ABSTRACT

Scoring is the most frequently used method of sensory testing of food for quality, especially by the dairy industry. The diversity and apparent simplicity of score cards are primary reasons for this widespread use. Dairy product score cards provide a numerical standard by which several quality parameters can be measured. Beginning students or amateur judges have established more consistent judging skills by following a definite, orderly routine through the application of score cards. Dairy product score cards provide the benefits: 1) educate, 2) develop improved judging procedures and habits, 3) minimize mechanical errors in scoring, 4) save time, 5) provide permanent records, and 6) serve as a guide to maintenance and improvement of dairy product quality. Revision of score card format could improve cards as educational tools for students.

INTRODUCTION

The title of this symposium topic suggests to the author a sense of foreboding, and it tends to flash a warning that there may be controversy just around the corner. Since flavor perception requires the direct participation of sensory receptors, this inherently qualifies this human process as an individual and variable characteristic (7). Hence, we conclude that flavor and flavor perception are necessarily highly subjective.

Behavioral sciences have changed substantially the food industry's outlook on the aforementioned principles. By the same token, the quality image of US dairy products has benefited significantly from application of dairy product score cards within production and quality control activities.

QUALITY DEFINED

Several definitions for food quality have been suggested. Kramer (4) defined food quality as "the composite of those characteristics that differentiate individual units of a product and have significance in determining the degree of acceptability of that unit by the user." An operational definition for quality (1) is "a composite response derived from all the sensory properties of a specific food that cause it to be judged superior by users who have been exposed to a random selection of the product over a period of time."

Briefly, food quality can be defined as the "relative degree of excellence." How close is the given food or beverage to the ideal or the perfect product that so-called "experts" or experienced judges may have in mind? This is the question that users of dairy product score cards frequently are asking.

Amerine et al. (1) have stressed that "consistent and reproducible evaluations of quality can be made only by judges who have a complete frame of reference, derived from experience and prolonged exposure to the specific commodity." Such individuals could be referred to as "prospectors for quality"; some notable experts are consultants on quality concerns. When untrained or inexperienced judges make quality evaluations, the results are often difficult to duplicate. An examination or sensory analysis based on the absence of defects, rather than a descriptive definition of quality, is the criterion generally applied in the score card approach to quality evaluation.

QUANTITY -- QUALITY EVALUATIONS

Laboratory panels, with a range in training from high to low, frequently are used in the food industry and food research either to: 1) measure differences in sensory properties

Received August 1, 1980.
between treatments, 2) establish the intensity of a sensory characteristic, or 3) assess the overall quality of a food. Amerine et al. (1) refers to these as laboratory studies for quantity — quality evaluations.

Sensory intensity generally is not difficult to assess; however, "quality" is more elusive and poses considerable difficulty in establishing the frame of reference, definition, measurement, and interpretation. The specific use of the product, the condition of testing, and the frame of reference of the judge are important factors that determine quality standards.

Scoring (facilitated through standardized score cards) is one of numerous basic types of laboratory studies or panels employed in an effort to accomplish quantity-quality evaluations of foods.

**SCORING OF FOODS (JUDGING)**

The most frequently used of all sensory testing systems is scoring. This is reportedly because of its diversity, apparent simplicity, and relative ease of statistical treatment. The dairy industry has used extensively scoring or judging methods for products such as butter and milk since the late 1800's.

**Score Cards**

Nelson and Trout (5) defined a score card as "a tabulated list of the factors contributing toward the quality of a product with a numerical value assigned to each factor."

Typical factors for scoring or evaluation of dairy products include: 1) flavor, 2) body and texture, and 3) color and appearance. Other factors may include package appearance, melting characteristics (for frozen dairy deserts), microbiological aspects, etc. These factors usually are arranged on the score card in the order of their relative importance to the overall quality of the given product. In essence "a dairy product score card becomes a numerical standard by which the quality of various dairy products can be measured" (5).

Traditionally, the sum of the numerical ratings or evaluations of all essential factors have combined to make a theoretically perfect or "ideal" product. This usually has been assigned the score of 100. The educational value of various types of score cards in different segments of the dairy industry long has been recognized and practiced. Examples are dairy farm and dairy plant inspection forms used by local, state, and federal sanitarians. These rating forms are essentially score cards; many of them are based on 100 as theoretically perfect.

A score card for a specific dairy product has provided a definite systematic means of arriving at a concise quality of the product(s) scored (5). The card is divided into component parts; each part has an assigned numerical score and range; hence, the comparative importance of the different items that should be considered are apparent.

The use of properly designed score cards has enabled the beginning student or amateur judge to establish more accurate and consistent judging habits by following a definite, orderly routine. This technique of evaluation requires a minimum of time and effort to achieve or derive results. The score card serves as a convenient list, a reminder of the possible defects within each sector of the card.

The score card also serves as a means of keeping an accurate, detailed record of the different samples that have been examined. The use of a score card (5) appears to provide these advantages: 1) educate, 2) develop improved judging procedures and habits, 3) minimize errors, 4) save time, 5) provide a permanent record, and 6) serve as a guide to quality improvement.

Additional advantages that can be cited for the application of dairy product score cards, as quoted from Nelson and Trout (4), are:

1. Standardized format, factors, and terms.
2. Enjoy wide use by industry, government, trade associations, and educational institutions.
3. Very useful in quality assurance programs — tend to be specific, relate to cause(s) of defect(s).
4. Simplicity of application and use.
5. Provide meaningful, permanent records of sensory perception of products.
7. Provides opportunities for students to develop and maintain competency in sensory evaluation.
Application of Dairy Products Score Cards

The primary users of dairy product score cards are:

1. Dairy processors.
   a. In-plant comparisons of manufactured products with competitor brands.
2. Dairy products trade associations (local, state, national).
   a. As an evaluation tool for educational clinics and demonstrations.
   b. Systematic scoring for quality contests.
3. Governmental agencies.
   a. USDA grades for butter, cheddar, and Swiss cheese etc.
   b. State Dept. of Agriculture marketing programs (quality grades for various dairy products).
4. Educational institutions.
   a. University students frequently initiated to sensory evaluation of dairy foods through use of score cards.
   b. Approximately 25 to 30 Universities, Junior Colleges, and Agricultural and Technical schools continue to participate in the annual Collegiate Dairy Products Evaluation Contest.
5. Research activities.
   Occasionally employs score cards (or modified cards) to help develop sensory evaluation data for:
   a. Product development.
   b. Quality improvement.
   c. Affects of processing on sensory characteristics.
   d. Affects of storage time on sensory properties and shelf life (palatability).

Format of Dairy Product Score Cards

Since advent of the score card for evaluation of dairy products, some differences or variation in the form or arrangement of the score card are notable. The card format is largely dependent on the use to be made of it.

Plank (6) stated that the fundamental supposition of rational quality grading is that the numerical score that expresses the grade is proportional to the property to be measured. In essence, the various developers of dairy product score cards generally have achieved this objective, as witnessed by the following weighting.

<table>
<thead>
<tr>
<th>Property</th>
<th>Previous Score</th>
<th>Current Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flavor</td>
<td>45</td>
<td>10</td>
</tr>
<tr>
<td>Body and texture</td>
<td>30</td>
<td>5</td>
</tr>
<tr>
<td>Color and appearance</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Melting characteristics</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Prerequisites for Effective Scoring

Anderson (2) outlined certain prerequisites for an effective scoring system as quoted here:

1. Development of a realistic score card with quality factors properly weighted to reflect their importance. Factors to be scored should be placed in logical order. First those estimated by sight, then odor, and finally those judged orally.
2. The scale should be such that a difference in score reflects a reproducible variation in the factors being scored (i.e., the scale range should be limited).
3. Agreement is necessary between judges as to standards of perfection so that the entire scoring range can be utilized and so that variability in scoring due to judges will be minimal.
4. The scaling system should lend itself to statistical analysis, but still recognize that a complex, highly sensitive, mathematical treatment does not and cannot overcome any deficiencies or lack of sensitivity in the original scoring or presentation of samples.
5. Physical and chemical analysis of the product should be employed to supplement sensory evaluation whenever possible.

There is a general tendency to be too arbitrary and to give a false sense of exactness (1) in scoring food quality through the use of score cards. Evaluators also frequently attempt to gather too much information from too many samples at one time.

Guthrie (3) reported that 50 to 100 samples of a dairy product could be scored more accurately than 5 to 10 samples in a given session. The author's experience has been that more precise scoring can be achieved with a limit of 10 to 30 samples, but it is partially dependent on the dairy product. The ability of a judge to detect small differences or to identify
replicate samples is considered by Amerine et al. (1) to be a better measure of persistence of taste acuity over prolonged periods of testing. The student judge of dairy products generally achieves recognition for his quality scoring prowess on the basis of his ability to duplicate the opinions of experts. However, this established “grading of students” method in competitions may serve to penalize students who have greater taste sensitivity than the experts.

The American Dairy Science Association and the Dairy and Food Industry Supply Association have sponsored judging contests for college students since 1916 to stimulate interest and foster uniform scoring of dairy products. The annual Collegiate Dairy Products Evaluation Contest is possible as an event of national scope primarily because of the extensive standardization of terminology applicable to specific defects for each product.

The success of scoring processes for quality assessment depend on judges that have standardized frames of reference and descriptive flavor terminology. More effort should be directed at successfully communicating specific experienced flavor perceptions and word descriptions of the flavor note, the defect, etc.

Communication Through Score Cards

When two or more individuals are attempting to communicate, a sequence of steps is helpful. The person (or future judge) to whom comments are addressed should understand: 1) the problem that is involved (basic flavor defect), 2) that the provided word description (semantic map) fits the actual flavor impression (the territory), and 3) that commonly familiar terms are used.

Several examples may help explain application of the above concepts in matching “common” terminology with a specific flavor note or defect. When description of a flavor note is doubtful, a phrase such as, “ten days on a troop train” or a “gross of unwashed athletic socks” will evoke a vivid image in the other person’s imagination and simultaneously suffice to describe an objectionable unclean or psychrotrophic off-flavor in market milk. “Graham cracker-like” fairly adequately describes the salty, slightly gritty, stale condensed milk off-flavor associated with the excessive use of whey solids in vanilla ice cream. Persons who have experienced the aroma of a “swamp” or “decaying potatoes in a damp cellar” quickly get the odorous picture of the “musty” flavor defect in dairy products.

In essence, word descriptions based on mutually acceptable experiences and objects, however clumsy or silly they may seem, can lead to productive communication. One can make the argument that flavor description should be approached from a common point of reference as people, rather than as scientists, if maximum communication is an objective. Educational goals also have some bearing on our choice of terms and the order in which the factors and terms are presented to the user on the score card.

**IMPROVED SCORE CARD ORGANIZATION AND FORMAT**

Most unfortunately the ADSA official score cards have listed the possible defects (within each factor) alphabetically. This makes it convenient to locate the “word map” or term to describe the noted defect(s) in samples, but it does not optimize the educational and quality assurance benefits of score cards. Students of quality judging most likely would benefit (and probably more rapidly develop their quality judging skills and competencies) if the given defects were organized by type or category. The following score card revisions suggest how this format might improve the score card format for vanilla ice cream and flavored yogurt.

<table>
<thead>
<tr>
<th>Revised Format – ADSA Ice Cream Score Card</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Flavoring material</strong></td>
</tr>
<tr>
<td>Lacks flavoring</td>
</tr>
<tr>
<td>Too high flavor</td>
</tr>
<tr>
<td><strong>Sweetener(s)</strong></td>
</tr>
<tr>
<td>Lacks sweetener</td>
</tr>
<tr>
<td>Too sweet</td>
</tr>
<tr>
<td><strong>Dairy ingredients and/or processing</strong></td>
</tr>
<tr>
<td>Cooked</td>
</tr>
<tr>
<td>High acid</td>
</tr>
<tr>
<td>Lacks freshness</td>
</tr>
<tr>
<td>Metallic</td>
</tr>
<tr>
<td>Old ingredient</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Storage</td>
</tr>
</tbody>
</table>
Revised Format – ADSA Flavored Yogurt Score Card

<table>
<thead>
<tr>
<th>Flavoring</th>
<th>Lacks flavoring</th>
<th>Too high flavoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unnatural</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lacks fine flavor</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Acidity level and microbiological characteristics

<table>
<thead>
<tr>
<th>High acid</th>
<th>Low acid</th>
<th>Fruity/fermented</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sweetener

<table>
<thead>
<tr>
<th>Lacks sweetness</th>
<th>Too sweet</th>
<th></th>
</tr>
</thead>
</table>

Dairy ingredients

<table>
<thead>
<tr>
<th>Cooked</th>
<th>Oxidized</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lacks freshness</td>
<td>Rancid</td>
<td></td>
</tr>
<tr>
<td>Old ingredient</td>
<td>Unclean</td>
<td></td>
</tr>
</tbody>
</table>

Other

<table>
<thead>
<tr>
<th>Bitter</th>
<th>Stabilizer</th>
<th></th>
</tr>
</thead>
</table>

The appearance section of the yogurt score card also might benefit from a revised format as follows:

Flavored Yogurt Appearance Defects

1. Surface appearance
   - Foamy (nonhomogeneous) Shrunken
   - Free whey Surface growth

2. Color or hue
   - Atypical color Too light
   - Color "bleeding" Too dark

3. Fruit frequency
   - Excess fruit Lacks fruit

It also may benefit efforts in scoring cottage cheese if the following score card format were applied:

Revised Format – ADSA Cottage Cheese Score Card

1. Dairy ingredients
   - Bitter Metallic
   - Cooked Oxidized
   - Feed Rancid

2. Processing, sanitation, handling
   - Fermented/fermented Musty
   - Lacks freshness Unclean
   - Malty Yeasty

3. Lactic bacterial aspects
   - Acid Coarse
   - Bitter Flat

4. Processing, other ingredients
   - Foreign Stabilizer
   - High salt Coarse

EFFORTS OF ADSA COMMITTEES

An ADSA subcommittee on Flavor Nomenclature of Milk and Reference Standards toiled for approximately 4 yr to review and update off-flavor terminology. A comprehensive review of their findings, conclusions, and recommendations was published by Shipe et al. (8). Essentially, a classification of the types of off-flavor based on causes was suggested. However, the Dairy Products Evaluation Committee of the ADSA did not comply with their recommendations for revision of the milk score card. Resistance to change, tradition, and a general unwillingness to adapt a new schedule of terms was primarily responsible for the nonacceptance of the revised nomenclature by the aforementioned committee.

The ADSA Committee on Dairy Products Evaluation has introduced several revisions for the cottage cheese, milk, butter, and ice cream score cards within the last 5 yr to reflect changes in industry trends and technological developments. First, "lacks uniformity" was deleted from the cottage cheese score card since it readily was confused with the "shattered curd" defect on the color and appearance section.

"Whey" was added as a listed flavor defect for the ice cream and butter score cards. This off-flavor resulting from use of excessive amounts of dry cheese whey solids in ice cream (salty, stale, heated, and gritty texture) is described simply by the term "whey" (cause). To define better the harsh or slight cheese-like odor in much whey cream butter, the same term applies on the butter score card.

The more recent score card revisions (1979) deal with several recent technological changes in milk processing and distribution. For instance, "fruity/fermented" relates to specific off-flavors resulting from growth of psychrotrophic bacteria in milk. Allowance for light activated off-flavors in milk has been covered by a notation for "light induced", copper or iron induced off-flavors in milk are notated by checking for "metallic" or "oxidized".

Several years ago the ADSA Committee on Dairy Products Evaluation revised the so-called official score cards to reflect more appropriate qualities associated with the use of score cards. The respective quality classifications are now indicated and assigned points as follows:
<table>
<thead>
<tr>
<th>Quality classification</th>
<th>Flavor</th>
<th>Body and texture or color and appearance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent (no criticism)</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Good</td>
<td>8 to 9</td>
<td>3 to 4</td>
</tr>
<tr>
<td>Fair</td>
<td>6 to 7</td>
<td>2</td>
</tr>
<tr>
<td>Poor</td>
<td>5 or less</td>
<td>1 or less</td>
</tr>
</tbody>
</table>

To summarize, the use of dairy product score cards in sensory evaluation activities must be: 1) Continuously reevaluated to avoid obsolescence. 2) Updated for objectives and mechanics. 3) Simplified with respect to numerical scoring systems. 4) Modified to reflect changes in flavor nomenclature due to new knowledge in flavor chemistry. 5) More receptive to the application of science to evaluation, i.e., greater use of flavor standards and visual aids. 6) Better standardized through use of an updated glossary of terminology.

I believe that if we keep the above criteria in mind, the appropriate use of dairy product score cards will be consistent with dairy industry efforts to maintain and improve the sensory qualities of US dairy products. Score cards have demonstrated their value in helping train personnel in the skill of recognizing various defects in dairy products. Furthermore, score cards and the associated score guides have been useful communication tools for quality control personnel. The American Dairy Science Association through the Committee on Dairy Products Evaluation is in the best position with the best expertise and industry liaison to continue the task of development of improved methods for sensory evaluation of dairy products for quality.

REFERENCES