Expert Panels, Consumers, and Chemistry

Thomas K. Rehfeldt

> We must next consider what account we are to give of any one of them; what, for example, we should say color is, or sound, or odor, or savor; and so also respecting [the object of touch. . . The point of our present discussion is, therefore, to determine what each sensible object must be in itself, in order to be perceived as it is in actual consciousness.
>
> Aristotle, (c330 B.C.) "On Sense and the Sensible"

IT'S THE ECONOMY

Large amounts of time, resources, and money are spent each year in the development of consumer products. Very large expenditures spent needlessly if the consumer does not like the products once in the market place. Thus, many more dollars are spent on consumer research to learn if the products will be embraced when on the market.

The process of chemical development is usually followed by expert panel evaluation, then one or more small consumer surveys, followed by a full-scale market research study. Anything that can be done to make the process more efficient, and, particularly, to make the testing predictive of consumer behavior is extremely valuable.

PREDICT WHAT?

Conventional market research testing makes use of methods such as factor analysis, multi-dimensional scaling, discriminant function analysis, and such like complex statistical methods. The value and advancement in these techniques, especially since the advent of cheap computing, has greatly increased in recent years. But, given the value and prevalence of these methods, one item is still lacking. These methods are not measurements and thus are not predictive, but solely descriptive of the most recent data.

Thomas K. Rehfeldt

Thomas K. Rehfeldt is Senior Research Statistician at the Unilever Home and Personal Care Innovation Center in Rolling Meadows, IL. After many years as an analytical chemist and developer of high tech. aircraft paint, he decided that he could have just as much fun solving problems that did not necessitate getting contaminated with smelly chemicals. He started in the statistics program at the University of Chicago where a fortuitous meeting with Ben Wright started him on the measurement path. Since then, he has tried to apply Rasch measurement tools in product development and marketing applications. He is currently working on the measurement of sensory perception as it relates to consumer choice and definition. His research interests include defining perception variables, relating expert panel judgment, consumer, marketing, and chemical data, and developing useful predictive models of consumer behavior. All of these entail analysis of rating scale, ranks, and dichotomous responses in many facets. When not chained to his computer he can be found building a new house in the north woods of Wisconsin or in the company of the two dogs to whom he belongs.

In the parlance of marketers, the predictions needed are "What are the key drivers of product acceptability?" and "What change in key drivers will produce a proportional change in acceptability?" The key drivers are those attributes, out of all possible product properties, that are the ones that are necessary for product acceptability, e.g., a shampoo may clean hair, but it will not sell if it does not lather. The key drivers may also be the complementary attributes; those that will cause the product to be rejected independently of the others, e.g., a shampoo may do everything well but have an undesirable fragrance.

In this context, the objective is to know what can be measured that will inform us of the effect of these key drivers, and what other facets may predict the level of acceptance.

OUR EXAMPLE

The example presented is for an examination of the attributes properties and acceptance of anti-perspirant products. Fourteen commercial products were tested in the consumer test; 400 consumers used 3 products, sequentially, for 2 weeks each. In the expert descriptive panel test, each of 14 panelists tested all products and 2 replicate trials were made. Analytical instrumental testing measured lightness, friction and rate of application for 14 products.

The objective is to identify the key drivers for the consumers. From history and experience, the drivers would be the efficacy, i.e., how it protects from odor and wetness, and application, i.e., how it feels when applied.

THE MEASUREMENT

Three types of data were collected: the analytical data, the expert panel data and the consumer data. The analytical data is a continuous scale. The whiteness was measured with a spectrophotometer; this is the L–value. The force to pull the anti-perspirant stick across a test material was measured as the dynamic friction. The consumer data was from a 10-point categorical scale, i.e., subjects were not allowed to mark fractional values but would check boxes at each of the scale marks. The expert panel data was collected as a 10 point continuous scale, i.e., subjects were allowed to mark the scale at any place on the line from 1 to 10. The direction of the consumer scale was cast as level of approval, so the direction was the same for all attributes. The expert panel data is collected as amount of the attribute so the direction of preference is not the same for all attributes.

This set of conditions illustrates the power of the Rasch model. Based on the set of common products, the expert panel data and the consumer data can be combined. The difference in how we ask the questions is of little concern. Since linear continuous measures are calculated, the analytical data is easily combined with the measures.

THE RESULTS

First, a FACETS analysis was performed on the consumer data. The FACETS program was used to account for the different attributes, the different products, the subjects and the replication or order of presentation. The measures for the attributes were examined. In the consumer study, all of the questions are worded so that all of the attribute scores progressed in the same direction. The questions were generally "How did you like the attribute?"

It was found that negative attributes scored high, along with positive attributes, indicating that the approved rating was related to lack of something (like greasiness).

The next step was to run a FACETS analysis on the expert panel data. The results obtained were similiar to the consumer data, with the products serving as the common link. The expert panel is trained to report the amount of an attribute on the 10 point continuous scale; no distinction is made for undesirable attributes. Low greasiness was reported as 'less grease'.

The first comparisons found some of the attributes, on opposite ends of the scale, due to different form of the questions, i.e., greasiness was generally low for commercial products, so the expert panel reported low greasiness, which produced the consequent low measures. In contrast, low greasiness is seen as desirable to the consumer so they approved this and gave a high approval score.

By judicious choice of the centering and anchoring, the expert panel measurements and the consumer measurements are on the same scale and in the same direction. The final measurement scales are shown in Figure 1.

One can observe which expert attribute assessments relate to the consumer assessments. For example, the expert assessments of 'slippery' and 'washability' will predict the consumer assessments of not greasy, doesn't stain clothes, and washes off easily.

In addition, one will note that 'force to apply' and 'force to spread' assessed by the expert panel will predict 'initial comfort' for the consumer. It is observed that the physical measurement of dynamic friction will predict force to spread which in turn may predict consumer acceptance of 'initial comfort'.

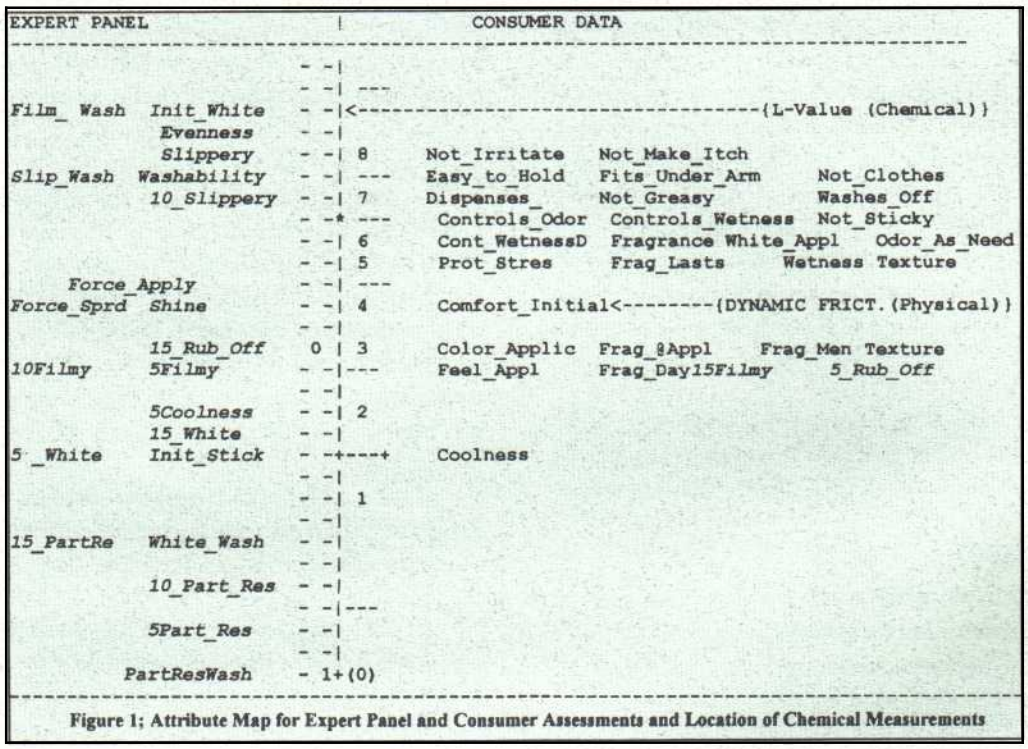

Figure 1; Attribute Map for Expert Panel and Consumer Assessments and Location of Chemical Measurements

The order of importance for acceptance is lack of irritation and lack of itchiness, followed by feel attributes, such as greasiness and stickiness. Next are the performance attributes of controls wetness and controls order. Attributes like coolness and color of the applicator are less important.

CONCLUSION

The Rasch model can provide the tool necessary to combine data from several sources, to relate several kinds of data and clear interpretation of assessments. We also have demonstrated the potential to decrease the number of tests, attributes, and the amount of time and money spent in development.

Our district has found that the Lexile Framework is proving to be a valuable way to allow us to coordinate the variety of instructional materials and programs that are presently in place in our county. As the Reading Specialist for grades 3-8, I have found that the Lexiles allow us to have another resourceful tool to assist teachers in customizing the reading programs in their own classrooms and to further link their instructional effectively to the end of grade testing in our state. The Lexile Framework also meshes well with our district's Balanced Literacy Program.

Kathy Bumgardner
Reading Specialist
Gaston County Schools