6598 E Canton S. Boston Rd. | Salem, IN 47167| USA

Document ID: ASTM-002

Engineering Consulting Product Research Field Testing and Inspections

Phone: 812.528.2743 Fax: 866.331.0045 www.asetservices.com

ASTM F2569 Adopts Repeatability and Bias Levels:
What this means in the real world

In the fall of 2011 ASTM F2569, The Standard Test Method For Evaluating the Force Reduction Properties of Surface for Athletic Use, was modified to include a precision and bias statement. A precision and bias statement contains information regarding the repeatability and reproducibility of a particular test method. This inclusion of this information is significant as it is the first time an international standard has addressed the variability of Force Reduction, commonly called 'Shock Absorption,' test results. It is also significant because the repeatability and reproducibility levels presented in this ASTM standard directly apply to a wide variety of commonly used standards such as DIN 18032-2, EN 14904, and industry standards such as MFMA PUR©. This document defines repeatability and reproducibility, explains why they are significant, offers example comparisons using the precision limits, and offers examples of how to incorporate this information into project specifications.

Introduction:

Many sports surfaces have been evaluated for force reduction or shock absorption. Force reduction and shock absorption are used to describe the same property within the industry (this document will use Force Reduction in later references). A number of standards are currently used within North American to present this property: DIN 18032, EN 14904, ASTM F2772, and a newly adopted industry standard from the MFMA, a manufacturer's association.

While all of the above standards are used to evaluate this property, they all use the same method. In fact they all use the same calculations with the exception of DIN 18032-2. The DIN standard uses a correction factor determined from a calibration sample that is generally not available for labs outside of Germany. So, quite often all of the standard produce identical results. The differences between these standards is how they interpret these results. But this document is not focused on this aspect.

Because all of these standards use the same methods, they have identical error levels. Within the testing and standardization world, error levels are represented by repeatability, reproducibility, and bias measurements. These error levels are important and they have been ignored in every standard published until this most recent publication by ASTM.

Project specifications which treat force reduction as an absolute number and do not account for any inter-lab variability need to consider this new information.

Definitions:

Force Reduction (Shock Absorption): The ability of a surface to reduce impact forces when compared to a rigid surface, such as concrete.

Precision is a measure of how close readings of identical products are to one another. In a context of golf, precision is a reflection of how tightly grouped a collection of shots are on the green.

Bias is a measure of how accurately measurements are able to capture the measured property. In the context of golf, bias is a reflection of where the shots are grouped with relation to the cup, or put simply how far away from the cup the average ball comes to rest.

Repeatability is a measure of how much

deviation exists within a single lab. This level of precision is evaluated by having a single lab measure the property of interest one day and then measure the same property on the same sample a second day.

Reproducibility is a measure of the differences that exist between labs. This level of precision is evaluated by having samples that are closely controlled sent to and evaluated by multiple labs.

Target Value is the force reduction level that has been selected for an individual project. For example the surface of choice has a force reduction level of 35%, thus 35% is made the target level and is placed in the center of the repeatability and reproducibility ranges.

Bias measures how close measurements are to the exact actual value. A bias statement is not included at this time.

Statistical Significance:

In simple terms, values that fall within the reproducibility and repeatability ranges around a target value can not be concluded to be different from the target value because the difference could be due to random errors within the test method.

Generally speaking repeatability and reproducibility limits are calculated using a 95% confidence interval. Meaning that if a value falls outside of the target value there is at least a 95% probability the difference is not due to random test errors and that the materials are actually different.

The following tables summarize the repeatability and reproducibility information presented in ASTM F2569.

Table 1: Force Reduction Repeatability(%)

Surface	Drop 1	Drop 2
SS1	1.57	1.42
SS2	4.28	4.45
ST3	1.38	3.63
Average	2.79	

Table 2: Force Reduction Reproducibility (%)

Surface	Drop 1	Drop 2
SS1	5.48	5.16
SS2	4.94	5.3
ST3	3.33	4.82
Average	4.84	

As you can see the repeatability limit is smaller than the reproducibility limit. This is expected as reproducibility greatly expands the variables involved to include different equipment, technicians and samples.

It is ASET's position that the repeatability limit has little application within the sports surface industry. The primary reason is that repeatability allows only one variable and that is time. Even in cases where the testing is performed by the same company, there are different samples, temperatures, humidity, and perhaps different technicians and locations involved..

Example Case

This section provides an example of how to use reproducibility limits to evaluate differences in performance levels. It also illustrates the significance of setting a 'target value' for the

desired performance.

This example assumes the following:

- a target value force reduction of 50% was selected for the project
- the reproducibility limit of 4.8% will be applied to the target value

These two assumptions allow the upper limit and lower limit of acceptable systems to be calculated (upper limit = 50%+4.8% = 54.8%; lower limit = 50%-4.8% = 45.2%).

Next our example assumes that 4 samples were submitted for consideration. The samples and their force reduction levels are shown in the table below.

Sample	F-Reduction	Sample	F- Reduction
1	50.0%	3	54.6%
2	45.5%	4	42.0%

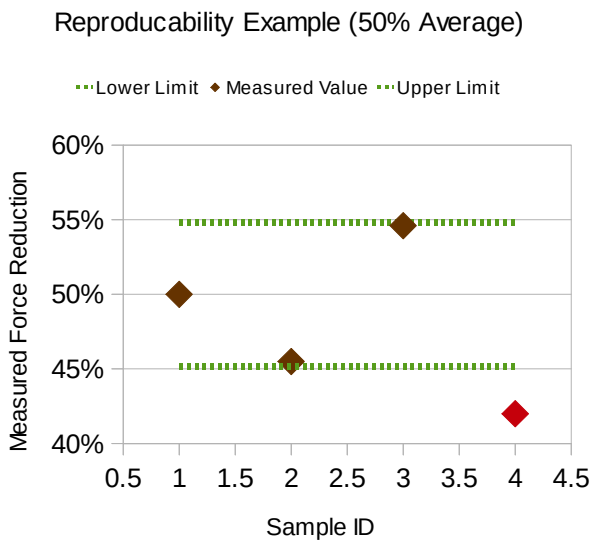

Figure 1: Reproducibility Example

Figure 1 shows the reproducibility limits plotted in green around the target value of 50%. It also shows the force reduction vales of the 4 systems under consideration. Note that all three systems within the reproducibility limit have been given the same color. Also, remember that the reproducibility limit is centered on the target value. There is a big difference between centering these limits on the target value and using the target value as the upper or lower limit, which should not be done.

Because Samples 1-3 all fall within the reproducibility limit as shown by the green lines in Figure 1, they are are considered to comply with the stated goal of providing a surface with a 50% force reduction level. Statistically, ASTM would consider all results within the green line in Figure 1 indistinguishable from the target value . Practically, all results within the green line in Figure1 should be considered equivalent.

It is important to note that the reproducibility limits are established around the target goal for the project or within the specification. In this case the target was 50%, thus the limits were set at 45.2% and 54.8%. And because the target in this goal was set at 50%, sample 1 at 45.5% and sample 3 at 54.6% are considered equal to the 50% target. Even though sample 2 and 3 are 9.1% different. However, at no time does this suggest that samples 2 and 3 should be considered to produce results equivalent to each other.

Sample 4 with a force reduction level of 42% (8% from the target value) falls outside the established limits and should not be considered to provide the desired performance, even though it differs by the target level of 50% by only 8%. It is not considered equal to the 50% target even though it is within 2 times the repeatability limit of the 50% goal.

Practical Application

There are a number of ways that owners and architects can incorporate this new information into their projects and specifications. The following section contains some suggestions from ASET Services. These suggestions apply

when specifications include specific target values (ex 50%), not when specifications include more generic language such as 'must pass DIN 18032' or 'must achieve class 2 rating according to ASTM F2772.'

Scenario 1: When any two systems with force reductions that vary by less than 4.8% are submitted, they should automatically be considered to provide statistically equivalent force reduction performance.

Scenario 2: When any two systems with force reductions that vary by more than 9.6% are submitted, they should be considered to provide significantly different force reduction levels.

Scenario 3: When any system is submitted with a force reduction level within 4.8% of the target level, it should not be excluded from consideration as the level falls within established reproducibility levels.

Application in Specifications:

Many sport surface specifications simply call for the submission to meet the requirements of a certain standard, or a certain class/type within that standard. In these instances, no modification is needed. Even though there is a known error in the test methods, labs are not allowed to simply apply that error factor to their results to make it easier for systems to pass a given standard.

However, some floor systems that fail to meet international guidelines include force reduction performance levels within their specifications, most often in a section title 'Technical Data' or something similar. The following is a force reduction line from such a standard:

Criteria	Level	Standard
Force Reduction	35%	(DIN 18032)

Such specifications fail to account for the variability of the testing method. The technical data contained in a specifications should be modified to contain an allowable range, and this should be for every property contained within it. An example is shown below

Criteria	Level	Range	Standard
Force Reduction	35%	+/- 4.8%	(DIN 18032)

Application in Field Testing:

Beyond specifications, owners and architects should consider that these reproducibility limits have significant implications in the field. At the time of writing this document, the North American sports surface industry does not validate actual performance against advertised performance.

This lack of verification means that actual and promised performance can, and probably do, vary greatly and that there is no active effort made to validate performance. With this new information, owners have technical support to require that their individual performance be no lower than 4.8% or 5% below the specified target value. This would ensure that the performance delivered was at least equal to what was specified. Such a specification would be a deterrent for promoting systems at the bottom edge of the reproducibility window.

More information on why your floor may not perform as advertised can be found in ASET Services document number Gen-002 within the ASET library.

For more information regarding repeatability and reproducibility, or on any of the test methods, please feel free to contact us. ASTE also offers pre-installation and post-installation testing services, as well as educational services.

Phone: 1.812.528.2743,
e-mail: info@asetservices.com,
web: www.asetservices.com.