

Malleability of taste perception: biasing effects of rating scale format on taste recognition, product evaluation, and willingness to pay

Antonia Mantonakis¹ · Norbert Schwarz² ·
Amanda Wudarczywski³ · Carolyn Yoon⁴

Published online: 9 January 2017
© Springer Science+Business Media New York 2017

Abstract Product-related cues, such as brand or price, can influence consumers' taste perception. Going beyond this observation, we examine the extent to which a stimulus-extrinsic factor, such as the format of the measurement tool on which consumers describe attributes of a taste sample, influences concurrent taste perception, and in turn, later taste recognition, overall product evaluation, and willingness to pay (WTP). The results of two experiments show that rating scale format (i) influences consumers' concurrent impression of a taste sample, (ii) systematically biases later identification of the sample in a taste recognition test, and (iii) affects overall product evaluation and WTP. However, scale format (iv) does not influence ratings and downstream judgments when consumers are highly knowledgeable in the product domain. These findings

✉ Antonia Mantonakis
amantonakis@brocku.ca

Norbert Schwarz
norbert.schwarz@usc.edu

Amanda Wudarczywski
awudarze@uwaterloo.ca

Carolyn Yoon
yoonc@umich.edu

¹ Department of Marketing, International Business, and Strategy, Goodman School of Business, Brock University, 1812 Sir Isaac Brock Way, St Catharines, ON L2S A31, Canada

² Department of Psychology, University of Southern California, 3620 S. McClintock Ave, Los Angeles, CA 90089-1061, USA

³ Department of Psychology, University of Waterloo, 200 University Ave West, Waterloo, ON N2L 3G1, Canada

⁴ Stephen M. Ross School of Business, University of Michigan, 701 Tappan St., Ann Arbor, MI 48109-1234, USA

demonstrate that the experience of taste is fleeting and not well represented in memory, and that like other subjective experiences, taste needs to be reconstructed based on accessible cues.

Keywords Taste perception · Rating scales · Sensory testing · Reconstructive memory · Consumer behavior

1 Introduction

The initial development of food products often involves consumer panels who taste and evaluate the products. To minimize the risk of biasing tasters' sensory evaluations, taste tests are conducted in a neutral setting and the tasters receive limited non-sensory information, thereby avoiding biases that may arise from the brand name or price (Eves 1994; Jaegar 2005). According to sensory testing experts, tasters in consumer product testing are typically asked to indicate their sensory experiences on rating scales (S. Kirkmeyer, personal communication, August 3, 2011). In addition to reporting their taste evaluations, tasters are expected to remember their sensory experiences and to assess them vis-à-vis the taste of other samples for comparative evaluation. Similarly, consumers who sample products need to remember their sensory experiences for comparison purposes and later purchase decisions. Whereas potential biases on sensory perceptions arising from non-sensory information in the form of brand names, product names, or prices have received attention in marketing research (as reviewed below), other likely sources of bias have gone unnoticed. One such neglected source is the format in which consumers take notes or report their taste experience to the researchers who conduct the taste test. Whereas a product's brand name or price is usually available before tasters experience the product, influences of the reporting format may also arise after tasters experienced the product.

Several studies have shown that post-experience exposure to product-related advertising can alter information learned from direct experience with the product (Braun 1999; Melcher and Schooler 1996). What is lacking in marketing research however are studies that shed more systematic light on what occurs post-consumption, with respect to consumers' retrospective evaluations of taste (see Köster 2003) after a bias has been introduced during encoding (e.g., from the measurement tool on which the initial sampling was recorded). Memory for taste is important because it affects repurchase decisions and can influence word-of-mouth communication. We conceptualize taste as a subjective experience (Bartoshuk 2000) that shares key characteristics with other subjective experiences, such as moods, emotions, and hedonic experiences of pleasure and pain. If so, the experience of taste is fleeting and must be reconstructed from memory traces (Robinson and Clore 2002).

As numerous prior studies have demonstrated, subjective experiences are generally fleeting and poorly represented in memory. From moods and emotions to pleasure, pain, and processing fluency, people can introspect on their subjective experiences while they have them, but need to resort to reconstruction strategies shortly after the experience (for reviews see Robinson and Clore 2002; Schwarz 2007; Schwarz et al. 2009). As in other domains of autobiographical memory (Bradburn et al. 1987; Ross 1989; Schwarz and Sudman 1994), people will draw on any accessible piece of

information that may be useful to reconstruct what the experience was, including general knowledge about the domain and memorable features of the episode related to the experience. We propose that the same logic applies to taste experiences: the experience itself is fleeting and not well represented in memory; hence, consumers can introspect on it while they are having the experience, but need to reconstruct it later on. This leaves the door open for systematic errors of memory that can affect downstream intentions and behaviors.

Suppose you are asked to rate how “sweet” a drink is. You receive an 11-point rating scale, ranging from “not at all sweet” to “very sweet”. Would it make a difference whether the numeric values run from 0 to 10, or from -5 to $+5$? This very example was encountered by one of the authors while visiting a winery: The managers of the winery had chosen a bipolar (vs. a unipolar) scale for which to collect data about the new featured wine at the tasting bar. While most research methods textbooks assume that these scales are equivalent (both are 11-point scales and have the same verbal end-point labels), previous research on social and emotional judgments (e.g., “how successful would you rate your life?”) indicates that the numeric values influence participants’ interpretation of the question (Schwarz et al. 1991b). Specifically, scales containing only positive values convey that the researcher has a unipolar dimension in mind that pertains to the presence of an attribute (10) and its absence (0). On the other hand, scales with negative and positive numbers convey that the researcher has a bipolar dimension in mind that runs from the presence of an attribute ($+5$) to the presence of its opposite (-5). These differential interpretations result in differential ratings; for instance, 13% of respondents rated their life as successful in the former case, whereas 34% did so in the latter case (Schwarz et al. 1991b; for a review and a grounding of the rationale in the Gricean logic of conversation see Schwarz 1996, 1999). In the case of our wine tasting bar example, a 0 to 10 scale conveys that the underlying dimension runs from “no sweetness” to “high sweetness”, whereas a -5 to $+5$ scale conveys that the underlying dimension runs from “high acidity” to “high sweetness”. Hence, a given drink may be rated as sweeter (i.e., closer to the “high sweetness” end anchor) on the latter than the former scale.

We posit that these kinds of explicit ratings of the taste experience are incorporated into one’s memory trace and are, after some delay, more accessible than the sensory details of the taste experience itself, which to our knowledge has not been examined in prior research. Accordingly, the previously made taste ratings are likely to bias consumers’ memory of their taste experience and influence subsequent decisions that are informed by this memory.

The foregoing considerations lead to the present novel research that generates the following predictions. First, consumers’ attribute ratings for a given drink will be higher when made along a bipolar (-5 to $+5$) than along a unipolar (0 to 10) scale (hypothesis 1). Second, these attribute ratings will inform consumers’ later reconstructive memory for the taste experience such that those who assigned higher fruitiness ratings, for example, will reconstruct the drink as having been fruitier. This biased reconstruction will be reflected in systematic recognition errors: when given several samples to compare, consumers will erroneously identify the sample that is most representative of their rating as the one they originally tasted (hypothesis 2). For instance, in the case of fruitiness ratings, those who rated the original drink on a bipolar rather than unipolar scale will later be more likely to misidentify a fruitier drink as the

one they sampled earlier. Third, this biased reconstruction will also influence later evaluations of the product (hypothesis 3), and affect consumers' WTP (hypothesis 4). Finally, we expect that these influences are less pronounced for experts than for novices in a given product domain (hypothesis 5). We develop the rationale for this final prediction after presenting the first experiment that tests hypotheses 1 to 4.

1.1 Study 1: the effect of rating scale format on recognition, and later overall evaluation and WTP

1.1.1 Participants

One hundred sixty-nine students at a large North American University were recruited to participate (81 males and 88 females, with an average age of 20 years). The experiment was run as part of a set of unrelated studies; participants were compensated with \$10 USD.

1.1.2 Design and procedure

Participants were randomly assigned to a unipolar scale ($n = 66$) or bipolar scale ($n = 62$) condition. The study was conducted in individual sessions and administered through MediaLab, a software tool that is used to present experimental stimuli to participants. Participants completed the study individually. All participants received an unlabeled styrofoam cup with a sample of the standard quality orange juice, which was Kroger's (store) brand frozen concentrate, prepared according to instructions (adding 3 cans, or 1065 ml, of water, to the 355 ml frozen concentrate). Immediately after tasting the sample, participants rated it on four attributes (sweet, concentrated, sugary, and fruity), which were presented in counterbalanced order. Depending on condition, participants provided all attribute ratings on a unipolar or bipolar scale. Participants were asked to "rate the orange juice according to its [attribute]" where each of the four attributes were presented one at a time at the top of the screen. The numeric values of the unipolar scale ranged from $0 = \textit{not at all}$ to $10 = \textit{extremely}$; for the bipolar scale, they ranged from $-5 = \textit{not at all}$ to $+5 = \textit{extremely}$.

After a five-minute break during which participants completed an unrelated filler task comprising simple math problems, participants answered additional questions, including "Overall, how much do you like this orange juice sample?" The ratings were administered on an unnumbered scale that consisted of 11 square boxes presented in a horizontal line. The far left box was labeled "*not at all*", and the far right box was labeled "*very much*". Participants were also asked, "How much would you be willing to pay for a 1 gal carton of this orange juice?" (Siegrist and Cousin 2009). At the time of the experiment, a gallon of orange juice usually cost between \$1.99 USD (generic) to \$3.99 USD (branded).

Finally, the experimenter provided each participant with three qualitatively different orange juice samples: *standard quality* juice, which was the 355 ml can of Kroger's (store) brand frozen concentrate prepared as per the instructions, with 3 cans, or, 1065 ml of water; *diluted quality*, which was the 355 ml can of Kroger's (store) brand frozen concentrate prepared with 4 cans, or 1420 ml of water; and *concentrated quality*, which was the 355 ml can of Kroger's (store) brand frozen concentrate prepared with 1.5 cans, or 532.5 ml of water.

The samples were provided in styrofoam cups and labeled with numbers "1" (diluted), "2" (standard), and "3" (concentrated). Participants were prompted by the

computer to sample all three juices and to identify the orange juice tasted earlier. They were asked, “Which one of all the orange juice samples that you just tasted is the one you sampled in the first half of the experiment?”

1.1.3 Results

Ratings on the four attributes (sweet, concentrated, sugary, fruity) were correlated ($\alpha = .67$) and combined into an initial evaluation index. To examine whether the scale manipulation affected initial evaluation on the four attributes, we coded both unipolar and bipolar scale responses to values ranging from 1 to 11, and examined mean differences between the two groups. As predicted, the attribute ratings were higher on the bipolar ($M = 6.04$, $SD = 2.42$) than on the unipolar ($M = 4.88$, $SD = 2.24$) scale, $F(1, 163) = 20.96$, $p < .001$. This supports hypothesis 1.

To examine whether the numerical values of the attribute ratings affected subsequent overall evaluations, we examined mean differences between the two groups on the 11-point unnumbered scale. As predicted, participants rated the standard sample of orange juice more favorably when they had rated its attributes on bipolar ($M = 6.00$, $SD = 2.38$) rather than unipolar ($M = 4.23$, $SD = 2.45$) scales, $F(1, 163) = 22.24$, $p < .001$. Moreover, the former participants were willing to pay more for the juice ($M = \$2.35$ USD, $SD = \$1.00$) than the latter ($M = \1.61 USD, $SD = \$0.86$), $F(1, 163) = 25.87$, $p < .001$. These findings support hypotheses 3 and 4, respectively.

Participants were more likely to correctly identify the sample they had previously tested after rating its attributes on unipolar (55.95%) than on bipolar (43.53%) scales, Z -contrast = 3.69, $p < .001$. Moreover, the error was systematic and the more concentrated juice sample was twice as likely to be misidentified as the one sampled earlier in the bipolar (29.41%) than unipolar (14.29%) rating conditions, Z -contrast = 4.26, $p < .001$. These observations support hypothesis 2.

1.1.4 Discussion

These findings show (i) that the numeric format of the rating scale can influence participants' attribute ratings, with downstream effects on (ii) overall evaluation and (iii) WTP for the product. Moreover, (iv) participants use their attribute ratings as a highly accessible input in reconstructing their earlier taste experience, resulting in a systematic bias in taste memory: when the scale induced participants to provide higher attribute ratings, they erroneously identified the sample that was most representative of their ratings as having been the sample they initially tasted.

Next, we address whether the observed influence of rating scales generalizes to products that are more complex and are assumed to require a higher degree of perceptual expertise, as is the case for wine (Melcher and Schooler 1996). We further assess how the observed findings are moderated by participants' expertise in the product domain.

1.2 Experiment 2: effect of rating scale format on later evaluation for high versus low knowledge consumers

Previous research suggests that consumers with high knowledge in a product domain are less influenced by contextual cues when they evaluate a product's taste. Whereas

consumers with little knowledge give considerable weight to taste-irrelevant cues, such as price, consumers with high knowledge are more likely to attend to the actual taste of the product when providing concurrent taste ratings (Lockshin and Rhodus 1993). This is well documented in the domain of wine tastings, where experts use twice as many terms to describe wine odor, and have better vocabulary, recognition, and discrimination ability than novices (Gawel 1997; Lawless 1984; Solomon 1990; Parr et al. 2002). While expertise may also be present among consumers who regularly engage in comparative orange juice tastings, their number is likely to be small. Accordingly, we use wine as the taste sample in experiment 2 and predict that consumers who are highly knowledgeable about wine (i.e., experts) will be less affected by the format of the rating scale than consumers with low knowledge in the product domain (i.e., novices) (hypothesis 5).

Several processes may contribute to the expected expertise effect. On one hand, experts in a domain may draw on highly accessible knowledge about the product class and related exemplars (Alba and Hutchinson 1987) when making their initial taste ratings. This knowledge may include an understanding of the dimension on which the product class is typically rated, relevant exemplars that anchor this dimension (e.g., icewine), and so on. Moreover, sensory scientists suggest that more knowledgeable consumers use a different “sensory-perceptual-cognitive strategy” when assessing a concurrent taste experience (Gawel 1997; Lawless 1984), relying less on any sort of cue that is not intrinsic to the product itself. Accordingly, experts’ initial ratings of their concurrent taste experience may be less affected by contextual cues, attenuating the effect of numeric scale values present on the initial taste rating. This, in turn, would reduce or eliminate any downstream influence of the initial ratings. On the other hand, experts may attend to a larger number of attributes, including attributes not included in the initial rating, consistent with their tendency to encode taste in more complex ways (Gawel 1997; Lawless 1984; Solomon 1990; Parr et al. 2002). In addition, experts may enrich their encoding of the taste experience with associations pertaining to other tastes, related samplings, and so on. Compared to novices, this would provide experts with a larger number of accessible cues for later reconstruction, which could dilute the impact of any given cue. From this perspective, the numeric values of the scale may influence experts’ initial ratings to the same extent as novices’, but the downstream effect of the initial taste ratings would be less pronounced for experts. Experiment 2 bears on these issues and addresses expertise differences in consumers’ initial taste ratings and their downstream effects.

1.2.1 Participants

One hundred twenty-nine students and members of the local community of a mid-sized North American University were recruited via advertisements to participate in a “Wine Tasting” study (60 males and 69 females, with an average age of 24 years). Participants were compensated with either \$5 CAD or course credit.

1.2.2 Design and procedure.

Participants were randomly assigned to either a unipolar (0 to 10; $n = 66$) or bipolar (−5 to +5; $n = 63$) scale condition. They sampled a glass of wine and rated its attributes

(crispness¹, freshness, fruitiness, and complexity) along unipolar (0 to 10) or bipolar (−5 to +5) rating scales, as in experiment 1. The wine was unbranded, made by students in the Oenology and Viticulture program of the University where the experiment was carried out. After a five-minute break, participants answered, “Overall, how much do you like this wine?” The ratings were administered on an unnumbered scale that consisted of 11 square boxes presented in a horizontal line; the far left box was labeled “not at all”, and the far right box was labeled “very much”. Participants were also asked, in an open response format, “How much would you be willing to pay for a bottle of this wine?” Finally, participants completed an objective measure of wine expertise, used in many prior studies (e.g., Mantonakis et al. 2009), which categorizes participants as having either high or low knowledge about wine based on accuracy at answering various questions pertaining to wine (Hughson and Boakes 2001). No taste recognition test was administered in this experiment.

1.2.3 Results

Ratings on the four attributes (freshness, crispness, complexity, and fruitiness) were correlated ($\alpha = .69$) and combined into an initial evaluation index. To examine whether the scale manipulation affected initial evaluation on the four attributes, we coded both unipolar and bipolar scale responses as values ranging from 1 to 11.

A 2 (scale: unipolar vs. bipolar) \times 2 (knowledge: low vs. high) factorial ANOVA of the initial evaluation index revealed a significant interaction, $F(1, 117) = 4.26, p < .04$. Simple effect tests showed that for the low knowledge participants, initial ratings were higher in the bipolar ($M = 7.34, SD = 1.58$) than the unipolar ($M = 6.27, SD = 1.18$) condition, $F(1, 117) = 6.92, p < .01$. This supports hypothesis 1 and replicates the results of experiment 1. In contrast, high knowledge participants were not affected by scale type ($M_s = 6.83$ and $6.95, SD_s = 1.79$ and 1.72 , for the bipolar and unipolar scale conditions, respectively), $F < 1$.

An identical ANOVA testing the effects of initial scale type and knowledge level on subsequent overall evaluation also revealed a significant interaction, $F(1, 117) = 6.34, p < .01$. Simple effect tests showed that for low knowledge participants, subsequent overall evaluation was higher in the bipolar ($M = 7.84, SD = 2.36$) than the unipolar ($M = 6.14, SD = 1.77$) condition, $F(1, 117) = 7.04, p < .01$. This supports hypothesis 3 and replicates experiment 1. However, high knowledge participants' overall evaluation of the wine was unaffected by initial scale type ($M_s = 6.00$ and $6.59, SD_s = 3.22$ and 2.51 , for the bipolar and unipolar conditions, respectively), $F < 1$. This supports hypothesis 5.

An ANOVA, examining the effects of initial scale type and knowledge level on subsequent WTP, showed a significant main effect of knowledge, $F(1, 117) = 10.27, p < .002$, with higher WTP for low knowledge ($M = \$15.84$ CAD, $SD = \$7.38$) than high knowledge ($M = \12.38 CAD, $SD = \$5.69$) participants. This was qualified by a marginally significant interaction, $F(1, 117) = 3.75, p = .055$. Simple effect tests showed that for the low knowledge participants, WTP was higher in the bipolar ($M = \$17.54$ CAD, $SD = \$8.68$) than unipolar group ($M = \13.96 CAD, $SD = \$6.36$), $F(1, 117) = 5.78, p < .02$. This supports hypothesis 4 and replicates experiment 1.

¹ We acknowledge that because crispness is related to acidity, it's not obvious that it is necessarily positive.

However, high knowledge participants' WTP was not affected by the initial scale manipulation ($M_s = \$12.11$ CAD and $\$12.62$ CAD, $SD_s = \$6.69$ and $\$4.36$, for the bipolar and unipolar conditions, respectively), $F < 1$. This supports hypothesis 5.

1.2.4 Discussion

As predicted, the numeric values of the rating scale on which consumers evaluated the taste of the sampled wine exerted more influence on the overall product evaluation and WTP of novices than of experts. This was the case because experts' initial taste ratings were unaffected by the format of the rating scale to begin with. This observation is compatible with the assumption that more knowledgeable consumers use a different "sensory-perceptual-cognitive strategy" when assessing their taste experience (Gawel 1997; Lawless 1984), relying less on any sort of cue that is not intrinsic to the product itself, including the scale on which the taste ratings are made. Note that this does not necessarily imply that more knowledgeable consumers have better memory for their actual taste experience. The present study primarily shows that more knowledgeable consumers are less affected by contextual cues when they give a concurrent report of their taste experience to begin with, which limits the possible downstream impact of the initial ratings on subsequent taste reconstructions.

2 General discussion

Consumer researchers have paid scant attention to memory for taste experiences. The present research seeks to provide an answer to the question of whether memory-based judgments about taste could be influenced by a stimulus-extrinsic factor, namely, the format of the measurement tool on which initial taste evaluations are reported. We showed that when a product (orange juice, wine) is sampled and initially evaluated using a measurement tool, that tool can influence how the taste is perceived. More important, these biased perceptions become part of the memory trace and influence later taste recognition, product evaluation, and WTP. Thus, our findings show that extra-product cues present during initial sampling can exert a persistent and non-trivial influence on subsequent judgment and behavior.

This research makes several novel contributions to the marketing research, sensory evaluation, and consumer science literatures. First, the influence of rating scale format on participants' initial taste ratings demonstrates that consumers' concurrent encoding of taste experiences can be profoundly influenced by contextual cues that do not bear on the product itself. Earlier research (Schwarz et al. 1991b; for a review see Schwarz 1996) showed that the numeric values of rating scales can change the meaning of identical verbal scale anchors by conveying that the underlying dimension is unipolar (0 to 10 scale) or bipolar (−5 to +5 scale). The resulting shift in scale meaning reliably elicits higher attribute ratings along bipolar than unipolar scales. To our knowledge, experiment 1 is the first to demonstrate this finding with respect to taste. In addition, experiment 2 is the first study that identifies an exception to this regularity: participants with high expertise in the product domain were unaffected by the scale manipulation. Presumably, these consumers had developed complex representations of the attributes of wine, as suggested by earlier research (Gawel 1997; Lawless 1984; Solomon 1990;

Parr et al. 2002), which limited their reliance on the dimensionality implied by the research instrument. This is in line with previous research showing that experts rely more on intrinsic than extrinsic cues (Gawel 1997; Lawless 1984). We conjecture that these consumers' initial ratings of their concurrent taste experience were based primarily on the actual taste of the wine.

Second, and more important, our findings highlight that the experience of taste is fleeting and not well represented in memory; like other subjective experiences (Robinson and Clore 2002; Schwarz et al. 2009), taste needs to be reconstructed based on accessible cues. Earlier explicit encodings of the taste experience serve as inputs into this reconstruction and bias memory even under conditions where the testing situation provides the best recall cues possible: a repeated taste of the original product itself. In experiment 1, participants who had given the initial sample higher attribute ratings because they were presented with a bipolar scale, subsequently misidentified a more concentrated orange juice as the one they had originally tasted. Thus, the scale-guided encoding of the initial taste experience was sufficient to override the influence of full exposure to the same taste experience in a recognition test, within only a few minutes.

Third, and equally important, given the orienting impact of scale-guided encoding on taste recognition, we report the novel finding that the same manipulation also affects other judgments that draw on consumers' taste (sensory) memories. In both studies, consumers subsequently rated the product more favorably, and were willing to pay more for it, when the scale format of the initial taste report elicited more favorable ratings. This influence of rating scale format extends previous work on the influence of frequency scales on a wide variety of consumer judgments (e.g., Menon et al. 1995; for a review, see Schwarz 1996), including product evaluation and WTP (Hamilton et al. 2011).

Fourth, the initial taste ratings of experts were not affected by scale format and hence no downstream effect on evaluation and WTP was observed for these participants (experiment 2). This suggests that professional tasting panels are unlikely to be affected by scale format. In contrast, consumers recruited for product taste testing or who attend food or wine tastings may be more likely to render favorable taste evaluations and subsequently purchase the items when asked to record their responses along bipolar as compared to unipolar attribute scales.

Fifth, our results add to the very limited body of knowledge that demonstrates that the numerical format of the rating scale on which judgments (social, emotional, taste, etc.) are elicited can have a large impact on consumer judgment, decision making, and memory, as in the case of the wine-tasting bar example we mentioned earlier. This is an important consideration for consumer and marketing researchers, who may not consider the format of rating scales when designing a research instrument about not only taste-related judgments, but other judgments as well. Similar considerations hold for the format of frequency scales, where the frequency values presented to participants influence question interpretation (Schwarz et al. 1988), frequency reports (Schwarz et al. 1985), comparative judgments (Schwarz et al. 1991a; Schwarz and Scheuring 1988) as well as product evaluation and WTP (Hamilton et al. 2011).

Finally, the observed findings differ from other contextual influences on taste perception and evaluation insofar as they are driven by a variable that is extrinsic to the product itself, namely the format of the rating scale on which initial taste

impressions are reported. In contrast, other research has varied information that is substantively related to the product, such as the brand (Hoegg and Alba 2007), color (Garber et al. 2000; Hoegg and Alba 2007), price (Plassmann et al. 2008), or information about the ingredients (Lee et al. 2006). Both lines of work converge in highlighting the reconstructive and context sensitive nature of memory for sensory experiences.

Future research is needed to address whether the same reconstructive processes hold for other senses, including touch and smell. Odor memory is often assumed to be highly accurate (Cain 1979; Engen and Ross 1973), although this conclusion may primarily reflect that perception researchers rarely attempt to introduce potentially misleading context cues. In contrast, the interplay of product and context in consumers' consumption experience is of primary interest to marketing researchers and practitioners as indicated by the rapidly growing interest in sensory marketing (Krishna 2010).

References

- Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of consumer expertise. *Journal of Consumer Research*, 13(4), 411–454.
- Bartoshuk, L. M. (2000). Comparing sensory experiences across individuals: recent psychophysical advances illuminate genetic variation in taste perception. *Chemical Senses*, 25(4), 447–460.
- Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: the impact of memory and inference on surveys. *Science*, 236(4798), 157–161.
- Braun, K. A. (1999). Postexperience advertising effects on consumer memory. *Journal of Consumer Research*, 25(4), 319–334.
- Cain, W. S. (1979). To know with the nose: keys to odor identification. *Science*, 2(4379), 467–470.
- Engen, T., & Ross, B. M. (1973). Long-term memory of odors with and without verbal descriptions. *Journal of Experimental Psychology*, 100(2), 221–227.
- Eves, A. (1994). Sensory analysis—an alternative to wine tasting? *International Journal of Wine Marketing*, 6(2), 32–44.
- Garber, L., Hyatt, E., & Starr, R. (2000). The effects of food color on perceived flavor. *Journal of Marketing Theory and Practice*, 8(4), 59–73.
- Gawel, R. (1997). The use of language by trained and untrained experienced wine tasters. *Journal of Sensory Studies*, 12(4), 267–284.
- Hamilton, R. W., Ratner, R. K., & Thompson, D. V. (2011). Outpacing others: when consumers value products based on relative usage frequency. *Journal of Consumer Research*, 37(6), 1079–1094.
- Hoegg, J., & Alba, J. W. (2007). Taste perception: more than meets the tongue. *Journal of Consumer Research*, 33(4), 490–498.
- Hughson, A. L., & Boakes, R. A. (2001). Perceptual and cognitive aspects of wine tasting expertise. *Australian Journal of Psychology*, 53(2), 103–108.
- Jaegar, S. (2005). Consumer evaluation of novel kiwifruit: willingness-to-pay. *Journal of the Science of Food and Agriculture*, 85(15), 2519–2526.
- Köster, E. P. (2003). The psychology of food choice: some often encountered fallacies. *Food Quality and Preference*, 14(5–6), 359–373.
- Krishna, A. (2010). *Sensory marketing: research on the sensuality of consumers*. New York, NY: Routledge.
- Lawless, H. T. (1984). Flavor description of white wine by “expert” and nonexpert wine consumers. *Journal of Food Science*, 49(1), 120–123.
- Lee, L., Frederick, S., & Ariely, D. (2006). Try it, you'll like it: the influence of expectation, consumption, and revelation on preferences for beer. *Psychological Science*, 17(12), 1054–1058.
- Lockshin, L. S., & Rhodus, W. S. (1993). The effect of price and oak flavor on perceived wine quality. *International Journal of Wine Marketing*, 5(2–3), 13–25.
- Mantonakis, A., Rodero, P., Lesschaeve, I., & Hastie, R. (2009). Order in choice: effects of serial position on preferences. *Psychological Science*, 20(11), 1309–1312.
- Melcher, J., & Schooler, J. W. (1996). The misremembrance of wines past: Verbal and perceptual expertise mediate verbal overshadowing effect. *Journal of Memory and Language*, 35(2), 231–245.

- Menon, G., Raghuram, P., & Schwarz, N. (1995). Behavioral frequency judgments: an accessibility-diagnostics framework. *Journal of Consumer Research*, 22, 212–228.
- Parr, W. V., Heatherbell, D., & White, G. K. (2002). Demystifying wine expertise: olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. *Chemical Senses*, 27(8), 747–755.
- Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing actions modulate the neural representation of experienced pleasantness. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, 105(3), 1050–1054.
- Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, 96(2), 341–357.
- Robinson, M. D., & Clore, J. L. (2002). Episodic and semantic knowledge in emotional self-report: evidence for two judgment processes. *Journal of Personality and Social Psychology*, 83(1), 198–215.
- Schwarz, N. (1996). *Cognition and communication: judgmental biases, research methods, and the logic of conversation*. Hillsdale, NJ: Erlbaum.
- Schwarz, N. (1999). Self-reports: how the questions shape the answers. *American Psychologist*, 54(2), 93–105.
- Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition*, 25, 638–656.
- Schwarz, N., & Scheuring, B. (1988). Judgments of relationship satisfaction: inter- and intraindividual comparison strategies as a function of questionnaire structure. *European Journal of Social Psychology*, 18, 485–496.
- Schwarz, N., & Sudman, S. (1994). *Autobiographical memory and the validity of retrospective reports*. New York, NY: Springer-Verlag.
- Schwarz, N., Hippler, H. J., Deutsch, B., & Strack, F. (1985). Response scales: effects of category range on reported behavior and subsequent judgments. *Public Opinion Quarterly*, 49, 388–395.
- Schwarz, N., Strack, F., Müller, G., & Chassein, B. (1988). The range of response alternatives may determine the meaning of the question: further evidence on informative functions of response alternatives. *Social Cognition*, 6, 107–117.
- Schwarz, N., Bless, H., Bohner, G., Harlacher, U., & Kellenbenz, M. (1991a). Response scales as frames of reference: the impact of frequency range on diagnostic judgment. *Applied Cognitive Psychology*, 5, 37–50.
- Schwarz, N., Knäuper, B., Hippler, H., Noelle-Neumann, E., & Clark, L. (1991b). Rating scales: numeric values may change the meaning of scale labels. *Public Opinion Quarterly*, 55(4), 570–582.
- Schwarz, N., Kahneman, D., & Xu, J. (2009). Global and episodic reports of hedonic experience. In R. Belli, D. Alwin, & F. Stafford (Eds.), *Using calendar and diary methods in life events research* (pp. 157–174). Newbury Park, CA: Sage.
- Siegrist, M., & Cousin, M. E. (2009). Expectations influence sensory experience in a wine tasting. *Appetite*, 52(3), 762–765.
- Solomon, G. E. (1990). Psychology of novice and expert wine talk. *American Journal of Psychology*, 103(4), 495–517.