FOOD FOR THOUGHT

Topical Insights from our Subject Matter Experts

GAINING DEEPER REFORMULATION INFORMATION FROM PRODUCT GUIDANCE TESTS

The NFL White Paper Series
Volume 1, January 2012

Overview

Product guidance tests typically address two key objectives: 1) assessing the overall liking or acceptability of a product versus a designated benchmark; and 2) determining, and then prescribing, reformulation direction (or optimization opportunities) when warranted. Answering the first objective is generally more straight-forward than answering the second, despite several tools available to the researcher and/or development team. At the end of the day, though, prescription of reformulation direction should be clear, should pertain to the areas most related to (or most influential on) overall liking, and be in a format that developers can readily understand and take action upon – essentially honing them in on the problematic areas. Processes that adhere to these criteria – especially those that augment learnings from one area to another – maximize the opportunity to correct or optimize any deficiencies, thus leading to efficiencies for the development team and the development process.

A Logical Starting Point

The default and well-regarded '**J**ust **A**bout **R**ight' scale (i.e., 'JAR' scale) has been a staple of product guidance tests for decades. Its basic structure typically consists of 5-scalar ratings for each of several attributes; for example, overall flavor strength, sweetness, color, or mouthfeel. Respondents, upon tasting or using a product, use the scale to describe their perception of each attribute: is the product too strong, too weak, or just about right? The scales offer an intuitive means for generating such perceptions, and can provide initial direction for diagnostic reformulation efforts by focusing on the attributes that deviate from the 'just about right' portion of the scale (i.e., those with distributions skewed to the 'too strong/too much' or 'too weak/not enough' portions).

Figure 1 below provides a simplified example of typical JAR-scale output for a guidance test on a strawberry yogurt consisting of 10 attributes. In this hypothetical study, 2 products were assessed: a new offer from a manufacturer of yogurt, and a target competitive product. The action standard was for the new yogurt product to achieve at least parity performance versus the competitor for overall liking. Results showed the new product to be significantly less-liked, meaning that the action standard was not met, and that reformulation would be needed to improve the product. From the output in Figure 1, an analyst would likely conclude that 5 areas had distributions of potential concern: color (new yogurt was too light); overall flavor strength (new yogurt was too weak); strawberry flavor strength (new yogurt was too weak); sweetness (new yogurt was not sweet enough); and viscosity (new yogurt was too thin). Based just on this information, the development direction could be hypothesized as darkening the color of the yogurt, increasing its overall and strawberry flavor strength, making it sweeter, and adjusting the viscosity to make it thicker. This gives the developer a good idea on the starting points that require modification and which, in theory, should improve the overall acceptance or liking of the product to a level needed to achieve the desired action standard of parity performance versus the target competitor.

Figure 1: Typical 'JAR-scale' Output and Potential Reformulation Direction

New Yogurt Attribute	JAR-scale Distribution			Hypothesized Reformulation Direction
	Too Much	Just Right	Not Enough	
Color (too dark-too light)	8%	47%	**35%**	Darken color Increase overall flavor Increase strawberry flavor Make sweeter Make thicker
Smooth Appearance	10%	84%	6%	
Overall Aroma	10%	78%	12%	
Strawberry Aroma	8%	76%	16%	
Overall Flavor	12%	56%	**32%**	
Strawberry Flavor	10%	54%	**36%**	
Sweetness	15%	57%	**28%**	
Tartness	20%	62%	18%	
Smooth Mouthfeel	9%	81%	10%	
Viscosity	13%	60%	**27%**	

Further Directing/Prioritizing the Starting Point

Output from the above example does, in fact, steer the analyst/developer toward the potentially problematic areas, or areas that warrant optimization; case in point, 5 of the 10 attributes have distributions that suggest deficiencies. In this situation, the analyst has a choice to make: address all 5 of the areas based on judgment and interpretation of the data, or apply a more rigorous assessment to determine which area(s) are creating the most concern for consumers. Addressing the latter prioritizes where development efforts can have the most positive impact.

Various statistical techniques, referred to as 'Penalty Analyses,' exist, and are employed to prioritize and then hone in on the most problematic areas. Penalty Analyses collectively share a common thread – they identify the attributes (in our case, the JAR-scales) that most closely align with a key variable (in our case, overall liking of the new yogurt product). As such, their output can categorize problematic attributes into top, secondary or bottom tiers, thus giving more prioritized prescriptive direction to the analyst and his/her development team. Using the same attribute data from above, Figure 2 represents a simple example of output from a Penalty Analysis. In this example, the analysis clearly designates that 3 of the 5 attributes (sweetness, strawberry flavor strength and overall flavor strength) penalize the overall liking rating more so than the remaining attributes (color and viscosity). With this information in hand, the analyst can focus the development team on 3 top areas, rather than 5, with greater likelihood that attention to the sweetness and flavor areas will positively impact overall liking.

Figure 2: Example Penalty Analysis Output and Refined Reformulation Direction

Importance of Attributes	% Stating	Total Penalty*	Prioritized Reformulation Direction
Top Tier			
Not sweet enough	28%	0.59	
Strawberry flavor too weak	36%	0.57	Make sweeter
Overall flavor too weak	32%	0.51	Increase strawberry flavor
Second Tier			Increase overall flavor
Viscosity not thick enough	27%	0.37	
Bottom Tier			
Color not dark enough	35%	0.12	

* Total Penalty = (% not 'just right') x (Mean Overall Liking Rating Drop)

The Final Detail: Honing the Reformulation Direction with NDA

Applying Penalty Analyses to diagnostic JAR-scale data brings clarity and focus to required reformulation efforts, but should such analyses be the final tool to aid Product Developers? In some cases, yes. In others, perhaps not. Some diagnostic attributes, including those from Penalty Analyses, can be too vague or general, leading to interpretation challenges for even the most adept developers. A good example is the output in Figure 2, where 'overall flavor strength' was identified as one of the Top Tier attributes and perceived as being too weak by target consumers. A similar finding exists with 'strawberry flavor strength' in our example, along with 'sweetness.' The logical conclusion with these data is to increase the concentration of these components, but even with this direction there are subjective judgments that need to be made or agreed upon. For example, should the modifications be related to a 'jammy' strawberry flavor, or a 'sweet candy' strawberry flavor, or perhaps a 'fresh' strawberry flavor? Developers faced with the challenge of screening/testing several different strawberry flavor variants could certainly benefit from knowing which strawberry character to focus on.

Narrative Descriptive Analysis (NDA) is a tool that provides such focus by further and objectively 'dissecting' the products in question – in our case, the new yogurt. NDA relies on the use of descriptive sensory panelists – individuals who have undergone extensive training so that they can characterize product attributes in very specific ways, using terminology that developers can readily understand – and thus act on. Figure 3 represents a brief summary of typical NDA output. As you can see, the output is verbal in nature and offers specific prescriptive detail about how the new yogurt product deviates from the reference competitive product across all key dimensions (e.g., appearance, aroma, flavor, texture). In particular, it builds upon what has been determined in our previous example figures by further describing some of the flavor and strawberry character differences. Once available to a developer, they can now focus on reformulation efforts pertaining to fresh/ripe and possibly candy-like strawberry flavor systems, as well as reducing the cultured

dairy notes (the latter of which may be interacting with strawberry flavor or sweetness) – certainly more granular information that could guide the re-development efforts to success more efficiently and effectively.

Figure 3: Example Narrative Descriptive Analysis (NDA) Output

	Target Competitive Yogurt (Reference)	New Yogurt Test Product (Deviation from Reference)
Color/Appearance	Bright pale pink, thick, small darker pink specks, some small red pieces of strawberry	Slightly lighter pink, slightly thinner
Aroma	Moderate strawberry, slight to moderate sweet aromatic, slight cultured dairy	Moderately less strawberry, slightly less sweet aromatic
Flavor	Moderate strawberry (fresh fruit character), moderately sweet, slightly sour, slight cultured dairy	Moderately less sweet, **slightly more cultured dairy, slightly less strawberry, slightly different strawberry character (more candy-like/less fresh ripe fruit-like)**
Texture	Moderately thick, smooth	Slightly thinner

Additional Advantages of NDA/Trained Panelists

Beyond the ability to 'drill down' the reformulation direction, several other advantages exist with using and/or relying on trained tasters with Guidance Testing programs.

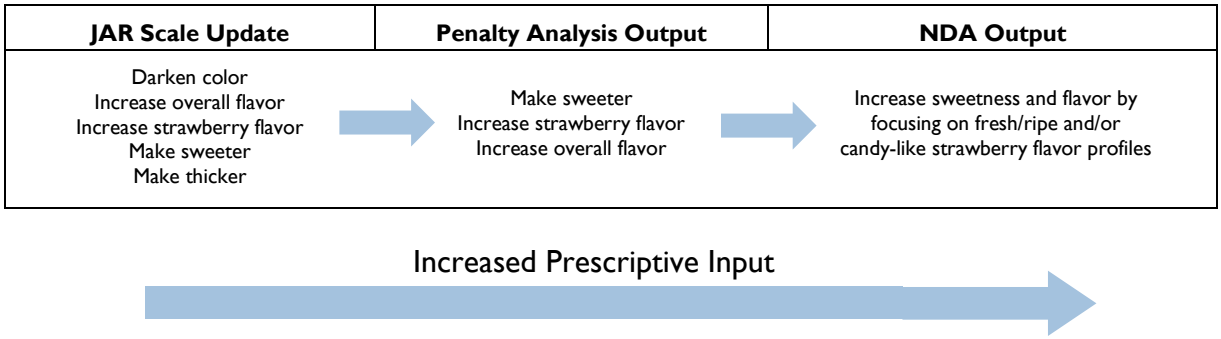
- Ability to use language in a predictable way compared to consumers. A classic example relates to two of the basic tastes – bitterness and sourness – which can often be confused by consumers, or used as a means of expressing a 'dislike' of the product, thus creating interpretation challenges. This situation is minimized when objective tools, such as NDA, are applied.
- Adding clarity when polarized JAR-scale data exists. There are often situations where 'too strong/too much' and 'too weak/not enough' are both chosen by large percentages of consumers. This could mean that different preference segments exist, or it could mean that consumers did not care for the flavor character and do not know how to answer the question (i.e., it isn't 'just right,' but it may not necessarily be 'too strong' or 'too weak'). The addition of NDA or related tools can tease out such insights.

- Efficiencies in thoughtful questionnaire development. Survey questionnaires have always been scrutinized (and in many cases chastised) for their excessive length. The use of JAR-scales adds to this phenomenon and asking too many of them can dilute the quality of the data. Scaling back their use, and relying on NDA to 'fill in the details,' is ultimately more effective, efficient, and actionable.
- An aid for record keeping. NDA output provides an objective summary of what the products tasted like in a Guidance Test, and over time can be used to develop a 'knowledge center' or record of what was tasted. This becomes especially useful if products become unavailable for tasting in the future, or if some stakeholders on the project have not tasted the products themselves.

In The End: Progression That Focuses Action

Beyond the added benefits of NDA mentioned above, a thorough analysis such as that described in this paper leads to greater prescriptive advice, focus and efficiencies for developers faced with product reformulation challenges. The summary diagram in Figure 4 demonstrates this, and shows that while conclusions can be made based on one element of the process outlined in this paper, better conclusions are likely when learnings from one analysis can be augmented with effective sub-analyses, especially those that include input from trained tasters.

Figure 4: Summary Output Conclusions for Different Diagnostic Analyses



For more information please contact:
Christie Hoyer at HoyerC@TheNFL.com, 925.551.4262 or
Dawn Chapman at ChapmanD@TheNFL.com, 925.551.4243