

# The Missing Link in Product Development

By Ton Teerling, Quest International, Bussum, The Netherlands

You're looking at the missing link in product development when you stand with your product in one hand and the consumer report in the other and wonder: "So this is what consumers think of my product. But what do I do now? Where do I go from here?"

Very often consumer reports do not contain the kind of information that helps directing product development. The results are not "actionable." By letting sensory research and consumer research each have their specific, but closely interwoven input into the process of product development, we at Quest believe we have filled this gap by supplying the missing link in product development.

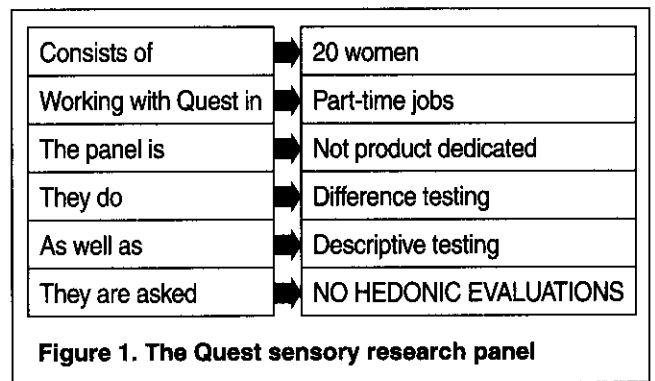
To explain their specific roles, sensory research and consumer research are dealt with in detail separately first. Then their close relationship will be explained.

## Sensory Research

Basically, all questions that enter the product development department come down to the question: "Are there differences among the given products?" This question may be answered either by difference tests such as triangle tests or by more complex descriptive techniques such as profiling or quantitative descriptive analysis. For difference tests, any group of subjects with minimal sensitivity will do; for profiling, an expert panel is required.

**The Panel:** The Quest sensory research panel (Figure 1) consists of 20 women who have been recruited externally. It is a deliberate choice not to recruit from within the company to prevent the panel's bias. (Another consideration is cost. Internally recruited panels may be several times more expensive when working on a regular basis.)

A very strict house rule concerns the distinction between difference or descriptive tasks and liking-related tasks. This means that the Quest sensory research panel will not do any preference tests or liking tests. One reason for this is that the panel cannot be considered to be representative of any target group or population. Another reason is that the panel needs a clear and unambiguous task definition; the analytical task of detecting and scoring intensities should not be



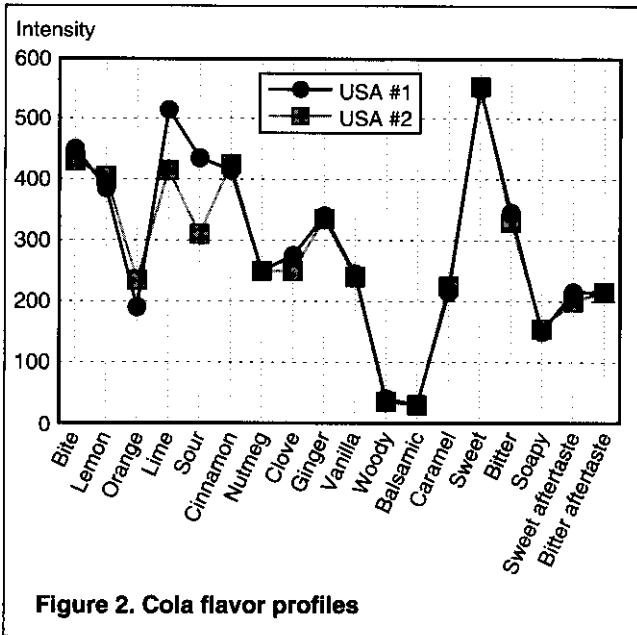
contaminated by emotional evaluations about pleasantness.

The panel is an analytical instrument. At Quest, the panel has easy access to flavorists and application technicians. This accessibility has advantages that become clear during training.

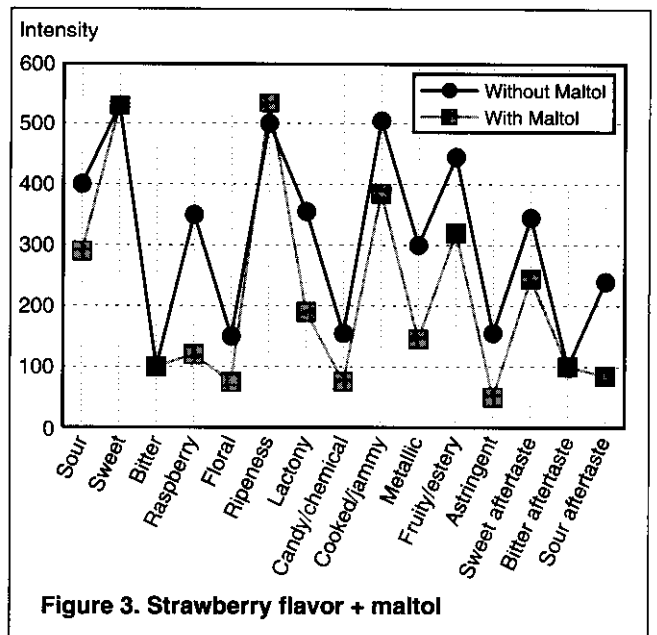
**Training:** The procedures of the descriptive technique (profiling) can best be explained by an example of the training procedures for a particular type of product. Given a wide range of products of the relevant type, the panel starts generating spontaneously as many attributes as possible describing the sensory properties of these particular products. As the panel is rather uninhibited (many members have been doing this for about nine years now) the number of descriptors can amount to over forty. In subsequent sessions this number is reduced by means of two actions:

- Panel members discuss the attributes in order to eliminate overlap in meaning and to reach agreement upon the exact meaning of descriptors.
- Flavorists involved with the products help determine the exact meaning of any still ambiguous descriptors by supplying examples of those descriptors. These examples are given on smelling strips or as drinkable solutions.

For example, when some panel members say "I taste a



raspberry note in this strawberry yogurt” but others say “So do I, but I’m not sure whether it is indeed raspberry” and again others say “What are you talking about?”, it is time to invite the flavorist to supply several examples of what raspberry in strawberry yogurt might be until the panel says



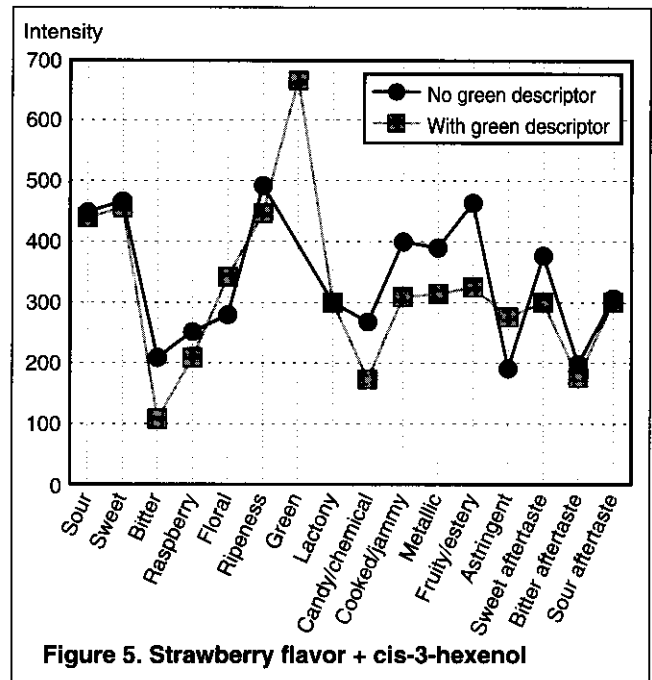
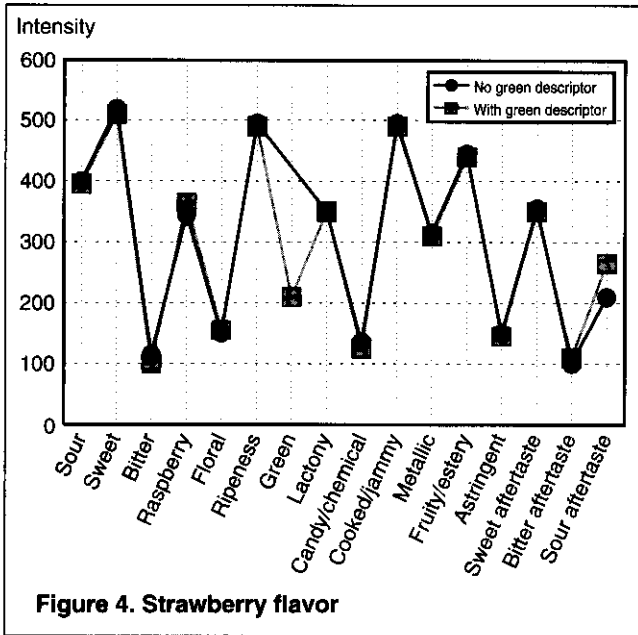
“Yes, this is what we mean when we say raspberry!” (and others say “So this is what you were talking about!”).

This procedure has two major advantages. First, all members of the panel now know what is meant by “raspberry in strawberry yogurt;” there is agreement about its meaning. Second, the flavorist involved knows exactly what to do when he later gets information back from the Sensory Research Department showing that in a particular development product the raspberry note is too high or too low; after all, he knows which physical component is responsible for this sensory attribute. (This, by the way, explains why a descriptor set cannot simply be transferred from one panel to the next; the descriptor set will contain its full and unambiguous meaning only with the group (panel and flavorists) that built it.)

Essential in all this is the close feedback loop between descriptor generation by the panel and creation/application by the flavorist. What is crucial for the resulting descriptor set is that it must be actionable.

It should be clear that only a panel with a high level of training will be able to accurately demonstrate differences between products. Consider, for example, the two well-known U.S. colas profiled in Figure 2. Although many people believe that they can tell these colas apart because of their alleged difference in sweetness, it can be seen that they are actually perceived equally sweet by the Quest sensory research panel; panel results show differences in “sourness” and “lime” taste. Instrumental measurements do, indeed, show nearly identical sugar levels. (The results shown in Figure 2 were obtained during training sessions at a time when the descriptor set for colas was still incomplete.)

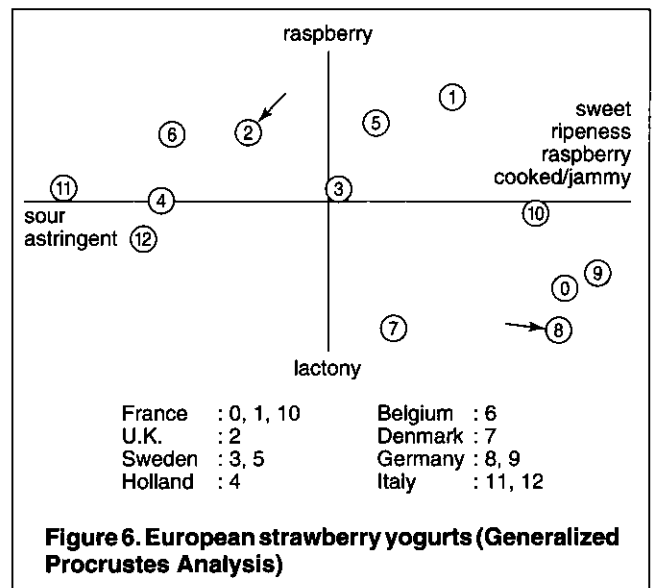
**Descriptor Sets:** A well-defined descriptor set is a prerequisite for demonstrating the influence of a component on a flavor profile. Consider, for example, the small but not obvious effects of a low concentration of maltol on strawberry yogurt as shown in Figure 3. Maltol tastes like



candy floss. One would expect the “candy” character to increase when maltol is added, but something quite different occurs. In low concentrations maltol appears to suppress the “candy” character and, even more dramatically, to reduce the raspberry note in the flavor. However, this effect can be demonstrated clearly only when a well-defined descriptor set is available. For example, the introduction of the descriptor “green” does not seem to make much difference in the case of a straight strawberry flavor (Figure 4), but with a “green” type of strawberry flavor the difference is clear (Figure 5).

A descriptor set should not only be unambiguous and well understood by the panel members, but it also must be complete, covering all relevant attributes of a product. During training for strawberry flavors, it was initially assumed that “Ripe” would be the opposite of “Green.” Products could be very “Ripe” (towards sulphury) and consequently be “Not green” or the other way around but not simultaneously “Green” and “Ripe.” This resulted in a descriptor set that contained “Ripe” but did not contain “Green.” However, after several training sessions it became apparent that profiling data with respect to “ripeness” were not as actionable as the data from other descriptors; in other words, reformulations on the basis of “Ripe” data did not lead to expected corresponding profiles. Subsequent training on the descriptor “Ripe” resulted in more actionable data; this appeared true for all descriptors.

**Combining Profiles:** Demonstrating the varied relations among the profiles of different products calls for a different way of handling data. You can’t combine several profiles by simply superimposing them in one figure: it takes different techniques such as Generalized Procrustes Analysis or Principal Component Analysis. Figure 6 shows combined data on thirteen European strawberry yogurts. There it can be seen that, for example, Italian products are found at the “sour” end of the horizontal axis; while German



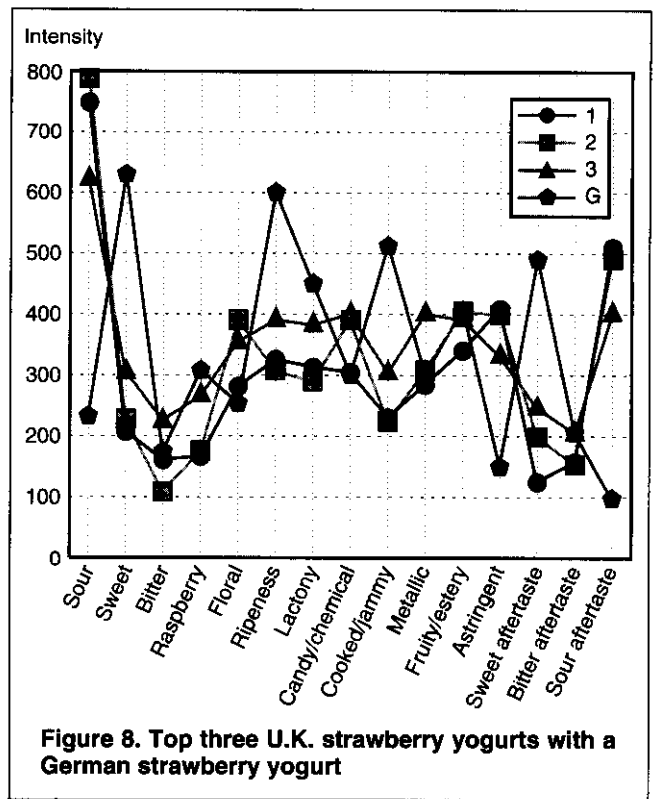
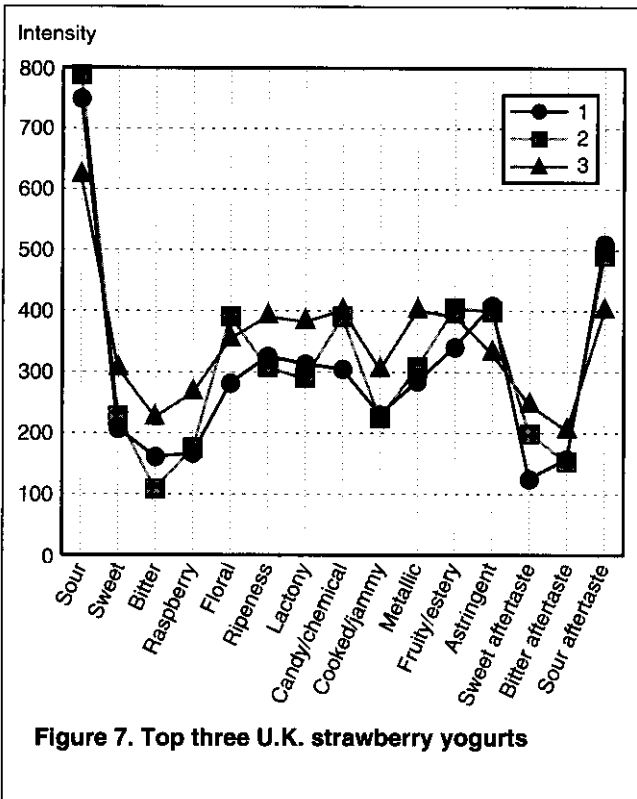
products are more towards the “sweet, ripe and jammy” end of this axis. The Danish product is somewhere in between the extremes on the horizontal axis, but it lies close to the “lactony” end of the vertical axis. Mark that these data are about the intensities of attributes, definitely not about the liking for them.

What would you do if you were a German producer of strawberry yogurts and wanted to penetrate the U.K. market? (See arrows in Figure 6.)

Sensory research information gives you at least two options:

- You try to match the existing products, or
- You deliberately distinguish from them.

Figure 7 shows the top three strawberry yogurts (by volume) in the U.K. They are very similar products. Figure 8



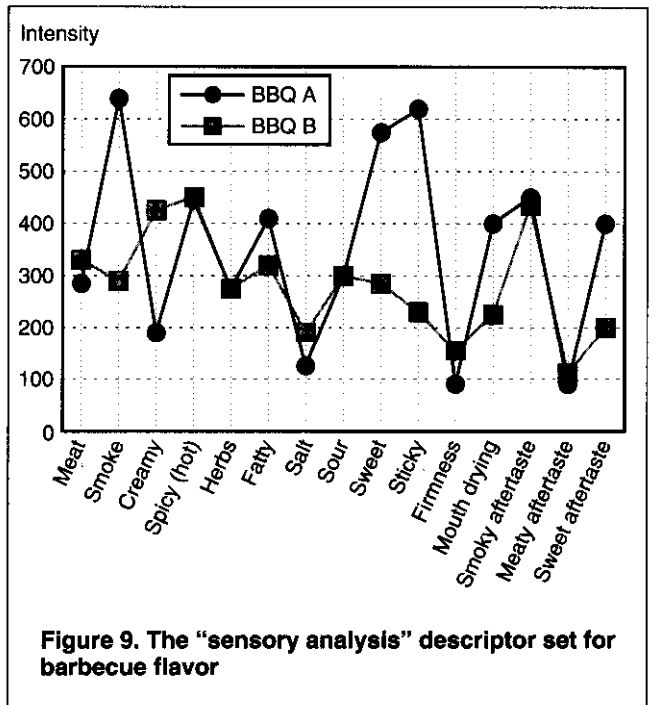
overlays a German strawberry yogurt. It is a totally non-U.K. product and has proven to be very successful in the U.K. This demonstrates that at some locations the liking for a particular product may be present, but the appropriate product is missing.

**Consumer Research**

Sensory and consumer research may show similarities in basic structure, but the differences become apparent when one realizes that in sensory tests we make life relatively easy for the subjects. We give them extensive training on descriptive tasks and the answers in difference testing are valid for any perceptible difference they can point out to us. (And we can shield them from distractions. For instance, if we do not want them to be distracted by color, we turn on the red lights). In consumer tests subjects' lives are more difficult. During discrimination tests (Do you prefer A to B?) subjects are permitted to report differences between products with respect to one and only one aspect: liking. In descriptive tasks they are asked to tell us about product characteristics they have not been trained in and may very well misunderstand. All this makes results from regular consumer tests rather unreliable. Basically, with consumer tests the problem comes down to two issues:

- If questions about liking were easy to answer, the predictive value of consumer tests would be higher.
- However silly the question, you will always get an answer.

The message is clear: consumer tests must be as



simple as possible but the results must still be actionable. Somewhere between the complexity of the sensory research panel's descriptor set and the simple but uninformative vocabulary of consumers there must be a level of detail that is intelligible to consumers and actionable to flavorists.

Taking cola flavors as an example, it is obvious that the descriptors as defined within sensory analysis are too

much on an expert level for consumers to deal with them. Consumers would probably be perfectly willing to tell us that there is too much lemon or too little clove in a particular sample; the problem would be that the results from such a complex task will have no validity. On the other hand, oversimplification would lead to loss of valuable information. How to deal with this problem?

This gap between consumer test design and product development is bridged in the following manner. Sensory analysis tells us which descriptors are the most important ones for a particular type of product; statistical techniques like the mentioned Principal Component Analysis or Generalized Procrustes Analysis are of help here. Results from these statistical techniques suggest the existence of meaningful groups of descriptors, such as a "citrus" group (lemon, orange, lime) and a "spicy" group (cinnamon, nutmeg, clove) in cola flavors. In this way a shorter and more simple descriptor set is constructed from the original detailed one from sensory analysis. But statistics are not enough. The new set must be valid. Here again the close collaboration between the Sensory Research Department and flavorists ensures that the interpretation procedures are set before the actual consumer test is done. Do the more simple descriptor groups still contain the required information for the flavorist to know what to do in terms of recipe

reformulation? Of course, a flavorist will be able to determine this only if he or she was involved in the original definition of the descriptors used by the sensory research panel.

What remains is to ensure that the eventual "consumer descriptor set" will indeed make sense to consumers and that they will know, for example, the difference between "fresh" and "green" if these descriptors turn up in the consumer descriptor set. This validation is achieved by running a pilot test in which a random group of consumers evaluates a series of products in which each product has deliberately been modified to represent one of the specific characteristics that appear in the consumer descriptor set. In such a test consumers do not evaluate for liking, but they score the intensities of each descriptor for each of the products. From this it should become clear whether they perceived the product with increased greenness as being indeed high on "green" but not on "fresh."

One typical example comes from our savory group. In Figure 9, two profiles of barbecue-flavored snacks show two totally different products. Of course, this figure does not contain any information about which of the two would be preferred. From the sensory analysis database on barbecue flavors, descriptors were combined according to statistical procedures into descriptor groups most appropriate for this type of flavor. Flavorists helped determine if this set was

---

adequate. The final "consumer test" descriptor set for barbecue flavor was the following:

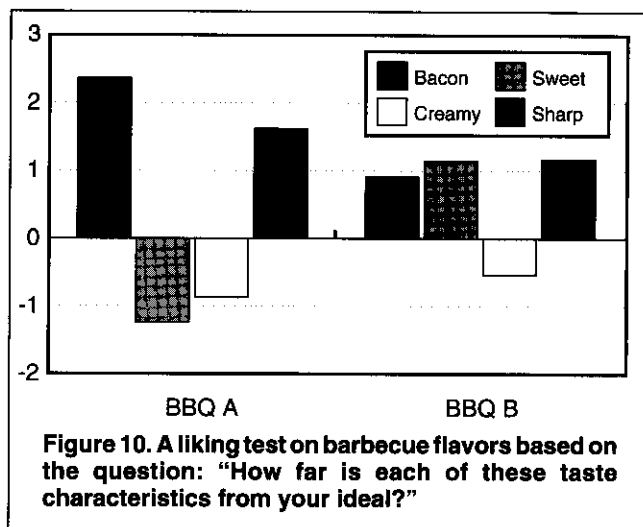
- Bacon
- Creamy
- Sweet
- Sharp

In a subsequent pilot study, consumers evaluated four barbecue-flavored snacks of which one was high on "sweet," one was high on "sharp," and so forth. This was not a liking test, but people scored intensities of descriptors for all four products. From this it could be learned that consumers did indeed understand the descriptors and that the current set was valid.

Finally, in the actual consumer test it became clear in what direction product development should go to increase the acceptance of the flavor coded A in Figure 9.

One additional test, a simple liking test, was run in order to collect liking data for both products (Figure 10). It should be noted here that Quest consumer tests never require that consumers compare products; each product is evaluated individually on its own merits. From fundamental research into the validity of consumer test designs it could be demonstrated that in comparative tests people are strongly inclined to show a preference for the product that stands out among the others because of any characteristic, not necessarily liking; taste intensity or color may equally well appear attractive to people. For this reason comparative liking tests are avoided as much as possible.

Moreover, liking for a product at one particular moment does not necessarily reflect liking over time or even future liking. This is why a liking test should mirror the dynamics of liking—the development of liking in time. Some products happen to become tedious rather soon while others continue to be appealing. Although it is difficult to tell exactly what the elements are that cause these qualities, the dynamics of liking should be accounted for in a consumer test. One way of doing this is to perform repeated measurements in a relatively short period of time in order to see whether liking increases, remains stable or decreases. It can be demonstrated that an increasing liking for a product is a more powerful predictor of future acceptance than the numerical average on the liking scales. In other words, an increase from 3 to 5 or 6



**Figure 10. A liking test on barbecue flavors based on the question: "How far is each of these taste characteristics from your ideal?"**

would suggest a more promising product than one that gets a score of 6 on all four measurements.

### Summary

Consumer research should play a more integrated role in product development and it can by using these techniques:

- Base the consumer tests on the sensory analysis database (the descriptors and their definitions).
- Encourage close collaboration across departmental lines so people from various departments (sensory analysis, consumer science, flavorists) help design the consumer tests.
- Verify that consumer test results can be translated back into "product development language."

The synergy with sensory research, recipe formulation work and psychology will render consumer test results actionable. This multidisciplinary approach has proved effective in leading product development further away from the traditional trial and error procedures and closer to a more cost-effective and goal-directed way of working.

### References

Address correspondence to Ton Teerling, Quest International, PO Box 2, 1400 CA Bussum, The Netherlands.

