VOLUME 20 • NUMBER 1 • FEBRUARY 2005

Journal of Sensory Studies

Maximo Gacula, Jr., Editor
ARIZONA STATE UNIVERSITY

Blackwell Publishing

ISSUES AND VIEWPOINTS

CONDUCTING SENSORY RESEARCH WITH CHILDREN

RICHARD POPPER[1] and JEFFREY J. KROLL

Peryam & Kroll Research Corporation
6323 N. Avondale Ave.
Chicago, IL 60631

Accepted for Publication September 29, 2004

ABSTRACT

Sensory testing with children can provide important insight into their likes and dislikes. Sensory professionals need to use methods that are appropriate for different age groups. This article reviews the current state of knowledge in this area and stresses the importance of considering the sensory, cognitive and social factors that may impact how best to conduct testing with children.

CHILDREN'S BEHAVIOR, LIKES AND DISLIKES

There is no argument that kids have become one of the largest markets in many parts of the world. While it may be difficult to put a finger on any precise amount, the purchasing influence of kids has been estimated at $300 billion in the United States alone. Food and beverages represent as much as 60% of that impressive youth market. Small wonder that food and beverage manufacturers are continuously scrambling to discover what tickles the palates of kids and teens.

Today's kids have more choices and are more in control of their diet than ever before. Parents, in many cases, are more than ready to give in to what their kids want, especially as their kids grow older. For the child, the power to choose can be confusing and conflicting. A choice might be made to express real personal preference, to exercise control of themselves and of their environment or to be viewed as older and more mature. Urbick (2002) suggests that another key driver of a child's behavior can be the craving to create excitement and to stretch boundaries. He says that the stated desire of youths is to push the limits and to avoid "the same old thing."

[1] Corresponding author. TEL: 972 769 0001; FAX: 972 769 1172; EMAIL: Richard.Popper@pk-research.com

While the children's need for self-expression and the marketers' wish to capture their attention are important variables in kids' food choices, early childhood reveals a more basic interplay of nature and nurture in the development of food likes and dislikes. Some aspects of food preference are innate. We are born with a liking for sweet and an aversion to bitter. The fact that these preferences are hardwired in our brains makes evolutionary sense – in scavenging for food, our ancestral forebears would have been guided by their food preferences to seek foods high in caloric energy (such as sweet fruits) and avoid bitter, potentially poisonous plants. Sour tastes are also rejected by newborns – they grimace when tasting sour substances. A genetic disposition toward liking salt has not been so clearly established – newborns are indifferent to salt, but infants, at 4 months, show a liking for moderate levels of salt, possibly the result of a natural maturation process.

While basic tastes such as sweetness and bitterness may be intrinsically pleasant and unpleasant, the preferences for specific foods are largely learned. The diversity of world cuisines attests to the role of culture and environment in shaping what we like to eat (Rozin 1984). Exposure, in and by itself, plays a key role in the acquisition of food preferences. We like to eat foods which we are familiar with. Conversely, unfamiliar foods are often rejected, a culturally universal phenomenon referred to as food neophobia (Pliner 1982).

Zajonc (1968) was the first to identify the "mere exposure effect" across a variety of domains, demonstrating that the repeated exposure to a stimulus (such as a sound or a shape) can enhance liking. In a study on the food preferences of young children, Birch (1979) found that familiarity was a key factor in explaining what foods children liked and Birch and Marlin (1982) experimentally demonstrated the "exposure effect" in the development of food preferences. In the latter study, the authors observed the change in food preferences among a group of 2-year-olds as they were exposed to unfamiliar foods over a period of several weeks. How often the child experienced a food determined how much the child eventually liked it. Children who were given frequent opportunity to taste a food over the course of the study grew to like it, unlike those kids who were offered the same food less frequently.

The importance of familiarity has several implications for companies trying to introduce new food products to the market. The fact that repeated exposure is likely to build acceptance (rather than breed contempt) means that ensuring repeated exposure can be a key marketing strategy in introducing a new food product. Urbick attributes Procter & Gamble's successful introduction of Sunny Delight into the U.K. market to its decision to offer quantities of free product, giving families the opportunity to try the product lots of times.

The role of exposure in developing food preferences also has implications on how novel foods are tested. Most sensory research protocols expose a child only once to a novel food, usually providing only a small sample, not

even a representative portion of the product. Urbick has pointed out that new foods may require repeated testing to accurately assess the product's true potential. Meanwhile, marketers should be conscious of the need to balance novelty and innovation (something that is likely to give them a competitive edge) with children's propensity to prefer the familiar. Urbick suggests that by combining familiar and unfamiliar elements in the same product, marketers may be able to achieve both goals. The flavor may be familiar, but the product's color may be unexpected (e.g., green ketchup) and the packaging may be innovative.

The role of peers has often been noted as a key influence on what children like. A study by Birch (1980) provides experimental evidence for the role of peers in children's food preferences. According to Birch's research, 3- to 5-year-old children will change their preferences depending on what they see other children eat, e.g., choosing vegetables that they initially did not like after seeing other children eat them. This behavioral change was not just the result of momentary peer pressure. The shift in food choice was also reflected in liking ratings collected weeks after the experiment and in the absence of any peers, suggesting that the change was relatively long lasting and reflected a true change in preference.

Again, there are implications of peer influence for marketers as well as for sensory researchers. The fact that children influence other children, even at a very young age, suggests that finding ways to leverage peer influence can be an important element in growing the market for a new product. For example, as part of a grassroots marketing campaign, a company might enlist cool kids to champion their products.

The researcher must make a decision on how to handle the potential for the interactions among children. The purist approach, often advocated by researchers, is to attempt to minimize such interactions to assure obtaining an unbiased opinion. While the merits of this approach are self-evident to many, Urbick, who is a proponent of testing in schools, believes that capturing peer influence in a study can be advantageous and indicative of the real world. Probably all would agree that peer influences in a research setting must be carefully managed. Hemingway (2002) points out that testing in school environments poses many challenges, not the least of which is that the existence of friendships and other social structures may not be immediately apparent to the researcher and may unduly influence the study results.

DESIGNING RESEARCH WITH CHILDREN

It is always important, in designing research with children, to ask what information we desire to obtain from children and what information children

are capable of providing. The Swiss psychologist Jean Piaget is well known for his description of the stages of a child's cognitive and linguistic development. Gollick (2002) describes some of the limitations of children that may affect their ability to answer research questions at any particular age. Young children, for example, have difficulty with concept formation (e.g., sweetness) and classification (e.g., like/dislike). Even when they understand the principles, their attention span may limit their ability to perform the task. For example, $3^1/_2$-year-old children can understand a standardized sorting task, but only about half the children may have the attention span to remember the assignment and to successfully complete the task.

"Seriation," the ability to rank things in order of magnitude, is not fully mastered until age 7, according to Gollick, and this has implications on the reliability of any scaling results from younger children. In addition, children have limited memory skills, which may affect their ability to remember a succession of flavors that are presented for evaluation in a sensory test.

Young children also have limited linguistic skills, which will affect their ability to understand directions, and can have difficulty with the abstract nature of symbols or pictures. For example, children may respond to pictures, such as smiley faces that are often used in children's hedonic scales, based on what they show (a happy person), rather than based on what they are supposed to represent (how the food makes you feel). Gollick also notes the difficulty that children under six have in attending to more than one aspect of a situation at one time. For example, when viewing two rows with an equal number of pennies, young children will easily see that the number of pennies is the same, provided that the pennies are perfectly lined up. However, when one row of pennies is then spread out, children judge the row with the spread-out pennies to have more pennies, showing an inability to separate numerousness from linear extent in making their judgment. In judging foods, young children may attend to one dimension at the expense of another, unlike older children or adults, who may base their reaction on a simultaneous consideration of multiple aspects.

In school-age children, reasoning ability, memory and language skills are more mature and allow for more complex tasks. However, there is tremendous variation in skills among children of the same age. Gollick's experience with cognitive testing has shown that the age at which 10% of children can master a particular task, compared to the age at which 90% of children can do so, varies by as much as 4 years. Thus, assumptions regarding what a particular age group can do are often going to be true only approximately and researchers need to take into account the considerable variation in children's abilities, even at similar ages.

Given the cognitive and linguistic limitations of children at any given age, it is not surprising that research on children has focused attention on what

test methods are most appropriate for different ages (see review by Guinard 2001). The younger the age group, the more challenging it is to devise valid, reliable test methods. Therefore, when products are expected to appeal to a wide age range, it is often convenient to test older children (above age 12), who require far fewer special considerations compared to adults than younger children. However, when the target age for the product is specifically younger children, it may not be appropriate to focus on the older age group.

The taste preferences of newborns and infants have been studied using behavioral measures (e.g., Beauchamp and Moran 1984). Using a procedure adapted from the baby food industry, Bovell-Benjamin *et al.* (1999) obtained data on the food preferences of infants and toddlers by asking mothers to interpret the behavior of their child as the child tasted the food. Mothers rated their child's reaction on a traditional (adult-version) 9-point hedonic scale. Using this methodology, the authors were able to draw conclusions regarding the relative acceptability of different fortifications added to porridge.

The testing of children age 3 and above allows for more direct methods. Because consumer testing with kids is mostly concerned with measuring a child's liking of a product, it is especially important to know what the most appropriate hedonic methods are for testing with kids. Kroll (1990) introduced a liking scale for testing children age 5 and older. It is similar to the traditional 9-point hedonic scale, except that the verbal anchors associated with the scale are more child-friendly – instead of using terms such as "like extremely" and "dislike extremely," for example, it employs the terms "super good" and "super bad." This so-called Peryam and Kroll (P & K) scale, now widely used in the industry, was shown to perform better than the adult scale (and better than a variety of alternative children-oriented scales) in determining liking among children.

While older children may be able to use verbal scales effectively (even if they may require some interviewer assistance), there is still no consensus on how best to test preschoolers. Pictorial scales continue to be popular in testing preliterate children, based on the rationale that preschoolers cannot read and may not fully understand complex words or phrases but can more accurately deal with facial expressions. Besides, pictures are entertaining and are thought to inspire closer attention to the task. Kroll (1990), however, found that a 9-point face scale actually discriminated less well than the 9-point "super good/super bad" scale.

One of the reasons that face scales may be difficult for children may be that the faces themselves may be ambiguous. In her discussion of pictorial scales, Cooper (2002) makes the point that a face, intended to show a degree of "dislike," can be interpreted by a child as saying "I am angry," whereas the child might not feel anger in response to the food but might dislike it. Cooper has found that the eyes and the mouth are particularly important to the

interpretation of the facial expression and are more likely than other elements to lead to misinterpretations of the scale unless carefully chosen.

Despite some of the difficulties with pictorial scales, Cooper has had success in applying them and even extending them to measuring not just acceptability but also sensory intensity. Working in conjunction with a graphic artist, she has devised scales for measuring the level of fruit or chocolate flavor, as well as more conceptual attributes, such as stickiness.

Cooper also reports on her research on devising pictorial scales for testing in different cultures. Working with Asian and Pacific Island groups of children, she has attempted to create culturally relevant acceptance scales. She notes the difficulty of judging cultural relevance from a Western perspective and the importance of developing the scale from within a culture, piloting it to ensure appropriateness. Certain expressions are appropriate in some cultures but not others (e.g., showing the tongue to indicate liking is inappropriate for Thai and Malay children).

TWO PRESCHOOLER METHODS STUDIED

Popper *et al.* (2002) compared two different methods for measuring liking among preschoolers. They chose not to include a face scale because of some of its potential problems but instead chose ranking and rating as two methods for eliciting liking judgments. They also assessed the effect that the interviewer has on preschoolers' responses. Preschool-age children are preliterate and must, by necessity, be interviewed one-on-one. Typically, research personnel (usually female) serve as interviewers. The study compared using the child's mother as interviewer with using a female interviewer who was not familiar to the child. On the one hand, children might feel more comfortable with their moms than with an unfamiliar interviewer, making it possible for them to better focus on the task. Children might also feel more at ease telling their moms that they do not like something, as compared to a researcher who, they might feel, expects them to like what they taste. On the other hand, the child's mom is untrained at interviewing and could introduce her own biases into the test.

The study involved three different formulations of a powdered orange drink, formulated with 0, 30, or 100% of the recommended sugar level. The respondents were 206 children ages 3–5, about an equal number of boys and girls.

There were two interviewer conditions – the children were either interviewed by their mom or by a trained female interviewer (a member of Peryam and Kroll's research staff). When interviewed by the research staff, the mom was not present in the room with her child. When the mom acted as the

interviewer, she did not taste the products herself, only her child did. In each interviewer condition, a child evaluated all three orange drinks, using two different procedures. The order of procedures was alternated across the children and the children were never explicitly told that the three formulations were repeated.

In the ranking procedure, children first tasted all three products and selected the one that they liked best, which was then set aside. Children then retasted the remaining two samples and selected the one that they liked better. The third sample, by default, was the one that they liked least.

The other procedure was a rating task using a bifurcated 5-point scale. In this procedure, the child was first asked if the sample was "good" or "bad" and, depending on the answer, was then asked whether the sample was "really good" (or "really bad") or "just a little good" (or "just a little bad"). If the child had trouble committing to whether the sample was good or bad, the answer was recorded as "neither." Only a few children had trouble expressing whether they liked or disliked a sample – even the 3-year-olds were quite outspoken and opinionated.

A comparison of the test methods showed that the ranking and rating procedures produced very similar results – both methods showed significant differences in liking among all three formulations. Not surprisingly, children liked the sweetest formulation the most, regardless of interviewer condition or test procedure.

Perhaps the most interesting finding concerned the effect of the interviewer. Figure 1 shows the difference in liking ratings when the mom was the interviewer, compared to when the unfamiliar researcher was the interviewer.

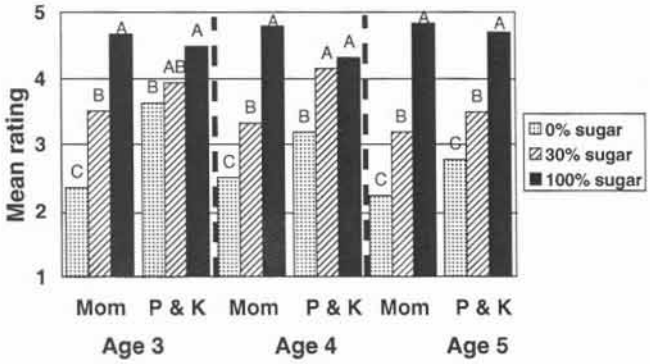

FIG. 1. THE EFFECT OF INTERVIEWER (MOM VERSUS PERYAM & KROLL RESEARCHER) ON CHILDREN'S LIKING RATINGS (5 = "REALLY GOOD") OF POWDERED ORANGE DRINKS DIFFERING IN PERCENT SUGAR

Means sharing a common letter are not significantly different from one another ($P < 0.05$) (Popper *et al.* 2002).

When the mom did the interviewing, the average ratings for the three formulations spanned a larger range than when the researcher did the interviewing and the differences were more likely to be statistically significant. This effect varied by age – the benefit of the mom as the interviewer was evident at ages 3 and 4 but was largely absent by age 5. The ranking data showed similar interviewer effects.

While this study showed that using the mom as the interviewer can increase the sensitivity of the test, especially when working with 3-year-olds, it is important to keep in mind that this study posed little risk that the mom could introduce her own biases about the products tested. All formulations looked the same and the mom was intentionally asked not to taste the samples. In situations when the appearance of the samples might suggest something about the quality or the nutrient content of the foods, or when brand information is provided as part of the test, the mom's role in the interview would need to be reevaluated.

Simple methods of scaling, such as the ranking procedure or the bifurcated liking scale, appear to work well with preschoolers. Children ages 5–7 are probably capable of longer scales (7- or 9-point), as some studies have demonstrated (Kroll 1990; Kimmel *et al.* 1994). At that age, kids will still require assistance and mom is probably best for that purpose. With children over eight, the test can be self-administered, although the research staff will, on occasion, need to assist the children.

DIFFERENCES BETWEEN KIDS AND ADULTS

There is indication (see Guinard 2001) that sensory thresholds are higher in children than adults, suggesting reduced sensory sensitivity. Zandstra and de Graaf (1998) report that in children ages 6–12, the perception of sweetness increases less rapidly with increasing sucrose concentration than in older children and adults, although the same was not found with regard to the perception of sourness in response to changes in citric acid. Of course, any determination of threshold sensitivity and supra-threshold perception are complicated by the fact that measuring sensory perception in children is difficult – differences between adults and children in sensory perception may reflect, in part at least, differences in how children interpret the questions that they are asked and in how they use the intensity scales on which the research is based.

How do children and adults differ in their *preferences*? It is well documented that children prefer sweeter foods than adults, a difference which seems to diminish in mid to late adolescence. Also, as many parents of young children know and as Urbick (2002) reminds us, children like simple, smooth textures. Even the appearance of bits (e.g., specks, flecks) can turn young children off.

While the prevailing influence of culture is bound to lead to many broad similarities between the food preferences of adults and kids (Pliner 1983; Rozin and Millman 1987), it is also the case that the process of optimizing a product for children often results in a formulation that differs from one that is optimal for adults – and from one that adults *think* would be optimal for children (Moskowitz 1994).

OTHER SENSORY RESEARCH METHODS

Swaney-Stueve (2002) undertook the challenge of attempting to conduct a descriptive analysis with children. Using six different brands of creamy peanut butter, she trained four sensory panels differing in age – a panel of fourth graders (ages 9–10), eighth graders (ages 13–14), high school students (ages 16–18) and a college-age panel (ages 18–22) – and compared the results with those that she obtained from using an experienced adult panel (ages 24–56).

Surprisingly, even the fourth graders performed exceptionally well throughout the process. The fourth-grade, eighth-grade and high school panels generated many terms similar to those generated by the college-age and experienced adult panels, and all panels were quite consistent in how they rated the differences in appearance, flavor and texture among the products (although the fourth-grade, eighth-grade and high school panels tended to agree more with each other than with the adult panels). In some respects, among the three children's panels, the fourth graders were the most consistent and reliable.

Since the products were perceived similarly across different ages, the results support the current practice of using adults for the descriptive analysis of children's products. It is uncertain how generalizable this similarity of adults and kids is – as noted above, children's sensory perceptions differ in certain respects from those of adults and children may differ from adults in the importance that they place on various product dimensions (e.g., attending to appearance more than flavor). These age differences may not always yield to panel training and the use of reference standards.

But Swaney-Stueve's results are also noteworthy because they show that even 10-year-olds are capable of performing sensory tasks that are cognitively quite demanding. Crucial to the success of such tasks is providing training and keeping the children engaged. Urbick makes a similar point with reference to conducting his creative discovery work with children – keeping kids involved in the process and excited about working on a real project are the key elements of success.

In addition to quantitative testing, qualitative research with children is also a very important tool in product research. Hemingway (2002) identifies

several qualitative techniques that she has found to have worked well with children, such as drawing (where/how the child figures in the picture), collages (children form montages from material provided to them), concept boards (identify which icons are cool or gross for each age) and storytelling (the child tells a story or finishes one off).

IMPORTANCE OF TEST CONDITIONS

Part of making any sensory testing with children successful includes not only using age-appropriate measurement scales, instructions and wording of questionnaires but also making the child feel comfortable in the research environment (e.g., the benefit of using moms with preschoolers). The research staff needs to create a friendly and inviting atmosphere. Things to avoid, according to Hemingway, are an authoritarian style, criticizing comments that kids make or using a hands-up or other classroom-style behavior (especially in a qualitative research context). Hemingway (2002) and others (e.g., Kimmel *et al.* 1994) have argued for the merits of warm-up exercises as a way of introducing children to the research methods employed.

Several researchers have mentioned the importance of the time of day in children's research. According to Gollick (2002), depending on the time of day, a child's IQ score on a standardized test can vary by as much as 30 points. Urbick (2002) advocates conducting consumer tests with children in the morning, when kids are most alert and avoiding the after-school hours, when children are mentally tired and need unstructured playtime and a chance to be physically active.

Another aspect of the time of day that is relevant to sensory testing is whether it is appropriate to test morning foods (e.g., breakfast cereals) at times other than the morning and lunch/dinner foods at times other than around lunch or in the afternoon. Birch *et al.* (1984) showed that children as young as 3 years old have already learned to categorize foods as "for breakfast" or "for dinner." In their study, both adults and children were asked to taste breakfast and dinner foods at two different times of day and showed a significant preference shift for foods with the time of day, with breakfast items being more preferred when tasted in the morning than in the afternoon and dinner items more preferred in the afternoon than in the morning (the shifts, however, were larger for adults than for children).

The issue of how the time of day affects test results is interesting and deserves further research in children. In contrast to the results of Birch *et al.* (1984), Kramer *et al.* (1992), working with adult subjects, did not find an effect of the time of day on liking ratings or food intake when testing breakfast and lunch foods at different times of day. It has also yet to be determined

whether the time of day affects the relative differences in liking among similar formulations of a product (e.g., several versions of the same breakfast cereal), in addition to affecting the level of liking for very different foods (e.g., cereal versus pizza). Most sensory tests are concerned with relative differences among similar products and the appropriateness of the time of day may be less of a factor in that situation.

The time of day is only one context that may be important to consider. Urbick (2002) notes that for kids, products are a "holistic experience" and that kids are less able than adults to respond meaningfully to a product that is lacking its real-world context, such as packaging, concept or brand identity. Urbick recommends testing a product in a form "as close to the real thing" as possible. For example, if the product is yogurt in a tube, Urbick says that the sensory testing should be done using the tube. Sensory researchers are prone to think about the product itself, but, especially for kids, the success of a new product may depend not just on how it tastes but also on how it handles, its play value and the image it projects (e.g., whether it satisfies kids' aspiration to be viewed as older).

MORE RESEARCH NEEDED

Kroll (1990) noted that while sensory testing with children was becoming increasingly important to the food industry, sensory research itself was not keeping pace with the need. "Testing with children is in an embryonic stage," she wrote. "Over the years, a few sensory researchers have considered the problems involved in applying their science to this special population, but for the most part, the field has been static. The need for serious investigation is pointed up by how little research has been done in this area." Some of Kroll's concerns remain true today. Researchers have taken up the challenge during the ensuing dozen years and several have been the focus of this article. But given the size of the market and the potential that reliable kid testing has for the food industry, there is still a need for more research to help maximize the insights that research with children is able to provide.

ACKNOWLEDGMENTS

This article is reprinted with modification from *Food Technology*, 2003, *57*(5), pages 60–65.

REFERENCES

BEAUCHAMP, G.K. and MORAN, M. 1984. Acceptance of sweet and salty tastes in 2-year-old children. Appetite *5*, 291–305.

BIRCH, L.L. 1979. Dimensions of preschool children's food preferences. J. Nutr. Educ. *11*, 91–95.

BIRCH, L.L. 1980. Effect of peer models' food choices and eating behaviors on preschoolers' food preferences. Child Dev. *51*, 489–496.

BIRCH, L.L., BILLMAN, J. and RICHARDS, S.S. 1984. Time of day influences food acceptability. Appetite *5*, 109–116.

BIRCH, L.L. and MARLIN, D.W. 1982. I don't like it; I never tried it: Effects of exposure on two-year-old children's food preferences. Appetite *3*, 353–360.

BOVELL-BENJAMIN, A.C., ALLEN, L.H. and GUINARD, J.X. 1999. Toddlers' acceptance of whole maize meal porridge fortified with ferrous bisglycinate. Food Qual. Prefer. *10*, 123–128.

COOPER, H. 2002. Designing successful diagnostic scales for children. Presented at Ann. Mtg. Institute of Food Technologists, Anaheim, CA, June 15–19.

GOLLICK, M. 2002. Asking kids questions: Possible pitfalls. Presented at Ann. Mtg. Institute of Food Technologists, Anaheim, CA, June 15–19.

GUINARD, J.X. 2001. Sensory and consumer testing with children. Trends Food Sci. Technol. *11*, 273–283.

HEMINGWAY, M. 2002. Effective techniques for consumer research in a challenging market. Presented at Ann. Mtg. Institute of Food Technologists, Anaheim, CA, June 15–19.

KIMMEL, S.A., SIGMAN-GRANT, M. and GUINARD, J.X. 1994. Sensory testing with young children. Food Technol. *48*(3), 92–99.

KRAMER, F.M., ROCK, K. and ENGELL, D. 1992. Effects of time of day and appropriateness on food intake and hedonic ratings at morning and midday. Appetite *18*, 1–13.

KROLL, B.J. 1990. Evaluating rating scales for sensory testing with children. Food Technol. *44*(11), 78–80,82,84,86.

MOSKOWITZ, H.R. 1994. *Food Concepts and Products: Just-in-Time Development*. Food and Nutrition Press, Trumbull, CT.

PLINER, P. 1982. The effects of mere exposure on liking for edible substances. Appetite *3*, 283–290.

PLINER, P. 1983. Family resemblance in food preferences. J. Nutr. Educ. *15*, 137–140.

POPPER, R., SCHRAIDT, M. and KROLL, J. 2002. Testing with pre-school children: The effect of the interviewer. Presented at Ann. Mtg. Institute of Food Technologists, Anaheim, CA, June 15–19.

ROZIN, P. 1984. The acquisition of food habits and preferences. In *A Handbook of Health Enhancement and Disease Prevention* (J.D. Matarazzo, S.M. Weiss, J.A. Herd and N.E. Miller, eds.) pp. 590–607, Wiley, New York.

ROZIN, P. and MILLMAN, L. 1987. Family environment, not heredity, accounts for family resemblance in food preferences and attitudes: A twin study. Appetite *8*, 125–134.

SWANEY-STUEVE, M. 2002. Can children perform descriptive analysis? Presented at Ann. Mtg. Institute of Food Technologists, Anaheim, CA, June 15–19.

URBICK, B. 2002. Kids have great taste: An update to sensory work with children. Presented at Ann. Mtg. Institute of Food Technologists, Anaheim, CA, June 15–19.

ZAJONC, R.B. 1968. Attitudinal effects of mere exposure. J. Pers. Soc. Psychol. *9*(2), 1–27.

ZANDSTRA, E.H. and DE GRAAF, C. 1998. Sensory perception and pleasantness of orange beverages from childhood to old age. Food Qual. Prefer. *9*, 5–12.

COMMENTARY: SENSORY RESEARCH WITH CHILDREN

The purchasing "pester" power and influence of children in today's food and beverage industry cannot be overlooked by major corporations. In such a highly competitive environment, where consumer packaged goods companies compete for a share of the market, the likes, dislikes and opinions of kids can drive trial, repeat, purchase and sustainable volume for innovative, new products. Children have more control in purchase decisions today than ever before, and often have the final say in whether or not a product becomes a household staple.

Testing with children can provide invaluable insights regarding key product attributes, how to market and talk about product benefits, and what is "cool" and what "is not." However, to gain these insights, the sensory scientist, market researcher and marketer must know the right way to gather this information from kids. This includes the stages of cognitive and physiological development of a child, the role of peer influence, how repeated exposure to a product impacts perception, key differences between boys and girls, the best methods to gather data from them at any given age, how their likes and dislikes differ from adults, and how to put kids at ease in a testing environment that will allow for the best "data collection." For example, young children (3–5 years of age) have limitations in their abilities to answer research questions based on cognitive development. Even simple rating and rating exercises may be too complex or abstract for the average three-year-old. Teens (above age 12), on the other hand, are quite capable of performing the same tasks as an adult, but their responses to a given question may be highly influenced by their peers, or may reflect a desire to show individualism and free choice.

Common pitfalls that occur when testing with children are usually the result of not using the right tools and designing the correct study. Often, a product targeted at a younger audience will be validated with older children. This is done due to the ease of gathering data from the older versus younger children. However, sensory researcher, beware!!! These results may steer the company to design a product that is too complex in sensory dimensions (multicomponent, variable texture, etc.), or has a marketing campaign not appropriately targeted for the younger audience. Putting a child in a foreign environment can often result in false or misleading results. Facilitating data collection in a strict "classroom" setting, or using a facilitator unknown to the kids can often be intimating and overwhelming. Letting the kids have fun with the data collection task and putting them at ease are key to gathering their real insights and opinions.

The accompanying paper, "Sensory Research with Children" by Popper and Kroll provides a comprehensive overview of key things to consider when

Journal of Sensory Studies **20** (2005) 88–89.

testing with kids. It will stimulate thought and provoke many questions yet unanswered within the sensory community about the best ways to use children in sensory and market research testing. As a working practitioner in a company that has many kid targeted products, I hope that the future of testing with children continues to evolve as the sensory community focuses on developing methods and techniques to gather the richness of insights available from this vast, important segment of the consumer population.

Dana Craig-Petsinger
Vice President
Product, Consumer and Technology Understanding
The Kellogg Company
TEL: 269 961 2997
EMAIL: dana.craig-petsinger@kellogg.com

COMMENTARY: CONDUCTING SENSORY RESEARCH WITH CHILDREN

Children today are more involved than ever in choosing what they buy, use and eat. With an increasing number of parents, guardians and older siblings working outside the home, many children return home from school each day to an empty house. Often it is the children's responsibility to decide what to do, what to play with, what to eat and how to prepare it. A Gallup Survey of fourth through eighth graders revealed that 83% make their own snacks, 80% make their own breakfasts and 57% are involved in food purchases. This in itself is a prime reason that children's responses to products must be assessed.

The article by Popper and Kroll gives us an excellent feel for the general state of the art in children's testing. Unfortunately, as the article clearly shows in the reference list, that state of the art often is unpublished, untested opinion rather than peer-reviewed, published science. That is not a reflection on Popper and Kroll, nor on the authors cited in the reference list – it is just simple fact.

Howard Moskowitz's review, which I also have read, focused extensively on the business sense of using children for testing. I will not take that path because I agree with his comments and have little to add. Instead, I suggest that a major problem in the field of sensory testing with children (and to some degree with adults) is that the science often gets lost in the milieu of personal experiences, professional disagreements and belief systems – things that are not necessarily based on good science. I am pleased that Popper and Kroll specifically commented that more research is needed, because that certainly is true. It is especially the case for children, because the age at which children learn various concepts, are exposed to products and ideas and interact with measurement tools and technology changes with each succeeding generation. Much of the original research on Piaget's cognitive development theory and the ages associated with various aspects of that theory are now decades old. Although many of the innate psychophysiologic development issues have not changed a great deal, children are maturing earlier, and the cognitive demands and processing skills to meet those demands have changed. It is essential that we continually test the methods that we use to determine at what age they can be used, how they need to be adapted for certain age groups and how societal influences may have affected the techniques. Personal experiences and anecdotal evidence are insufficient to meet our needs.

Popper and Kroll provide an excellent example of adaptation of a sensory technique (ranking) to make it more useful with children. Instead of being asked to "rank all the samples," the child is asked only to pick the best product. That product is removed, and then the child is asked to pick the best sample

Journal of Sensory Studies **20** (2005) 90–92.

of those remaining. We also use a similar method in our work with children, except that we ask them to pick the best and worst, remove those and proceed accordingly. Such adaptations are addressed in the ASTM International E2299-03 *Standard Guide for Sensory Evaluation of Products by Children* (ASTM 2003), but I do not know of anyone who has actually published research showing that the technique is appropriate. We would not dream of believing the data from a new technique for measuring the chemical or physical composition of a product without validation of that method, but we do not hesitate to make decisions that will influence millions of dollars of consumer spending based on a new sensory test, as long as the test: (1) makes logical sense; (2) got the OK from a consultant who said his or her experience shows that it works; (3) costs less to do than what we are doing now; and (4) seems to give us the answer we expect when we use it on our product.

Popper and Kroll mention another example, this time an example that has been researched, of how changing societal influences on language can allow adaptation of sensory techniques in the so-called Peryam and Kroll (P & K) scale for children. I doubt that, a generation ago, a scale using words such as "super good" or "super bad" would have been possible because the connotation of "super" was "best, bigger or greater" (e.g., supermarket or superman), not necessarily "extreme," which is its connotation in the P & K scale. However, in the late 1980s and early 1990s, when the scale was in the development and testing stage, the word "super" was very much in use by children to mean "extreme." Changing language, especially the vernacular of children, and a more global focus in many product categories means that we will always need to revalidate and update sensory methods and techniques that depend heavily on vocabulary.

One area of particular importance that Popper and Kroll address is that of differentiating testing abilities among children. Certainly preschoolers are different in their testing abilities from 12-year-olds. But what about children in between those ages – the kindergarten to 5th grade crowd? Are they more like adults or young children in their ability to conduct sensory tests? Spaeth *et al.* (1992) showed that 8- to 10-year-olds could use traditional 9-point hedonic scaling techniques as well as any of the other scales tested, but those authors pointed out that they used children clearly in the age group associated with Piaget's concrete operational stage. Popper and Kroll discuss some issues of facial scales but do not provide concrete information about the age groups for which they are necessary and appropriate. In most published research studies, researchers have been careful to steer clear of the "cusp years" – those years that typically serve as the end of one developmental stage and the beginning of another. For example, age 6 is the end of Piaget's preoperational stage and 7 is the beginning of the concrete operational stage. Thus, most research is done with 4- to 5-year-olds or 8- to 10-year-olds. That occurs

because there is too much chance that the cognitive development of any given child is slightly different from the "norms." Thus, it is easier to exclude those children who might be too old or too young developmentally (e.g., a bright 6-year-old who is already in the concrete operational stage) for a given task. However, we know that children have different abilities to do various tasks and we know that all the age groups are involved in decision making for a variety of products. Instead of excluding groups of children, would not it make more sense to develop a carefully researched screening tool to determine whether a given child had the ability to do a certain task or use a certain procedure than to arbitrarily set age limits that make no scientific sense and ultimately no business sense either?

Popper and Kroll have done an admirable job of touching on many key points in the discussion of testing with children. It is now up to us to move ahead with logic and sound science into a new era of evaluating children's needs and wants.

Edgar Chambers IV
Sensory Analysis Center, Department of Human Nutrition
Kansas State University
Justin Hall, Manhattan, KS 66506-1407
TEL: 785-532-0156
FAX: 785-532-0176
EMAIL: eciv@ksu.edu

REFERENCES

ASTM INTERNATIONAL 2003. *E2299-03 Standard Guide for Sensory Evaluation of Products by Children, Vol. 15*.08, ASTM International, West Conshohocken PA. Available at http://www.astm.org.

SPAETH, E.E., CHAMBERS, E. IV, and SCHWENKE, J.R. 1992. A comparison of acceptability scales for use with children. In *Product Testing with Special Consumer Populations for Research Guidance* (L.S. Wu, and A.D. Gelinas, eds.) pp. 65–77, ASTM, Philadelphia, PA.

COMMENTARY: WHITHER NOW THE GRAND SENSORY "PROJECT" IN AN AGE OF IMPROVING METHODOLOGY?

A controversy has smoldered in the sensory field for the past 40 years, ever since sensory analysis became respectable in the 1960s through Rose Marie Pangborn's pioneering research and teaching efforts, and through the seminal book on which she collaborated, *Principles of Sensory Evaluation of Food* (Amerine *et al.* 1965). The controversy concerns what the "sensory" field should comprise, and what its limits should be. Many of its researchers who come from other content-rich disciplines, such as psychophysics, feel that sensory analysis ought to concentrate on creating a repository of knowledge, and, along with that repository, ensure that the methods work. In light of this controversy, we can now proceed to one person's analysis of the Popper and Kroll (2004) paper.

Popper and Kroll have written an excellent review on the limits of testing with children. Their review now firmly establishes that children have limited cognitive abilities that must be considered in the choice of any test. Furthermore, as Popper and Kroll point out, the child is a developing organism, so children at different ages pose different problems because of their varying and evolving capabilities. Popper and Kroll's (2004) review is noteworthy because it provides a strong foundation for sensory measurement in a field that too often has been dictated by fads.

Popper and Kroll (2004) also provide an additional benefit for the reader, albeit perhaps a benefit not realized in their original project. What becomes increasingly clear from their comprehensive paper and from many of the research papers in sensory analysis is the picture of sensory analysis as a *testing system*, rather than as a system to build knowledge. The reader almost gets a sense of the field of "Sensory" as having evolved from a promising future three and four decades ago into "Tests Are Us." Those of us working in the field over the past 20 to 40 years know that all too frequently, our field of "Sensory" has been relegated to a testing service, low cost at that, much to the dismay of practitioners and scientists who have wanted to establish the field as more content driven rather than method driven.

What is missing from the sensory world is the same type of review provided by Popper and Kroll (2004), but looking at content. From Popper and Kroll (2004), but really from other papers as well, we get only a partial glimpse about children's perception of products. The public data are limited when it comes to content. We have every reason to believe that Popper and Kroll have a very vast base of knowledge about children's perception because they have conducted extensive work in the field for almost 40 years through their company, Peryam & Kroll, Inc. But this data is not available for publi-

Journal of Sensory Studies **20** (2005) 93–95.

cation, because much of it was funded through contract research projects. Perhaps there is a corpus of such work in the files of companies, but the public is not privy to those files. Consequently, the overwhelming message is method and procedure, test and result, statistics and proper conclusions.

STEPPING BACK AND JUMPING FORWARD, THE NEXT "PROJECT" OF "SENSORY"

Historically, we can locate analogies to the Popper and Kroll (2004) paper in the early days of psychophysics. When psychophysics began in the late 1800s, there was the belief that this area of psychology specializing in the perception and measurement of external stimuli by the senses would provide a rich vein of knowledge about human sensory processes. The first 50 years, or so, however, comprised the laborious, meticulous and ultimately quite fruitful study of method to measure perception. It was in those early years that psychophysicists developed the different methods to measure perception. Those early years of psychophysics further saw controversies about what the human perceiver could do, what the perceiver could not do and numerous attempts to create viable methods for testing in a field that hitherto was unexplored and often deemed unexplorable by philosopher-skeptics. The Popper and Kroll (2004) paper follows that tradition of method in a field that up to now has remained *terra incognita*. With their masterful review of the literature, combining the best of psychophysics with a sprinkling of today's well-known practitioners and theories, the authors have presented us with a repository of information about methods that help us know how to measure sensory responses of children, as well as reiterate the limits of such measurement.

Now where do we go? Does sensory analysis remain "Tests Are Us" with increasingly sophisticated methods to develop information about children's responses? Do practitioners in the field check off yet another set of tests that can or cannot be run? Do practitioners simply have new prescriptions for testing with children? *Is that all there is*? Have Popper and Kroll (2004) given us pause, allowing us to express a collective sigh of relief that now we know how to work with children?

I think not. Rather, the Popper and Kroll (2004) paper calls now for the publication of more substantive work, not just better tests. The covert message in the paper is that yes, there is a science underlying the tests, and that the tests must be adjusted to match the child's cognitive level. The danger, however, is that we stop with this review, pat ourselves on the back (those of us who can reach that far without arthritic twinges) and stop thinking. This is always the problem with a good review; it leads inevitably to the same type

of satisfaction that we get from a rich meal. We have arrived – we know the compass of the testing methods, the strengths, the weaknesses. We are *au courant*, propped up by the knowledge that there is science behind us.

Rather than become "Tests Are Us," the sensory community must proceed forward to create substantive knowledge as did psychophysics a century ago, to create databases about the children's preferences, about their reactions to concepts, about the changes in their sensory-oriented and scale-oriented mindset with maturation. It is vital to move forward with the project to build a science-comprising data, not just prescriptions of how to better test with children in light of their cognitive abilities. If we stop now, basking in this wonderful review rather than taking proper inspiration from it to move forward, we remain "testers" all our lives, albeit better testers, more knowledgeable about the science, more facile in the statistical treatment, more precise in our conclusions. But we miss the knowledge base. We become better, science-based technicians and lose our chance to found a new science.

ACKNOWLEDGMENTS

I would like to thank Jackie Beckley, Edgar Chambers IV and Richard Popper for exceptionally constructive criticism and insight. Their help in formulating the point of view was invaluable.

Howard R. Moskowitz
Moskowitz Jacobs Inc.
1025 Westchester Ave
White Plains, NY 10604
EMAIL: mjihrm@sprynet.com

REFERENCES

AMERINE, M., PANGBORN, R.M. and ROESSLER, E. 1965. *Principles of Sensory Evaluation of Food.* Academic Press, New York.

POPPER, R. and KROLL, J.J. 2003. Conducting sensory research with children. Food Technol. *57*(5), 60–65.