparse. Therefore, if helmets for children under age 5 were made with softer liners, there are insufficient data to estimate either (1) the level of protection that might be gained at the lower-severity impacts, and (2) the protection that might be lost at the severe impact conditions that completely crush the liner.

III. RECOMMENDATION

Based on the items discussed above, CPSC staff'recommends that there be no special provisions in the final standard for headform mass and peak-g criteria for young children's helmets. The staff recommends this approach because of insufficient data to justify the changes and the consideration that these changes could provide less protection in the most severe impacts which could result in more serious head injuries to children. However, should future studies provide evidence that young children, or bicyclists of any age, could benefit from decreased liner stiffness, the Commission could consider revisions to the bicycle helmet standard at that time.

TAB E

United States
CONSUMER PRODUCT SAFETY COMMISSION
Washington, D.C. 20207

MEMORANDUM

DATE: July 16, 1997

TO : Scott Heh, Project Manager
Directorate for Engineering Sciences
Division of Mechanical Engineering

Through: Andrew G. Ulsamer, Ph.D. AGU
Associate Executive Director
Directorate for Laboratory Sciences

FROM : George F. Sushinsky 301-413-0172
Mechanical Engineering Team Leader
Division of Engineering

SUBJECT: Response to Comments on the 12/6/95 NPR for a Mandatory Bicycle Helmet Standard

Background

Staff at the Division of Engineering of the Directorate for Laboratory Sciences (LSE) reviewed and responded to written comments containing testing issues that were received as a result of the Commission's publication on December 6, 1995, of a notice of proposed rulemaking (NPR) for a mandatory bicycle helmet standard. The majority of the comments addressed by LSE staff relate directly to specific procedural issues in test protocols contained in the proposed standard. The sections of the draft standard covered in the LSE response are:

§1203.5 Construction Requirements - projections
§1203.8 Conditioning environments.
§1203.10 Selecting the test headform.
§1203.11 Extent of impact protection - marking the test line.
§1203.13 Test schedule - conditioning
§1203.13 Test schedule - retention system testing
§1203.17 Impact attenuation test - impact velocity
§1203.17 Impact attenuation test - Test Procedure - Impact sites.

The comments and responses for each section are in Attachment A. Based on the findings in attachment A, LSE staff recommend certain changes to the proposed CPSC standard for bicycle helmets. These changes deal with sections 1203.5, .8, .10, and .13 of the proposed CPSC standard. A summary of the suggested changes is in Attachment B. Attachment C addresses comments regarding the use of a curbstone anvil.

cc:

R. Garrett, LSE
N. Caballero, LSE
R. Hundemer, LSE
H. Lim, LSE

ATTACHMENT A

**RESPONSE TO COMMENTS ON THE NPR of December 6, 1995
FOR A MANDATORY BICYCLE HELMET STANDARD**

§1203.5 Construction Requirements - projections.

Comment: Two respondents [CC96-1-2 and 6] addressed the proposed standard's requirements for rigid projections. The proposed standard provided that "Rigid projections on the inner surface shall not exceed 2 mm (0.08 in) and shall not make contact with the headform after testing in accordance with §1203.17." One respondent [2] is concerned with the definition of the term "rigid." The other respondent [6] wants an objective way to determine if the projection makes contact with the headform.

Response: Engineering staff (ESME) are recommending that the requirement for projections be changed to wording similar to that contained in the most recent Snell standard (B95). This requirement prohibits "fixtures" (projections) on the inner surface of the liner that project more than 2 mm into the helmet interior. LSE staff agrees with this recommendation since it limits qualification criteria to a quantity that can be measured objectively and eliminates the need to define or interpret qualitative issues such as rigidity.

§1203.8 Conditioning environments.

Comment: One respondent [CC96-1-2] requested expansion of the range of the cold environment for conditioning helmets before testing from -16 to -13°C to -18 to -8°C to be consistent with the wider ranges specified for other conditioning environments in the proposed standard. Also, he claimed that the narrower range was difficult to maintain with reasonably priced conditioning chambers.

A second respondent [CC96-1-26] stated that immersion was unrealistic and recommended spray conditioning of the helmets.

Response: Staff notes that the temperature range in the NPR apparently contained a typographical error. The range should have been (-17 to -13°C). This tolerance range is consistent with existing ANSI, ASTM, Snell B 95, and CSA standards. No change is recommended other than the correction of the typographical error.

The subject of wet immersion was discussed in the previous comment/response memorandum (Sushinsky to Heh, August 3, 1995). No new information has been received by CPSC staff since that time to address wet-conditioning of helmets. No change is recommended.

§1203.10 Selecting the test headform.

Comment: One respondent [CC96-1-5] questioned the need for two additional helmets for tests on a larger headform. This respondent, along with one other [CC96-1-29], felt that the definition of "fit" in the proposed standard is inadequate in its specification of the compression of the foam fit pads.

Response: In testing to the proposed standard, staff used a separate helmet to test for positional stability (§1203.15) when testing a set of five helmets. Impact tests were not run on this separate helmet. Although this was evident in the Test Schedule shown in Table 1303.13, it was not explicitly defined in the requirements for the additional helmets. A similar requirement for positional stability testing on a separate helmet was maintained when a helmet fit more than one headform. Recommendations in the requirements involving the test schedule have been made in response to several comments on the draft standard. The revised test schedule presented in the project manager's redraft of the standard eliminates the requirement to test on the larger headform. This change simplifies the test procedure by testing on a single headform size and is consistent with current interim standards.

With regard to the issue of fit, staff previously recommended adoption of the definition for "fit" from the ASTM Standard F1446-94 Section 3.1.7.1. Staff reviewed their practice in fitting a helmet to a headform. Based on that review, it is concluded that the respondents' comments have merit. Staff recommends that the proposed definition of fit be amended to reflect current practice.'

Proposed wording to reflect the recommended changes is provided in Appendix B.

§1203.11 Extent of Impact Protection - Marking the test line.

Comment: One respondent [CC96-1-28] to the proposed CPSC standard submitted a lengthy comment concerning the practical problems in certifying helmets-when only a test line is specified, The respondent requested that the standard be amended to require additional coverage below the test line, particularly at the front and rear of the helmet. Without this change, the respondent states that bias and conflict will be inescapable.

Response: As addressed in the 1995 response to similar concerns, staff recommends that only the one line be specified, measured, and drawn on the helmet. This singular line is the test line or center of impact line. This recommendation is based primarily on the fact that coverage does not imply impact protection. The only area on the helmet required to pass impact protection requirements is the area above the test line. A coverage line may also be design restrictive. Therefore, staff does not recommend specifying additional coverage below the test line.

§1203.13 Test schedule - conditioning.

Comment: One respondent [CC96-1-2] noted that, as written, there is potentially no upper limit to the exposure time to recondition a helmet once it is removed from the conditioning environment for more than three minutes. He suggests a change in the wording to specify an upper limit to reconditioning by insertion of the phrase "or 4 hours, whichever is shorter" at the end of the last sentence of 51203.13 (c).

Response: The requirement, as currently written, requires five minutes of reconditioning for each minute beyond three minutes that a helmet is removed from its conditioning environment. As worded, a helmet would have minimum reconditioning requirements that equal or exceed the original requirements (four hour minimum) if it is conditioned initially as required and then, for whatever reason, is removed from that environment for more than 51 minutes. At this interval of unconditioned exposure, a helmet would need more than four hours of reconditioning plus an additional five minutes of reconditioning for each additional minute the helmet is unconditioned. To eliminate this possibility, staff recommends revising the standard as suggested by the respondent.

§1203.13 Test schedule - retention system testing.

Comment: One respondent [CC96-1-8] wants the retention system test (§1203.13(d)) done after impact testing. He reasons that an accident can damage a helrnet and severely compromise the retention system. The retention system must ensure that the helmet remain on the head during an accident sequence. The respondent also recommends that the "zero" position for measuring elongation be established without pre-tensioning the straps with a 4-kg mass as called for in the standard.

Response: Staff recommends that no changes be made to the sequence for retention system testing. The test sequence issue raised by the respondent was addressed during the prior comment period. ASTM standard F1447 and Snell standards B-90 and B-95 test the retention system prior to impact attenuation testing. ANSI standard 290.4 does :not specify clearly a test sequence. LSE staff has no evidence that the test sequence specified in the ASTM and Snell standards would allow helmets that do not have adequate retention systems to pass the retention system test.

Staff also recommends that no changes be made to the procedure for establishing the pre-test "0" position. There is no evidence that pre-tensioning the straps prior to performing the retention system test would allow helmets to pass the retention system test that do not have adequate retention systems.

§1203.17 Impact attenuation test - impact velocity.

Comment: One respondent [CC96-1-8] suggested that the impact velocity tolerance be changed from ± 3 percent to -0, +5 percent to insure that impact testing is done at no less than the specified velocity. He notes that the Federal standard for motorcycle helmets (FMVSS 218) specifies tolerances for impact velocities in this manner.

Response: The proposed mandatory standard specifies impact velocity tolerances common to existing voluntary standards for bicycle helmets. The difference between a tolerance of ± 3% and -0%, +5% has little practical significance. LSE staff has no concerns with permitting an impact velocity of up to 3 percent lower than the target velocity, and recommends no change to the proposed rule.

§1203.17 Impact attenuation test. - (b) ***Test Procedure-(2) Impact sites.***

Comment: Two respondents [CC96-1-27 and -29] commented that the minimum spacing between the centers of impact should be 150 mm. One of these respondents [27] felt that the CPSC had lowered the impact spacing from other voluntary standards' requirements.

Response: The selection of 120 mm in the proposed standard is based on ongoing discussions in the ASTM subcommittee to revise impact location spacing. Snell standard B-95 specifies 120 mm minimum impact spacing, and 120 mm is also consistent with provisions of 1/6th of the maximum circumference of the helmet in the Snell B-90 standard. Impact spacing of 150 mm limits the flexibility in choosing impact sites, especially on small helmets. LSE staff recommends no change to the proposed requirement.

ATTACHMENT B

SUGGESTED CHANGES TO PROPOSED RULE of December 6, 1995 FOR A MANDATORY BICYCLE HELMET STANDARD

§1203.5 Construction Requirements

(b) *Projections.* ... . ~~Rigid projections on the inner surface shall not exceed 2 mm (0.08 in) and shall not make contact with the headform after testing in accordance with §1203.17.~~ Replace with wording similar to that provided in Snell B 95, as suggested by ESME staff.

§ 1203.8 Conditioning environments.

(b) *Low temperature.* This is a temperature of ~~-16~~ _-17°_ C to -13° **C** (3 1° F to 9" F). The helmet shall be kept in this environment for 4 to 24 hours prior to testing.

§ 1203.10 Selecting the test headform.

(a) A helmet shall be tested on the smallest of the headforms appropriate for the helmet sample. In fitting the helmet to this headform, all of the helmets's sizing pads are partially compressed when the helmet is equipped with its thickest sizing pads and positioned correctly on the headform.

...

§ 1203.13 Test schedule.

(c) Testing must begin within 2 minutes after removal of the helmet from the conditioning environment. If the helmet is returned to the conditioning environment within 3 minutes after removal for testing, it shall be reconditioned for a minimum of 2 minutes before testing is resumed. If the helmet is out of the conditioning environment for more than 3 minutes, it shall be reconditioned 5 minutes for each minute it is out of the conditioning environment beyond the allotted 3 minutes~~, or 4~~ hours (whichever is shorter) before testing is resumed.

ATTACHMENT C

Response to Comments on the NPR for a Mandatory Bicycle Helmet Standard - Section 1203.13 Schedule of Tests (Curbstone impact tests)

United States
CONSUMER PRODUCT SAFETY COMMISSION
Washington, D.C. 20207

MEMORANDUM

DATE: July 14, 1997

TO : Scott Heh, Project Manager
Directorate for Engineering Sciences
Division of Mechanical Engineering

Through: Andrew G. Ulsamer, Ph.D AGU
Associate Executive Director,
Directorate for Laboratory Sciences

FROM : Han Lim, Mechanical Engineer (301) 413-0158
George F. Sushinsky Mechanical Engineer (301) 413-0172
Division of Engineering

SUBJECT: Response to Comments on the NPR for a Mandatory Bicycle Helmet Standard - Section 1203.13 Schedule of Tests (Curbstone impact tests)

Background

The first CPSC draft helmet standard, published in a Notice of Proposed Rulemaking (NPR) in August 1994, contained provisions for a single curbstone impact in an ambient environment and reflected the consensus test schedule of ASTM F1447-94 - "Standard specification for Protective Headgear Used in Bicycling." One respondent to that NPR suggested that the curbstone anvil be included in the impacts in all conditioning environments [CC94-2-3a]. Similarly, two other respondents [CC94-2-3 and 8] requested a revision to section 1203.17(b)(2) of the draft standard to include the curbstone anvil in impact tests for all conditioning environments.

Engineering Division (LSE) staff considered these comments and agreed with them after limited testing on toddler helmets at LSE. This testing suggested that the curbstone anvil impacts typically result in lower peak G readings than the flat anvil impacts and are similar to the G levels of the hemispherical anvil impacts. Based on the different footprints for the three anvils, LSE staff recommended that helmets be tested on the three anvils under each environmental condition with the fourth impact anvil selected at the discretion of the test analyst. This also eliminated the need for impact testing on a fifth helmet which had been required for impact testing on a curbstone anvil under ambient conditions. Publication of this change to the August 1994 draft standard in December 1995 resulted in new comments.

In response to publication of proposed 16 CFR Part 1203 "Safety Standard for Bicycle Helmets; Proposed Rule" on December 6, 1995, CPSC received 31 comments on various aspects of the proposed rule. Staff at LSE was requested to respond to issues in the test protocols in the proposed standard. This memorandum deals with the general issues raised by respondents regarding Section 1203.13 of the proposed rule and specifically with respect to the use of the curbstone anvil during impact testing. The comments received on this issue are summarized below followed by the response.

Comments

Six respondents [CC96-1-5, 12, 27, 29, 30, and 31] submitted comments requesting changes in Section 1203.13 Test Schedule regarding the use of the curbstone anvil. All of the respondents expressed concern over using two curbstone impacts on a single helmet. As proposed, section 1203.13(d) and Table 1203.13 do not define the conditions of the fourth impact on a helmet. The fourth impact, left to the discretion of test personnel, could be a second curbstone impact. There also was concern about impacting the helmet with the curbstone anvil after the helmet was conditioned in a wet environment [CC96-1-12]. There also was concern about the curbstone footprint overlapping other impact sites and violating the "single impact" principle of testing [CC96-1-27 and 31]. The length of the curbstone anvil restricts the location of impact sites that can be used without overlap. The use of a second curbstone anvil, and the damage caused by curbstone impacts can restrict the selection of test sites further to the point where only three impacts may be possible on a small helmet without overlap.

The respondents* provided suggestions to amend the proposed rule by :

(1) Specifying a particular (non-curbstone) anvil for the fourth impact [CC96-5, 12, 29, and 30]
(2) Using the curbstone anvil only in a single impact and only in the ambient condition [CC96-1-29 and 30]
(3) Testing according to the schedule specified in the ASTM F 1447 bicycle helmet standard [CC96-1-27], or the original CPSC draft standard [CC96-1-31]. These are essentially similar suggestions to (2) requesting only a single curbstone impact testing in only an ambient environment.
(4) Using a curbstone-only impact as one of four impacts in an ambient environment. All other test environments

would consist of impacts using the hemispherical and flat anvils. [CC96-1-12]

Response

LSE Tests

In response to these 'comments, LSE staff conducted a series of 19 impact tests at LSE on 6 models of adult helmets in the $40 to $100 retail price range. The helmets tested used headforms ranging from sizes E to M. All of the impacts used a drop height of 1.2 m and a headform assembly weighing 5000 gm. The impact velocity criterion of 4.8 ± 3% m/s was met for all impact runs. Fifteen of the tests were conducted on helmets that had been conditioned by immersion in water for a period between 4 and 24 hours. Four ambient tests were conducted for comparison purposes.

LSE and ESME staff selected the smallest solid area between vents in each of the helmets as the target area for initial helmet impact sites. Normally, a left or right front vent was chosen. If several subsamples of a helmet were available, impact tests on that type of helmet were conducted in an ambient environment on a second subsample under nominally identical test conditions.

All of the helmets experienced at least one crack along the impact vent line after the initial curbstone impact. From the total of 19 impact test runs, one helmet failed to meet the 300-G acceleration limit when it was impacted once on the curbstone anvil. For this helmet, the curbstone anvil wedged open the vent of the helmet and split the helmet in half. A different model helmet also experienced a similar splitting effect, but did not fail to meet the 300-G limit. Both of these helmets had been wet-conditioned.

Four helmets were impacted a second time after the initial curbstone impact by dropping them from 1.2 m onto a hemispherical anvil. Three of these were helmets conditioned in the wet environment. None of the helmets that were impacted both on a curbstone anvil and a hemispherical anvil split in half or experienced acceleration levels of 300 G's or more. No helmets were tested with two curbstone impacts.

Discussion:

There are three major concerns addressed in the comments received on Section 1203.13 of the proposed bicycle helmet standard. They are:

(1) the possibility of two curbstone impacts on a single helmet,

(2) the effect of testing a wet conditioned helmet with a curbstone anvil.

and

(3) violation of the "single impact" principle of testing or restriction the location of impact sites that can be used without overlap if a second curbstone anvil is used.

The test schedule proposed in the NPR was developed'to allow test flexibility and economy of testing. It was not intended to require or forbid the use of a second curbstone impact in a test sequence. However, because of the damage seen in the testing of helmets on a single curbstone anvil and the problem with overlap, LSE staff agrees with the comments that two curbstone anvils should not be used in a test sequence on one helmet. The final standard should be modified to preclude this possibility.

Staff, however, disagrees in general with the comments requesting that the test specify which anvil would be used for the fourth impact in a four impact test sequence, except that it shall not be a curbstone anvil. Specification of the fourth anvil limits the discretion of test personnel to use the most appropriate conditions for worst-case testing based on their expertise.

LSE staff also disagrees with comments to limit environmental conditioning of the helmet to ambient conditioning when impacting a curbstone anvil. In the LSE tests, there was greater damage to helmets conditioned in environments other than ambient.

LSE staff agrees in general with the comments requesting that tests using a curbstone be limited to a single impact on a helmet and that no other anvils would be used to-impact that helmet. However, in LSE tests, two helmets split almost in half on a single curbstone impact to each helmet, but only one exceeded a 300-G deceleration limit. The split helmet presents a potentially unsafe result. Snell standards (N 94 and B 95) provide for sample rejection if the test personnel conclude that the headgear has been compromised by breakage. LSE staff does not recommend this approach because it is too subjective to employ as a mandated safety criterion. In such cases **where the** helmet may "marginally" **pass the standard, the Commission could elect to collect one or more additional** samples for **retesting in** order to verify the initial test results. Notice of the possibility of further compliance testing activity to address such situations should be given to the industry in the final CPSC standard.

Recommendation:

LSE staff offers the following recommendation for changes to the impact test schedule.

> One helmet each is tested in each of the four environments (ambient, hot, cold, and wet-immersed). Each helmet is impacted twice on the flat anvil and twice on the hemispherical anvil. The order of impacts is at the discretion of the technician. A second set of four helmets (one helmet for each of the four environments) Incorporates a single curbstone anvil test. Each helmet is impacted once. Eight helmets are needed for impact testing under this schedule.

It is further suggested that language be included in the FR notice to make it clear that the Commission may elect to test additional samples in cases where helmets meet the 300-g criteria, but show a significant amount of damage after testing. To address this and other marginal passing results the following language is proposed:

> Test experience using the curbstone anvil shows that it is possible for a helmet to show significant structural damage (to the point of nearly splitting in half) and still remain under the 300-g failure criteria. In such cases where the helmet may "marginally" pass the standard, the Con-mission may elect to collect one or more additional samples for retesting in order to verify the initial test results. Other conditions that may prompt the Commission to undertake verification testing on additional helmet samples include (but are not limited to) peak-g readings that are very close to the 300-g failure criteria.

LSE staff also recommends that the peripheral vision test of §1203.14 and the positional stability test (roll-off resistance) of §1203.15 be performed on a single helmet in the impact test matrix. This helmet would be conditioned to the ambient environment prior to environmental conditioning for impact testing.

This recommendation provides an equal or greater degree of protection than many of the existing voluntary standards that research has shown to reduce the risk of head injury by 69 percent, brain injury by 65 percent, and severe brain injury by 74 percent [Rivara, et. al Harborview, 1996]. This recommendation combines the basic test matrix of the most commonly used voluntary standards (two impacts each with flat and hemispherical anvils in each environment) with curbstone anvil impacts on additional helmets.

The recommendation applies to §1203.12 and 1203.13 of the proposed draft standard.

cc:
R. Garrett, LSE
R. Hundemer, LSE
N. Caballero, LSE

TAB F

United States
CONSUMER PRODUCT SAFETY COMMISSION
Washington, D.C. 20207

MEMORANDUM

DATE: JUL 1 8 1997

TO : Scott Heh, ESME
Project Manager, Bicycle Helmet Project

Through: Mary Ann Danello, Ph.D., Associate Executive Director
Directorate for Epidemiology and Health Sciences
Arthur K. McDonald, Acting Director
Hazard Analysis Division (EHHA)

FROM : Terry L. Kissinger, Ph.D. , EHHA

SUBJECT: Report on Interlaboratory Bicycle Helmet Study

Attached is the report on the interlaboratory bicycle helmet study.

AN ANALYSIS OF THE INTERLABORATORY BICYCLE HELMET STUDY

July 14, 1997

Terry L. Kissinger

US. Consumer Product Safety Commission
4330 East West Highway
Bethesda, MD 20814

Executive Summary

This report provides an analysis of data from an inter-laboratory study on bicycle helmets. The main purpose of the study was to determine if there are statistically significant mean differences in test results when using monorail and twin-wire test rigs under standardized testing conditions.

The study featured a repeated measures design with two impacts on each helmet specimen. Between-subject factors were those that varied between helmet model specimens; within-subject factors were those that varied between the impacts on a single helmet model specimen. The between-subject factors were:(l) rig type; (2) anvil; (3) helmet model; (4) laboratory; and (5) impact location permutation. The within-subject factor was impact location.

For the two rig types, with results summed appropriately over laboratories and compared, it was found that the monorail rig yielded a significantly greater mean than the twin-wire rig at each impact location on the second impact when testing a curbstone anvil on one specific helmet model. Thus, it is believed that specifying the test rig to be used for bicycle helmet testing or modifying the test procedure to ensure that the two test rigs give similar results would provide improved standardization of the test conditions.

I. Introduction

This report provides an analysis of data from an interlaboratory study on bicycle helmets. The main purpose of the study was to determine if there are statistically significant mean differences in test results when using monorail and twin-wire test rigs under standardized testing conditions. Also of interest was testing for statistically significant mean differences in test results at different laboratories.

In this report, the experimental design is discussed; the results of an instrument systems check are given; and an analysis of the data is presented. Special statistical features of the experimental data are discussed in the Appendix.

II. Experimental Design of Study

In this study, specimens of bicycle helmet models underwent impact testing using both monorail and twin-wire rigs. Impact testing was conducted by dropping the helmet model specimens on two types of anvils at controlled velocities. Each helmet model specimen was impacted twice, at separate locations (the crown and rear). For each specific impact location of a helmet model specimen tested on a rig with a particular anvil in a laboratory, a recording of the "peak G" (a measure of acceleration) imparted to the headform was made. This may be seen as a repeated measures experiment, with independent variables, or factors, that vary between subjects and within subjects. Between-subject factors were those that varied between helmet model specimens; within-subject factors were those that varied between the impacts on a single helmet model specimen. The specific categories, or forms, of a factor are known as the levels of that factor.

The dependent variable was the "peak G" measurement, the maximum acceleration imparted to the headform during impact, which is a continuous variable. The between-subject factors were (1) rig type (two levels: monorail and twin-wire); (2) anvil (two levels: flat and curbstone); (3) helmet model (three levels); (4) laboratory (five laboratories tested both rig types, and two laboratories tested the monorail rig type only); and (5) impact location permutation (two levels, corresponding to the two possible permutations of impact locations). The within-subject factor was impact location (two levels: crown and rear). All testing was performed under ambient conditions. Details of standardization of the experiment may be found in the test plan prepared by Mr. Heh dated November 19, 1996.

The objective of the study was to test for statistically significant mean differences in peak G measurements among groups defined by levels of the factors. Specifically, it was of interest to see if statistically significant mean differences were obtained with monorail and twin-wire rig types, and among the different participating laboratories. It was expected that statistically significant mean differences would be obtained with different anvil types, different helmet models, and different impact locations. It was unknown if there would be statistically significant mean differences between different impact location permutations. Including these factors in the study permitted an evaluation of interaction effects involving these factors and provided test data under a wide range of test conditions used by laboratories to conduct impact-attenuation tests of bicycle helmets.[1]

At each combination of levels of the five between-subject factors (henceforth called a "treatment"), there were two helmet specimens tested, except for some missing values due to inadvertent testing under inappropriate conditions (such as using the wrong anvil) and two laboratories not having the twin-wire test rig. Thus, analyzing the data at each of the two impact locations separately, this may be seen as two five-factor analyses of variance with replication in some cells and with some cells empty (which is discussed in more detail in the Appendix).

Factors in an experiment may be considered fixed or random. Fixed factors are those whose levels chosen for inclusion in the experimental design are the levels of specific interest to the experimenter. Random factors are those whose levels are randomly chosen for inclusion in the experimental design from a larger population of levels. All factors in this experiment should be considered fixed. For test rig, anvil type, and impact location permutation, this seems natural, since the levels included in the experiment are specifically those of interest. Laboratory, helmet model, and impact location are considered fixed because their levels were not randomly chosen.[2]

III. Instrument Systems Checks

Prior to testing at each laboratory, instrument system checks were performed by dropping two impactors (an IS0 J size magnesium impact headform and a spherical impactor) on a modular elastomer programmer (MEP). Each impactor was dropped 13 times on each test rig at the seven laboratories, with the first three drops considered "warm-up" drops and the corresponding data discarded.

It was found that the 10 retained impactor drop measurements were strongly correlated. A principal component analysis of the correlation matrix of the 10 drop measurements, done separately by impactor type and rig type and considering data from different laboratories as replication, showed that nearly all of the total variance was accounted for by an average of the 10 measurements.[3] The means of the 10 measurements, by laboratory, impactor type, and rig type, may be seen in Table 1.

Table 1: Means of 10 Peak G Measurements at Seven Laboratories, with Two Impactors, and on Two Test Rigs

	ISO J Impactor		Spherical Impactor	
Laboratory	Monorail	Twin-Wire	Monorail	Twin-Wire
A	449.7	421.1	421.1	402.2
B	413.2	424.0	358.9	375.2
C	424.7	438.0	427.1	418.8
D	428.0	423.4	411.3	395.7
E	450.0	408.0	387.0	393.6
F	427.8	---	395.2	---
G	420.8	---	401.1	---

Source: CPSC Interlaboratory Study on Bicycle Helmet Testing

While there were some large mean differences between the two test rigs on a given impaotor at a given laboratory, for the five laboratories with both test rigs, the mean peak G was roughly similar for the two test rigs on each impactor. The means are given in the following table:

Laboratories A,B,C,D, & E	ISO J Impactor		Spherical Impactor	
	Monorail	Twin-Wire	Monorail	Twin-Wire
Mean	433.1	422.9	401.1	397.1

It may be noted, however, that the mean peak G was lower for the twin-wire test rig than for the monorail test rig using each of the two impactors.

IV. Mean Contrasts

Special statistical features of the experimental data were taken into account in the analysis. These features are described in technical detail in the Appendix. As a result of these features, to better satisfy basic assumptions underlying the analysis of variance, the analysis was conducted using the natural logarithm of the peak G measures, instead of the peak G measures themselves (i.e., a natural logarithmic transformation was used, as explained in Section A of the Appendix).

Additionally, due to the presence of high-order interaction effects, tests were performed on differences of pairs of means, called "contrasts,"[4] to see if they were significantly different from zero. Such tests of contrasts were done to provide comparisons for levels of a between-subject factor at specific combinations of levels of other between-subject factors (e.g., testing for a mean difference between test rigs for a specific combination of anvil type, helmet model, laboratory, and impact location permutation). These tests were performed separately for each impact location. Tests of contrasts were also performed after appropriately summing over laboratories. These contrasts gave insights to the nature of some of the high-order interaction effects and also to the effects of the main factors.

A. All Treatments with Nonempty Cells

For each of the five between-subject factors, tests were conducted on mean contrasts, taking into account patterns of empty cells (e.g., for comparing the two test rigs, comparisons were made excluding data from laboratories not having one of the two test rigs) and controlling overall error rates for multiple comparisons.[5] Results of the tests for each of the between-subject factors were as follows, with antilogarithms given for ease of interpretation.[6]

1. Test Rigs

There were 56 contrasts tested for the crown measurements, and 54 contrasts tested for the rear measurements. There was a statistically significant difference found for the two test rigs with three contrasts for crown measurements. The factor levels at which these statistically significant differences occurred may be seen in Table 2. Figures l-3 show the antilogarithm of the mean logarithm for each test rig at these three combinations of factor levels.

Table 2: Factor Levels at Which There Was a Statistically Significant Difference between Test Rigs in the Mean Natural Logarithm of the Peak G Measurements

Impact Location	Laboratory	Anvil	Model	Order	Figure
Crown	A	Curbstone	I	Rear-Crown	1
Crown	A	Curbstone	I	Crown-Rear	2
Crown	C	Curbstone	III	Crown-Rear	3

Source: CPSC Interlaboratory Study on Bicycle Helmet Testing

Figures 1-3
Factor Level Combinations for Which There Was a Statistically Significant Difference between Test Rigs

Figure 1

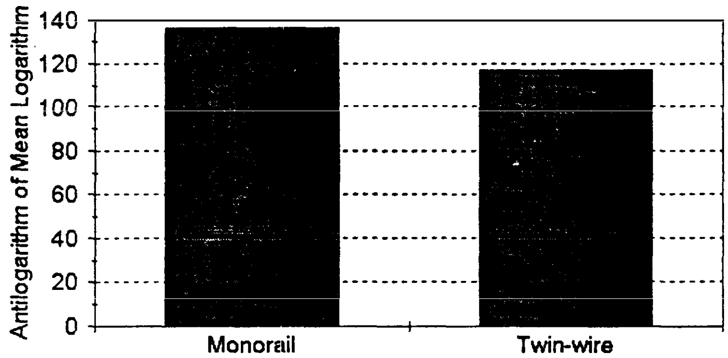


Figure 2

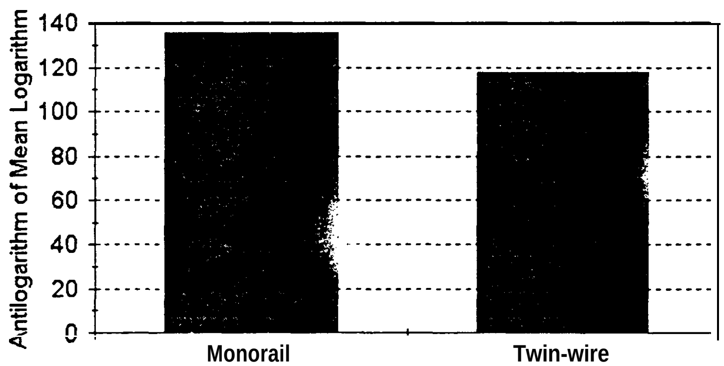
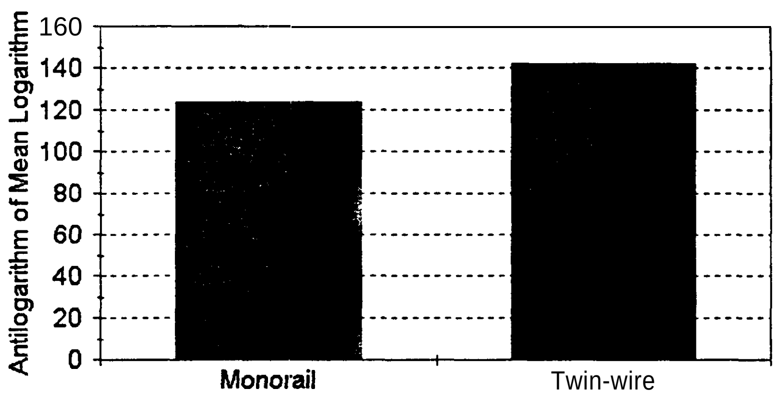


Figure 3



Note that all three of the statistically significant differences involved the crown impact locations, all three involved the curbstone anvil, two involved Model I helmets, and two involved the second impact on the helmet. The importance of these findings will be clearer after testing for significant differences after summing appropriately over laboratories.

2. **Laboratories**

There were 356 contrasts tested for the crown measurements, and 348 contrasts tested for the rear measurements. There was a statistically significant difference found among the various laboratories at two combinations of between-subject factor levels for crown measurements and at five combinations of between-subject factor levels for rear measurements. The factor levels at which these statistically significant differences occurred may be seen in Table 3. Figures 4-10 show the antilogarithm of the mean logarithm for each laboratory at these seven combinations of factor levels.

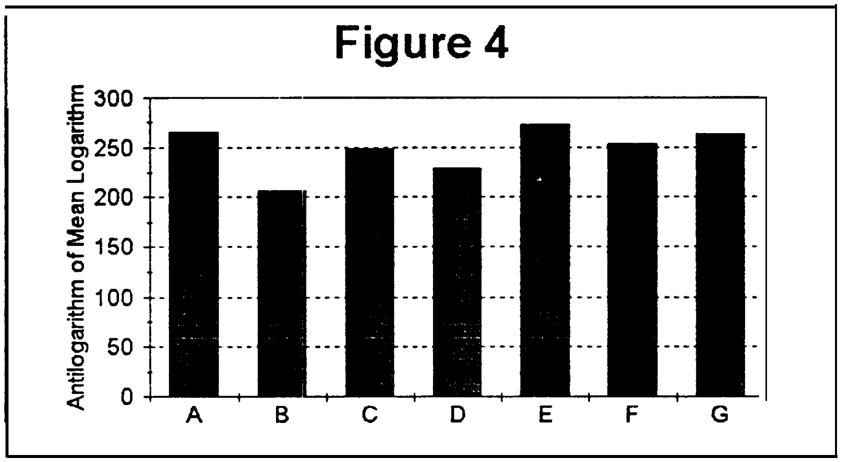
Table 3: Factor Levels at Which There Was a Statistically Significant Difference between Laboratories in the Mean Natural Logarithm of the Peak G Measurements

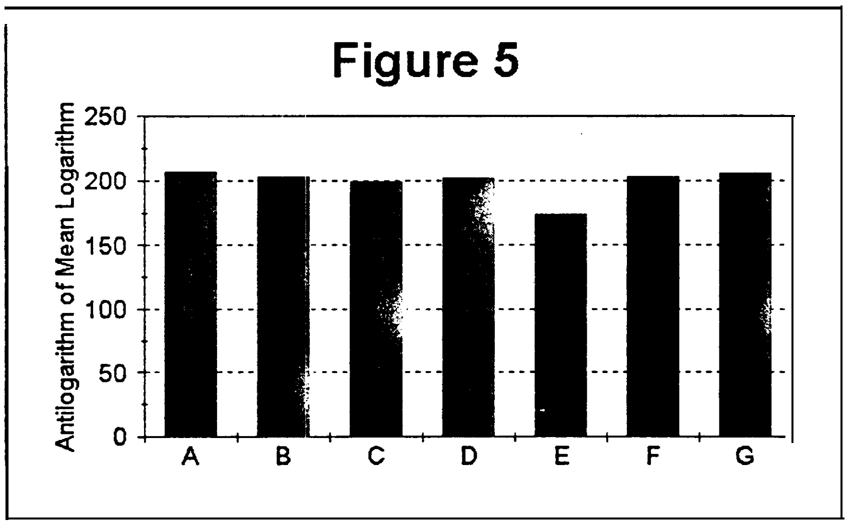
Impact Location	Rig Type	Anvil	Order	Model	Figure
Crown	Monorail	Flat	Rear-Crown	I	4
Crown	Monorail	Flat	Rear-Crown	II	5
Rear	Monorail	Flat	Rear-Crown	I	6
Rear	Twin-wire	Flat	Rear-Crown	II	7
Rear	Monorail	Flat	Crown-Rear	I	8
Rear	Twin-wire	Flat	Crown-Rear	I	9
Rear	Twin-wire	Flat	Crown-Rear	II	10

Source: CPSC Interlaboratory Study on Bicycle Helmet Testing

Figures 4-10

Factor Level Combinations for Which There Was a Statistically Significant Difference among Laboratories





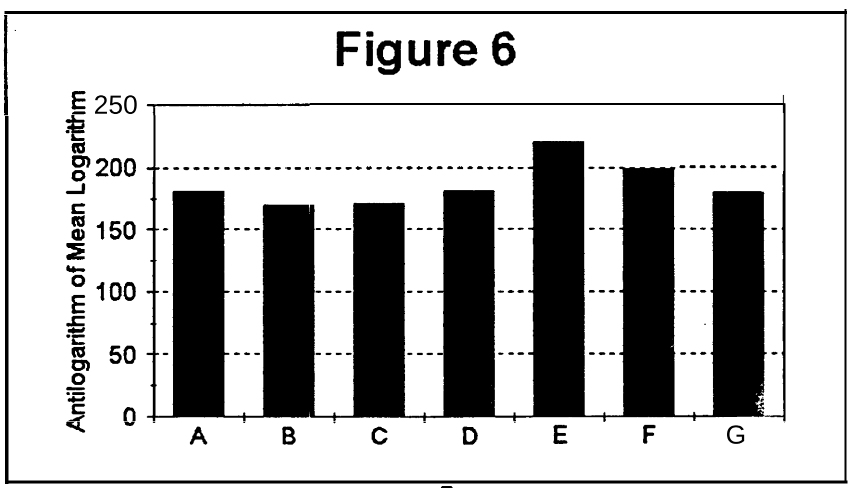


Figure 7

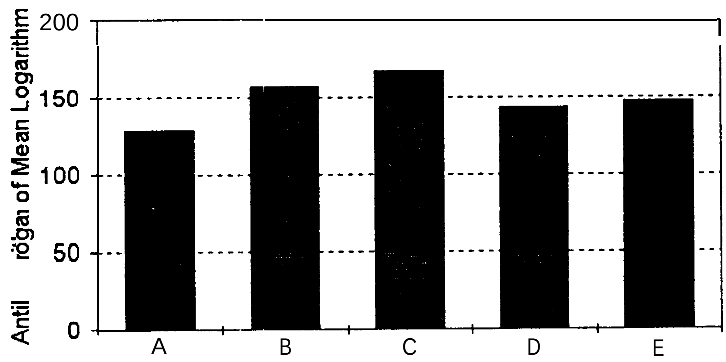


Figure 8

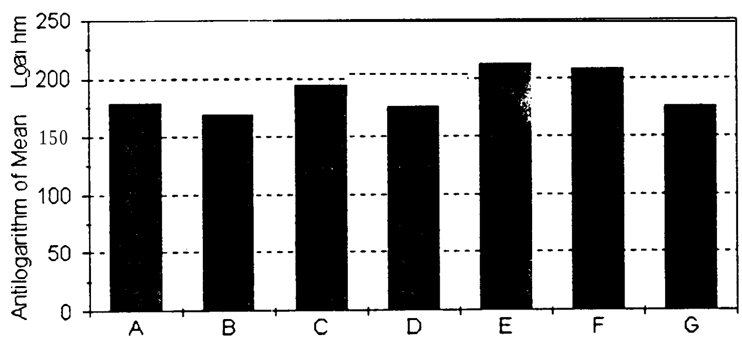


Figure 9

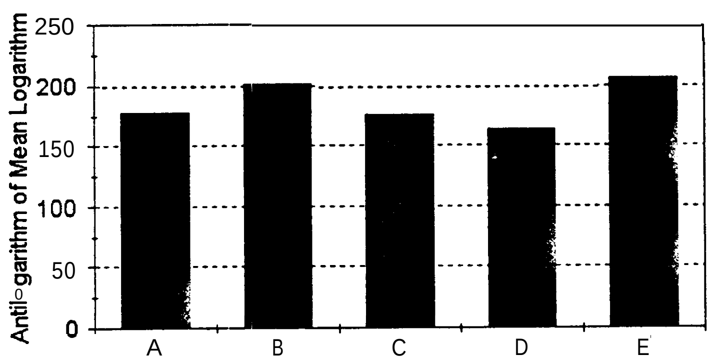
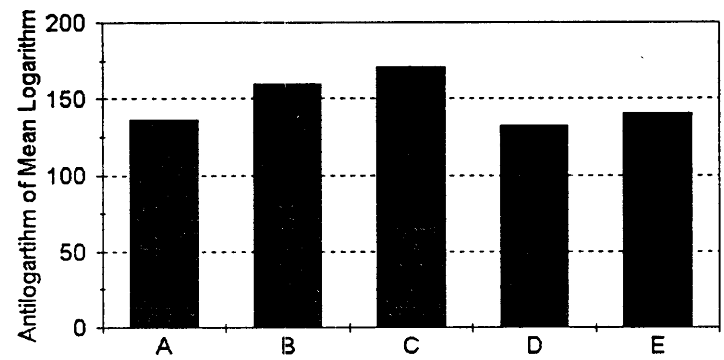


Figure 10



Of the seven combinations of between-subject factor levels for which there were statistically significant differences, all seven involved either Model' I or II helmets; all seven involved a flat anvil; five involved the second impact on the helmet; and five involved the rear impact location.

3. Impact Location Permutations

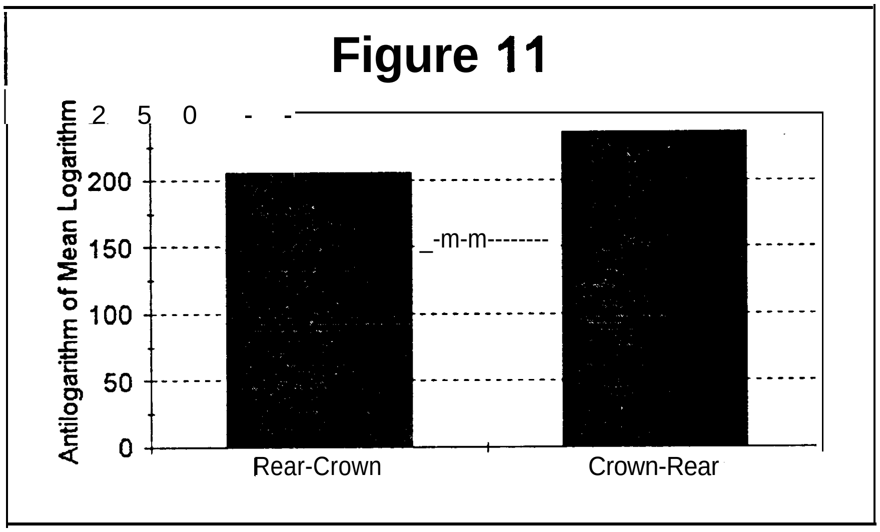
There were 69 contrasts tested for the crown measurements, and the same number tested for the rear measurements. There was a statistically significant difference found for the two permutations with two contrasts for crown measurements. The factor levels at which these statistically significant differences occurred may be seen in Table 4. Figures 11-12 show the antilogarithm of the mean logarithm for each permutation at each of these two combinations of factor levels.

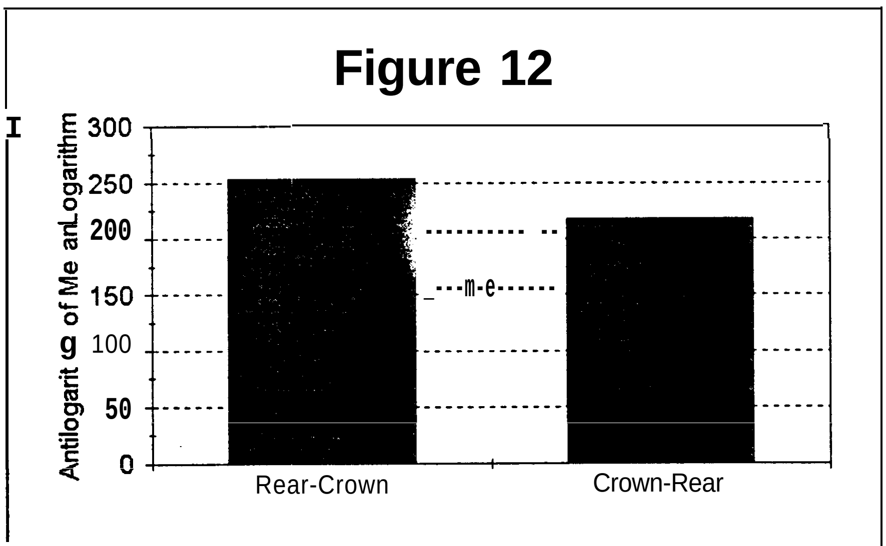
Table 4: Factor Levels at Which There Was a Statistically Significant Difference between Permutations in the Mean Natural Logarithm of the Peak G Measurements

Impact Location	Rig Type	Laboratory	Anvil	Model	Figure
Crown	Monorail	B	Flat	I	11
Crown	Twin-wire	B	Flat	I	12

Source: CPSC Interlaboratory Study on Bicycle Helmet Testing

Figures 11-12
Factor Level Combinations for Which There Was a Statistically Significant Difference between Impact Location Permutations





It may be noted that both statistically significant differences involved the crown impact location, flat anvil, and Model I helmets.

4. Models

There were 139 contrasts tested for the crown measurements, and 138 contrasts tested for the rear measurements. There was a statistically significant difference found among the three helmet models at nearly every combination of between-subject factor levels for each impact location. Typically, Model III helmet testing produced the highest mean, and Model II helmet testing produced the lowest mean. Often the difference in means for Models I and III helmets was relatively small. Statistical significance was usually achieved with contrasts involving Model II helmets and one (or either) of the other two helmet models.

5. Anvils

There were 68 contrasts tested for the crown measurements, and 66 contrasts tested for the rear measurements. There was a statistically significant difference found for the two anvils with contrasts at each combination of between-subject factor levels for each impact location. The flat anvil clearly produced higher means of the natural logarithm of the peak G measurements than the curbstone anvil.

B. Factor Level- Combinations after Appropriately Collapsing over Laboratories

After summing over laboratories, for each of the four remaining between-subject factors, tests were conducted on mean contrasts, controlling overall error rates for multiple comparisons (as discussed in the previous section). Results of the tests for each of the between-subject factors were as follows, again with antilogarithms given for ease of interpretation.

1. Test Rigs

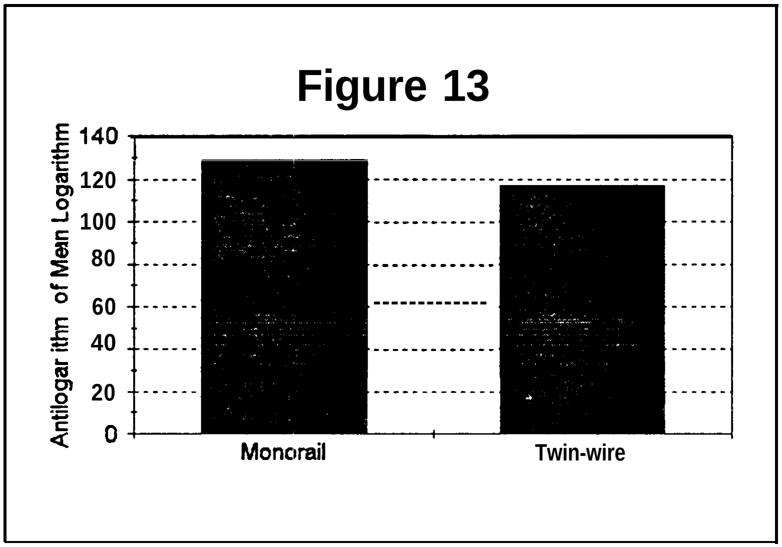
There were 12 contrasts tested for the crown measurements, and the same number tested for the rear measurements. There was a statistically significant difference found for the two test rigs with one contrast for crown measurements and one contrast for rear measurements. The factor levels at which these statistically significant differences occurred may be seen in Table 4. Figures **13-14** show the antilogarithm of the mean logarithm for each test rig at each of these two combinations of factor levels

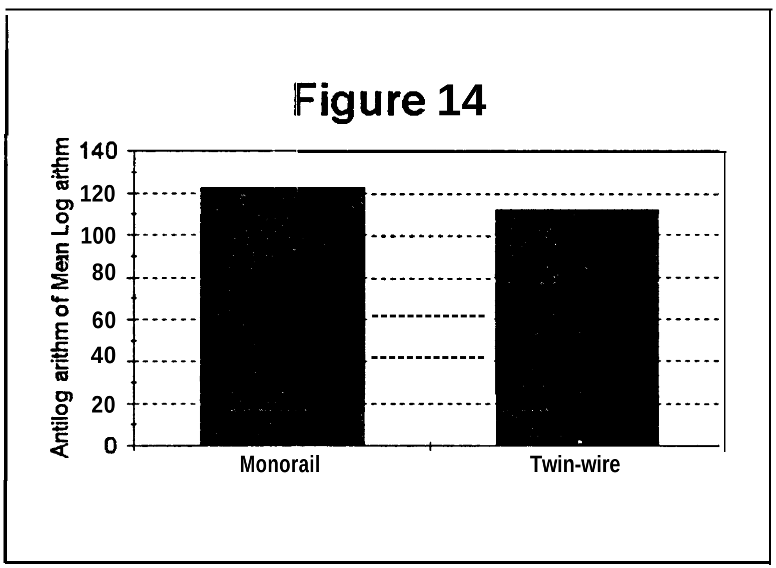
Table 4: Factor Levels at Which There Was a Statistically Significant Difference between Test Rigs in the Mean Natural Logarithm of the Peak G Measurements

Impact Location	Anvil	Model	Order	Figure
Crown	Curbstone	I	Rear-Crown	13
Rear	Curbstone	I	Crown-Rear	14

Source: CPSC Interlaboratory Study on Bicycle Helmet Testing

Figures 13-14
Factor Level Combinations, Summed Appropriately over Laboratories, for Which There Was a Statistically Significant Difference between Test Rigs





Each of the two contrasts yielding a statistically significant difference involved the Model I helmet striking the curbstone anvil on the second impact. Thus, for this particular combination of anvil and helmet model on the second impact, the monorail rig type produced a significantly greater mean logarithm of the peak G measurements than the twin-wire rig type at both impact locations.

2. Impact Location Permutations

There were **12** contrasts tested for the crown measurements, and the same number tested for the rear measurements. None of the tests yielded statistical significance.

3. Models

There were 24 contrasts tested for the crown measurements, and the same number tested for the rear measurements. At each combination of between-subject factor levels for either impact location, the mean for Model II helmets was significantly less than that for either Model I or Model III helmets. At some combinations of between-subject factor levels, the mean for Model III helmets was significantly greater than that for Model I helmets (particularly for the rear measurements).

4. Anvils

There were 12 contrasts tested for the crown measurements, and the same number tested for the rear measurements. There was a statistically significant difference found for the two anvils with contrasts at each combination of between-subject factor levels for each impact location. The flat anvil clearly produced greater means of the natural logarithm of the peak G measurements than the curbstone anvil.

V. Discussion and Recommendations

It would be preferable to test for the effects of a main factor simply with one test, but it is not appropriate to do so because of the complex interactions present in these experimental data, as described in the Appendix. The choice of test rig did not have an appreciable effect on test results in most situations, but did have an effect when the curbstone anvil was struck by Model I helmets on the second impact, with the monorail rig yielding a significantly greater mean logarithm of the peak G measure at each impact location.

Hence, since it is possible to have the two test rigs yield significantly different results under specific testing conditions, it seems advisable to specify the test rig in the test procedure or make some modification to the test procedure to ensure that the two test rigs give similar results. It is believed that this would provide improved standardization of the test conditions.

Additionally, while the main purpose of the present study was to examine the effects of using two different test rigs on test results, other recommendations on the basis of analysis of the data include the following:

(1) In the instrument systems check procedure, include provisions for accuracy as well as precision.

(2) Modify the test procedure in such a manner as to provide guidance to the tester in the selection of impact sites and order of use of anvil types.

(3) Test an appropriate sample of helmet model specimens under each test condition (instead of just one helmet model specimen) to take statistical variability into account.

Recommendation (1) pertains to ensuring that the data obtained in the instrument systems check procedure are sufficiently similar in different laboratories (instead of just ensuring that the data are sufficiently similar in repeated tests within the same laboratory, as the test procedure currently requires). Recommendation (2) deals with eliminating potential tester bias in the selection of impact locations, anvil types, and order of use on a helmet model specimen. Recommendation (3) ensures that more reliable results are obtained than those obtained by testing just one specimen under each test condition in CPSC Compliance testing.

Appendix
Special Statistical Features of Experimental Data

To fully understand the analysis of the data and the conclusions drawn, some special statistical features of the data need to be described. In particular, it is important to note that a basic assumption of the analysis of variance was not satisfied without transforming the data, and the presence of high-order interactions and empty cells resulted in the analysis being conducted differently from the way it would have been conducted otherwise.

A. Heteroscedasticity and the Need for the Logarithmic Transformation

A basic assumption underlying the analysis of variance is that there is homoscedasticity (equal variances) of the experimental errors at all treatments. As will be seen shortly, this assumption was clearly not satisfied with these data. An approach often used to deal with this problem of heteroscedasticity (unequal variances) is to transform the observations and apply the analysis of variance to the transformed data,[7] as was done with these data.

First, upon inspection of the data, it was clear that a positive correlation existed between the sample mean and the sample standard deviation. Nonparametric (Spearman rank) correlation coefficients of the sample mean and sample standard deviation were used to test for this, using data from each treatment with two observations (Pearson correlation coefficients were also used, but the validity of such correlation coefficients is questionable here because they require the assumption of normality, and the sample standard deviation would not be expected to have a normal distribution*).

The nonparametric correlation coefficients of the sample mean and sample standard deviation on crown and rear impacts, respectively, were 0.42 and 0.20, both significantly different from zero.[9] To try to reduce this dependence between the sample mean and sample standard deviation, the square root and (natural) logarithmic transformations were used. The nonparametric correlation coefficients of the sample mean and sample standard deviation on crown and rear impacts, respectively, were **0.28** and 0.06 after the square root transformation and 0.14 and -0.07 after the logarithmic transformation. Of these four rank correlation coefficients, only that on the crown for the square root transformation was significantly different from zero.[10]

Thus, since the logarithmic transformation appeared more effective than the square root transformation in reducing the dependence between the sample mean and sample standard deviation, the logarithmic transformation was chosen as more appropriate.

The logarithmic transformation is the one commonly used when the error standard deviation is proportional to the mean?

B. Interaction

Interaction is present in an analysis of variance whenever the difference in mean response between the levels of one factor is not the same at all levels of the other factors. When interaction is present, the effects of the factors are said to be nonadditive (as opposed to additive).

The presence of empty cells (discussed in C. of this Appendix) presents difficulties in testing for interaction effects. Analysis of variance was conducted, separately for data corresponding to the two impact locations, using data from four laboratories with no empty cells for any treatments. It was found that complex high-order interactions among the factors were present. Specifically, two four-factor interactions were present in the crown impact location data, and two four-factor interactions were also present in the rear impact location data (along with other lower-order interaction effects for both impact locations). Although variance-stabilizing transformations also tend to eliminate many interaction effects[12], most of the interaction effects were still statistically significant after using either the square root or logarithmic transformation on the data.

When factor effects are additive, tests for the effects of a main factor can be conducted without taking into account the levels of other main factors. When factor effects are not additive, the presence of interaction can mask the the significance of main factors.[13]

Due to the presence of complex, high-order interaction effects, tests were performed for appropriate contrasts of treatment means, as discussed in Section IV of the report.

C. Empty Cells

When there are no observations for some treatments in an experiment, it is said that there are empty cells for these treatments. In this experiment, empty cells arose because (1) two laboratories did not have the twin-wire test rig and (2) a substantial amount of inadvertent testing under inappropriate conditions occurred at one laboratory (with the data from such inadvertent testing discarded).

It turns out that when the analysis of variance is conducted on data with empty cells, computational difficulties are encountered using the customary effects model. Such difficulties are described in texts, recommending that a means model be used instead of an effects model.[14]

As a result, in the present study, the mean squared error was computed for the transformed data corresponding to each of the two impact locations. (This involved interpreting each treatment for which there was at least one observation as a level of one large factor; there was a total of 140 such levels at the crown impact location and 138 at the rear impact location.) The mean squared error for the transformed data at each impact location was then used to perform tests for appropriate mean contrasts, as discussed in section IV of the report.

Endnotes

[1]It may be noted that the mandatory bicycle helmet test procedure proposed by CPSC involves testing under all these conditions and more. It was not feasible to include all levels of all factors potentially of interest due to sample size and time limitations. Four test conditioning environments are given in the proposed test procedure, including low temperature, high temperature, and water immersion, in addition to ambient, as was used in this study. Testing also involves a hemispherical anvil, in addition to the flat and curbstone anvils. Additionally, any impact locations can be used above a prescribed test line, as long as the impact location is at least 120 mm from any prior impact location on a helmet model specimen (there is a total of four impacts on each helmet model specimen tested). For more information on conditions specified for the proposed test procedure, see the Federal Register Notice published December 6, 1995.

[2]As cautioned on p. 617 of *Applied Linear Statistical Models* by John Neter and William Wasserman (1974), a random effects model should be used only if the levels of the different factors do indeed represent random samples from the populations of interest.

[3]The main purposes of principal component analysis are data reduction and interpretation. Principal component analysis was performed here with correlation matrices instead of covariance matrices. When principal component analysis is performed with a correlation matrix, it may be interpreted as yielding the principal components of standardized variables (see, e.g., pp. *367-368* of *Applied Multivariate Statistical Analysis* by Richard A. Johnson and Dean W. Wichern, 1982). In a principal component analysis of a correlation matrix, the variance-covariance structure is explained through a few linear combinations of the original standardized variables. Principal components, and the percent of the total population variance they represent for standardized variables, are estimated by computing eigenvectors and eigenvalues of the correlation matrix.

Principal component analysis was performed four times, for each possible combination of an impactor type and a rig type. The correlation coefficients in each of the four correlation matrices were strongly positive, most of them 0.96 or greater. In each case, it was estimated that the first principal component represented at least 97 percent of the total population variance of the standardized variables. Also, each time, the first principal component was a linear combination giving nearly equal weight to each of the 10 standardized observations, suggestive of a population correlation matrix with equal correlation coefficients of any two of the 10 observations. Hence, it was concluded that it was appropriate to sum the 10 observations at each combination of an impactor type, a rig type, and a laboratory.

It may be noted that relatively small sample sizes were used to estimate correlation coefficients in this analysis, since there were only five laboratories that used the twin-wire rig and seven laboratories that used the monorail rig. If observations are accumulated over impactor type and rig type to give a total sample size of 24, similar results are obtained (all correlation coefficients are nearly equal to one; the first principal component is estimated to account for over 98 percent of the total population variance of the standardized variables; and the first principal component gives nearly equal weight to each of the **10** standardized observations). Accumulating observations this way involves assuming that different impactor types and rig types provide testing under sufficiently similar conditions. Principal component analysis performed with the covariance matrix instead of the correlation matrix also gives very similar results, regardless of whether observations are accumulated over impactor type and rig type.

[4]A mean contrast may be defined as a linear combination of treatment means such that the multiplicative constants defining the contrasts sum to zero (see, e.g., pp. *468* & *594* of *Applied Linear Statistical Models* by John Neter and William Wasserman, 1974). Only contrasts consisting of pairwise differences of treatment means were used here (i.e., they were simply the difference of two treatment means).

[5]The Bonferroni technique was used to control error rates for multiple comparisons. The error rate did not exceed 0.05 for the set of contrasts consisting of all pairwise differences of between-subject factor level means of a given factor under similar test conditions (i.e., all levels of the other between-subject factors held constant) in the analysis of variance conducted for a given impact location. Thus, e.g., the error rate did not exceed 0.05 for the set of contrasts consisting of all pairwise differences of the three helmet model means examined at each set of similar test conditions (holding levels of the four other between-subject factors constant) in the analysis of variance conducted at the crown impact location. A parallel statement could be made for the rear impact location and for any one of the other four between-subject factors.

[6]Since the natural logarithm transformation was used, tests were conducted to find statistically significant differences in means of the natural logarithm of the peak G measurements. The antilogarithm of the mean of a group of logarithms is the geometric mean of the group of original measurements.

assumptions, see texts on analysis of variance, such as *Design and Analysis of Experiments* by Douglas C. Montgomery (1976).

[8]As seen in theoretical texts, such as on p. 14 of The

assumptions, see texts on analysis of variance, such as *Design and Analysis of Experiments* by Douglas C. Montgomery (1976).

[8]As seen in theoretical texts, such as on p. 14 of The *Theory of Linear Models and Multivariate Analysis* by Steven F. Arnold (1981), when sampling from a normal population, as is assumed in the analysis of variance, the sample variance multiplied by the appropriate constants has a chi-squared distribution, not a normal distribution. The sample standard deviation would then be the square root of a chi-squared random variable, multiplied by constants.

[9]A test given on p. 301-302 of *Nonparametrics: Statistical Methods Based on Ranks* by E. L. Lehmann (1975) was used to test if Spearman rank correlation coefficients were significantly different from zero. The test statistic includes adjustments for ties and makes use of a normal approximation. The p-values were less than 0.001 and 0.024, respectively. The corresponding Pearson correlation coefficients here were 0.43 and 0.28. For the sample mean and sample variance, the corresponding Pearson correlation coefficients were 0.40 and 0.29.

[10]The p-values for the tests on the crown and rear, respectively, were 0.001 and 0.484 after the square root transformation, and 0.107 and 0.447 after the logarithmic transformation. The corresponding Pearson correlation coefficients were 0.29 and 0.11 after the square root transformation and 0.10 and -0.06 after the logarithmic transformation. For the sample mean and sample variance, the corresponding Pearson correlation coefficients were 0.27 and 0.10 after the square root transformation, and 0.07 and -0.07 after the logarithmic transformation.

[11]See, e.g., p. *507* of *Applied Linear Statistical Models* by John Neter and William Wasserman (1974).

[12]See, e.g., p. 61 of *Design and Analysis of Experiments* by Douglas C. Montgomery. Transformations used to stabilize variance and make error term distribution closer to normal often also reduce interaction effects.

[13]See p. 123 of *Design and Analysis of Experiments* by Douglas C. Montgomery.

[14]In *Linear Models for Unbalanced Data* by Shayle R. Searle (1987), a discussion of the use of Type I through Type IV sums of squares in SAS is given on pp. 461-465. It is noted that for Type IV sums of squares with data having empty cells, the sums of squares are not necessarily part of any traditional analysis-of-variance partitioning of the total sums of squares; they do not necessarily involve all the data; and altering the coding of levels of the variables can lead to different sums of squares.

Throughout the text, and in particular in the chart on p. 9, it is recommended that a cell means analysis be conducted when interaction is present with some cells empty.