Why “Shipping Tests” Don’t Work

Herbert Schueneman,
Founder & Board Chairman

v2.1

WESTPAK
TESTING. OWNERSHIP. INTEGRITY.

83 Great Oaks Blvd
San Jose, CA 95119
800.830.8021

10326 Roselle Street
San Diego, CA 92121
800.830.8021

westpak.com

Abstract

"Shipping tests" or "trial shipments" are often used to determine if a product/package system will survive the distribution environment without damage to the product. The reality is that a single shipping test is statistically insignificant; utilizing this approach will often lead to incorrect conclusions unless further testing is conducted. This paper will explore the reasons for this while recommending a much more feasible approach to evaluating product/package systems.

Background

A typical product development cycle includes the design and evaluation of the product/package system. In the case of consumer products, the display package is often an integral part of the product and may be as significant. For non-consumer or industrial products, the package is sometimes an afterthought often delegated to a packaging supplier or perhaps a junior level staff member. Hopefully someone at a product development review will ask the question, "How do you know this product will arrive to the customer undamaged?" If the response is, "We did a ship test and everything looks fine" more questions must be asked!

During the early years of single parcel shipment, the default shipping test was almost always to ship the product via UPS through the Chicago hub. For whatever reason, this facility had the reputation of being able to chew up and quickly dispose of almost any package that wasn't rugged, so it was a good worse-case scenario. Another favorite ship test alternative was shipping via parcel post through the central New York area and back again. Lots of other scenarios existed as well.

By the mid-1960s, considerable effort was directed toward defining exactly what the shipping environment consisted of, how it might be quantified, and most of all how it might be brought into the laboratory in a controlled and reproducible manner.

The simple fact is that a shipping test is the best possible hiding place for significant product, packaging or distribution-related problems, regardless of the product, the packaging, the shipping environment, or almost anything else.

Introduction

The problem with a single shipping test is probability; significant variability exists with hazards in the distribution environment involving impacts, vibration, and similar. In the single parcel case, impacts are mainly a function of humans; sometimes people handle packages gently and sometimes not. With the results of a single shipping test, one would be making conclusions based on the people who handle that particular package - their characteristics, mood, time of day, and a wide variety of other factors, few of which will be consistent from one shipment to the next.

Vehicle induced vibration represents another large variable in the distribution environment. In the best scenario the vehicle in which the product travels is well equipped, properly loaded, and driven on well-maintained roads throughout the journey. In non-optimum scenarios, the vehicle might be poorly maintained, have unbalanced wheels, lightly loaded, and traveling over marginal roadbeds, all of which may cause vastly increased levels of vibration induced stress on the product.

Figure 1 – Single Parcel Delivery

To properly characterize these and other hazards, a large sample size is necessary combined with environmental extremes. As the sample size increases, so does the cost of obtaining this data. In addition, determining the validity of data at the extremes of the statistical distribution can be very difficult.

With proper data collection techniques and sample size, all of the hazards of the common distribution requirements can be successfully quantified in terms of their statistical probability as well as the mean and extremes of the data. Specifically, the statistical significance of the data collected can be determined from the following standard deviation equation where (n) is equal to the number of samples and (x) indicates the actual measured values:

$$S = \sqrt{\frac{\Sigma (X - \bar{X})^2}{n - 1}}$$

Equation 1 – Standard Deviation

It is apparent from the above equation that when the number of samples (n) is one, the standard deviation is infinite (dividing by 0) meaning the data has no significance. Herein lies the largest single issue with a shipping test: it is totally insignificant unless a large number of samples make up the data. Therefore, a single shipping test is just as likely to steer you in the wrong direction as anything else, regardless of the test results.

> **"Everyone I have been in communication with are always friendly, willing to help, and most importantly are really open to communicate and provide updates to the testing."**
>
> M.P.

A Better Approach

Beginning in the 1950s, considerable effort was expended toward a unified quantification of the distribution environment hazards. Numerous studies, many from military sources, attempted to quantify inputs such as vibration, impacts, temperature and humidity extremes, altitude or negative pressure levels, top load compression levels, and similar. Resulting from this effort was a series of papers and articles that, when taken in total, offered a reasonably good quantification of distribution environment hazards. Specifically, impact levels were quantified in terms of drop height as a function of package weight, orientation of the impacts, and total number of impacts. Vibration levels in vehicles were quantified in terms of their input spectra, duration, and overall intensity. Temperature and humidity extremes were quantified in terms of their duration and intensity. Altitude levels and top load compression were similarly quantified.

Figure 2 – Setting Up a Vibration Test

When put together in a coherent format, these quantified hazard levels will uniquely define the input levels for both mean and extreme values as well as the sequence of the likely hazard inputs. When this data is correlated to a package system of similar size and mass as the anticipated one, then a reasonably good statistical probability exists that a laboratory test using these inputs will indeed duplicate the hazards of the intended distribution environment.

Thus, one can take these hazard inputs, properly quantified, and assemble them into a sequence representative of the anticipated distribution environment and result in what is commonly referred to as a **Package Integrity Test** protocol. This can be thought of as an attempt to **bring the distribution environment into the laboratory** in a controlled and repeatable format. When the product/package system in question is subjected to this input, the product development team will have a high level of confidence that successful completion of the laboratory testing will enable successful distribution of the product a majority of the time.

Another significant advantage of laboratory testing is the speed in which the tests can be conducted compared to field studies. The typical mechanical portions of a package integrity test (impacts, vibration, and compression) will take less than one day to complete in the lab. If temperature and humidity exposures are included in the protocol, a two or three-day test is not unusual due to the time necessary for the package to equilibrate to the extreme conditions.

Figure 3 – Pallet Compression Test

In addition, laboratory testing is very repeatable, a highly desirable characteristic that should not be overlooked. The repeatability of the test inputs will allow the engineering team to evaluate trade-offs in the product/package design including different materials, suppliers, processes, palletization techniques, and a variety of other variables including whether or not to proceed with the project in the first place. This is perhaps the best reason for using the laboratory testing technique over the “tried and true” shipping test.

The term Package Integrity Test has been used often and refers to a procedure whereby a package system is subjected to a test protocol simulating a known distribution environment. It is a qualitative test in that no measurement of product response is called for. The end result of the test is a pass/fail analysis based on examination of the product. There is no quantitative data available. A variation of this test, referred to as a **Package Performance Test**, will often use instrumentation attached to the product itself to measure the level of the hazard transmitted through the package and into the product. This information is useful for margin analysis (over-packaging) if the sensitivity of the product to that particular input is known. This technique is valuable for cost reduction projects and other product/package optimization efforts.

The cost of a typical Package Integrity Test protocol will likely be approximately three to five times that of a typical shipping test, depending largely on the mode of transportation and the duration of the shipping test. The confidence level and value of the data collected in the Package Integrity Test will be many times higher, however.

Figure 4 – Rotational Impact Testing

WESTPAK finds that many companies still prefer the use of a ship test over other techniques including laboratory analysis. Surprisingly, the ship test is often used in addition to a laboratory test protocol. Much of this, we find, is for the personal satisfaction of managers who do not understand the background of a well-designed Package Integrity Test conducted in the laboratory. The concept of a ship test is bone-simple and understandable with a minimum of explanation. The potential danger of using only a shipping test for decisions regarding the successful distribution of products is substantial. The inherent limitations of a shipping test dictate that the data be restricted to validating laboratory testing.

"I have used Westpak for various testing and have always been satisfied with your customer service, handling of my test articles, and your quick-turn reports. I will continue to be a customer of yours."

J.C.

Conclusions

For the vast majority of products, the distribution environment represents the highest level of hazards the product will experience in its lifecycle. The product development process must include an evaluation of the product/package systems' performance against these distribution environment hazards.

It has been shown that the use of a single shipping test alone for the verification of packaging adequacy is a statistically invalid and dangerous approach. A highly recommended methodology is to include a laboratory-based Package Integrity Test as a component of the product/package design cycle. The relatively quick, repeatable nature of laboratory testing enables an economically viable means of evaluating alternative designs, materials, processes, or suppliers, along with high confidence levels of the test results.

It is suggested that the use of a shipping test be restricted for validation of laboratory testing.

Thorough testing of the product/package system is necessary to increase the likelihood the customer will receive a damage-free product. A laboratory-based Package Integrity Test is one of the best tools available to simulate the effects of distribution environment hazards on both consumer and industrial products.