FEATURE

Bundle test evaluates detergency

The following is based on a talk given by Jay R. Brummer of FMC Corp., Princeton, New Jersey.

The bundle test is a test method that allows a detergent formulator to compare cleaning and brightening performance of any two home laundry detergents or procedures. The method uses naturally soiled paired clothing and linens that are given to "typical" families and subsequently washed and visually evaluated under controlled laboratory conditions for a specified number of wash cycles.

The bundle test method was first published as an ASTM Standard Method in 1971 by ASTM Committee D12 on Soaps and Detergents. The test is entitled "Controlled Laundering Test Using Naturally Soiled Fabrics and Household Appliances." The method D-2960 can be found in the **1987 Annual Book of ASTM Standards, Volume 15.04**, page 588. Although the exact evolution of the method is unclear, large detergent companies used this technique as a cost-effective step in their product evaluation process and to substantiate advertising claims. ASTM committee D12 became involved as a medium to develop and publish a standardized test procedure for the detergent industry. The members of ASTM subcommittee D12.25 that worked on the standardization of this method represented a wealth of experience in bundle testing and other detergency evaluation procedures. Bundle testing is still used extensively in the industry as a detergency evaluation tool.

The detergency evaluation process consists of a number of steps. As a detergent chemist develops new formulations, raw materials and delivery systems, the need to evaluate detergent performance is critical to his success. He may use the Terg-O-Tometer to screen different formulations in the laboratory until the data indicate that a promising product has been developed. The next step would be to test the product under more realistic and practical conditions that more closely correlate with actual field experience. This is where the bundle test is a valuable tool to use as it is a limited, controlled performance evaluation test method using naturally soiled fabrics and one that is relatively inexpensive compared to larger scale panel or consumer tests. The cost per test generally runs around $1,000 for time and materials per family. In addition, the bundle test can be used for laboratory screening, competitive analyses and substantiation of advertising. The extent to which the bundle test is used for these and other applications depends on the needs and resources of the user.

The ASTM method recommendation for a 10-family, 10-week bundle test includes multiple hardnesses, concentrations, temperatures and soil levels to compare two detergent formulations over a wide range of detergency conditions. As a general rule, the bundle test method published by ASTM is seldom practiced as written. At FMC, we use a modification of the ASTM procedure, as do many other companies, to perform our bundle testing although the basic principles of the ASTM method are not compromised. The essence of our bundle test method is given in the following sections.

Bundle test description

The bundle test method used by FMC is a relatively simple procedure that can be easily used by any laboratory to compare the cleaning or brightening/whitening performance of any two detergent formulations. Two "typical" families are chosen from a pool of families at our facility who have at least one child under the age of 12. To minimize differences in soiling for a two-family test, the families chosen have moderate soiling characteristics. The families are given 40 items each, consisting of cotton and cotton/polyester clothing and linens. The items are white to allow for maximum contrast during visual evaluation. The paired bundle items are matched visually to minimize any inherent differences in the articles and marked for identification.

The families use each item as they normally would for a week and return the bundle at the end of the week. Normally, the test is run for 10 weeks to allow time to observe any trends that develop. Each bundle is then split into two separate wash loads weighing approximately six pounds each and washed separately in the two detergents being evaluated. The typical conditions for washing a bundle are given in Table 1. It should be noted that hardness, temperature, detergent concentration and load weight should be carefully measured and controlled to obtain valid results. In addition, it is essential that the washers used are matched for fill volume, agitation force and cycle time. If this is not possible, the same washer may be used to wash both loads. The wash loads are then dried and folded for the evaluation process.

The evaluation is done under two different light sources to distinguish between cleaning performance and the effects of the optical brightening system. The light sources used include incandescent light and north daylight (ultraviolet). The evaluation is done against a gray background to provide maximum contrast.

During the first nine weeks of the 10-week cycle, an experienced four-judge panel rates each pair of items for preference or no preference. The 160 evaluations per bundle are compiled each week to monitor the progress of the test. Any staining that occurs should be noted but the stained area should not be included as part of the normal evaluation process. If any extensive staining occurs, the item should be discarded from the bun-

dle and replaced to maintain load weight. At the end of the 10-week period, a panel of 10 judges—in addition to the four-member panel—evaluates the paired bundle items, and this data is included in the overall performance evaluation for the two detergents. The evaluation period may be decreased if a clear trend develops after at least five weeks that one detergent is superior to the other. The minimum is five weeks, to ensure that redeposition is not occurring with one of the two test detergents.

The final data are treated by standard statistical methods to determine if the differences observed between the two detergents are significant. We have found a preference ratio of approximately 1.5:1 to be significant for our bundle test evaluations; however, it is important to consider other data available from the test to get a true indication of performance.

In addition to the overall preference ratio from the four-judge panels, the preferences of the final 10-judge panel should be part of the overall evaluation. The final evaluation may not agree with the weekly evaluations, thus indicating further tests are required. In addition, the preferences by fabric type are important for overall detergency evaluations and should be considered separately as well. Consideration of the data from the north daylight and incandescent light sources for each fabric type also is necessary to evaluate the whitening enhancement from any optical brighteners in the detergent formulations.

Finally, it is essential to observe any trends that develop during the test period that may indicate a longer-term preference for one of the detergent formulations that is not evident from the overall preference ratio.

It is important to remember that the bundle test is not an analytical determination. The method is a practical performance evaluation using realistic soiling conditions and controlled laboratory washing conditions to compare two detergent formulations. It is an effective method to compare the relative performance of two detergents and can discriminate between small differences in detergency. It should be noted, however, that the test method does not rate the absolute performance level of a detergent and the formulator should always keep this in mind when drawing conclusions from the data. Care should also be taken when comparing the results of one bundle test to another unless some statistical basis for cross comparisons has been established.

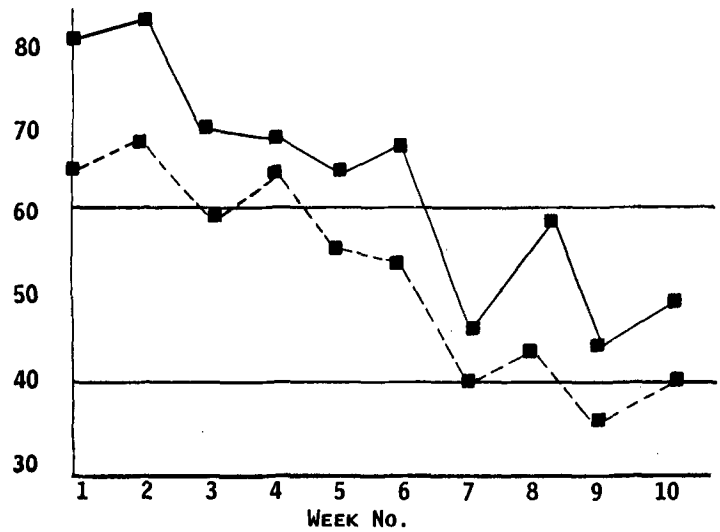

FIG. 1. Change in weekly preference for Detergent A when viewed under different light. Solid line, UV light; broken line, incandescent light. Detergent A was an STPP built/mixed surfactant with bleach and enzymes.

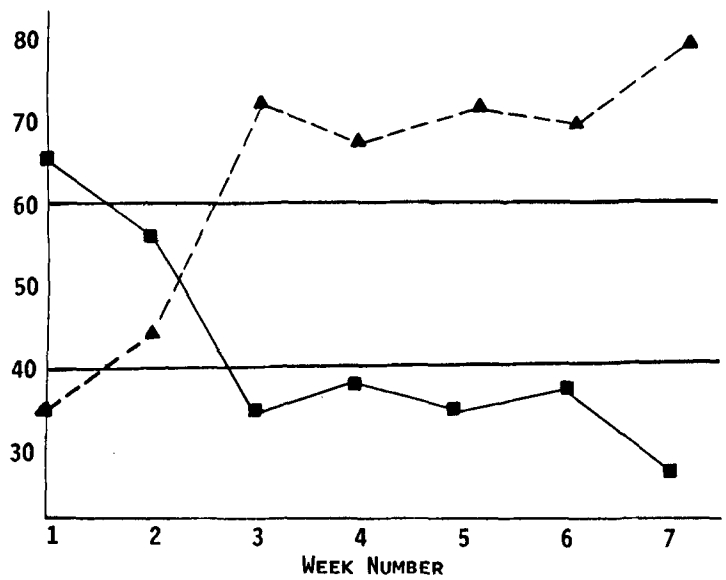

FIG. 2. Change in weekly preference, Detergent C vs Detergent D. Detergent C, an STPP/anionic, is indicated by the broken line. Detergent D, an unbuilt liquid/mixed surfactant system, is indicated by the solid line.

Bundle test in practice

As mentioned previously, the bundle test is only one step in evaluat-

ing detergency during the product development cycle. It may give completely different results from initial screening results obtained on the Terg-O-Tometer Test, indicating the importance of including practical testing during the product evaluation process.

To demonstrate this, three examples are given where two proprietary detergent formulations were compared using the Terg-O-Tometer method and the bundle test method. The conditions used for a typical Terg-O-Tometer test are given in Table 2.

These examples highlight some valuable information that can be obtained from the bundle test method. In the first comparison, an STPP-built mixed surfactant system with bleach, enzymes and a unique delivery system (detergent A) was run against an STPP-built anionic system with enzymes (detergent B). The results of this comparison are shown in Table 3. As

TABLE 1

Typical Bundle Test Washing Conditions

Temperature	100°F (wash and rinse)
Water hardness	150 ppm (moderate stress)
Detergent concentration	0.15% or manufacturer's recommendation
Washers	Whirlpool LA7800XM (toploader) using fill level closest to 17 gal
Dryers	Whirlpool LE7680XM or equivalent
Weight of each half-bundle	6 lb.

TABLE 2

Typical Terg-O-Tometer Test Washing Conditions

Temperature	100°F
Water hardness	(34, 68, 102, 136 ppm) (2 Ca:1 Mg)
Detergent concentration	Recommend use (~0.15%)
Soiled fabric	Clay and/or Spangler soiled cotton and polyester/cotton
Load	Two 4″ × 4″ soiled and unsoiled swatches of each fabric type, totaling 8 swatches
Evaluation	Hunter Labscan Reflectometer (Model SN12714)

TABLE 3

Comparison of Two Proprietary Powder Detergents

Detergent	Type	Fabric	Terg-O-Tometer reflectance vs standard (100%)[a] Clay	Terg-O-Tometer reflectance vs standard (100%)[a] Spangler	Bundle % preference
A	STPP built/mixed surfactant powder with bleach and enzymes	Cotton	100	60	77
		PE/cotton	97	66	51
		Overall			64
B	STPP built/anionic powder with enzymes	Cotton	102	103	23
		PE/cotton	99	99	49
		Overall			36

[a]AATCC standard with 50% STPP and 17% anionic.

TABLE 4

Comparison of Liquid and Powder Detergent

Detergent	Type	Fabric	Terg-O-Tometer reflectance vs standard (100%)[a] Spangler soil	Bundle % preference[b]
C	STPP built/anionic powder	Cotton	100	63
		PE/cotton	86	50
		Overall		59
D	Unbuilt/mixed actives liquid	Cotton	105	30
		PE/cotton	87	38
		Overall		32

[a]AATCC standard with 50% STPP and 17% anionic.
[b]9% no preference.

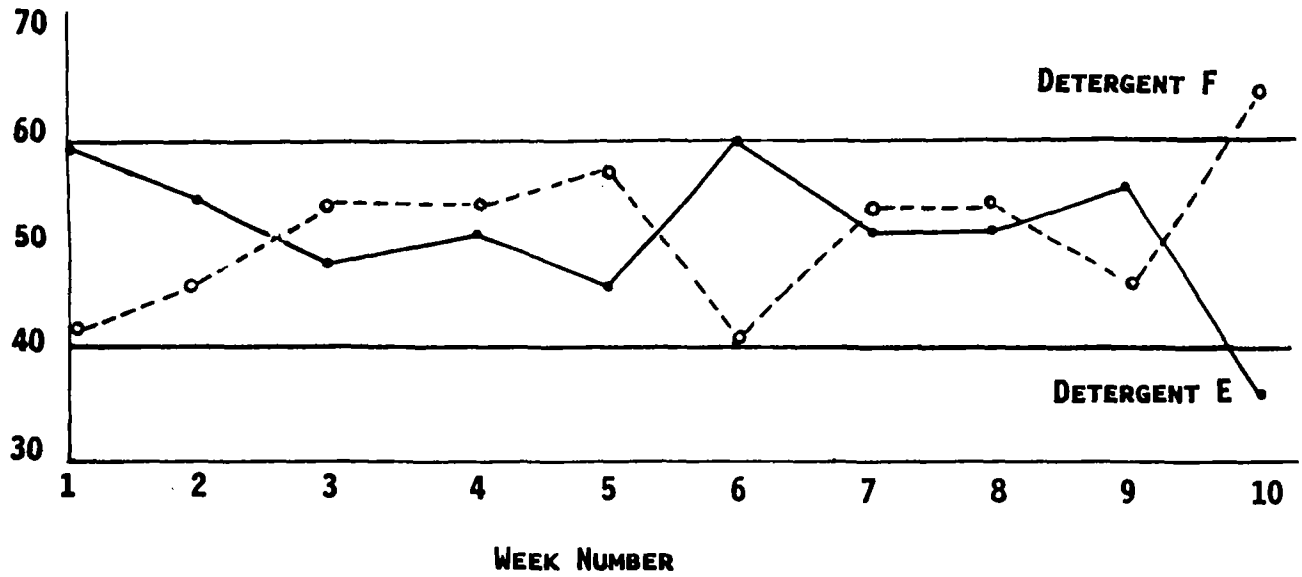

FIG. 3. Change in weekly preference, Detergent E vs Detergent F under UV light. Detergent E, an STPP built/anionic with bleach and enzymes, is indicated by the solid line. Detergent F, an STPP built/anionic with enzymes, is indicated by the broken line.

can be seen from the Terg-O-Tometer data, detergent B was significantly superior to detergent A on Spangler soiled fabric, while similar performance was measured on clay soiled fabric. The overall bundle test results, however, showed detergent A to be clearly superior to detergent B. This anomaly may be due in part to the type of soils chosen for the Terg-O-Tometer method versus the mix of naturally occurring soil types encountered on the bundle items. Whatever the reason, this comparison demonstrates that the bundle test yields valuable data which may indicate the need for further testing of a formulation before additional resources are committed to product development.

The results of the first product comparison also demonstrate two other advantages of the bundle test method. The results of preferences for detergent A were plotted on a weekly basis for incandescent light and north daylight. As can be seen from Figure 1, a significant difference in visual preference was observed for detergent A under north daylight versus incandescent light. This is due to the optical brightener present in detergent A which enhances the apparent whiteness under north daylight. The preference for detergent A under incandescent light would give the actual cleaning performance without the enhancement from the optical brightener. This figure also reiterates that using only the final preference ratio is not sufficient to obtain an overall performance evaluation. While the preference for detergent A was 64%, a weekly downward trend was observed over the 10-week test period. This may be due to redeposition over time or other factors and must be noted to get a complete picture on the performance of detergent A. This trend would not be apparent in the traditional Terg-O-Tometer Test.

A second comparison was made between an STPP-built anionic detergent powder (C) and an unbuilt liquid detergent (D) with a mixed surfactant system. The Terg-O-Tometer results are based on an average of Spangler soiled cotton and cotton/polyester blend fabric. The Terg-O-Tometer results given in Table 4 show the unbuilt liquid (D) to be slightly superior in performance on cotton fabric to the STPP-built powder (C) and equivalent on the cotton/PE blend. Once again, the overall bundle test results show a preference for the STPP-built powder (C). A look at the weekly preferences for each detergent, given in Figure 2, indicates an initial preference for detergent D in the first two weeks, with a significant preference for the STPP-built detergent C for the remainder of the test. Using only the Terg-O-Tometer results in this case would give a misleading conclusion which would be similar to the conclusion reached after only one week of the bundle test. This comparison once again demonstrates the usefulness of the bundle test to allow trends to develop that may indicate problems with the detergent formulation.

The final example compares the performance of an STPP-built powder with anionic surfactant, enzymes and bleach (E) to an STPP-built powder with an anionic surfactant system and enzymes (F). The Terg-O-Tometer and bundle test results are shown in Table 5. The Terg-O-Tometer results indicate that the two detergent formulations perform approximately the same on clay and Spangler soiled cotton and cotton/polyester blend fabric. In this case, the bundle test results also show that the two formulations are not significantly different in performance. When the

TABLE 5

Comparison of Two Proprietary Powder Detergents

Detergent	Type	Fabric	Terg-O-Tometer reflectance vs standard (100%)[a] Clay	Terg-O-Tometer reflectance vs standard (100%)[a] Spangler	Bundle % preference (UV light)
E	STPP built/anionic powder with bleach and enzymes	Cotton	100	104	53
		PE/cotton	99	102	42
		Overall			48
F	STPP built/anionic powder with enzymes	Cotton	99	103	47
		PE/cotton	101	99	58
		Overall			52

[a]AATCC standard with 50% STPP and 17% anionic.

visual preferences of detergents E and F were plotted on a weekly basis as shown in Figure 3, the two products were essentially equivalent in performance during the first nine weeks of the test period with no preference trends developing for either detergent. The preference for detergent F in the tenth week may indicate that the test should be extended for a few more weeks to determine if the preference was just noise or the beginning of a significant performance trend. Once again, this demonstrates the necessity to evaluate all of the data available.

Bundle testing provides a realistic evaluation of performance that bridges the gap between laboratory screening tests and extended consumer testing. The bundle test has proven to be an invaluable cost effective tool in the new product evaluation cycle. Information obtained from the bundle test often shows the necessity for additional laboratory development work before proceeding on to more extended and expensive consumer testing.

Predicting a bundle test winner

The following, based on a talk given by Paul X. Riccobono, was prepared by Riccobono and Richard Polanski, both of Colgate Palmolive Co., Piscataway, New Jersey.

The ultimate objective of product development is introducing profitable new products into the marketplace. Usually it is known in the early stages of development whether a product is new or not. Profitability is not as easy to determine and is dependent on a number of interrelated factors, not the least of which is consumer acceptance. This is particularly true in the laundry products area, where the consumer's ability to discern a point of difference in performance between a new or improved product entry and a product already in the marketplace at times rests on rather tenuous differences in sensory perception.

Unfortunately, the ability to accurately predict by quantitative laboratory tests the effects of new laundry detergent compositions on consumer perception has remained largely an unfulfilled goal of research workers in this area. Usually, the evaluation of a new laundry formulation involves a progression of testing methodologies, from simple laboratory determinations of detergency utilizing the Terg-O-Tometer, to complex consumer-testing involving hundreds of participants. The ultimate are the sales tests, in which entire cities or regions of the country are involved (1-3). At some point in the process, the new product must be taken out of the laboratory and handed over to consumers for their judgment. The decision to consumer-test a new product is a critical point in the product development process for it involves large sums of money and considerable amounts of time and human effort. At Colgate, the decision to consumer-test a new laundry product is often made only after results of a bundle test have been evaluated and factored into the decision-making process.

In the arsenal of laboratory test methods available today, the bundle test is generally acknowledged by the detergent industry as closest to typical consumer response (4). Formalized as an ASTM method in 1972 (5), it today is the principle "bridge" between the closely controlled laboratory-testing of the formulation chemist and the variability of the real world.

Useful as it is, the bundle test is a rather long and tedious procedure. A decision to proceed with the bundle testing of a formulation—which typically takes six to eight weeks—is itself a decision of some significance. Thus, the ability to accurately predict the outcome (i.e., the visually preferred product) of a bundle test in one or two days would be of considerable value to a product development staff. It is this problem which is the subject of this paper.