

Alternatives to Just About Right Scales for Obtaining Feedback on Product Diagnostics



C. Ledeker, T. Aldredge, A. Weigand; McCormick & Company, Inc., MD, USA
 Email: christie_ledeker@mccormick.com

INTRODUCTION

The objective of this study was to identify alternatives to the Just About Right (JAR) scale that provide more information on product diagnostics by:

Expanding product terminology

Diagnostic questions were developed that allowed consumers to rate attributes with the option to indicate that certain terms did not apply to the samples, differing from traditional JARs which force ratings for each attribute. These alternatives to JARs allow consumers to select attributes which they associate with each sample, allowing for more attribute options within a ballot.

Using priming to increase differentiation among samples

Priming was also investigated to determine if eliciting memories of previous experiences with the products assisted in differentiating sample ratings.

METHODS

Six ballots were evaluated in Central Location Tests on 2 gravy samples served over mashed potatoes in balanced order. Unique respondents were used to evaluate each ballot, and each ballot was evaluated by 80 to 107 respondents. Respondents were screened to like and eat brown gravy on mashed potatoes.

Ballot diagnostics included the following:

Standard JAR: 5-point JAR (6 attributes total)
 Please rate this Brown Gravy sample for the following characteristics:

Color

To light	Slightly too light	Just about right	Slightly too dark	Too dark
----------	--------------------	------------------	-------------------	----------

Rate All: 3-point JAR with an option to indicate that attributes did not apply (23 attributes total)
 Please rate this Brown Gravy sample for each of the following APPEARANCE characteristics.
 If the characteristic listed does not apply to the sample, please select "Does not apply."

Color

To light	Just about right	Too dark	Does not apply
----------	------------------	----------	----------------

RATA: 2-point ratings presented as rate all that apply (RATA) to identify attributes too low or too high in intensity (23 attributes)
 For this Brown Gravy sample please rate the following APPEARANCE characteristics.
 If a characteristic is just about right or if a characteristic does not apply to the sample, do not rate the characteristic.

Color

To light	Too dark
----------	----------

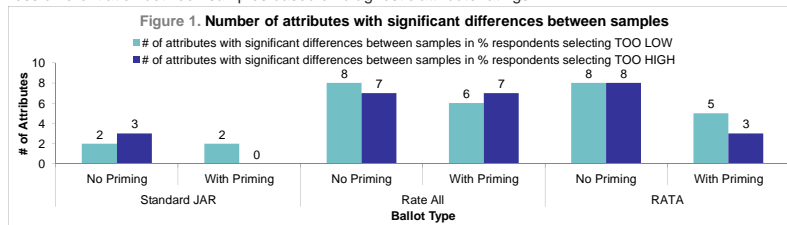
ATTRIBUTES ON BALLOTS	Standard JAR	Rate All & RATA
APPEARANCE ATTRIBUTES		
Color	X	X
Brown Color		X
Amount of Seasoning		X
Specks		X
FLAVOR ATTRIBUTES		
Total Flavor	X	X
Spice/Seasoning Level	X	X
Beef Flavor	X	X
Roasted Flavor		X
Fatty Flavor		X
Onion Flavor		X
Garlic Flavor		X
Black Pepper Flavor		X
Floury / Starchy Flavor		X
Richness		X
Artificial Flavor		X
Aftertaste		X
Sourness		X
Saltiness		X
Bitterness		X
Sweetness		X
TEXTURE ATTRIBUTES		
Consistency		X
Smoothness		X
Clumpiness	X	X
Creaminess	X	X

Each ballot was presented with and without priming:

Priming: Prompted respondent to think about the last time they ate dish
 Please take a moment to try the Brown Gravy sample.
 Think about the last time you ate Brown Gravy as you answer the following questions...

RESULTS & DISCUSSION

As shown in Figure 1, results showed that broader attribute selections in the alternative diagnostic questions led to more significant differences ($p \leq 0.05$) between samples in diagnostic feedback. Priming tended to show less differentiation between samples based on diagnostic attribute ratings.

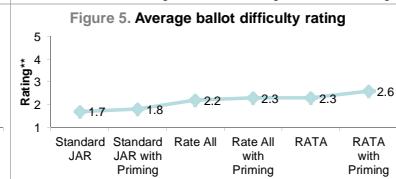
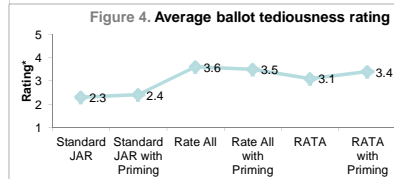
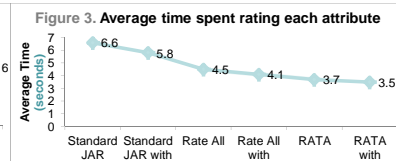
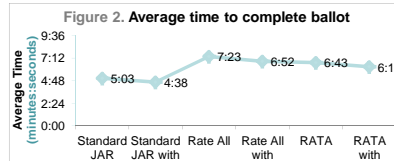


Significant differences at 95% confidence level

Results showed similar % respondents selected terms as too low or too high in intensity among terms presented in all ballots, as seen in Table 1.

	Standard JAR	Standard JAR with Priming	Rate All	Rate All with Priming	RATA	RATA with Priming
Average % of attributes rated as "do not apply" to the sample	N/A	N/A	13%	14%	N/A	N/A
Average % of attributes rated as "too low"	17%	19%	17%	16%	18%	14%
Average % of attributes rated as "too high"	13%	9%	8%	9%	8%	7%
Average % of attributes rated as "just about right"	69%	72%	62%	61%	N/A	N/A

As seen in Figures 2-5, Standard JAR ballots tended to take less time to complete and were considered to be less tedious and difficult. However, more time was spent, on average, rating each attribute on Standard JAR ballots. Rate All ballots took on average 36 to 40 seconds more time than their RATA counterparts to complete.



*9-point rating (1 = not tedious; 9 = extremely tedious)

**9-point rating (1 = not difficult; 9 = extremely difficult)

Tables 2-4 show a comparison of feedback obtained on significant differences between samples in attribute ratings, focusing on Standard JAR vs. RATA comparison, which is an alternative to Standard JAR that provides detailed attribute feedback and slightly less time to complete the ballot than Rate All.

Comparison of results from Standard JAR and RATA diagnostics

Table 2. Significantly more respondents rated Sample A as "too low" in the following:

ATTRIBUTES	Standard JAR	RATA
Color ("too light")		X
Brown Color ("not brown enough")		X
Amount of Seasoning Specks		X
Spice / Seasoning Level	X	X
Beef Flavor		X
Garlic Flavor		X
Richness		X
Saltiness		X

Table 3. Significantly more respondents rated Sample B as "too high" in the following:

ATTRIBUTES	Standard JAR	RATA
Color ("too dark")	X	X
Total Flavor	X	X
Beef Flavor	X	
Brown Color ("too brown")		X
Amount of Seasoning Specks		X
Onion Flavor		X
Richness		X
Aftertaste		X
Sourness		X

Table 4. Significantly more respondents rated Sample B as "too low" in the following:

ATTRIBUTES	Standard JAR	RATA
Consistency ("too thin")	X	

Significant differences at 95% confidence level.

CONCLUSIONS

Alternative diagnostic questions:

- provide more flexibility with terminology, allowing respondents to only rate attributes that apply to the sample being evaluated and providing the researcher opportunity to expand attribute terminology on ballots.
- tended to provide the most information about diagnostic attributes and led to more differentiation between samples when presented without priming.
- led to longer ballot completion times and heightened perceived tediousness due to additional attributes; however, less time was spent rating each attribute.
- when presented as RATA questions, ballot completion times were slightly reduced.

Alternative diagnostic questions can increase feedback available for formula optimization.

REFERENCE

Ares, G., Bruzzone, F., Vidal, L., Cadena, R.S., Gimenez, A., Pineau, B., Hunter, D.C., Paisley, A.G., Jaeger, S.R. (2014). Evaluation of a rating-based variant of check-all-that-apply questions: Rate-all-that-apply (RATA). Food Quality and Preference, 36, 87-95.