Improved sensitivity testing of explosives using transformed Up-Down methods

Geoffrey W Brown
High Explosives Science & Technology (WX-7)
Los Alamos National Laboratory, Los Alamos, NM, 87545, USA

E-mail: geoffb@lanl.gov

Abstract. Sensitivity tests provide data that help establish guidelines for the safe handling of explosives. Any sensitivity test is based on assumptions to simplify the method or reduce the number of individual sample evaluations. Two common assumptions that are not typically checked after testing are 1) explosive response follows a normal distribution as a function of the applied stimulus levels and 2) the chosen test level spacing is close to the standard deviation of the explosive response function (for Bruceton Up-Down testing for example). These assumptions and other limitations of traditional explosive sensitivity testing can be addressed using Transformed Up-Down (TUD) test methods. TUD methods have been developed extensively for psychometric testing over the past 50 years and generally use multiple tests at a given level to determine how to adjust the applied stimulus. In the context of explosive sensitivity we can use TUD methods that concentrate testing around useful probability levels. Here, these methods are explained and compared to Bruceton Up-Down testing using computer simulation. The results show that the TUD methods are more useful for many cases but that they do require more tests as a consequence. For non-normal distributions, however, the TUD methods may be the only accurate assessment method.

1. Introduction
The goal of explosive sensitivity testing is to characterize the response of a material to an external stimulus. The stimulus can be impact, friction, spark, shock, or any other action that imparts energy. The response of the explosive is statistical due to variabilities in the material, the sample preparation, and the stimuli. As a result, “characterization” is the process of determining parameters describing the statistical response function. These parameters are then used for relative comparison of explosives or to assess the likelihood of an energetic reaction at stimulus levels that were not specifically tested.

Typically some initial assumptions are made about the response of the material to reduce the number of required tests or to provide a simpler characterization. Often the first assumption is that the explosive exhibits a Gaussian response using the chosen stimulus level spacing. Another common assumption is that the spacing between stimulus levels is close to some optimal value (e.g. the standard deviation of the response function). Assumptions such as these can cause problems with interpretation of the results, especially when the goal is to estimate response at stimulus levels outside the normal testing range. If the assumptions are not checked, or cannot be checked adequately given the details of the testing, erroneous conclusions may be obtained.
There are many different sensitivity testing methods and popular approaches that undergo continual modification include Bruceton Up-Down, Neyer D-optimal, Probit, and Robbins-Monro [1-4]. Each of these methods suffers from either a requirement that a probability distribution be assumed, limitations on the probability levels that can be probed, or the need for a large number of tests. For example, typical application of the Robbins-Monro method is only good for determining the 50% reaction level while the Probit test requires many tests at many different levels.

An optimum method for the type of sensitivity testing required by many laboratories today would not require an assumed distribution, not require a step size, be simple to implement, only require a moderate number of tests and provide the ability to access a range of probability levels.

A method that has many of these characteristics is Transformed Up-Down (TUD) testing. TUD testing was developed in the early 1960’s for psychometric applications – measuring the statistical response of living subjects to external stimuli such as sound, smell, etc [5]. In this application, the response function is often non-Gaussian, step size adjustment may be limited, and subjects fatigue or “learn” if too many tests are carried out. Since these overlap some of the issues noted above, many of the optimizations of TUD testing can greatly benefit explosive sensitivity testing.

2. Transformed Up-Down Testing

The origins of Transformed Up-Down methods were developing in 1963 [6], outlined in detail in 1965 [5], and have been applied in many different psychometric arenas since then [7]. Simply, TUD testing is an Up-Down method in which multiple tests at a single level are used to make a decision about increasing or decreasing the stimulus. This provides access to probability levels other than the 50% reaction level, especially when the decisions about different possible sequences of results at a single level are assigned in different ways. Bruceton Up-Down testing can be viewed as one case of TUD testing in which only a single level is used for decision making.

A simple example shows TUD methods. Table 1 shows all cases for TUD testing with a single test at each level and all cases when three tests at each level are used. Reactions are labeled with G for “Go” and lack of reaction is N for “No-Go”. Table 1 is only for cases where levels must be changed up or down by a single level when a decision is made. The last column of table 1 shows the probability probed by each case, with the calculation of these results described below. By examination, the single test case is just the Bruceton Up-Down method.

| Move Up if observe | Move Down if observe | Probability Level Probed |
|-------------------|---------------------|--------------------------|
| N                 | G                   | 0.5                      |
| NNN               | Anything else       | 0.206                    |
| NNN, NNG          | Anything else       | 0.293                    |
| NNN, NNG, NGN     | Anything else       | 0.405                    |
| NNN, NNG, NGN, GNN| GGG, GGN, GNG, NGG | 0.5                      |
| Anything else     | GGG, GGN, GNG       | 0.595                    |
| Anything else     | GGG, GGN            | 0.707                    |
| Anything else     | GGG                 | 0.794                    |

Considering that rules can be assigned based on number of tests and changing the number of steps to be moved, there can be many rules per case and there are then a very wide range of TUD methods available that could be tailored for specific purposes.

The probability probed by a given method can be determined by noting that the testing must stabilize around a stimulus level that produces up and down steps with equal probability. By
construction, when tests are carried out at a level that produces one or the other result more often, the
stimulus level is adjusted in the direction to equalize the results. Testing far from the stable level in
either direction will produce steps in a direction appropriate to bring the testing back to the stable
level. With that in mind, for any set of \( n \) rules within a single case whose probabilities sum to unity,
the general equation for calculating the probed probability is

\[
\sum_{n, \Delta_n \neq 0} \frac{1}{\Delta_n} P_n = 0.5
\]  

(1)

Where \( P_n \) is the probability of a reaction at the stable stimulus level, \( \Delta_n \) is the change in stimulus
level and rules that have \( \Delta = 0 \) are not included in the sum since they do not contribute to the level
adjustment. Note that for this general equation, \( \Delta \) can be any value (including non-integer) and must
have a sign, either + or – , that denotes the direction of adjustment. (equation (1) would be the same if
\( P_n \) is taken as the probability of no-reaction.) In either case, all possible outcomes must be considered
so that a total probability of 1 is available. As an example, for the case of the third row in table 1,

\[
(1 - P)^3 + P(1 - P)^3 = 0.5
\]  

(2)

Where the first term is the probability of three consecutive non-reactions at the stable level and the
second term is the probability of two non-reactions and one reaction at the stable level. Solving this
equation produces \( P = 0.293 \), which is the probability level that produces equal numbers of Up and
Down steps for the case in the third row of table 1. From the entries in table 1, seven different
probability levels may be probed.

Finally, since the testing stabilizes around the stable stimulus level that produces equal Up and
Down steps, the actual stable level may be determined by averaging either the levels used in testing or
the levels of the local extrema where there is a reversal in the Up and Down motion of the sequence.
Wetherill claims in [5] that the latter case is more efficient than maximum likelihood estimates.

From table 1 we see that many of the desirable testing characteristics are achieved with TUD
methods. No distribution need be assumed if several probability levels are tested – the form of the
distribution may actually be evident in such a case. Any step size is valid since levels or extrema are
averaged to produce a result. The testing can be carried out with a simple spreadsheet and calculator.
Many probability levels can be accessed since many rules are available.

3. Comparison with Bruceton

We can use computer simulations to compare TUD to Bruceton Up-Down. We have done this by
modelling the response of an explosive with a probability distribution and simulating its testing by
statistical sampling. In the sections below, we examine step size, number of steps, and the case of
assuming the wrong distribution as variables in test methods.

3.1. Step Size.

Figure 1 shows the effects of varying the step size for both TUD and Bruceton-Up Down. In this case
a normal distribution with a mean of 2 and a standard deviation of 0.4 was used to model the response.
Then, using each method, the level corresponding to \( \mu + \sigma \) was evaluated. 100 tests were used for each
Bruceton trial. For a TUD trial, 100 tests were also used but they were split between the 70.7% and
29.3% levels. These levels can be simply transformed to the \( \mu + \sigma \) levels. In each case, 10,000 trials
were simulated. In figure 1, for TUD the \( \mu + \sigma \) level vs. step size is plotted as the red line and the
orange lines are plotted using the standard deviation of the results of 10,000 trials. For Bruceton, each
trial results in a \( \mu, \sigma \) pair which are added together. The average level is plotted as the blue line and
the green lines are constructed using the standard deviation of the 10,000 pair results. From this
figure, both methods have similar accuracy but different behavior with step size – the TUD method is
better for low step size and the Bruceton is better for larger step size that approaches $2\sigma$. Bruceton is known to work best for step sizes between $0.5\sigma$ and $2\sigma$.

![Step Size Comparison for Gaussian Response](image)

**Figure 1.** Step size comparison using 10,000 trials of 100 tests each for TUD and Bruceton.

### 3.2. Number of tests.
Using a step size of 0.3, the effect of the number of tests per trial was examined for both methods. This is shown in figure 2 where again, the red/orange lines are the $\mu+\sigma$ range for TUD and the blue/green lines are the $\mu+\sigma$ range for Bruceton. The 0.3 step size was chosen since both methods have similar behavior at that size. For comparison, the individual points plotted at 15 tests per trial in figure 2 are calculated for a step size of 0.8 and assigned the same color scheme. Even though Bruceton performs better with a step size of 0.8 at 100 tests per trial, both methods are similar when 15 tests are used.

### 3.3. Non-normal distribution.
The differences due to assuming an incorrect distribution shape were tested by modelling the response with a Weibull distribution with parameters $\lambda = 1, k = 1.5$. The TUD method was used to test the 70.7%, 50%, and 29.3% levels with 16 tests at each while the Bruceton was allowed 48 tests per trial. The results are shown in figure 3 with the red line as the Weibull cumulative distribution function, the green points as the Bruceton results and the blue points as the TUD results. Overall the TUD method is close on all three levels, missing slightly low away from the mean. The Bruceton method is only accurate at the $\mu-\sigma$ level, missing significantly at and above 50%.

### 4. Summary
The details provided above suggest a useful general algorithm that can probe explosive sensitivity without making limiting assumptions or succumbing to restrictive requirements common to many methods employed today. The implementation of the method is simple and interpretation requires only a calculator. Some additional tests per trial may be required depending on the goals of the sensitivity testing, but the benefits may outweigh the extra effort when predictions are required.
Figure 2. Number of tests comparison using 10,000 trials at 0.3 step size for TUD and Bruceton.

Figure 3. TUD and Bruceton estimates of probability levels using a Weibull response.
References
[1] Dixon W J and Mood A M 1948 *J. Am. Stat. Assoc.* **43** 109
[2] Neyer B T 1994 *Technometrics* **36** 61
[3] Stromsoe E 1992 *Prop. Expl. Pyro.* **17** 295
[4] Schilperoord A A, Biilsma M, Verweii H D, and Bruckman H W L 1982 *Prop. Expl. Pyro.* **7** 46
[5] Wetherill G B and Levitt H 1965 *The British J. Math. And Stat. Psych.* **18** 1
[6] Wetherill G B 1963 *J. Royal Stat. Soc.* B **25** 1
[7] Brown L G 1996 *Perception and Psychophysics* **58** 959