A Regression Approach to Penalty Analysis to Assess the Relative Importance of JAR Attributes

Jason C. Parcon[1]
[1]PepsiCo Global R&D, 617 W. Main Street, Barrington, IL 60010

Abstract
An important aspect of product development is to verify if a prototype is perceived by consumers as having a "just-about-right" (or JAR) level of an important attribute or if it has "too much" or "not enough" of it. It is assumed that target consumers penalize a product for not achieving JAR, and the penalty is in terms of a drop in overall liking. In particular, it is the difference between the mean liking scores from those who perceived the prototype as JAR versus having "too much" or "not enough" of an attribute. A *t*-test is then conducted to assess the statistical significance of the penalty. In a typical consumer test, JAR assessments are made on several attributes, and penalty analysis is performed on each attribute. The individual analyses pose two problems: (1) the possibility of multiplicity and (2) the relative importance of the various JAR attributes on liking is not measured. This paper proposes a regression procedure that addresses the above-mentioned issues. Examples from real data showed the current version of penalty analysis leads product developers to focus on attributes that may be relatively unimportant. The proposed method provides a clear differentiation between important and unimportant JAR attributes.

Key Words: Just-about-right, multiplicity, regression, penalty analysis

1. Introduction

A product development process is not complete without an assessment of test products by target consumers. In experimental studies with target consumers as subjects, various data are collected with the aim of obtaining diagnostics that indicate if a product is ready for launch or if further modifications need to be considered.

One technique used by sensory and consumer research scientists in making such assessments is to determine if a test product is perceived as having “too much,” “not enough,” or “just-about-right” levels of product attributes of interest. It is assumed that a perception of “too much” or “not enough” of an attribute is associated with a drop in product acceptability or overall liking. If the drop is statistically significant, then product developers may need to consider making modifications to the product.

Depending on the product category under investigation, a list of pertinent attributes may include overall flavor, sweetness, tartness, bitterness, saltiness, stickiness, etc. A primary interest is to establish the link between just-about-right, or JAR, perceptions on these attributes and overall liking of test products.

Information on overall liking is obtained by asking subjects to rate test products using a 9-point overall liking scale, i.e., 1 = Dislike Extremely to 9 = Like Extremely. See, for example, Popper et al (2004).

Perceptions on levels of product attributes, on the other hand, are obtained from so-called just-about-right, or JAR, scales (Rothman and Parker, 2009) given in Table 1.

Table 1: The Just-About-Right, or JAR, Scale

Too weak	(1)
Somewhat too weak	(2)
Just-about-right	(3)
Somewhat too much	(4)
Too much	(5)

Establishing the association between JAR and overall liking data is not trivial, primarily because of the differences in the structure of both scales. The 9-point overall liking scale is monotonic with the ideal value at the upper end of the scale. The 5-point JAR scale, however, is bipolar. The end anchors of a JAR scale are semantic opposites and the ideal value is located at the midpoint. Thus, linking JAR with overall liking using techniques like the correlation coefficient may provide misleading conclusions.

The bipolar nature of JAR scales is among the primary reasons penalty analysis has gained wide usage in the field of sensory and consumer science. Penalty analysis is a method for determining if, relative to JAR perceptions for a specific attribute, perceptions of "too much" or "too weak" result to a drop in overall liking score. If the drop is statistically significant, then further refinement on the product is needed. The details of its calculations are discussed in the next section, where it will become apparent that penalty analysis has the following limitations:

(1) Penalty analysis tends to give "false-positive" results. As discussed in Section 2, penalty analysis performs *t*-tests on all attributes, making it subject to statistical multiplicity.
(2) Even if the results of penalty analyses on all attributes were not false-positives, penalty analysis, in its current form, does not provide information on which of the significant skews are more impactful on overall liking.

Statement (1) implies that penalty analysis may prompt product developers to rectify a non-existent product defect, the business implications of which need no further explanation. Statement (2), on the other hand, implies that penalty analysis does not provide product developers information on which attributes to prioritize in cases where several JAR skews are statistically significant.

The objective of this paper, hence, is to discuss modifications to penalty analysis to mitigate the effect of multiplicity and make results more meaningful by providing information on the relative importance of product attributes on overall liking.

Section 2 discusses how penalty analysis is calculated in practice and provides mathematical arguments on the need to propose a modification to penalty analysis; Section 4 discusses the proposed modification.

2. Penalty Analysis Calculations

Penalty analysis compares the mean overall liking scores between the following groups of respondents:

(1) Respondents who perceived a test product as JAR on an attribute vs. those who perceived test product as having "too weak" of an attribute; and
(2) Respondents who perceived a test product as JAR on an attribute vs. those who perceived test product as having "too much" of an attribute.

Consider an example where a product was evaluated on overall liking and assessed on sweetness level. A summary of the results is in Table 2.

Table 2: Penalty Analysis of Sweetness

	Number of Respondents, *n*		Mean Overall
	#	%	Liking Score
Too weak (1 or 2)	36	28%	5.4167
JAR (3)	70	55%	6.4857
Too much (4 or 5)	21	17%	6.0952

Penalty for "too weak" = 6.4857 – 5.4167 = 1.0690
Penalty for "too much" = 6.4857 – 6.0952 = 0.3905

For the above study, n = 127 respondents were recruited based on specific recruitment criteria. More than half stated that the test product at hand was JAR on sweetness, and they provided an overall liking score of about 6.5 on the product, on the average. A considerable number of respondents perceived the test product as having too weak or somewhat too weak levels of sweetness, from whom a mean overall liking score of 5.4 was obtained. Thus, the penalty for having too weak sweetness is estimated to be at least a full hedonic point (~1.1). This penalty is statistically significant (2-sample t-test assuming equal variances, p-value < 0.01), implying the need for further refinement on the sweetness level.

Note also that 17% of respondents perceived the product as having "too much" sweetness. A minimum percentage skew for "not JAR" is often employed as a means of eliminating smaller, less impactful attributes from consideration. This percentage is typically 20% (Schraidt, 2009). Thus, for the above example, perceptions on the test product as having "too much" sweetness are not considered a primary concern.

2.1 Penalty Analysis as a Regression Model

It can be easily demonstrated that penalty analysis, as described above, is inherently a regression model.

Let

(a) n = number of consumer respondents;
(b) $\boldsymbol{Y}_{n\times 1}$ = vector of overall liking scores on a specific test product;
(c) $\mathbf{1}_{n\times 1}$ = vector of 1's;
(d) For j = 1, ..., n, let $\mathbf{X} = [\boldsymbol{x}_1\ \boldsymbol{x}_2]$ be an $n \times 2$ matrix $\ni$
- $\boldsymbol{x}_1 = [x_{1j}]_{n\times 1}$, where $x_{1j} = 1$ if respondent j perceived attribute X as "too weak," 0 otherwise; and

- $\boldsymbol{x}_2 = [x_{2j}]_{n\times 1}$, where $x_{2j} = 1$ if respondent j perceived attribute X as "too much," 0 otherwise;
- Note that if $x_{1j} = x_{2j} = 0$, then this implies that the j^{th} respondent perceived attribute X as JAR;

(e) $\boldsymbol{\beta} = [\beta_1 \quad \beta_2]'$, where β_1 and β_2 are the repective penalties of "too weak" and "too much" in the target population;

(f) $\boldsymbol{\varepsilon}$ be an $n \times 1$ random vector $\ni E(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $V(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}_{n\times n},\ \sigma^2 > 0$.

Formulate a regression model of the form

$$\boldsymbol{Y} = \beta_0 \mathbf{1} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

and reparameterize it so that

$$\boldsymbol{Y} = \mathbf{X}^* \boldsymbol{\beta}^* + \boldsymbol{\varepsilon}$$

where $\mathbf{X}^* = [\mathbf{1} \quad \mathbf{X}]$ and $\boldsymbol{\beta}^* = [\beta_0 \quad \boldsymbol{\beta}']'$**.** Penalty analysis, then, is just a least squares regression of $\boldsymbol{Y}$ against $\mathbf{X}^*$, i.e., penalty estimates for not achieving JAR can be obtained from the usual least squares estimator of $\boldsymbol{\beta}^*$ given by

$$\widehat{\boldsymbol{\beta}}^* = [\hat{\beta}_0 \quad \hat{\beta}_1 \quad \hat{\beta}_2]' = (\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1}\mathbf{X}^{*\prime}\boldsymbol{Y}. \tag{2}$$

<u>Proposition</u>: *The expressions $\hat{\beta}_1$ and $\hat{\beta}_2$ in equation (2) are the sample penalties for "too weak" and "too much," respectively. $\hat{\beta}_0$, on the other hand, is the sample mean overall liking score among respondents who perceived the test product as JAR on attribute X.*

The above proposition is easily shown, as follows. For attribute X, let n_1 = number of respondents who perceived the test product as "too weak," n_2 = number of respondents who perceived the test product as "too much," and n_3 = number of respondents who perceived the test product as JAR. Hence, $n = n_1 + n_2 + n_3$. Likewise, let

$\bar{Y}^{(1)}$ = mean overall liking score of test product among "too weak" respondents;
$\bar{Y}^{(2)}$ = mean overall liking score of test product among "too much" respondents;
$\bar{Y}^{(3)}$ = mean overall liking score of test product among JAR respondents.

For brevity, the required matrix algebra is not shown, but the components of equation (2) can be expressed as:

$$(\mathbf{X}^{*\prime}\mathbf{X}^*)^{-1} = \begin{bmatrix} \frac{1}{n_3} & -\frac{1}{n_3} & -\frac{1}{n_3} \\ -\frac{1}{n_3} & \frac{n-n_2}{n_1 n_3} & \frac{1}{n_3} \\ -\frac{1}{n_3} & \frac{1}{n_3} & \frac{n-n_1}{n_2 n_3} \end{bmatrix}$$

and

$$\mathbf{X}^{*\prime}\boldsymbol{Y} = \begin{bmatrix} \mathbf{1}'\boldsymbol{Y} \\ \boldsymbol{x}_1'\boldsymbol{Y} \\ \boldsymbol{x}_2'\boldsymbol{Y} \end{bmatrix} = \begin{bmatrix} \text{sum of all overall liking scores} \\ \text{sum of overall liking scores from "too weak" respondents} \\ \text{sum of overall liking scores from "too much" respondents} \end{bmatrix}.$$

The above expressions imply that:

$$\hat{\beta}_1 = -\left(\bar{Y}^{(3)} - \bar{Y}^{(1)}\right) = \text{penalty for "too weak,"}$$
$$\hat{\beta}_2 = -\left(\bar{Y}^{(3)} - \bar{Y}^{(2)}\right) = \text{penalty for "too much,"}$$

and

$$\hat{\beta}_0 = \bar{Y}^{(3)} = \text{mean overall liking score among JAR respondents.}$$

The raw regression output for the example in Table 2 is given in Table 3. Note that results in both tables coincide. $\hat{\beta}_1 = -1.0690$ and $\hat{\beta}_2 = -0.3905$ correspond to penalties for "too weak" and "too much" sweetness in Table 2, respectively. The negative signs are expected: if the product was perceived as "too weak" or "too much" in sweetness, then there is an expected decrease in overall liking. Finally, note that $\hat{\beta}_0 = 6.4857$ corresponds to the mean overall liking score of the test products among respondents who perceived sweetness as JAR.

Table 3: Penalty Analysis Using a Regression Approach

Predictor	Coef	SE Coef	T	P	VIF
Constant	6.4857	0.2357	27.52	0.000	
Sweetness TOO WEAK	-1.0690	0.4044	-2.64	0.009	1.085
Sweetness TOO MUCH	-0.3905	0.4906	-0.80	0.428	1.085

2.2 Penalty Analysis is an Underspecified Regression Model

Expression (1) assumes that overall liking is impacted by only one attribute, i.e., X, and other attributes that may affect overall liking are thrown into the error term.

It should be reasonable to assume that overall liking of a test product is impacted by more than one attribute. Consider another attribute, W, and for $j = 1, \ldots, n$, let $\mathbf{W} = [\boldsymbol{w}_1\ \boldsymbol{w}_2]$ be an $n \times 2$ matrix ∋

- $\boldsymbol{w}_1 = [w_{1j}]_{n\times 1}$, where $w_{1j} = 1$ if respondent j perceived attribute W as "too weak," 0 otherwise; and
- $\boldsymbol{w}_2 = [w_{2j}]_{n\times 1}$, where $w_{2j} = 1$ if respondent j perceived attribute W as "too much," 0 otherwise.

Similar to attribute X in Section 2.1, if $w_{1j} = w_{2j} = 0$, then this implies that the j^{th} respondent perceived attribute W as JAR.

If attributes X and W both significantly impact overall liking, then the true regression model is

$$\boldsymbol{Y} = \beta_0 \mathbf{1} + \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\theta} + \boldsymbol{\varepsilon}. \tag{3}$$

where $\boldsymbol{\theta} = [\theta_1 \quad \theta_2]'$ is the regression coefficient of $\mathbf{W}$. Under model (3), the expected value of penalties obtained via traditional penalty analysis in Section 2.1 is given by:

$$\begin{aligned} E(\hat{\boldsymbol{\beta}}) &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\boldsymbol{Y}) \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\beta_0\mathbf{1} + \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\theta} + \boldsymbol{\varepsilon}), \end{aligned}$$

that is,

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\beta_0\mathbf{1} + \mathbf{W}\boldsymbol{\theta}). \tag{4}$$

Hence, the penalties $\hat{\beta}_1$ and $\hat{\beta}_2$ for attribute *X* are not unbiased for β_1 and β_2. The bias of the estimation is given by the second addend in (4): $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\beta_0\mathbf{1} + \mathbf{W}\boldsymbol{\theta})$. Traditional penalty analysis is thus an underspecified regression model.

Since the elements of **X** and **W** are either 1's or 0's, it can be easily shown that

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{1} = [1 \quad 1]'$$

and

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\boldsymbol{\theta} = \begin{bmatrix} n_1^{-1}(\boldsymbol{x}_1'\boldsymbol{w}_1\theta_1 + \boldsymbol{x}_1'\boldsymbol{w}_2\theta_2) \\ n_2^{-1}(\boldsymbol{x}_2'\boldsymbol{w}_1\theta_1 + \boldsymbol{x}_2'\boldsymbol{w}_2\theta_2) \end{bmatrix}.$$

Since $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ are penalties, it should be reasonable to assume that their elements are less than or equal to 0. The above expressions imply that

$$\left|E\left(\hat{\beta}_k\right)\right| > |\beta_k|, \quad k = 1, 2.$$

Therefore, if more than one attribute has an impact on overall liking, then sample penalties obtained from traditional penalty analysis will over-estimate the true penalties.

The study described in Table 2 examined not just one but seven product attributes. Hence, depending on the percentage skews, traditional penalty analysis will conduct a considerable number of two-sample *t*-tests on the data. Table 4 provides a summary of the results for the example at hand. There were 8 "not JAR" skews that met the guideline for the minimum percentage skew, resulting to 8 individual two-sample *t*-tests.

Except for one "not JAR" skew, the *t*-tests conducted at $\alpha = 0.10$ yielded significant penalties for all "not JAR" skews.

Several issues arise with the said results, to wit:

(a) From a practical standpoint, it is disconcerting to learn that a test product has so many "defects" even after going through several layers of product development. Table 4 results imply that the test product at hand requires an overhaul.
(b) The mathematical arguments in this section clearly illustrate the high likelihood that penalty estimates in Table 4 may be inflated. Hence, results of significance testing may be false-positives, a consequence of multiple *t*-testing. How should these values be adjusted so that they reflect an unbiased estimate of the penalties?
(c) Even if, for example, the estimates in Table 4 are not unbiased, traditional penalty analysis does not provide information on which of the significant skews should be prioritized when fixing the product defects.

Section 3 discusses an adjustment to traditional penalty analysis to address the above issues.

3. Modifying Penalty Analysis

It is clear from Section 2 that, if several attributes are being investigated for their possible impact on overall liking, then they should be analyzed in the same regression model. For example, if two attributes *X* and *W* are being investigated, then model (3) should be used;

Table 4: Penalty Analyses on Eight Product Attributes

	Overall Flavor			Flavor X		
	Mean Overall Liking	*n*	%	Mean Overall Liking	*n*	%
Too Weak	5.5357	28	22%	5.3696	46	36%
JAR	7.0678	59	46%	6.9063	64	50%
Too Much	5.1250	40	31%	5.1765	17	13%
Penalty for TOO WEAK	1.5321	*		1.5367	*	
Penalty for TOO MUCH	1.9428	*		1.7298	Not tested	

	Sweetness			Saltiness		
	Mean Overall Liking	*n*	%	Mean Overall Liking	*n*	%
Too Weak	5.4167	36	28%	5.3871	31	24%
JAR	6.4857	70	55%	6.3444	90	71%
Too Much	6.0952	21	17%	6.5000	6	5%
Penalty for TOO WEAK	1.0690	*		0.9573	*	
Penalty for TOO MUCH	0.3905	Not tested		-0.1556	Not tested	

	Tartness			Mouthfeel		
	Mean Overall Liking	*n*	%	Mean Overall Liking	*n*	%
Too Weak	5.8421	38	30%	5.6042	48	38%
JAR	6.4359	78	61%	6.4667	75	59%
Too Much	4.8182	11	9%	5.7500	4	3%
Penalty for TOO WEAK	0.5938	n.s.		0.8625	*	
Penalty for TOO MUCH	1.6177	Not tested		0.7167	Not tested	

	Flavor Y		
	Mean Overall Liking	*n*	%
Too Weak	5.4694	49	39%
JAR	6.8833	60	47%
Too Much	5.3333	18	14%
Penalty for TOO WEAK	1.4139	*	
Penalty for TOO MUCH	1.5500	Not tested	

*Significant at $\alpha = 0.10$
n.s. = Not significant at $\alpha = 0.10$

traditional penalty analysis would instead formulate versions of model (1) for both X and W, and these models result to inflated penalty estimates.

Therefore, it is proposed that traditional penalty analysis be modified as a multiple regression of overall liking against all attributes under investigation. For $i = 1, \dots, p$ ($p =$ number of attributes), extend the notation defined earlier as:

(a) $\boldsymbol{Z}^{(i)} = \begin{bmatrix} \boldsymbol{z}_1^{(i)} & \boldsymbol{z}_2^{(i)} \end{bmatrix}$ be ∋
- $\boldsymbol{z}_1^{(i)}$ = 1 if attribute $Z^{(i)}$ was perceived as "too weak," 0 otherwise;
- $\boldsymbol{z}_2^{(i)}$ = 1 if attribute $Z^{(i)}$ was perceived as "too much," 0 otherwise;

(b) $\boldsymbol{\beta}^{(i)} = \begin{bmatrix} \beta_1^{(i)} & \beta_2^{(i)} \end{bmatrix}'$ be the regression coefficient of $\boldsymbol{Z}^{(i)}$.

The proposed adjustment to traditional penalty analysis is to consider the linear model

$$\boldsymbol{Y} = \beta_0 \boldsymbol{1} + \boldsymbol{Z}^{(1)}\boldsymbol{\beta}^{(1)} + \cdots + \boldsymbol{Z}^{(p)}\boldsymbol{\beta}^{(p)} + \boldsymbol{\varepsilon} \tag{5}$$

with the estimated penalties for "not JAR" the least squares estimates of the regression coefficients. This scheme has several advantages over traditional penalty analysis, namely:

(a) If the $\boldsymbol{Z}^{(i)}$'s have significant impact on overall liking, then Section 2 indicates that the estimated penalties are unbiased estimates of the true penalties.
(b) The above multiple regression model will provide information on the relative importance of the $\boldsymbol{Z}^{(i)}$'s on overall liking, allowing product developers to prioritize which attributes to refine first.
(c) The estimate of the intercept, β_0, is the estimated mean overall liking score if all attributes were perceived as JAR. Note that traditional penalty analysis provides no insight on overall liking if all attributes were perceived as JAR.

Table 5: Multiple Regression Approach to Penalty Analysis

Predictor	Coef	SE Coef	T	P	VIF
Constant	7.3048	0.2607	28.02	0.000	
Overall Flavor TOO MUCH	-1.6511	0.4096	-4.03	0.000	1.436
Overall Flavor NOT ENOUGH	-0.9936	0.4780	-2.08	0.040	1.558
Flavor X NOT ENOUGH	-0.8247	0.5376	-1.53	0.128	2.649
Flavor Y NOT ENOUGH	0.1478	0.5205	0.28	0.777	2.547
Sweetness NOT ENOUGH	-0.7157	0.3826	-1.87	0.064	1.180
Saltiness NOT ENOUGH	-0.4711	0.4256	-1.11	0.271	1.326
Tartness NOT ENOUGH	0.5076	0.4079	1.24	0.216	1.384
Mouthfeel NOT ENOUGH	-0.1054	0.3833	-0.28	0.784	1.371

Table 5 is the raw regression output under model (5) after "not JAR" skews that did not meet the minimum 20% response rate were removed. Table 6 summarizes the results of traditional penalty analysis alongside the results in Table 5. Recall that the traditional penalty analyses in Table 4 pointed to all but one of the "not JAR" skews tested as significant, while the proposed multiple regression approach resulted to only three significant "not JAR" skews (α = 0.10).

Finally, if all "not JAR" skews are 0, then this implies that all attributes were perceived as JAR. Thus, the estimate of the intercept in Table 5 also gives us information on the overall liking score if JAR was perceived on all attributes. For the given example, an estimated mean overall liking of 7.3 is achieved if JAR was perceived on all attributes.

Table 6: Traditional Approach and Proposed Modification to Penalty Analysis

		Penalties	
		Traditional Approach	Modified Approach
Overall Flavor	Too weak	-1.5321^{S}	-0.9936^{S}
	Too much	-1.9428^{S}	-1.6511^{S}
Flavor X	Too weak	-1.5367^{S}	-0.8247
	Too much	-1.7298^{NT}	
Flavor Y	Too weak	-1.4139^{S}	0.1478
	Too much	-1.5500^{NT}	
Sweetness	Too weak	-1.0690^{S}	-0.7157^{S}
	Too much	-0.3905^{NT}	
Saltiness	Too weak	-0.9572^{S}	-0.4711
	Too much	0.1556^{NT}	
Tartness	Too weak	-0.5938	0.5076
	Too much	-1.6177^{NT}	
Mouthfeel	Too weak	-0.8625^{S}	-0.1054
	Too much	-0.7168^{NT}	
Intercept		Not Applicable	7.3048

^{NT} not tested since response rate was below 20%

^{S} significant at $\alpha = 0.10$

4. Summary and Conclusions

The above discussion illustrated the advantages of utilizing a multiple regression approach to penalty analysis compared to individual analyses of attributes. Traditional penalty analysis is likely to provide inflated results since the effect of other attributes on overall liking is not accounted for. This was shown mathematically and via an example. Further, multiple *t*-testing subjects traditional penalty analysis to multiplicity, increasing the chances of obtaining false-positives.

There are other advantages that a regression approach to penalty analysis offers that were not discussed in this paper. Note that overall liking data are typically not normally distributed, an important least squares regression requirement. Hence, more efficient estimators of penalties may be obtained by exploring other regression methods, such as R regression (Hettmansperger and McKean, 2011). Also, in cases where there is a considerable number of attributes being investigated, one may explore elimination approaches (stepwise, backward, forward, all possible regression, etc.) if model parsimony is a primary concern.

References

Hettmansperger, T. and McKean, J. 2011. Robust Nonparametric Statistical Methods, Second Edition. Boca Raton, FL: Chapman Hall.

Popper, R. et al. 2004. The Effect of Attribute Questions on Overall Liking Ratings. *Food Quality Preference*. Vol. 15. 853-858.

Rothman, M. and Parker, M. 2009. Structure and Use of Just-About-Right Scales. In *Just-About-Right (JAR) Scales: Design, Usage, Benefits, and Risks*. West Conshohocken, PA: ASTM International. 1-13.

Schraidt, M. 2009. Appendix L: Penalty Analysis or Mean Drop Analysis. In *Just-About-Right (JAR) Scales: Design, Usage, Benefits, and Risks* (Rothman, M. and Parker, M., editors), West Conshohocken, PA: ASTM International. 50-53.

Disclaimer

The opinions expressed in this manuscript are those of the author and do not represent a statement of strategy or investment by PepsiCo.