ABSTRACT

Evaluation of the sensory characteristics of food products has been, and will continue to be, the ultimate method for evaluating product quality. Sensory quality is a parameter that can be evaluated only by humans and consists of a series of tests or tools that can be applied objectively or subjectively within the constructs of carefully selected testing procedures and parameters. Depending on the chosen test, evaluators are able to probe areas of interest that are intrinsic product attributes (e.g., flavor profiles and off-flavors) as well as extrinsic measures (e.g., market penetration and consumer perception). This review outlines the literature pertaining to relevant testing procedures and studies of the history of sensory analysis of fluid milk. In addition, evaluation methods outside of traditional sensory techniques and future outlooks on the subject of sensory analysis of fluid milk are explored and presented.

Key words: sensory analysis, flavor, fluid milk

INTRODUCTION

Sensory evaluation is critical for every application of milk. It is necessary to understand the sensory qualities of milk in part because of the widespread familiarity of fluid milk and its typical sensory profile. Sensory evaluation of the flavor or at least the aroma of raw milk can identify handling or production problems before milk is processed. In the processing and preparation of commercial milk products, fluid milk may be exposed to multiple unit operations at varying temperatures. In turn, sensory evaluation of the finished milk product helps identify deviations in processing or handling. In many cases, the deviation of quality may not be significant day-to-day changes but rather a drift over time, which requires frequent sensory evaluation and strong documentation of evaluations to successfully address areas of concern.

Though formal sensory analysis as we know it today is a relatively new practice, sensory measures of food quality have been practiced and documented throughout history (see Appendix Table A1). As early as the 1800s, studies focused on understanding human psychometrics (the study of quantitatively explaining human perceptions and decision making) and psychology as well as the statistical relevance on which those stimuli should be judged (Fechner, 1860; Thurstone, 1931). Eventually, those theoretical practices gave way to a practical desire for understanding consumer perceptions, especially as they applied to food. By the 1940s, affective consumer testing approaches, in conjunction with the US Army Corps of Engineers’ 9-point hedonic scale methodologies for measuring acceptability (Peryam and Pilgrim, 1957), had become a regular practice among many US food companies. Sensory evaluation of milk traditionally has been based on the identification of off-flavors or defects. The dairy product scorecard for fluid milk quality, which was based on defect identification, was first proposed by the Federal Dairy Division in the early 1900s, although several scorecards relating to milk handling and cleanliness were in circulation well before (North, 1917; Harding, 1921). Newer mainstream sensory approaches have been applied to fluid milk research and investigations into predicting and preserving acceptable milk quality. This review addresses a holistic view of the sensory history of fluid milk as well as the constituents, processes, and other factors that have contributed, and continue to contribute, to the sensory properties of fluid milk.

SENSORY EVALUATION TECHNIQUES

Quality Judging

The first standardized method for the sensory evaluation of dairy products was dairy product judging and the American Dairy Science Association (ADSA) scorecard system (Clark and Costello, 2008). As branding became an established concept in the early 20th century, companies began to turn to officially recog-
nized standards of quality to promote their quality to consumers. In 1917, when the first National Collegiate Dairy Products Evaluation Contest was held for milk, a USDA-developed, ADSA-approved scorecard was used that considered flavor as well as bacterial content, sediment, temperature, acidity, and bottle and cap appearance (Clark and Costello, 2008). In addition to the ASDA scorecard system, several other scoring systems were used in the 1920s and 1930s, and there was often debate within the industry about how scorecards should be used. By the early 1930s, the ADSA scorecard had become the standard scorecard for judging fluid milk (Nelson and Trout, 1934; see Table 1).

Traditional quality judging techniques are defect oriented and use 1 or 2 trained judges to document defects rather than profile the intensities of sensory attributes. By this approach, a large number of samples can be rapidly screened for recognized sensory defects. Early sensory studies on milk used quality judging techniques because better techniques were not yet available (Weaver et al., 1935; Hening and Dahlberg, 1939; Kratzer et al., 1987). These tests were designed to link a designated sensory defect with a specific root cause. Quality judging techniques are useful for on-the-spot evaluations of quality in industrial settings and for judging dairy competitions, but they are of limited utility for research for numerous reasons that have been reviewed previously (Drake, 2004, 2007; Alvarez, 2009). The modern ADSA milk and cream scorecard grades milk on a 0-to-10 scale, placing milks into categories of excellent (10), good (7–9), fair (4–6), poor (1–3), and unacceptable (0; see Table 2; Alvarez, 2009). Points are deducted for specific defects and their perceived intensities. In cases where a milk sample exhibits multiple defects, it typically is assigned a flavor score based on the most serious defect (Alvarez, 2009).

Many of the sensory defects found on the modern scorecard have remained unchanged from the 1934 version, although some changes have been made. Due to modern dairy sanitation measures, cowy, barny, and unclean flavors are rarely found in samples to be judged and therefore are not usually printed on the scorecard, and musty/stale has been removed as a defect entirely. Astringent, a defect added after 1934, is also rarely encountered and not usually printed (Alvarez, 2009). Cardboard and disinfectant attributes have been removed from the scorecard and are now considered part of the defect foreign, a defect term used to refer to atypical off-flavors or aromas from varied sources not commonly found in milk. The metallic defect has been clarified as metal oxidized, and light oxidized has been added as a defect. The defect cooked, once perceived as a severe defect before the widespread pasteurization of milk, is now viewed with far less criticism (Alvarez, 2009). When dairy judging contests first began, raw whole milk was evaluated. Eventually, pasteurized whole milk and, subsequently, pasteurized 2% milk replaced raw whole milk in dairy judging contests (Clark and Costello, 2008). Lacking richness, a defect from the 1934 scorecard associated with skim milk, has been removed from the modern scorecard, perhaps due to the change to evaluating reduced-fat milks instead of whole milks. Flat, a term on the current scorecard, was added to recognize the rare adulteration of milk with water.

Descriptive sensory analysis, developed in the 1950s, has slowly replaced quality judging techniques for all published research due to its versatility, specificity, and statistical robustness. Descriptive analysis uses

<table>
<thead>
<tr>
<th>Table 1. Undesirable flavors in milk, 19341</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defect</td>
</tr>
<tr>
<td>Barny, cowy</td>
</tr>
<tr>
<td>Bitter</td>
</tr>
<tr>
<td>Cardboard</td>
</tr>
<tr>
<td>Cooked</td>
</tr>
<tr>
<td>Disinfectant</td>
</tr>
<tr>
<td>Feedy, silage</td>
</tr>
<tr>
<td>Flat, watery</td>
</tr>
<tr>
<td>Garlic, leek, onion</td>
</tr>
<tr>
<td>Lacking richness</td>
</tr>
<tr>
<td>Malty</td>
</tr>
<tr>
<td>Metallic</td>
</tr>
<tr>
<td>Musty, stale</td>
</tr>
<tr>
<td>Rancid, strong</td>
</tr>
<tr>
<td>Salty</td>
</tr>
<tr>
<td>Sour</td>
</tr>
<tr>
<td>Unclean</td>
</tr>
</tbody>
</table>

1Data from Nelson and Trout (1934).
a group of 6 to 12 trained individuals to document intensities of the sensory attributes of a product (Table 3). Claassen and Lawless (1992) compared the ability of a defect-oriented system panel to detect differences in light-oxidized, metallic-oxidized, and rancid milk with that of a descriptive analysis panel. Given the same amount of training, the descriptive analysis panel was more sensitive to differences than the traditional judging panel. Although traditional quality judging methods imply hedonics, they are not good indicators

### Table 2. The 2005 ADSA scoring guide for off-flavors and defects of milk and cream

<table>
<thead>
<tr>
<th>Flavor defect</th>
<th>Description</th>
<th>Training reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acid</td>
<td>Sour off-flavor due to acid-producing organisms such as <em>Lactococcus lactis</em> ssp. cremoris</td>
<td>Add 6–7 mL of a 10% lactic acid solution to ~600 mL of milk.</td>
</tr>
<tr>
<td>Astringent(^2)</td>
<td>Puckering sensation on tongue and lining of the mouth; rare in milk but associated with rancidity</td>
<td>NA(^3)</td>
</tr>
<tr>
<td>Cowy/barny/unclean(^2)</td>
<td>Distinct cow breath–like odor and unpleasant, medicinal, or chemical aftertaste; suggestive of a poorly maintained barn</td>
<td>NA</td>
</tr>
<tr>
<td>Bitter</td>
<td>Persistent bitter taste detected at the base of the tongue; commonly caused by specific weeds consumed as part of roughage by cows or by proteolysis of milk proteins by microorganisms (especially psychrotrophic bacteria)</td>
<td>Add 2–2.5 mL of a 0.1% quinine sulfate solution to ~600 mL of milk.</td>
</tr>
<tr>
<td>Cooked</td>
<td>Sulfurous, heated, caramelized, or scorched flavors</td>
<td>Heat a working quantity of milk in a vessel to 80°C and hold for 1 min.</td>
</tr>
<tr>
<td>Feed</td>
<td>Aromatic taints resulting from cows consuming some feeds within a critical time frame before milking; have a characteristic cleanliness note and a mild aftertaste that disappears quickly</td>
<td>Add 4–7 mL of a prepared “tea” (brew alfalfa in water) to 600 mL of milk.</td>
</tr>
<tr>
<td>Fermented/fruity</td>
<td>May resemble the odor of sauerkraut, vinegar, pineapple, apple, or other fruit; commonly caused by the growth of microorganisms</td>
<td>Mix 6 parts pineapple juice and 1 part vinegar. Add 3–4 mL of mixture to 600 mL of milk.</td>
</tr>
<tr>
<td>Flat</td>
<td>Simulated by adding water to a sample of milk and noticing the alteration of mouthfeel of the mixture</td>
<td>Add ~20% water to 2% milk.</td>
</tr>
<tr>
<td>Foreign/chemical</td>
<td>Chemical flavor that may be caused by improper use of detergents, disinfectants, and sanitizers; exposure to gasoline or kerosene fumes; or contamination from insecticides or medicines</td>
<td>Add 2 mL of a 200-µL/L chlorine solution to 600 mL of milk immediately before presentation.</td>
</tr>
<tr>
<td>Garlic/onion</td>
<td>Weedy, pungent odors and somewhat persistent aftertaste</td>
<td>Add 2 mL of a 1% garlic powder mix (in water) to the milk; infuse a clove of garlic for 2 h and then either decant the milk or retrieve the clove. Open a carton of milk and store in the refrigerator for ≥7 d; use an unopened carton 1 wk beyond the pull date; add 10–15 g of skim milk powder to 600 mL of milk.</td>
</tr>
<tr>
<td>Lacks freshness</td>
<td>Stale, chalky flavor, lack of sweetness</td>
<td>Add 0.25–0.5 g of table salt in 600 mL of milk.</td>
</tr>
<tr>
<td>Malty</td>
<td>Suggestive of malt or Grape Nuts cereal; generally caused by growth of <em>Streptococcus lactis</em> ssp. <em>maltigenes</em></td>
<td>Add 1 g of malt powder to 1 L of warm milk; add 15 g of Grape Nuts to 100 mL of milk and infuse for 20–30 min before adding aliquots to ~600 mL of milk.</td>
</tr>
<tr>
<td>Oxidized, light</td>
<td>Burnt, burnt protein/feather-like, cabbage-like, medicinal, or chemical off-flavors resulting from light exposure</td>
<td>Transfer milk to a clear glass bottle and place on a windowsill exposed to direct, bright sunlight for a duration proportional to the intensity of the defect. Immerse a copper penny or wire in milk overnight; add several drops of 1% copper sulfate solution to 600 mL of milk and leave in a refrigerator for 24 h. Add 0.5 g of lipase powder to ~600 mL of milk, agitate, and hold at 21°C for 1 h; add a few drops of dilute butyric acid solution to ~600 mL of milk. Dissolve 0.25–0.5 g of table salt in 600 mL of milk.</td>
</tr>
<tr>
<td>Oxidized, metal</td>
<td>Metallic, oily, cappy, cardboard, stale, tallowy, painty, or fishy off-flavor commonly induced by the catalytic action of certain metals; characterized by a puckery mouthfeel</td>
<td>Add 0.5 g of lipase powder to ~600 mL of milk, agitate, and hold at 21°C for 1 h; add a few drops of dilute butyric acid solution to ~600 mL of milk. Dissolve 0.25–0.5 g of table salt in 600 mL of milk.</td>
</tr>
<tr>
<td>Rancid</td>
<td>Baby burp, feta cheese, or butyric acid aromas formed as a result of lipid hydrolysis</td>
<td>Add 0.5 g of lipase powder to ~600 mL of milk, agitate, and hold at 21°C for 1 h; add a few drops of dilute butyric acid solution to ~600 mL of milk. Dissolve 0.25–0.5 g of table salt in 600 mL of milk.</td>
</tr>
<tr>
<td>Salty</td>
<td>Commonly associated with milk from cows in advanced stages of lactation or with clinical mastitis, resulting in an increase of NaCl in the milk and a decrease of other mineral salts</td>
<td>Combine rancid, fruity, and bitter milks; mix spoiled milk (≥7–10 d beyond sell by date) with fresh milk.</td>
</tr>
<tr>
<td>Unclean</td>
<td>Offensive odor suggesting extreme staleness, mustiness, dirty socks, or foul stable air; may develop due to the action of certain psychrotrophic bacteria</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\)Data from Alvarez (2009).

\(^2\)Uncommon defect; therefore, not printed on the 2005 ADSA scorecard.

\(^3\)NA = not applicable or not listed.
of consumer liking (McBride and Hall, 1979; Bodyfelt, 1981; Sidel et al., 1981; Drake, 2004). Lawless and Claassen (1993) compared the correlation of consumer liking scores with data generated by traditional dairy judging and descriptive analysis panels and reported that descriptive analysis was more likely to correlate with consumer liking (although the method itself does not attempt to predict consumer liking). Defect quality judging remains a useful industrial quality test; its limitations and the role of modern sensory techniques have been reviewed elsewhere (Drake, 2004, 2007; Bodyfelt et al., 2008).

**Analytical Sensory Tools**

Analytical sensory tests are objective sensory tests that document sensory profiles, sample variability, or other product features free from liking considerations. These methods can be used to objectively profile products and have been applied extensively to fluid milk. Among the most common analytical sensory tests used are discrimination tests. Discrimination tests are simple tests that aim to identify whether a significant difference exists between 2 or more samples, and results can easily be ascertained by referencing published significance tables based on the binomial distribution (Lawless and Heymann, 2010). The most common discrimination tests used in the study of fluid milk have been paired-preference, duo-trio, and triangle testing, although tetrad testing has also been used (Carlisle, 2014). Bierman et al. (1956) used triangle testing methods to evaluate whether irradiation treatments resulted in noticeable differences in milk and cream products. In addition, Modler et al. (1977) used triangle tests to determine whether consumers could notice differences between milks with varying feed flavor intensities followed by paired-preference tests to determine whether noticeable differences were preferred. Consumers could distinguish control 2% milks from milks with pronounced feed flavors, and they preferred control milk samples. Difference testing is still widely used in quality control capacities for the milk industry as well as in academic research (Lee et al., 2016; McCarthy et al., 2017a; Yeh et al., 2017a).

### Table 3. Trained panel sensory attributes for fluid milk

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aroma intensity</td>
<td>The overall orthonasal effect of the sample</td>
<td>NA²</td>
</tr>
<tr>
<td>Sweet aromatic</td>
<td>Sweet aromatics generally associated with materials also having a sweet taste</td>
<td>Vanillin, caramelized sugar</td>
</tr>
<tr>
<td>Cooked</td>
<td>Aromatics associated with cooked milk</td>
<td>Skim milk heated to 85°C for 30 min</td>
</tr>
<tr>
<td>Sulfur/eggy</td>
<td>Aromatics associated with sulfur compounds</td>
<td>Boiled egg, H₂S bubbled through water, freshly struck match</td>
</tr>
<tr>
<td>Milk fat/lactone</td>
<td>Aromatics characteristic of milk fat, lactones, and coconut</td>
<td>Fresh coconut meat, heavy cream, δ-dodecalactone (40 mg/kg)</td>
</tr>
<tr>
<td>Caramelized</td>
<td>Aromatics associated with caramel</td>
<td>Sweetened condensed milk, burnt sugar</td>
</tr>
<tr>
<td>Butterscotch</td>
<td>Aromatics associated with butterscotch candies</td>
<td>Butterscotch candies</td>
</tr>
<tr>
<td>Feed/malty/silage</td>
<td>Aromatics associated with a mixture of grains and fermented hay and cattle feed</td>
<td>Corn silage, malt extract, freshly kilned malt, 2/3 methyl butanal</td>
</tr>
<tr>
<td>Grassy</td>
<td>Green, sweet aromatics associated with cut grass</td>
<td>Fresh-cut grass, hay, cis-3-hexanol (50 mg/kg)</td>
</tr>
<tr>
<td>Cowy/barny/phenolic</td>
<td>Aromas associated with barns and stock trailers; indicative of animal sweat and waste</td>
<td>Band-Aids, cresol (160 mg/kg)</td>
</tr>
<tr>
<td>Metallic/serum</td>
<td>Aromatics associated with metals or juices of raw or rare beef</td>
<td>Fresh raw beef steak or ground beef or juices from seared beef</td>
</tr>
<tr>
<td>Fruity</td>
<td>Aromas associated with fruit (e.g., pineapple, strawberry)</td>
<td>Fresh pineapple, fresh strawberries</td>
</tr>
<tr>
<td>Cardboard</td>
<td>Aromatics associated with the aroma of wet cardboard</td>
<td>Wet cardboard</td>
</tr>
<tr>
<td>Carroty</td>
<td>Aromatics associated with canned carrots</td>
<td>Canned carrots</td>
</tr>
<tr>
<td>Sweet taste</td>
<td>Fundamental taste sensation elicited by sugars</td>
<td>Sucrose (5% in water)</td>
</tr>
<tr>
<td>Bitter taste</td>
<td>Bitter basic taste</td>
<td>Caffeine (0.08% in water)</td>
</tr>
<tr>
<td>Metallic mouthfeel</td>
<td>The aftertaste or feeling on the oral surfaces associated with CuSO₄ solution and many nonnutritive sweeteners</td>
<td>Copper sulfate diluted (1%) in water or milk</td>
</tr>
<tr>
<td>Opacity</td>
<td>Visual term denoting the degree of opacity</td>
<td>Water = 0.0; whole-fat milk = 12</td>
</tr>
<tr>
<td>Yellow color</td>
<td>Degree of yellow color visible to the human eye</td>
<td>Behr paint chips: Ultra-Pure White (PPU18-06) = 0.0; Glass of Milk (P260-1u) = 3.5</td>
</tr>
<tr>
<td>Astringency</td>
<td>Chemical feeling factor on the tongue or oral cavity described as puckering or dry</td>
<td>Alum (1% in water)</td>
</tr>
<tr>
<td>Viscosity</td>
<td>Amount of force required to slurp 4.93 mL (1 tsp.) of liquid from a spoon over the lips</td>
<td>Water = 1.0; heavy cream = 3.2</td>
</tr>
</tbody>
</table>

¹Sources: Russell et al. (2006), Croissant et al. (2007), Brothersen et al. (2016), Lee et al. (2017), McCarthy et al. (2017a), and Yeh et al. (2017a).

²NA = not applicable.
Descriptive analysis techniques are valuable methods for quantitatively and objectively profiling the sensory attributes of fluid milk. Descriptive analysis requires training of panelists before data collection (Chambers et al., 2004). Panel training may take several hours; however, this extensive training is done in the hopes that the panel may produce objective data that are consistent and sensitive, analogous to an instrument. Many fluid milk studies have used descriptive analysis to evaluate and differentiate samples (Claassen and Lawless, 1992; Lawless and Claassen, 1993; Phillips et al., 1995a; Watson and McEwan, 1995; Chapman and Boor, 2001; Francis et al., 2005; Chung et al., 2008; McCarthy et al., 2017a; Lee et al., 2017). Furthermore, descriptive analysis data can be paired with consumer panel data in a technique known as preference mapping to better understand drivers of liking in a product (MacFie and Thomson, 1988; Thompson et al., 2004).

**Consumer-Focused Sensory Tests:**

**Affective Sensory Tests**

Evaluation of consumer acceptance has been integral to ensuring the acceptability of various fluid milk products and treatments since the inception of hedonic scaling methods in the 1940s. Consumer tests may be administered many ways with fluid milk, but central location tests and home usage tests are the tests most frequently used. True to their title, consumer evaluations are administered to untrained populations who represent the true consumer base of a product. Early studies on the hedonic qualities of fluid milk attempted to extrapolate consumer acceptance from trained panelists; however, deficiencies in the ability of trained panelists to predict the preferences of consumer populations were well documented (Bierman et al., 1956; Ellis, 1969). Practical uses of consumer tests include examining the effects of various processing methodologies (Horner et al., 1980; Gandy et al., 2008; Lee et al., 2017), flavor additions or fortifications (Campbell et al., 2003; Achanta et al., 2007), and shelf life (Hansen et al., 1980) of fluid milks to maintain adequate consumer acceptance and lead new product development.

Consumer studies may also focus on qualitative aspects of the consumer experience. Qualitative consumer data are often collected in the form of free-response comments, check-all-that-apply questions, or, in some cases, organized interview methods such as focus groups. In studies of light-induced oxidation effects on consumer liking of 2% milks, Walsh et al. (2015) used emotional check-all-that-apply questions and found that significantly higher frequencies of terms such as happy, safe, warm, and whole corresponded with higher hedonic scores. One particularly useful consumer testing question type pertaining to product development is the just-about-right scale. These scales are an excellent tool for understanding the influence of individual sensory attributes or product qualities on overall liking. They evaluate individual attributes of a product and deviation from just-about-right categorization to determine effects of those attributes on overall liking. Just-about-right scales have been used extensively in studies investigating consumer acceptance of functional or flavored milk beverages, such as probiotic milks (Villegas et al., 2010), regionally flavored milks (Zhi et al., 2016), and coffee-flavored milks (Li et al., 2014, 2015a), as well as different pasteurization methods (Chapman and Boor, 2001; Lee et al., 2017).

Though no technological advance or survey questionnaire can ever truly replace the human evaluation of food products in determining sensory qualities and preferences, surveys provide a cost-effective way to reach large numbers of consumers. Several studies have exposed that perception has a significant effect on the eventual sensory evaluation of food and beverage products (Liem et al., 2012; Norton et al., 2013; Fernqvist and Ekelund, 2014). Survey techniques can effectively probe these perceptions in several ways. Simple usage and attitude questions can be asked and liking and purchase intent related to concepts or photos can be examined to elucidate consumer tendencies, as exemplified in studies by Allen and Goddard (2012). In addition, surveys can apply complex statistical designs to design tradeoff exercises or product optimization exercises such as maximum difference scaling or conjoint analysis; these methods have been applied directly to fluid milk studies (Bai et al., 2007; Amadou and Baky, 2015; McCarthy et al., 2017b). These survey methods give insight into the relative importance of product attributes and help identify factors that significantly influence consumer choice.

Early initiatives in probing consumer attitudes of milk related strongly to understanding consumer beliefs as they pertained to milk nutritional qualities, with little focus on relation to sensory properties or hedonics (Martin et al., 2005). Although these insights mirror beliefs of the time, they ultimately convey more utilitarian information about milk consumption. With the growth of consumer-focused testing in the 1950s, survey objectives became more centered on the psychology of the consumer; however, there remained a perceived lack of focus on understanding flavor perceptions in milk (Trout, 1956). In a study of milk consumers, Swope and Nolan (1959) reported that taste, not nutrition, was the leading factor in consumption of milk following a questioning of 1,393 milk consumers, echoing the need for consumer insight in a changing market. Efforts to screen milk products for flavor defects played...
a large part in assuaging these concerns for the general consumer. By the late 1960s, sentiments about fluid milk had become increasingly influenced by factors beyond the milk itself. A marketing research survey report by the USDA (Fallert, 1971) revealed that milk consumers were primarily influenced by brand recognition and availability in their purchasing tendencies. Similar tendencies have been observed in subsequent decades with surveys focused on certain aspects of the final fluid milk product—for example, rBST free milk (Puetz, 2013), organic milk (Brooks and Lusk, 2010; Schroeter et al., 2016), and locally farmed milk (Pirog, 2004)—many of which do not play a significant role in increasing sensory acceptance when not explicitly revealed to the consumer (Grobe et al., 1996; Boppanna, 2007; Kouřimská et al., 2014).

Conjoint analysis can clarify the dichotomy between preconceived perceptions and actual sensory evaluation. Conjoint analysis uses discrete-choice data of tradeoff scenarios to build multiattribute utility models for the prediction of consumer choice in a hypothetical market (Green et al., 2001). Although conjoint analysis techniques have been used in marketing capacities since the 1970s, modern computational strength allows for more streamlined and even adaptive approaches at gauging consumer perception of different product attributes (Cunningham et al., 2010). Adaptive conjoint methods use algorithms to generally remove undesirable concepts for a given consumer as the ballot progresses. Based on previous answers, these adapted choices are weighted depending on their position in the survey and are designed to better understand differentiating features of product concepts using less time, a smaller number of questions, and a smaller number of panelists compared with traditional conjoint analysis. Using conjoint analysis, Bai et al. (2007) identified attributes such as low fat content, HTST pasteurization, and natural taste as drivers of liking for fluid milk. In a study by Boesch (2013), Swiss milk consumers preferred milk that was GMO free (a particular concern) and of local origin. More recently, McCarthy et al. (2017b) used a conjoint analysis survey with 702 dairy consumers followed by individual interviews with 172 consumers to probe consumer key drivers for purchase, beliefs, and values for milk compared with plant-based beverages. Tasting has not traditionally been a component of conjoint analysis in foods. Some recent studies have examined the validity of a tasting component in conjoint analysis, although none have yet done so in relation to fluid milk specifically (Vickers, 1993; Haddad et al., 2007). Use of innovative survey methods for understanding consumer perception, such as conjoint, is likely to expand in coming years because of the insight they give into the sensory experience of consumers.

**Instrumental Analysis of Flavors**

Sensory techniques have developed alongside instrumental techniques. Although instruments do not measure flavor, instrumental data can go hand in hand with sensory data and be applied to more clearly identify sources of flavors. As new instrumental methods for identifying and quantifying volatile compounds have developed, they have been applied as tools for understanding the mechanisms behind sensory differences. However, no instrumental method can stand on its own without sensory data because, ultimately, the human palate is more sensitive and able to grasp complex sensations far more effectively than any technology developed.

Volatile compounds are the source of aromas in fluid milk, so GC is an excellent tool for identifying possible sources of milk flavors. Partition chromatography was one of the predecessors to CG, developed in the 1940s. Prior to partition chromatography, paper chromatography methods were sometimes used to analyze volatiles in milk; however, these methods were extremely time consuming, required large sample volumes, and were difficult to reproduce (Wong and Patton, 1962). An early, pre-GC-olfactometry method for determining flavor significance of volatile compounds used a threshold test that involved spraying samples into the mouths of 5 trained panelists, with the threshold determined to be the 50% positive response level (Patton and Josephson, 1957).

Modern GC was developed in 1952. Methods papers published after the development of GC technology attempted to standardize sampling procedures and minimize contamination from sources outside the sample (Sundararajan et al., 1967). During the 1960s and 1970s, many dairy studies used GC or GC-MS technology to focus on feed flavors in milk, measuring the volatiles of the milk, blood, or urine of cows fed different feeds (Loney et al., 1963; Bassette et al., 1966; Gordon and Morgan, 1972). Milk fat flavor and heat-induced flavor changes, very relevant to industrial processing, were also studied (Scanlan et al., 1968). Other efforts sought to quantify levels of volatile compounds identified in milk to get a baseline for further research (Bassette and Ward, 1974).

At this time, descriptive sensory analysis was still very new, and there were no clearly established methods for combining instrumental and sensory data. As quality judging was still a common research practice, some studies attempted to relate instrumental data to the flavor scores of the milk assigned by trained dairy judges. Keller and Kleyn (1972) related total peak areas of gas chromatograms with haylage flavor (a feed flavor defect) scores, finding a correlation between total...
Milk components relating to sensory quality

Fluid bovine milk has been lauded throughout history for its nutritional qualities, especially in the adolescent demographic. However, the macronutrient and mineral components are not the only factors that help determine the sensory profile of milk. The process of bringing milk to grocery vendors includes several factors such as the feed for cattle, the cattle themselves, pasteurization techniques, vitamin fortification, and packaging. All of these factors may affect the flavor of fluid milk.

Macronutrient Components

The sensory perception of fluid milk is heavily influenced by the balance of its macronutrient components. Milk fat plays a critical role in the sensory perception of fluid milk. Milk fat is preferred by all consumer segments at varying levels and is considered to be a contributor to creaminess, which is positively correlated with product liking (Richardson-Harman et al., 2000; McCarthy et al., 2017a). However, many consumers purchase reduced-fat milk despite a preference for whole milk due to health reasons, and 2% reduced-fat milk has outsold whole milk every month since January 2005 (Brewer et al., 1999; Economic Research Service, 2014; Bakke et al., 2016).

Visual, texture, and flavor attributes of milk are all influenced by milk fat (Phillips et al., 1995a; McCarthy et al., 2017a). Descriptive sensory analysis of fluid milks of varying fat percentages demonstrated that opacity, thickness, mouthcoating, viscosity, milk fat flavor, and yellow color increased with fat content (Phillips et al., 1995a; Francis et al., 2005; McCarthy et al., 2017a). Nonfat milk was described as more chalky, less viscous, less sweet, and higher in sour aromatic flavor than whole milk (Francis et al., 2005). When milks were tasted without visual cues, panelists were not able to perceive texture differences as clearly (Phillips et al., 1995a; McCarthy et al., 2017a). Fortifying milk with nonfat dry milk powder increased the viscosity and mouthcoating of reduced-fat milks but did not increase the visual whiteness and opacity (Phillips et al., 1995b). Although texture may play a role in differentiation of milk products at very high fat levels, research suggests that visual cues are most important for determination of fat content at the lower fat levels typically encountered by consumers (0–4%; Pangborn and Dunkley, 1964a, b; Pangborn and Giovanni, 1984; Pangborn et al., 1985; Mela, 1988; Phillips et al., 1995b; McCarthy et al., 2017a).

Visual and texture cues may account for much of the consumer ability to distinguish between fat contents of milk, but research has also demonstrated that milk fat contributes to flavor (McCarthy et al., 2017a). Milk fat comprises 98% 26–54 carbon triglycerides and 2% volatile lactones, ketones, and aldehydes, which contribute most of the flavor (Schaap and Badings, 1990). Creaminess—a desirable consumer attribute for dairy products, including milk—is defined by high-fat dairy products with cream aroma, butter aroma, and sweet aromatic/vanilla flavor, indicating that desirable flavor attributes are closely associated with fat content (Richardson-Harman et al., 2000). Consumers rated 1% milks as thicker and creamier after the addition of vanilla extract (Lawless and Clark, 1992). Milk fat has been reported to improve the aftertaste of milk; this is associated with increased consumer liking because unpleasant aftertaste is a driver of dislike for fluid skim milk (Porubcan and Vickers, 2005). Higher fat milks had more sweet-related attributes in the aftertaste compared with lower fat milks as well as fewer cooked flavors (Francis et al., 2005). Tepper and Kuang (1996) used multidimensional scaling techniques to evaluate the effect of addition of natural cream flavor to a skim milk base with fat added as bland vegetable oil and reported that the addition of cream flavor provided the sensation of higher fat content. McCarthy et al. (2017a) conducted qualitative interviews with milk consumers; both skim and 2% milk drinkers reported 2% milk to be...
differences and that texture and surface gloss altera-
tions were more likely to be affected than flavor. When
protein content was altered using reverse osmosis, the
SNF content also changed, resulting in significant sen-
sory differences (lower SNF samples were perceived as
watery). The author’s primary concern, however, was
to investigate the feasibility of protein standardization,
not its effect on sensory qualities of the milk (Poulsen,
1978). Later studies investigated the effects of alter-
ing protein levels on physical properties of fluid milk
such as freezing point. Sensory differences between
samples were minimal and were primarily driven by
visual attributes rather than flavor (Rattray and Jelen,
1996; Quiñones et al., 1997, 1998). The vast major-
ity of research on ultrafiltration of milk over the past
50 yr has evaluated the resulting effects on milk as
an ingredient for other dairy products, such as cheese
and yogurt, or investigated technical aspects of milk
filtration (Covacevich and Kosikowski, 1978; Yan et al.,
1979; Kapsimalis and Zall, 1981; Trachoo and Mistry,
1998; Méthot-Hains et al., 2016). The sensory proper-
ties and consumer acceptance of ultrafiltered milk have
not been widely documented.

Increased availability of cheaper advanced filtration
systems has renewed interest in studies of filtered fluid
milk. Recent studies used sophisticated filtering meth-
ods to investigate differences in the type of protein in
fluid milk and protein beverage products. Misawa et
al. (2016) reported that increasing the level of CN as
a percentage of true protein in reduced-fat milk had
more effects on visual sensory attributes than increas-
ing the percentage of true protein in the milk. Trained
panelists detected increased mouthcoating and throat
cling in increased-CN 2% fat milks even when visual at-
tributes of the samples were masked. As the popularity
of protein-enriched products increases and individual
protein components of milk increase in value, it is likely
that research investigating the sensory changes that
result from changing protein levels in milk will increase.

Lactose is the primary carbohydrate component of
fluid milk. It is estimated that 70 to 90% of the
world’s population is lactose intolerant, creating a
large potential market for commercial lactose-free dairy
products (Harrington and Mayberry, 2008; Adhikari et
al., 2010). Lactose-hydrolyzed milk products became
commercially available in the 1960s and 1970s. Enzy-
matic hydrolysis of lactose via addition of lactase is
the process by which lactose reduction traditionally
is achieved and has been shown to produce milk with
less than 0.01% lactose in modern applications (Jelen
and Tossavainen, 2003). Lactose hydrolysis results in a
milk that is sweeter than traditional milk because glu-
cose and galactose are sweeter than lactose (Li et al.,
2015b). In an early study on lactose-hydrolyzed milk,
Paige et al. (1975) reported this increased sweet taste

creamier and to have a fuller flavor. In contrast, skim
milk was described as watery and having a high after-
taste intensity. In general, milk consumers preferred a
slightly higher percentage of milk fat than the milk fat
percentage they typically bought but began to dislike
the higher fat milk when the thickness increased too far
past that of their typically purchased milk.

In an attempt to understand the contribution of vari-
ous milk attributes to the perception of milk fat, Bom
Frost et al. (2001) used descriptive analysis data to
study the synergistic effects of additives to mimic milk
fat. They reported that the addition of thickener and
whitener and the homogenization of milk resulted in
an increased perception of fat content and suggested
that although whitener and thickener contributed to
fattiness, homogenization had only a very slight effect.
Adding aroma compounds increased the intensity of
crunchy smell and flavor attributes, which were also cor-
related with milk fat. They reported that crunchy smell
and flavor, sweet taste, thickness (visual and oral), glass
coating, and residual mouthfeel all correlated with total
fattiness of milk, as did yellowness to a small degree.

Protein also contributes to the flavor of fluid milk.
Protein content of fluid milk can be adjusted by using
the ultrafiltered permeate and retentate of skim milk.
Ultrafiltration separates milk components due to mo-
lecular size, with lactose, soluble minerals, and water
passing through a membrane to become the permeate;
larger particles such as CN and whey protein cannot
pass through and become the retentate (Yan et al.,
1979). Today, ultrafiltration can be used to increase
protein content while decreasing the lactose content of
milk products, resulting in dairy beverages that appeal
to consumer health concerns without added ingredients.
Other types of membrane fractionation such as micro-
filtration, nanofiltration, and reverse osmosis have been
used to fractionate bacteria and spores, somatic cells,
proteins, salts, and water in milk (Brans et al., 2004).
Adjusting the levels of milk components can have huge
effects on the overall flavor of milk and can significantly
affect consumer liking. As use of these filtration tech-
nologies has expanded, so has research into the sensory
qualities of filtered fluid milk. The sensory qualities of
adjusted protein levels are of particular interest due
to growing consumer demand for higher protein dairy
products (Özer and Kirmaci, 2010).

Early research on ultrafiltered fluid milk evaluated
the effects of protein content on sensory attributes of
milk. Poulsen (1978) noted that fat was standardized
for fluid milk products and investigated the feasibility
of standardizing fluid milk protein content through ul-
trafiltration. He found that varying the protein content
between 3.1 and 6.4% resulted in no significant sensory
differences and that texture and surface gloss altera-

as well as increased staleness and sourness. Rather than focusing specifically on sensory properties, however, the authors’ main focus was evaluating differences in blood sugar increase and lactose malabsorption when lactose-intolerant panelists consumed lactose-hydrolyzed milk as opposed to regular milk (Paige et al., 1975). The consumers compared the lactose-hydrolyzed milk with fresh whole cow milk using a 7-point scale in which they could specify only a single attribute in which the lactose-hydrolyzed milk differed from the fresh whole milk. The type of pasteurization used was not specified. The sensory methods used in this study were not validated or repeatable, thus weakening the sensory conclusions of the study.

In a more recent study using trained descriptive analysis panelists, Chapman et al. (2001) reported that the increased sweetness of lactose-free milks could be causing a halo effect, resulting in a perceived increase in viscosity. In the United States today, lactose-free milks are typically ultrapasteurized (UP), resulting in more cooked and sulfurous flavors than traditional HTST milk, which is a confounding factor when comparing the sensory attributes of regular and lactose-free milks (Adhikari et al., 2010). Compared with HTST milks, UP lactose-free milks have higher intensities of chalky texture, lack of freshness, light-oxidized flavor, and processed flavor, all of which are perceived as negative attributes by consumers (Adhikari et al., 2010). However, these differences are not necessarily attributable to lactose hydrolysis and may be due to the UP pasteurization process, packaging and storage, or a combination of these factors (Antunes et al., 2014). No recent published studies have evaluated the effect of lactose hydrolysis as an isolated component on the sensory profile of milk.

Trace minerals in milk impart a salty taste to milk and milk products. Minerals impart a background salty taste to fluid milk that is not directly noticeable but can be more clearly perceived in milk or whey permeates. Sodium and potassium, the largest contributors to salty taste, are found in milk at concentrations of 391 to 644 and 1,212 to 1,681 mg/kg, respectively (Gaucheron, 2005). Milk permeate has a salty taste intensity on par with that of Cheddar whey—less than that of cottage cheese whey permeates, Mozzarella whey permeates, or delactosed permeates but still clearly detectable (Smith et al., 2016). Li et al. (2015b) added lactose-hydrolyzed whey permeates to chocolate milk to increase the sweetness of the milk but found that the increase in saltiness that resulted from the permeate minerals was too intense to make this a viable alternative sweetening method. Other trace components such as calcium, magnesium, chlorine, and organic acids may also contribute to milk basic taste (Smith et al., 2016).

Fortification

Fortification is defined as the process of adding micronutrients such as essential vitamins to food. Fortification of fluid milk began in the 1930s and 1940s when milk provided 10% of American consumer food energy (Yeh et al., 2017b). Vitamin D fortification became common after the American Medical Association’s Council on Foods and Nutrition recommended the practice to decrease rickets in children (Stevenson, 1955). The popularity of vitamin D-fortified milk led to fortification of reduced fat and fat-free milks with vitamin A (Public Health Service, 1940). Today, fortification of reduced-fat milks with vitamin A and D is mandatory to replace the fat-soluble vitamins lost when the cream is skimmed from the milk. Vitamin D fortification of whole milk is optional yet widespread (PHS/FDA, 2015).

Fortification is often used to increase milk consumption during a period of declining dairy consumption, some companies have turned to fortification with other ingredients thought by consumers to be health promoting. Calcium, iron, antioxidant vitamins C and E, fiber, multivitamin mixes, and PUFA have been investigated for use in fortified milk beverages (Chandan, 1999). Kwak et al. (2003) fortified fluid milk with microencapsulated iron; sensory differences in astringency, metallic flavor, color, and overall quality scores were reported. In chocolate milk, iron fortification with a variety of iron compounds...
resulted in persistent off-colors (Douglas et al., 1981). The type of iron compound used for fortification affects the development of off-flavors: ferric compounds have been reported to produce few off-flavors, whereas ferrous compounds produce initial off-flavors that reduce across storage time (Douglas et al., 1981). van Aardt et al. (2005b) fortified reduced-fat milk with 0.025% α-tocopherol and 0.025% ascorbic acid. A triangle test did not show any significant sensory differences between the fortified and unfortified milks; however, after 10 h of exposure to fluorescent light, significant sensory differences were found between unfortified and unfortified milk. There were no significant differences between light-exposed unfortified milk and milk fortified with only α-tocopherol, suggesting that the combination of antioxidants rather than the α-tocopherol alone prevented the development of light-oxidized flavor in the milk.

Docosahexaenoic acid and eicosapentaenoic acid are long-chain PUFA found in fish and marine algal oils that have been reported to decrease the risk of cardiovascular disease and autoimmune inflammatory diseases (Nelson and Martini, 2009). Conjugated linoleic acid (CLA) is another fatty acid associated with reduced risk of cancer, obesity, and inflammation (Nelson and Martini, 2009). Fluid milk can be fortified with these fatty acids either by altering the diet of the cows to change the fatty acid composition of the milk or by adding oils rich in fatty acids directly to raw milk. Nelson and Martini (2009) reported that when eicosapentaenoic acid, docosahexaenoic acid, and CLA content in pasteurized whole milk was increased by replacing part of the cows’ diets with inert calcium salts of fish oil, a trained descriptive analysis panel could not detect any significant sensory differences between the treatments and control after 3 and 10 d of storage. Direct addition of fatty acid-rich oils to fluid milk has not been as successful. Campbell et al. (2003) reported that CLA-fortified milks exhibited a grassy or vegetable oil flavor. Omega-3 PUFA fortification in other dairy products has been reported to cause fishy and other undesirable off-flavors that can be masked only by strong added flavors, suggesting that it is a poor choice for fluid milk (Kolanowski and Weißbrodt, 2007).

Feed

Cattle feed is of obvious importance in the flavor profile of fluid milk and has been widely studied. Feed-related flavors generally appear in fluid milk within 2 to 4 h directly after cow ingestion of feed and traditionally have been explained using defect terms from dairy scorecard grading (Hedrick, 1955). Gamble and Kelly (1922) presented reports on the effect that corn silage had on sensory characteristics of fresh milk. They reported that silage feeds such as alfalfa and legumes had the ability to impart rank or undesirable qualities to milk when fed to cows before milking. The eventual use of GC in the 1960s became essential for quantitation of volatile compounds causing off-flavors that were specific to milks from certain feeds (Woods and Aurand, 1963). In addition, expanded descriptive analysis techniques provided more concise and quantifiable sensory profiles for studies focused on milk flavor and the role of cow feed (Croissant et al., 2007; Lawless and Heymann, 2010).

Chapman et al. (2001) used descriptive analysis terms such as malty/grainy and metallic to describe flavors commonly associated with cattle feed in UP milks. More modern studies examining the effect of feed on the flavor of milk have explored the relationship between specific volatile compounds in certain milks and the sensory ramifications. Croissant et al. (2007) examined the sensory properties and volatile compounds of milk from cows fed a pasture-based diet with those fed more conventional TMR diets and reported that milk from pasture-raised cows had higher intensities of grassy and cowy/barny flavors compared with milk from conventionally fed cows. With heightened demand in the US market for organic, locally farmed, and pasture-raised milk products, continued investigation into the influence of feed on sensory perception of milk will likely be an enhanced area of research (DuPuis, 2000).

Aside from the sensory effects that differences in feed have on fluid milk, the effect of cattle feed on milk autoxidation has persisted as an area of great concern. Autoxidation of fluid milk, or autoxidized off-flavor, is a sensory defect that is described as metallic or tallowy and is often associated with an increase in bitter taste (Alvarez, 2009). The flavor defect is purportedly caused by spontaneous oxidation of milk fat. The flavor defect and its cause are distinct from light-oxidized off-flavor (described later). Some debate exists over whether spontaneous oxidized flavors in fluid milk are attributable to enzymes, metal catalysts, or a combination of both (Day, 1960). It is widely speculated that metallic catalysts such as copper likely play a part (Gutierrez, 2014). Studies aimed at understanding the catalytic effect that common metals have on milk lipids first came to light in the early 20th century (Golding and Feilman, 1905; Hunziker and Hosman, 1917); however, these studies primarily addressed the contamination of trace metals resulting from processing surfaces instead of naturally occurring concentrations in bovine milk.

In experiments to quantify concentrations of naturally occurring metal in fluid milk, Supplee and Bellis (1922) reported that fluid milk naturally contained up to 0.4 mg of copper/100 g, although amounts were
Light Oxidation

Flavor changes in milk due to light exposure have been of interest to researchers since the earliest studies on fluid milk were published. Home refrigerators began entering the market in the early 1900s but did not become common until the mid-1940s, when new design elements and mass production made them a common household appliance (Higgins, 2001). During the winter months, consumers would often leave fluid milk and other dairy products outside to keep them cold. Early studies on light-oxidized flavor in milk were prompted by complaints from consumers whose milk developed off-flavors after storage outside in the sunlight. These early studies sought mainly to characterize off-flavors and to prove that they were directly related to light exposure and not the manufacturing practices (Frazier, 1928). The use of amber glass bottles to block sunlight exposure was also investigated (Hammer and Cordes, 1920). These early studies failed to describe their sensory methodology, leaving us to assume that the conclusions on flavors detected were derived from the sensory experiences of the authors alone. This is expected because modern objective sensory methods did not exist at this time.

Early attempts at vitamin D fortification in the late 1920s and early 1930s sought to increase vitamin D content through irradiation of milk, sparking renewed interest in investigating light-induced off-flavors in dairy products (Stull, 1953). Several studies published from 1930 to 1950 noted that there appeared to be a dual nature of light-induced flavor defects. The first type of flavor was a fatty, oxidized, tallowy note, whereas the second was a burnt, cabbage, mushroom note (Stull, 1953). As with earlier studies, studies published during this time were limited by their lack of sensory methodology. Trade journals made claims about temperature, bottle type, and season of milking affecting light-oxidized flavor, but results were often conflicting (Tracy and Ruehe, 1931; Doan and Myers, 1936). Because sensory methodology was not documented and often no uniform system was used, there was no way to objectively compare the results of the studies.

Once the off-flavors were linked to sunlight, subsequent studies sought to identify the components responsible for the light-activated flavor and investigate the effect of light exposure on other properties of milk, such as vitamin content (Greenbank, 1948; Stull, 1953). Weinstein et al. (1951) investigated the fact that homogenized milk was more susceptible to light oxidation than nonhomogenized milk, reporting that a whey protein fraction was responsible for the light-activated flavor rather than a lipid fraction. Riboflavin (vitamin B₂) was characterized in the 1930s and was determined to be an essential vitamin in 1939 (Northrop-Clewes and Thurnham, 2012). Studies throughout the 1970s investigated this water-soluble vitamin and eventually determined that photooxidation of riboflavin was the root cause of light-induced off-flavors in fluid milk (Korycka-Dahl and Richardson, 1978; Allen and Parks, 1979; Bradley, 1980). While searching for additional potential photosensitive components of milk, researchers realized that added vitamin A and D could both function as photosensitizers. However, the effect of vitamin fortification on light-oxidized milk flavor was not studied in depth at this time (Wishner, 1964). From here, researchers investigated whether artificial light sources produced the same detrimental effects, reporting that they did (Smith and MacLeod, 1955, 1957).

Later published studies began to incorporate modern sensory methodology. Studies often compared trained panel data against data from consumer preference or difference testing, reporting that both could detect differences between light-exposed and non-light-exposed milk, although the trained panels could detect it much sooner—often after less than 15 min of exposure.
(Coleman et al., 1976; Bray et al., 1977; White and Bulthaus, 1982). De Man (1978) measured the intensity of fluorescent light in retail dairy display cases before exposing milk in different packages to realistic light intensities of various Kelvin color temperatures of light. The Kelvin color temperature scale expresses the color of light in terms of Kelvin and indicates which wavelengths are emitted from the light source. A trained panel was used to quantify light-induced flavors. De Man (1978) found that warm white light and opaque packaging resulted in the least vitamin degradation and off-flavor development, but many grocery stores today use cool white light, and most milk is sold in transparent plastic cartons. Olsen and Ashoor (1987) used 25 untrained panelists to evaluate light-exposed milk on a 9-point hedonic scale. Although a consumer testing method is appropriate for evaluating consumer liking of light-exposed fluid milk, their sample size was far smaller than is recommended today. They reported that container type, container size, and fat content had no significant effect on the flavor and riboflavin content of milk, indicating that perhaps a more sensitive sensory method, such as descriptive analysis, would have been more appropriate. Other studies continued using the ADSA dairy judging system at the expense of sensitivity. Reif et al. (1983) judged 304 samples of fluid milk collected from retail stores in California on a 5-point scale for flavor. The frequency of criticism of light-induced off-flavor for milk in cardboard containers was compared with that for milk in plastic containers. The study deducted more defect points for the plastic containers for light-induced flavor; however, the study was limited in its ability to pull apart small variations in flavor.

Advances in instrumental techniques allowed for the association of specific chemical compounds with light-oxidized flavor, including methionine, acetaldehyde, n-pentanal, and n-hexanal (Bradley, 1980; Kim and Morr, 1996; van Aardt et al., 2005a). Within the past 15 yr, studies have combined sensory and this baseline instrumental volatile analysis to increase understanding of the effect of light exposure on fluid milk. Vitamin and lipid degradation and off-flavors increased with light intensity (Whited et al., 2002). Acetaldehyde, methyl sulfide, dimethyl disulfide, propanal, 2-methyl-propanal, 2-butanol, 2-pentanone, 2-hexanone, 2-heptanone, and 2-nonanone were associated with light-induced flavor (Webster et al., 2009). Triangle tests with untrained consumers (30–35 consumers) demonstrated that the consumers could detect light-induced off-flavors in fluid milk after 2 h of light exposure, although packaging materials, fat, flavorings, and antioxidants can delay the onset of detection (Chapman et al., 2002; Mestdagh et al., 2005; van Aardt et al., 2005a; b; Webster et al., 2009). Exposure to light from light-emitting diodes (purportedly less detrimental than fluorescent or sunlight exposure) also results in light-oxidized flavor and vitamin degradation in fluid milk (Brothersen et al., 2016; Martin et al., 2016). The type of plastic packaging used as well as the shade of packaging can affect light absorption and therefore light-oxidized flavor (Potts et al., 2017).

These studies on light-oxidation flavor integrated sensory methods to varying degrees. Mestdagh et al. (2005) described the use of a trained sensory panel screened for sensitivity to light-activated flavor to detect differences between milk stored in different varieties of polyethylene terephthalate bottles. Although the author described the use of a triangle test to determine sensory differences, the number of panelists and the training methods were not documented. Chapman et al. (2002) determined the duration of exposure needed to produce light-oxidized flavor detected by a trained panel and consumers. Training procedures for a 10-person sensory panel were documented, and samples were analyzed by untrained consumers using a paired difference method. They found that consumers were able to detect light-oxidized flavor after an exposure time of 54 min to 2 h, whereas a trained panel was able to detect light-oxidized flavor after as little as 15 to 30 min of exposure. Moyssiadi et al. (2004) used a trained panel of 17 people to scale off-flavors in light-exposed milk stored in several types of high-density polyethylene (HDPE) and polyethylene terephthalate bottles. The panel was trained to differentiate between only 2 attributes—burnt (light oxidized) and stale (lack of freshness)—but clear differences between samples were found. Brothersen et al. (2016) used a descriptive analysis panel to obtain a reliable picture of all the sensory characteristics of light-exposed milk. Further studies on light oxidation may probe deeper into protection offered by packaging, the effects of type and color temperature of light, and the interactions between vitamin fortification and light exposure.

**Microbial Considerations**

Microbes can have profound effects on milk flavor whether the contamination occurs before or after pasteurization. Raw milk quality can affect the rate of development of off-flavors (Ma et al., 2000; Barbano et al., 2006). Lipases and proteases from somatic cells and bacteria can cause free fatty acid flavor and bitter taste (Santos et al., 2003). In evaluations by trained descriptive analysis panelists, high-SCC raw milk taken from cows with mastitis (an inflammation of the udder due to bacterial infection) developed unpleasant sensory defects such as rancidity and bitterness up to 7 d sooner.
than milk taken from the same cows postinfection (Ma et al., 2000). Microorganisms allowed to grow in raw milk can affect the sensory quality of the milk before pasteurization. Malty flavors are produced by growth of Streptococcus lactis var. maltigenes, and musty potato aromas can be attributed to pyrazine compounds produced by Pseudomonas pelargonii (Morgan, 1976).

Bacterial growth is a common cause of premature milk spoilage. Off-flavors due to bacterial growth have been well documented. Early studies on bacterial spoilage lacked both the sensory and analytical methods needed to draw conclusions about the off-flavors caused by specific species of bacteria. Sadler et al. (1929) attempted to isolate the bacteria responsible for a feed flavor in milk. Although the researchers were able to isolate a culture of bacteria that they reported produced the same aroma as the feed flavor they were investigating, they were unable to identify the bacteria, and no sensory methodology was reported. Because modern sensory methodology did not exist at this time, one can assume that the similarity of the odor produced by the bacteria to the feed odor was judged only by the researchers during benchtop testing. In addition, the term feedy was used at the time to describe a variety of defects, so the results are not clearly linked to one specific off-flavor.

The genus Pseudomonas consists of psychrotrophic, gram-negative rod bacteria that are responsible for the majority of postpasteurization contamination of fluid milk (Walker, 1988; Ternström et al., 1993; Ralyea et al., 1998). Milk contaminated with Pseudomonas is characterized by a fruity (pineapple- or strawberry-like) off-flavor as well as lower levels of sour, rancid, and soapy flavors (Whitfield et al., 2000). Hayes et al. (2002) studied the aroma profile of pasteurized skim and whole milks after inoculation with 1 of 3 Pseudomonas species and storage at 4°C for 3 wk. They reported that specific spoilage aromas such as barny, fruity, shrimpy, cheesy, and rotten were influenced by species, milk fat, and time. Hayes et al. (2002) used a descriptive panel that allowed documentation of subtle differences in flavors produced by bacterial species that might have been missed if a quality judging system was used. Spores, typically from gram-positive rods such as Paenibacillus and Bacillus species, are also of concern for pasteurized milk (Fromm and Boor, 2004). Spores can survive HTST pasteurization and cause sensory spoilage within 25 to 30 d even when total bacteria counts are below 20,000 cfu/mL (Barbano et al., 2006).

**Processing**

Before fluid milk ever reaches the consumer, it is subjected to multiple processing steps to ensure consumer safety and quality standardization. Contact with any surface in processing can noticeably affect the flavor of fluid milk, but areas of chief importance in fluid milk processing, in sequential order, are milk fat separation, fat content standardization, pasteurization, homogenization, and packaging (Goff and Griffiths, 2006).

Pasteurization, named for French microbiologist Louis Pasteur, refers to the heating of food or beverage products to kill microbes that would otherwise jeopardize safety or shelf life. Following government mandates for pasteurization of the US milk supply in the early 1900s, the process of pasteurization has played a large role in dictating the flavor profile of fluid milk. A variety of pasteurization methods are acceptable for achieving legally pasteurized milk. Those methods generally are separated into the following 3 groups: vat pasteurization, HTST pasteurization, and ultrapasteurization or UHT pasteurization.

Vat pasteurization is defined under the Pasteurized Milk Ordinance (FDA, 2015) as the heat treatment of milk at a minimum of 63°C (145°F) for a minimum hold time of 30 min. Vat pasteurization is no longer a major method of commercial pasteurization; however, it is often used by smaller family farms. The minimal heat load of vat pasteurization has been described as having a notably lower cooked flavor compared with other pasteurization methods (Claeys et al., 2013).

High temperature, short time pasteurization is a continuous flow method of milk pasteurization and is defined under the Pasteurized Milk Ordinance as the heat treatment of milk at a minimum of 72°C (161°F) for a minimum hold time of 15 s, although treatments up to 100°C meet HTST standards. High temperature, short time pasteurization is widely regarded as the most common pasteurization method of milk in the United States. MacCurdy and Trout (1940) reported that compared with raw milk samples, HTST-pasteurized milks had higher cooked, oxidized, and heated flavors but that other off-flavors such as feed flavors were eliminated or masked. Other early studies on pasteurization sensory effects showed that pasteurization eliminated the existence of undesirable barny flavors in the milk (Tracy and Ruehe, 1931). Today, the flavor of HTST milk is widely accepted, and deviations are generally disliked by US consumers. In experiments comparing different pasteurization temperatures, Gandy et al. (2008) reported that increased cooked flavors in fluid milk due to increased HTST temperatures were associated with a marked decrease in overall consumer liking.

Ultrapasteurization is defined under the Pasteurized Milk Ordinance as the heat treatment of milk at a minimum of 138°C (280°F) for at least 2 s, resulting in a product with extended shelf life. Similarly, UHT pasteurization meets the same thermal standards of
ultrapasteurization and has the added features of aseptic transfer and packaging systems postpasteurization, resulting in a product that is commercially sterile and shelf stable (Burton, 1977; Boor and Murphy, 2002). The flavor of UP and shelf-stable milks is typically differentiated from that of traditional HTST milk by higher cooked, sulfur or eggy, and caramelized flavors; lingering aftertaste; stale flavor; and higher viscosity (Andersson and Öste, 1995; Chapman et al., 2001; Valero et al., 2001; Clare et al., 2005; Lee et al., 2017). Blake et al. (1995) reported that the increased cooked and caramelized character of direct steam and indirect plate-exchanged UHT whole milk treatments was generally undesirable to American consumers. Furthermore, subsequent studies by Valero et al. (2001) reported that the extent to which these off-flavors were perceived was higher in skim milk compared with reduced-fat milks. In consumer acceptance testing among American children, Chapman and Boor (2001) reported that traditional HTST milk was significantly preferred over UP and UHT milk. Suggested explanations for the observed preference of HTST milk over UP milk included greater familiarity with HTST milk flavor and the creation of off-flavors (Chapman and Boor, 2001; Valero et al., 2001). In a recent series of consumer tests (children and adults) comparing HTST skim and 2% milks with direct steam and indirect UP milks, Lee et al. (2017) reported that UP milk treatments received significant penalties to overall liking for having too much flavor and for being too thick. Adult and child US consumers still preferred HTST milk over UP milk. In addition, UP milks had significantly higher sulfur or eggy flavors, suggesting that the intensity of cooked flavors in UP milks is likely an influential detractor from consumer acceptance.

Besides flavor, a significant point of difference between HTST and UP processed milks is the viscosity or thickness. Chapman et al. (2001) used a trained panel to determine that the viscosity of nonfat UP milks was approximately the same as that of 1% HTST milks. Lee et al. (2017) confirmed this observation, suggesting that the differences in viscosity could be attributable to denaturation of proteins induced after UP pasteurization. Other studies suggest that these results bode well for UP milks in regard to consumer approval because milk consumers prefer the viscosity or thickness of 2% fat milks to those with lower or no fat content even if the consumers typically drink skim milk (Pangborn et al., 1985; McCarthy et al., 2017a).

Homogenization of fluid milk can also affect fluid milk sensory qualities, especially over shelf life. Homogenization is the application of pressure and shear to diminish fat globule size so that milk fat is homogeneously and stably dispersed throughout the final milk product. Since the inception of commercial homogenization in the early 1900s, it has been well documented that homogenization increases the oxidative stability of fluid milk (Thurston et al., 1936; Zahar et al., 1986). Furthermore, homogenization enhances sensory attributes in fluid milk such as color, creaminess, and mouthfeel (Richardson et al., 1993; Feng et al., 2011). These positive flavor and texture attributes are generally preserved in homogenized milk throughout shelf life, assuming that the milk is appropriately refrigerated and shielded from light (Alvarez, 2009).

Although the use of packaging materials for the final milk product is intended to slow the onset of off-flavors (e.g., light oxidation) in fluid milk, there are some flavor considerations related to packaging. Due to the direct contact between packaging surface and the fluid milk, it is reasonably likely that flavor components from packaging materials may be transferred to the milk. The oxygen permeability of selected packaging material can present as a limiting factor of milk quality over shelf life (Gilbert, 1985). The 2 most common packaging types for fluid milk are HDPE jugs and polyethylene-coated paperback cartons. In experiments on the sensory effects of polyethylene-coated paperback cartons, Leong et al. (1992) reported that noticeable packaging flavors were imparted to test milks within 1 d of filling when analyzed by a trained panel. The HDPE containers present minimal risk of flavor leaching; however, several studies have reported that HDPE packages are related to higher risk of light oxidation than other light-shielded package types (Cladman et al., 1998; van Aardt et al., 2001; Amin et al., 2016). Evaluation of packaging modifications by Johnson et al. (2015) has suggested that HDPE packaging with greater than 1.3% titanium dioxide is effective in diminishing the light oxidation risk of HDPE packaging.

SUMMARY

Sensory analysis tools are the definitive means for ensuring sensory quality, assessing acceptability, and identifying faults in fluid milk. Furthermore, the application of powerful sensory analysis techniques has significantly benefited research, marketing, and the general understanding of fluid milk properties throughout history. Application of sensory methodologies is extensive and can range from simple difference tests with untrained panelists to trained panel profiling of specific off-flavors, complex survey techniques, and much more. The majority of sensory tests are easy to conduct and provide powerful results when correctly applied. Studies on the sensory properties of milk, their sources, and consumer perception have evolved with these technologies to provide a broad understanding of...
fluid milk flavor. Future work with fluid milk should address the application of processing techniques such as membrane fractionation or packaging in conjunction with trained panels and consumer research to design fluid milk and fluid milk beverages with desirable sensory properties.

ACKNOWLEDGMENTS

Funding was provided in part by the National Dairy Council (Rosemont, IL). The use of trade names does not imply endorsement or lack of endorsement of those not mentioned.

REFERENCES


## APPENDIX

Table A1. Major milestones in sensory analysis of milk

<table>
<thead>
<tr>
<th>Date</th>
<th>Milestone</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1810</td>
<td>Cooperative dairying begins among butter producers in Goshen, Connecticut.</td>
<td>Weimar and Blayney, 1994</td>
</tr>
<tr>
<td>1842</td>
<td>New York receives its first railway shipment of milk, marking the first major transportation of fluid milk in the industrial United States.</td>
<td>Schweikart and Doti, 2009</td>
</tr>
<tr>
<td>1856</td>
<td>Following his experiments with pasteurization, Louis Pasteur announces his discovery that heating postpones milk souring.</td>
<td>Ligon, 2002</td>
</tr>
<tr>
<td>1865</td>
<td>First vacuum-type milking machine patented. Milking machines would not become widely accepted until the Mehring milking machine became popular in the 1890s.</td>
<td>Weimar and Blayney, 1994</td>
</tr>
<tr>
<td>1884</td>
<td>First glass milk bottles are patented.</td>
<td>Weimar and Blayney, 1994</td>
</tr>
<tr>
<td>1895</td>
<td>Commercial pasteurization begins with the introduction of commercial pasteurization machines</td>
<td>Weimar and Blayney, 1994</td>
</tr>
<tr>
<td>1900</td>
<td>The first homogenizer is presented by Auguste Gaulin at the Paris World's Fair.</td>
<td>Hayes and Kelley, 2003</td>
</tr>
<tr>
<td>1906</td>
<td>Founding of the National Association of Dairy Instructors and Investigators; the association's name was formally changed to American Dairy Science Association in 1916.</td>
<td>Rosati et al., 2007</td>
</tr>
<tr>
<td>1906</td>
<td>First single-serve disposable paper milk cartons are patented.</td>
<td>Weimar and Blayney, 1994</td>
</tr>
<tr>
<td>1908</td>
<td>Chicago becomes the first major US city to pass laws requiring pasteurization.</td>
<td></td>
</tr>
<tr>
<td>1916</td>
<td>First collegiate dairy judging contest is conducted on butter.</td>
<td>Clark and Costello, 2008</td>
</tr>
<tr>
<td>1917</td>
<td><em>Journal of Dairy Science</em> is first published. Milk is added to the collegiate dairy products judging contest.</td>
<td>Drake et al., 2008; American Dairy Science Association, 2017</td>
</tr>
<tr>
<td>1932</td>
<td>Paper milk cartons are introduced for commercial milks.</td>
<td></td>
</tr>
<tr>
<td>1940s</td>
<td>The triangle difference test is developed and introduced.</td>
<td>Drake et al., 2008</td>
</tr>
<tr>
<td>1944</td>
<td>The US Army Quartermaster establishes the Food Acceptance Research Branch in order to assess acceptability of various food products and rations.</td>
<td>Drake et al., 2008</td>
</tr>
<tr>
<td>1948</td>
<td>The modern plastic-coated paperboard carton is developed.</td>
<td>Weimar and Blayney, 1994</td>
</tr>
<tr>
<td>1949</td>
<td>The US Army Quartermaster Laboratory develops the hedonic scale.</td>
<td>Drake et al., 2008</td>
</tr>
<tr>
<td>1957</td>
<td>Tilgner publishes the first book on sensory analysis basics.</td>
<td></td>
</tr>
</tbody>
</table>

*Continued*
Table A1 (Continued). Major milestones in sensory analysis of milk

<table>
<thead>
<tr>
<th>Date</th>
<th>Milestone</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1957</td>
<td>Arthur D. Little Co. introduces the flavor profile method, the foundation for modern descriptive analysis.</td>
<td>Drake et al., 2008</td>
</tr>
<tr>
<td>1963</td>
<td>Rothe develops the odor activity value concept.</td>
<td>Mayol and Acree, 2001</td>
</tr>
<tr>
<td>1964</td>
<td>Fuller and colleagues publish the first design for GC-olfactometry equipment.</td>
<td>Mayol and Acree, 2001</td>
</tr>
<tr>
<td>1964</td>
<td>Plastic milk cartons gain popularity commercially.</td>
<td>Weimar and Blayney, 1994</td>
</tr>
<tr>
<td>1971</td>
<td>Green and Rao publish the first marketing-related conjoint study, the foundation of current consumer conjoint methods.</td>
<td>Green et al., 2001</td>
</tr>
<tr>
<td>1973</td>
<td>Institute of Food Technology forms the Sensory Evaluation Division.</td>
<td>Drake et al., 2008</td>
</tr>
<tr>
<td>1974</td>
<td>Tragon Corp. creates quantitative descriptive analysis.</td>
<td>Stone and Sidel, 2004</td>
</tr>
<tr>
<td>1979</td>
<td>Gail Civille presents the spectrum descriptive analysis method at Institute of Food Technology's sensory evaluation courses.</td>
<td>Society of Sensory Professionals, 2017</td>
</tr>
<tr>
<td>1993</td>
<td>The US Food and Drug Administration approves the use of Monsanto's recombinant bST for commercial milks.</td>
<td>US Food and Drug Administration, 2017</td>
</tr>
<tr>
<td>2002</td>
<td>Descriptive sensory analysis studies find no significant differences between the flavor profiles of organic and conventional milks.</td>
<td>Fillion and Arazi, 2002</td>
</tr>
</tbody>
</table>