GENERAL MILLS

PGI Product Guidance and Insights

The Journey To Tetrad: Tales From Industry

Jenny Kamerud
Associate Principal Scientist, Product Guidance and Insights
General Mills, Inc

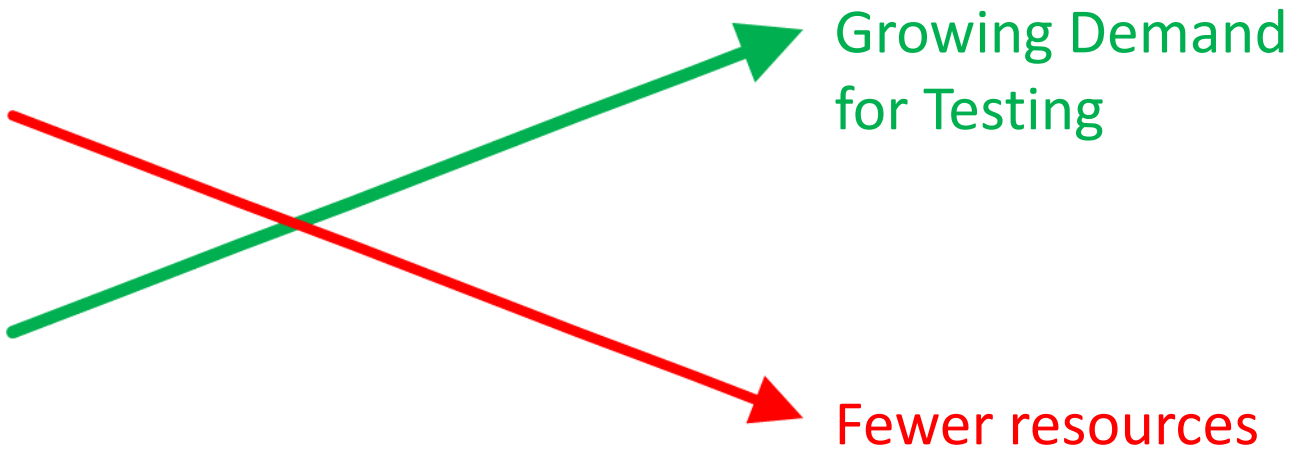
“If it ain’t broke, don’t fix it?”

- General Mills has a long standing history of using discrimination testing to guide product and processing reformulation changes.

- Triangle testing and degree-of-difference (DOD) testing had been used to manage risk for blind product changes.
  - One early mention of triangle test in GMI Technical Records is from 1966!!
  - GMI has two publications related to DOD (DOD-CV in 2006 and DOD-CTV in 2008)

But…

- Challenges to obtaining the **AMOUNT OF PRODUCT** and the **NUMBER OF PEOPLE** were ever-present and growing



One Question, Two Methods

Triangle Test for Homogeneous Products

- Panelists are served either 2 Controls and 1 Test, or 1 Control and 2 Tests
- Panelists choose the odd sample
- Outcome is %True Discriminators; Hurdle is 15% or less (historical GMI norm)



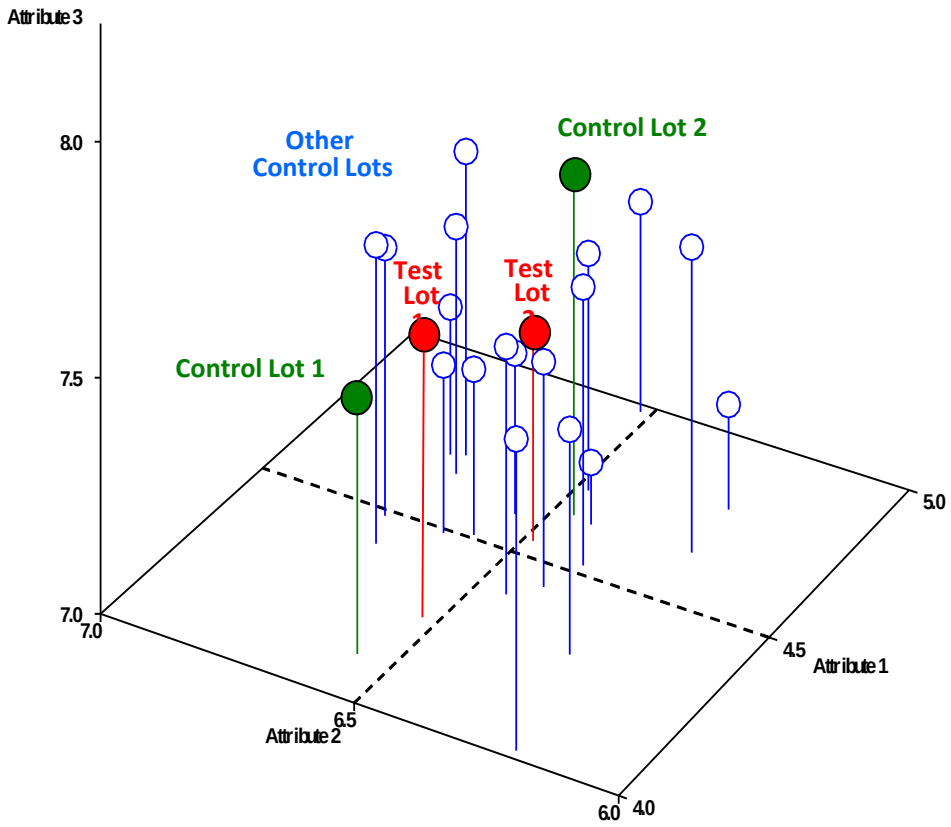
DOD for Heterogeneous or highly variable Products

- Panelists evaluated 3 side-by-side paired comparisons, a subset of 6 possible product combinations (assuming 2 lots of Control and 2 Lots of Test), on a 9-point difference scale
- Outcome is comparison of difference measurements from 6 product pairs ($\alpha<0.05$)

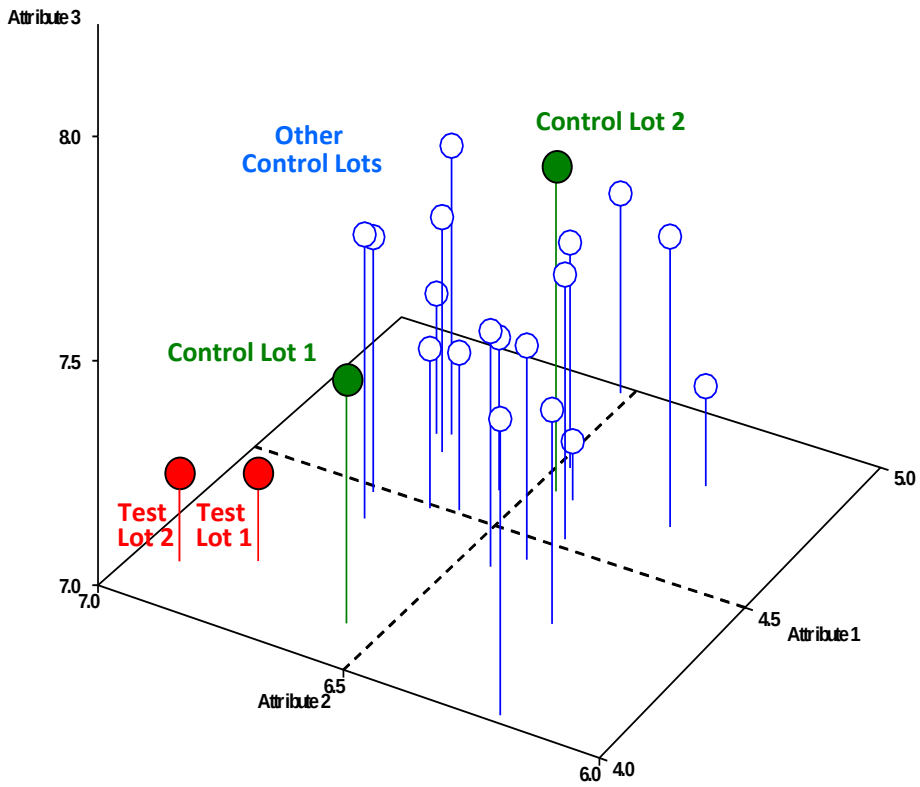


So why do DODs, anyway?

LOW RISK



HIGH RISK



In 2010… could Tetrad be a win-win?

- In the late 2000's, the current flurry of activity on Tetrad began, sparked by Ennis and Jesionka (J. Sensory Studies 2011, 26, 371-382).

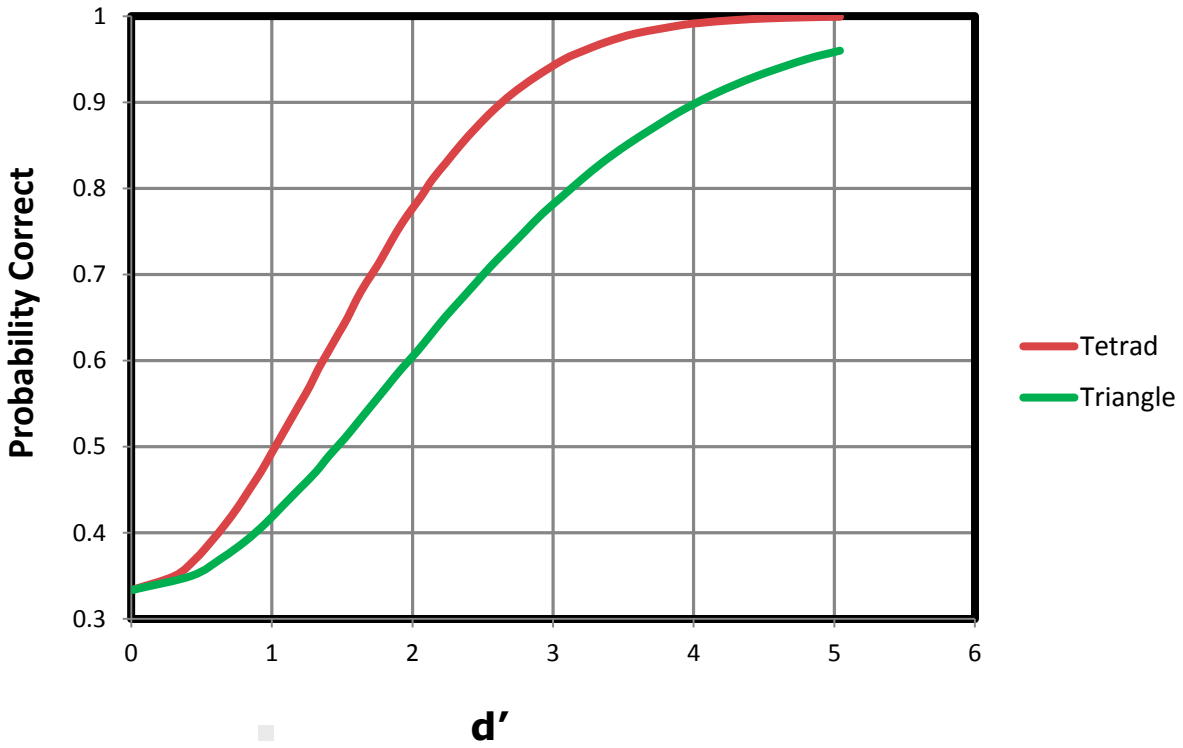
- General Mills Sensory Scientists had their curiosity piqued at SSP in 2010 and undertook internal empirical research conducted throughout 2011 and 2012.

- General Mills rolled-out Tetrad as our primary discrimination method in January 2013…

  …so how did we do it?

Tetrad Method Review

- Each panelist receives 2 Control samples and 2 Test Samples
- Panelists group 4 samples into 2 pairs based on similarity; p = 1/3 by chance
- Tetrad has greater power than triangle; so smaller sample sizes should deliver the same performance as a triangle

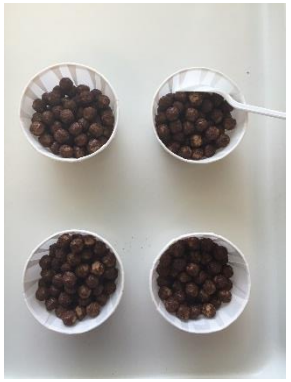


Could we consolidate to one method?

- We still have the product variability issue, but with its **four sample structure**, Tetrad provides a nice **analog to our DOD protocol**!
  - For Heterogeneous or highly variable products, the two Control samples are from **different production lots**, and the two Test samples are from different production lots



Triangle to Tetrad





DOD to Tetrad



Theory is great, but let's be PRACTICAL!

Would it work for **both homogeneous and heterogeneous or variable** products and therefore **replace both Triangle and DOD** testing at General Mills?

What's the **right risk profile for our business?**

Comparisons using GMI products showed Tetrad to be more sensitive in practice than Triangle or DOD

Tetrad vs. Triangle

10 comparisons across product categories (homogeneous products)



Tetrad showed a higher proportion of correct responses:

8 out of **10** times

Tetrad vs. DOD

13 comparisons across product categories (heterogeneous or variable products)



Tetrad and DOD reached the same conclusion:

10 out of **13** times

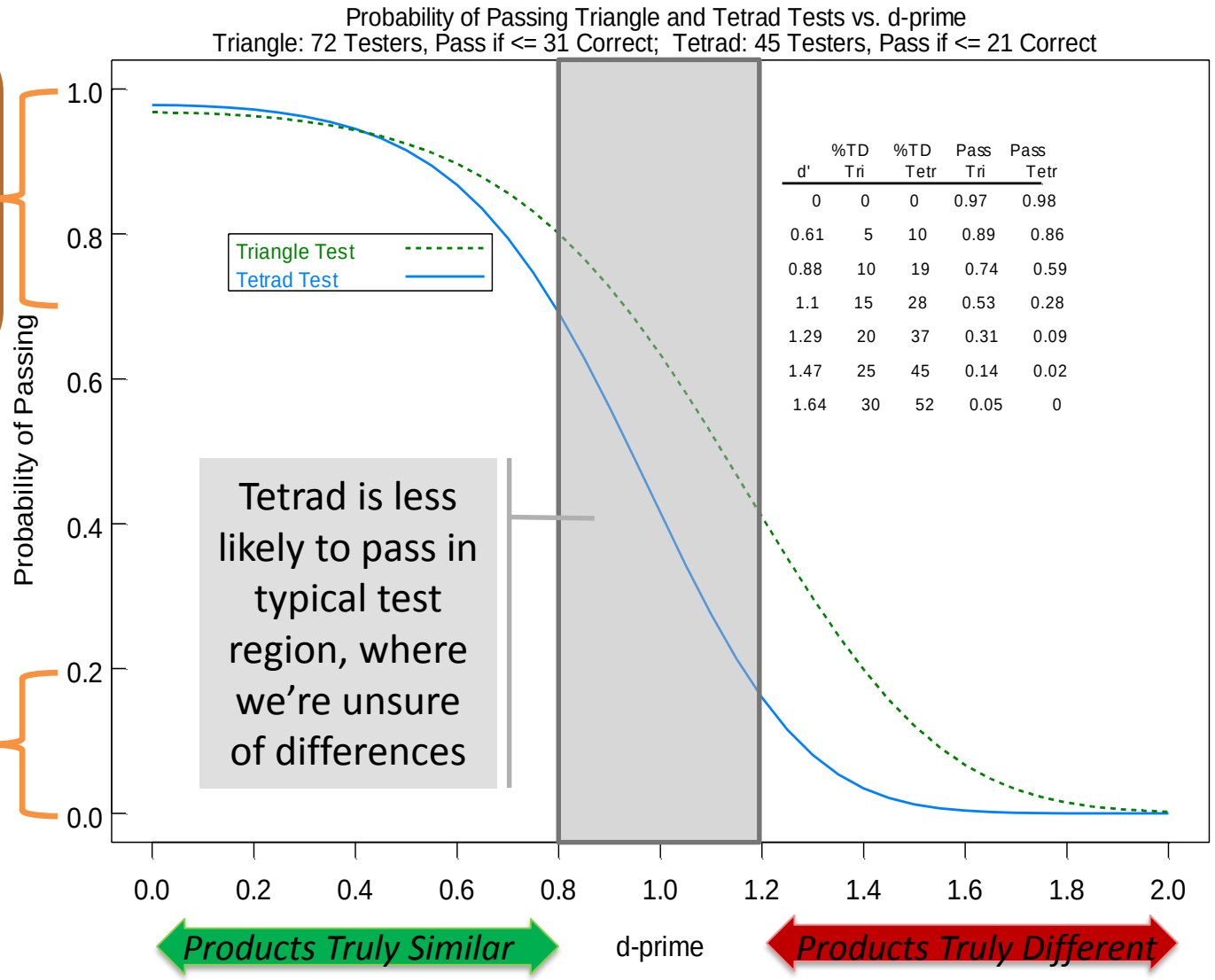
Tetrad reached a **more** conservative conclusion:

3 out of **13** times

We can manage risk to the business by carefully choosing the number of judges and decision criteria

Tetrad and Triangle show similar probability in sending R&D back to reformulate when they don't need to. (type I error)*

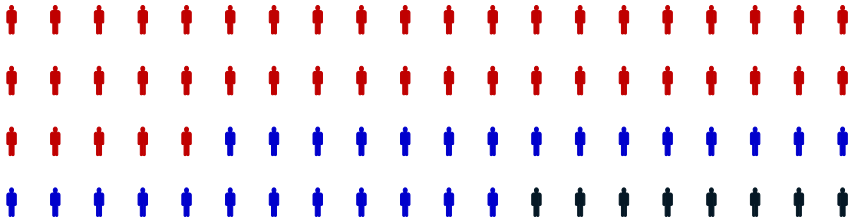
Tetrad is less likely to pass a change when it should have failed. (type II error)



d'	%TD Tri	%TD Tetr	Pass Tri	Pass Tetr
0	0	0	0.97	0.98
0.61	5	10	0.89	0.86
0.88	10	19	0.74	0.59
1.1	15	28	0.53	0.28
1.29	20	37	0.31	0.09
1.47	25	45	0.14	0.02
1.64	30	52	0.05	0

Practically speaking, it's a win-win-win

Improved Risk Profile with Fewer Panelists…



Tetrad n= 45 **Triangle n= 72** **DOD n= 80**

… lower quantities of product required from R&D...

DOD with N=80 required:	Triangle with N=72 required:	Tetrad with N=45 requires:
480	**216**	**180**
Total servings	Total servings	Total servings

… and simplicity and efficiency for the Discrimination Lab

- Quicker Test set-up and fewer serving staff required
- Faster completion due to fewer testers needed
- Enables increased capacity (more testing slots available each week)



GENERAL MILLS

PGI Product Guidance and Insights

THANK YOU!