Detergent performance in the home

The following, prepared by K.L. Mills, G. Gladstone and O.W. Neiditch of Lever Brothers, is based on a presentation given by Mills.

It is generally accepted that detergency testing should follow a progression from relatively simple to more complex methods. Thus, it is normal to begin evaluation of a detergent formulation in a bench-scale apparatus such as a Terg-O-Tometer using artificially soiled test cloths. The evaluation will then proceed to laboratory-scale tests using naturally soiled split articles or a matched-article bundle test. The final step in this progression is the presentation of the formulation to the consumer in a blind package panel test, in which a subjective assessment of its performance is solicited.

This approach is not without its problems.

Because of these problems, we have developed a method to actually measure the performance of detergent formulations as they are used by consumers. Although this method still has some of the problems—such as variability of results with naturally soiled articles—seen with laboratory test methods, it does have the virtue of being a direct measure of detergent performance. Signal differences are avoided, and no correlations are required. Instead of using complicated experimental designs, all of the factors of consumer practice are allowed to vary in a random manner. In this way, we emerge with an objective assessment of detergent performance which includes all of those variations found in consumer practice. It is even practical to control habits with this method should it be desired, for example, by requesting that panelists not use bleach during the test period.

The first step in developing this method was the recruitment of a panel of consumers who would actually perform the product testing. Since the plan was to visit panelists at the beginning and end of the test period, recruiting was restricted to a 50-mile radius of the Lever Research Center in Edgewater, New Jersey. We were fortunate that this area provides a wide range of demographics, consumer habits and, most importantly for testing detergents, water hardnesses.

Recruiting was conducted by mail, using both blind mailing and referrals from previously recruited panelists. Each potential panelist was asked to fill out a detailed questionnaire regarding family demographics, laundry and dishwashing equipment, product usage and user habits. Information from the questionnaires was coded and entered into our in-house computer. Our original target was to recruit a panel of approximately 3,000 families, a figure which took two years to reach. Once established, the panel has a 10-15% a year attrition rate, due mainly to families moving out of the testing radius. This meant we had to continue our recruiting effort, although at a lower level than the initial recruitment.

The computer data-base created from the questionnaire information allows us to review the habits and demographic make-up of the panel and determine how typical these are of national averages. Although this is valuable for establishing the validity of our panel, it is even more important for structuring a test panel based on selected demographics. After some initial range-finding experiments, we have decided 100 panelists are the minimum number required for testing each laundry detergent. This test panel can then be selected by water hardness, detergent form normally used (powder or liquid) and bleach usage (liquid, powder or none). Although it is possible to balance the panel by a wide range of variables, the aforementioned ones are the most critical in selecting a panel for testing laundry detergents.

From conception to completion, a typical in-home test takes approximately seven months. The test protocol states the test parameters and variables, such as water hardness balance, test articles and family size. A typical in-home test might compare two powdered laundry detergent formulations on a monadic basis. In this case, a test panel of 200 subjects would be chosen from the computer data-base—100 to test Formula 1 and 100 to test Formula 2. The 200 subjects would be selected from those who normally use powdered laundry detergents, with an equal number of panelists chosen from soft-. medium- and hard-water areas.

Although this methodology can be applied to any washable article on which reflectometer readings can be taken, we normally include articles commonly found in the consumer wash. These include hand towels, dish towels, pillowcases, t-shirts and other outer or undergarments. These articles can be supplied new to the panelists for the test period or the consumers' own articles can be used. In the latter case, it is necessary to pick up the articles prior to the test period, take reflectometer readings on them and return them to the panelists for the test period. Our experience has shown that maximum information can be gained by including both new and panelists' own articles in a test protocol.

Once the test panel has been recruited, the detergent made and the initial reflectance and fluorescence of the test articles read, the test is placed into the field. This is carried out by our own personnel who personally visit each panelist to deliver instructions, test articles and enough detergent for the period of the test (normally 12 weeks). During this visit, our technicians may also observe wash habits, take water samples or make more detailed explanations of the test procedure.

Once the test period begins, the panelists are asked to use the test detergent for all of their washes.

The test articles are worn or used in the same way the panelists would use their own articles and included in whichever wash their own articles would be. Panelists are requested to keep a diary recording how many times each article is washed. At approximately the midpoint of the test, each panelist is mailed a monitor cloth which may contain one or more test cloths including clean and artificially soiled swatches. This monitor cloth is included in a single wash and then mailed back to us. The use of this monitor cloth is not intended to measure the relative performance of the detergents being tested, but rather to determine if the products are performing in the field as they did in the laboratory. Results with these cloths are in no way predictive of results on the test articles and have, in some cases, given results contrary to those obtained on the test articles.

At the completion of the test period, our technicians visit the panelists to pick up test articles and make any required observations. The test articles are then returned to our laboratory, where reflectometer readings are taken again. At this point, visual observations may also be made on the test articles. Once the reflectance and fluorescence values of the articles have been read and the data analyzed, both the garments supplied by us and the panelists' own articles are shipped back to the panelist. Giving the articles which we have supplied back to the panelists increases panelist loyalty and acts as a source of suitable panelists' articles for future tests.

The determination of the relative performance of laundry products using this procedure involves hundreds of articles and does not lend itself to visual examination by expert judges or panelists for routine use. By necessity, rapid and repeatable instrumental readings of the appearance of the articles are required. No single equation or chart for whiteness can represent the many observing conditions and observer preferences which exist.

As with any other color, three numbers are necessary for the complete identification of any white. The uniform color-scale system most commonly used for whites is the Hunter L, a, b system. Because fluorescence effects can be confused with the reflectance response resulting from the cleaning compounds, all measurements are made both with and without the presence of an ultraviolet filter. Cleaning response (whiteness) is measured with this filter in place, and fluorescence effects can be determined from the difference in whiteness with and without the ultraviolet filter. We have seen good correlation between the two in cases in which we have compared whiteness readings to visual observations.

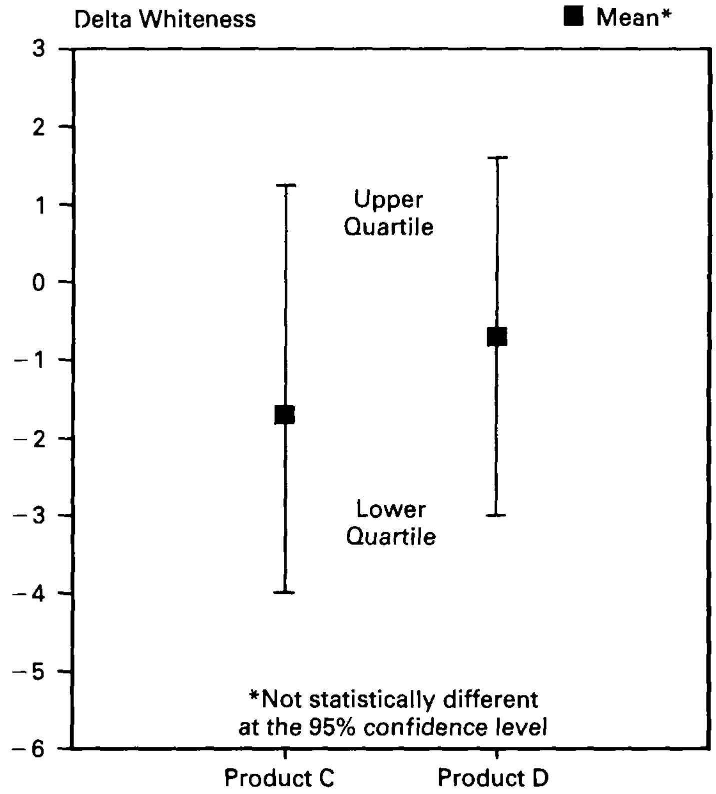

FIG. 1. In-home testing of articles.

Analysis of the data from in-home testing does not lend itself to standard quantitative statistical techniques. In most cases, calculation of an average and standard deviation for each detergent formulation will show that all formulations tested are statistically equal (Fig. 1). However, the use of descriptive statistical techniques can be very useful in interpreting the data. When the data from Figure 1 are normalized by calculating percentages and plotted as a cumulative distribution function from lowest to highest result, some interesting differences begin to appear (Fig. 2).

Detergent formulas which were statistically equal are now seen to be different. In particular, we have noticed in a number of tests that the 20–40% of the population which achieves the poorest cleaning results shows the greatest difference between detergent formulations. On reflection, this result is not surprising as that segment of population in which cleaning is most difficult offers the greatest opportunity for the chemical ingredients in a detergent formulation to function to their maximum effect. Thus, this segment offers the best opportunity to measure the relative effectiveness of these chemical ingredients.

Further insight can be gained from in-home testing data through the use of frequency distributions or histograms. We can take the data from Figure 1 and divide the population into cells depending on performance. In the case of panelists' own articles, we have grouped the population by those who obtained an improvement in performance, those who had a deterioration in performance and those whose articles did not change with use of the test formula. In the case of newly issued articles, only two cells are used as new articles will not normally show an improvement in performance.

The frequency of these cells is then expressed as a percentage of the test population. The frequency distributions may then be described as pie charts or bar charts (Figure 3). Expressed in this way, we can see that the majority of the test population shows no difference between these detergent formulations. However, the percentage of the population showing an improvement in performance can vary dramatically with detergent formulation. And the ratio of this frequency to the frequency of those panelists whose articles became dirtier is an important indication of a formulation's performance.

We have found that in-home testing is a valuable tool for evaluating the relative performance of laundry detergent formulations. It provides a direct comparison of formulations under conditions actually used by consumers and eliminates the need for performing complicated correlations between laboratory screening tests and subjective consumer tests. Although the test procedure is time-consuming and expensive, it has the virtue of yielding a true assessment of detergent performance in the hand of consumers. In some cases, it has prevented us from being misled by laboratory test results which were not duplicated in the home.

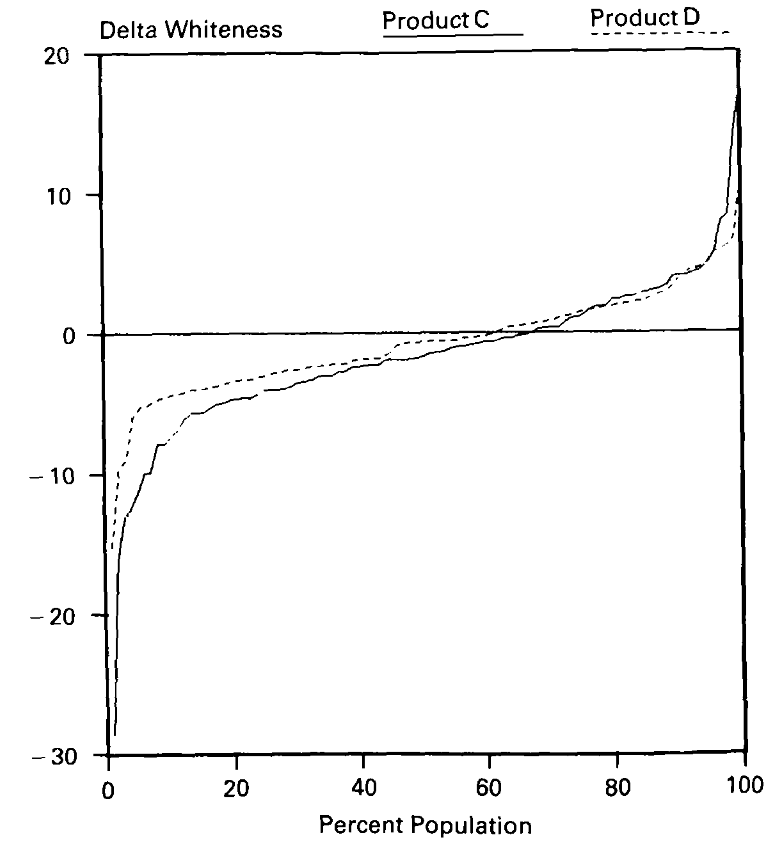

FIG. 2. In-home testing of articles.

Product	Percentage of panelists whose articles were:		
	Cleaner	No Change	Dirtier
Product C	11.3%	62.9%	25.8%
Product D	10.1%	76.4%	13.5%

FIG. 3. In-home testing of articles.