

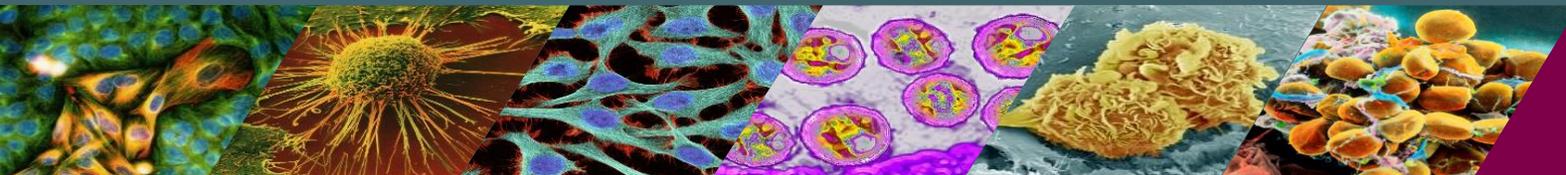
# Confessions of a Reformed MaxDiff Zealot

## *The Road from Advocacy to Absolution*

Stephan Benzekri, *Senior Group Director*

Debra A. Kossman, Ph.D., *Senior Vice President*

March 21, 2017



# Our Presenters



**Stephan Benzekri**

*Senior Group Director*

[sbenzekri@naxionthinking.com](mailto:sbenzekri@naxionthinking.com)

215.496.6835



**Debbie Kossman, Ph.D.**

*Senior Vice President*

[dkossman@naxionthinking.com](mailto:dkossman@naxionthinking.com)

215.496.6846



**Confession is  
cathartic**



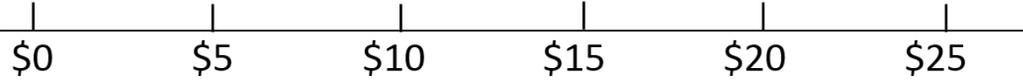
**Our confession today relates to ...**

# MaxDiff

**... a tool for  
differentiation  
that some of  
you may have  
experience  
with**



# But first, a few comments on measurement theory ...

Frequency	 <p>14 PEACH SINGLES GUM COOL MINT POTENTIALLY FLAVORED WITH Recaldent MINT GUMMED™</p>  <p>It's a fact! 4 out of 5 dentists surveyed would recommend Trident for their patients who chew gum</p>
Order	<p>IRV Ballot</p> <p>Alice <input type="text" value="2"/></p> <p>Bob <input type="text" value="1"/></p> <p>Charles <input type="text" value="3"/></p> 
Valence	  
Intensity	<p>Willingness to Pay</p>  <p>\$0 \$5 \$10 \$15 \$20 \$25</p>

# And is neutrality an opinion or a cop-out?

- Are people ever truly indifferent?

or ...

- Do they sometimes lack self-awareness about their own proclivities and preferences?

**Are there advantages to marketers if subtle differences in preferences are measured in a more granular way?**



# Common measurement tools in the market researcher's armamentarium



Tool	Pros	Cons
<b>Rating</b> (ties allowed)	Simple, univariate  Measures intensity & valence	Allows ties (doesn't always clearly ordinate preference)  Often lacks differentiation in results
<b>Ranking</b> (ties <u>not</u> allowed)	Simple, univariate  Provides ordinal data	Does <u>NOT</u> measure intensity or valence  Burdensome if many items to evaluate

# Common measurement tools ...



Tool	Pros	Cons
<b>Derived Importance</b>	Derives preference/importance Measures intensity as well as rank	Cannot distinguish empirical co-occurrence from causal relationships
<b>Conjoint or Discrete Choice</b>	Derives preference/importance Measures intensity as well as rank Aims to mimic real-world heuristics by placing attributes in context Modeling power/versatility	Advanced analytics required adding time, complexity and cost Limitation on # of attributes/levels that can be modeled (respondent burden) Attribute “levels” must be operationalized

# The ubiquitous survey battery of items to evaluate

Product Attribute	Not At All Important							Extremely Important
	1	2	3	4	5	6	7	
Overall survival (OS)								
Progression-free survival (PFS)								
Time-to-progression (TTP)								
Response rate (RR)								
Rate of immunologic side effects								
Rate of cytotoxic side effects								
Appropriate for all patients								
Appropriate only for patients w/relevant marker								
Requires pre-treatment antibody testing								
Requires pre-treatment mutation testing								
Route of administration								
Dosing schedule								
Out-of-pocket costs for patients								
Requires prior authorization from payors								
Preferred product on hospital formulary								



## Example:

Rating of attributes driving selection of treatment for non-small cell lung cancer (NSCLC)



# But ratings produce many ties, even in the aggregate ...

Product Attribute		Attribute Importance	
		Individual Level	Aggregate Level
1	Overall survival	7	6.8
2	Progression-free survival	7	6.6
3	Time-to-progression	7	6.3
4	Response rate	6	6.3
5	Rate of immunologic side effects	6	6.3
6	Rate of cytotoxic side effects	6	5.9
7	Appropriate for all patients	5	5.5
8	Appropriate only for patients w/relevant marker	5	5.5
9	Requires pre-treatment antibody testing	5	4.9
10	Requires pre-treatment mutation testing	4	4.5
11	Route of administration	4	4.1
12	Dosing schedule	4	4.1
13	Out-of-pocket costs for patients	3	4.0
14	Requires prior authorization from payors	3	3.3
15	Preferred product on hospital formulary	3	2.8



**In sports or market research, ties are frustrating**

# The challenge of tie-breakers ...

Product Attribute	
1	Overall survival
2	Progression-free survival <b>1</b>
3	Time-to-progression
4	Response rate <b>2</b>
5	Rate of immunologic side effects
6	Rate of cytotoxic side effects <b>3</b>
7	Appropriate for all patients
8	Appropriate only for pts w/relevant marker
9	Requires pre-treatment antibody testing
10	Requires pre-treatment mutation testing
11	Route of administration
12	Dosing schedule
13	Out-of-pocket costs for patients
14	Requires prior authorization from payors
15	Preferred product on hospital formulary

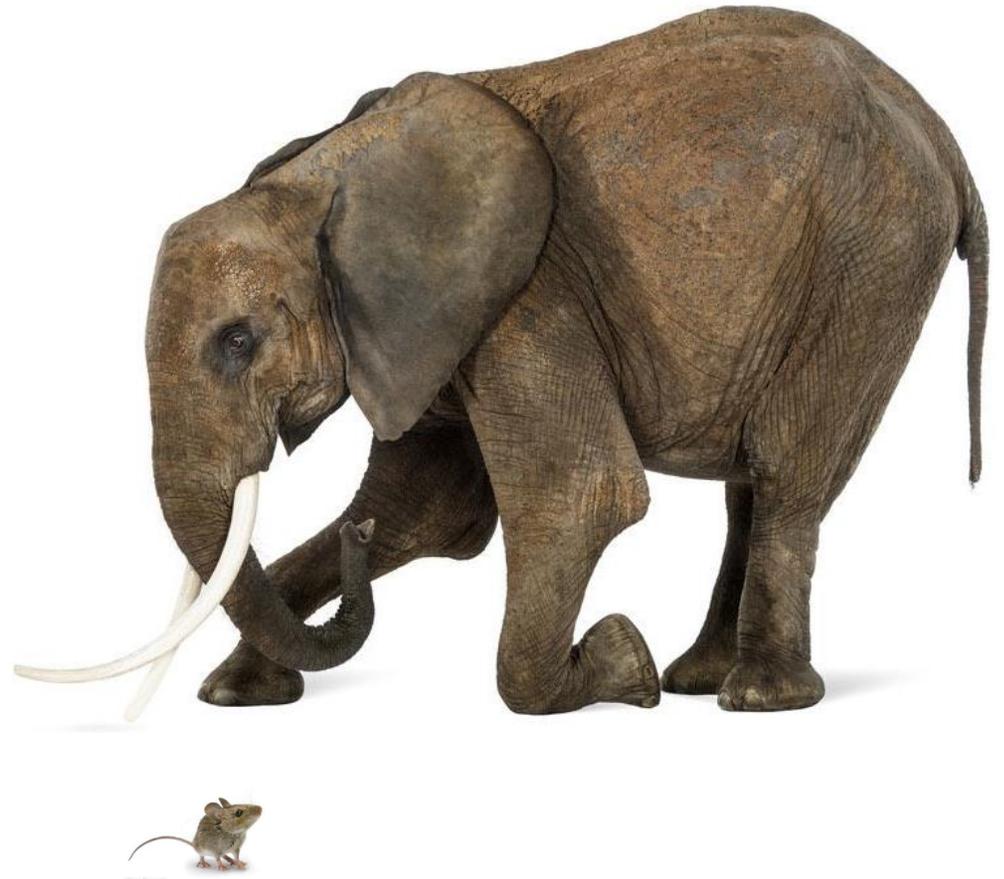
We can ask for rank ... but ranking 10+ items is tedious and time-consuming ...



... and still deprives us intensity or **distance**

# Enter **MaxDiff** ...

- A method for estimating preferences
- Seeks to **Maximize Differentiation** among items being evaluated
- Forces choice even where no true preference
- An established technique, but 21<sup>st</sup> century brought to wider use



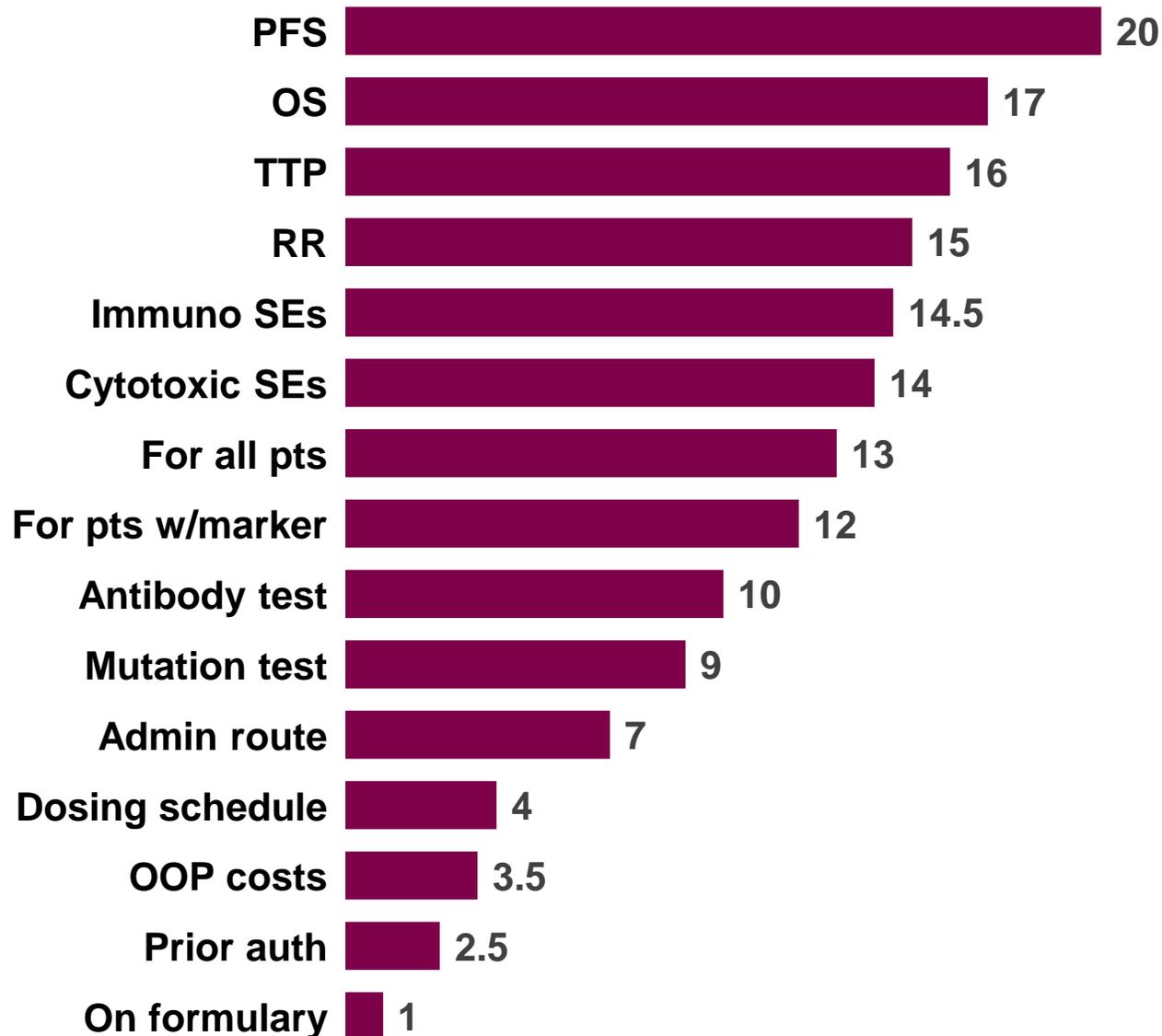
# The MaxDiff Task

## *Sample Scenario*

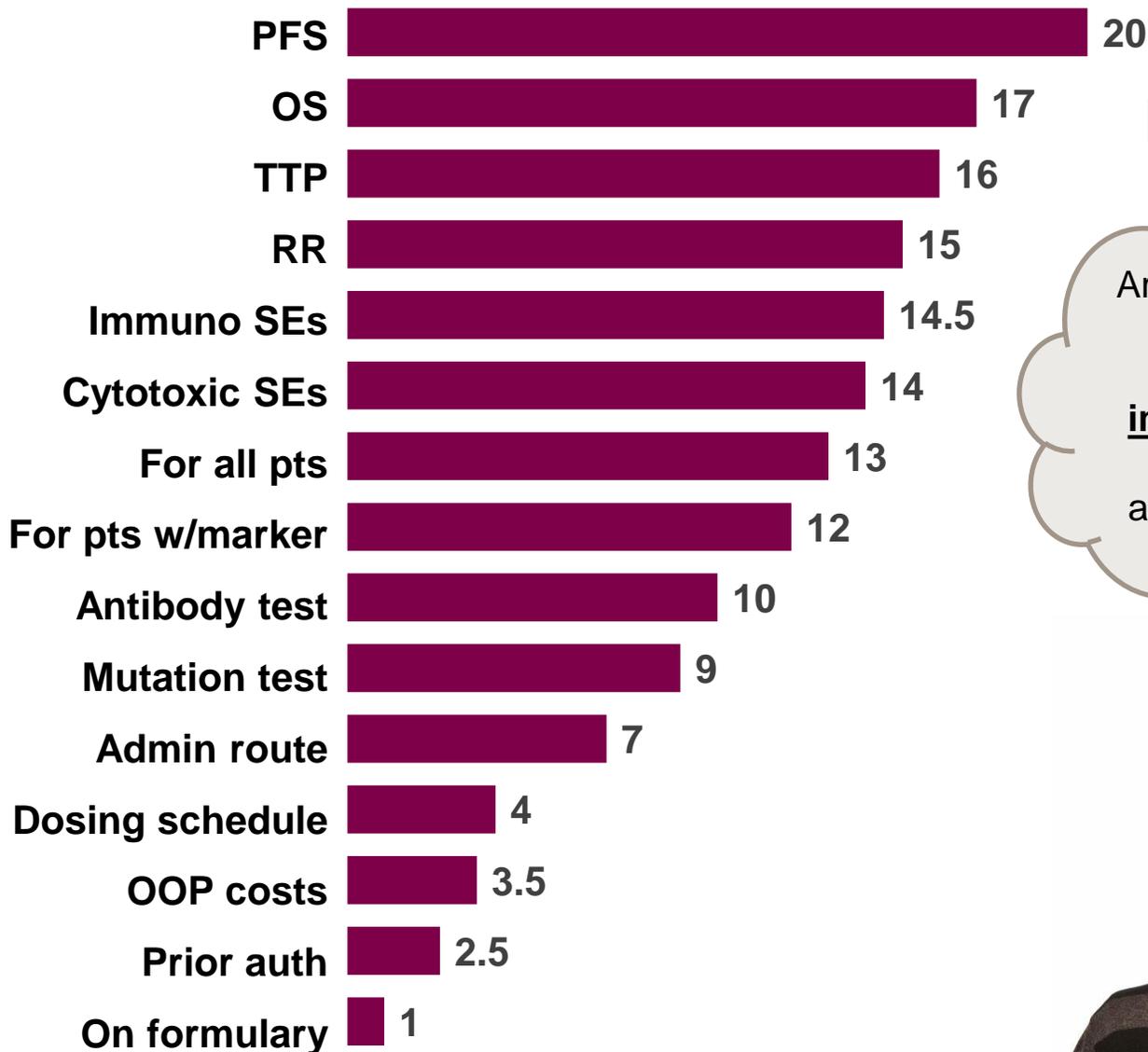
Least Important	Product Attribute	Most Important
<input type="radio"/>	Overall survival	<input type="radio"/>
<input type="radio"/>	Dosing schedule	<input type="radio"/>
<input type="radio"/>	Time-to-progression	<input type="radio"/>
<input type="radio"/>	Rate of cytotoxic side effects	<input type="radio"/>

Respondents choose the **most** and **least** important in a series of random subsets of attributes

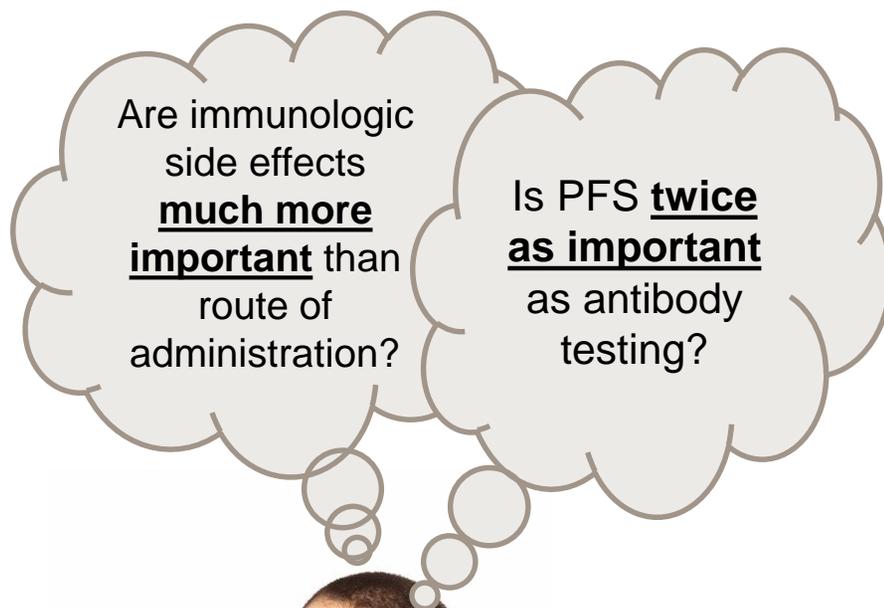
# The **MaxDiff** model generates so-called “utilities” for each attribute ... with no ties



# What can we say based on MaxDiff utilities?



It's tempting to ask ...



# To assess whether these questions can be answered with **MaxDiff** utilities, we ran an experiment:

## Objective:

Determine if MaxDiff conserves both rank and intensity of underlying preferences

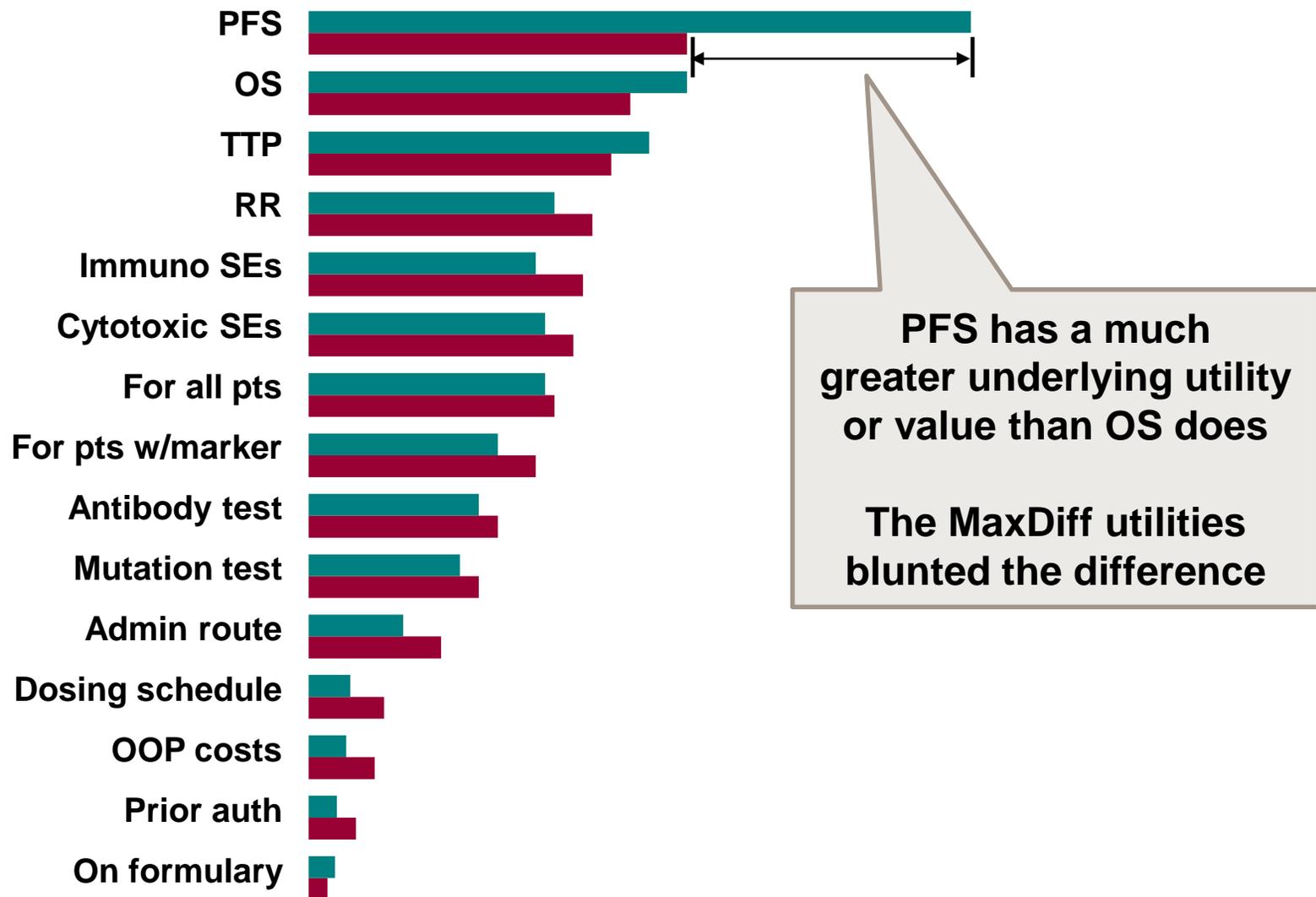
## Procedure:

1. We pre-assigned an importance value or “utility” to each attribute for “n” respondents resulting in a decisive “winner”: PFS
2. The underlying utilities were used to simulate the choices respondents would make when completing a MaxDiff exercise
3. The generated MaxDiff utilities were then compared to the pre-assigned underlying utilities

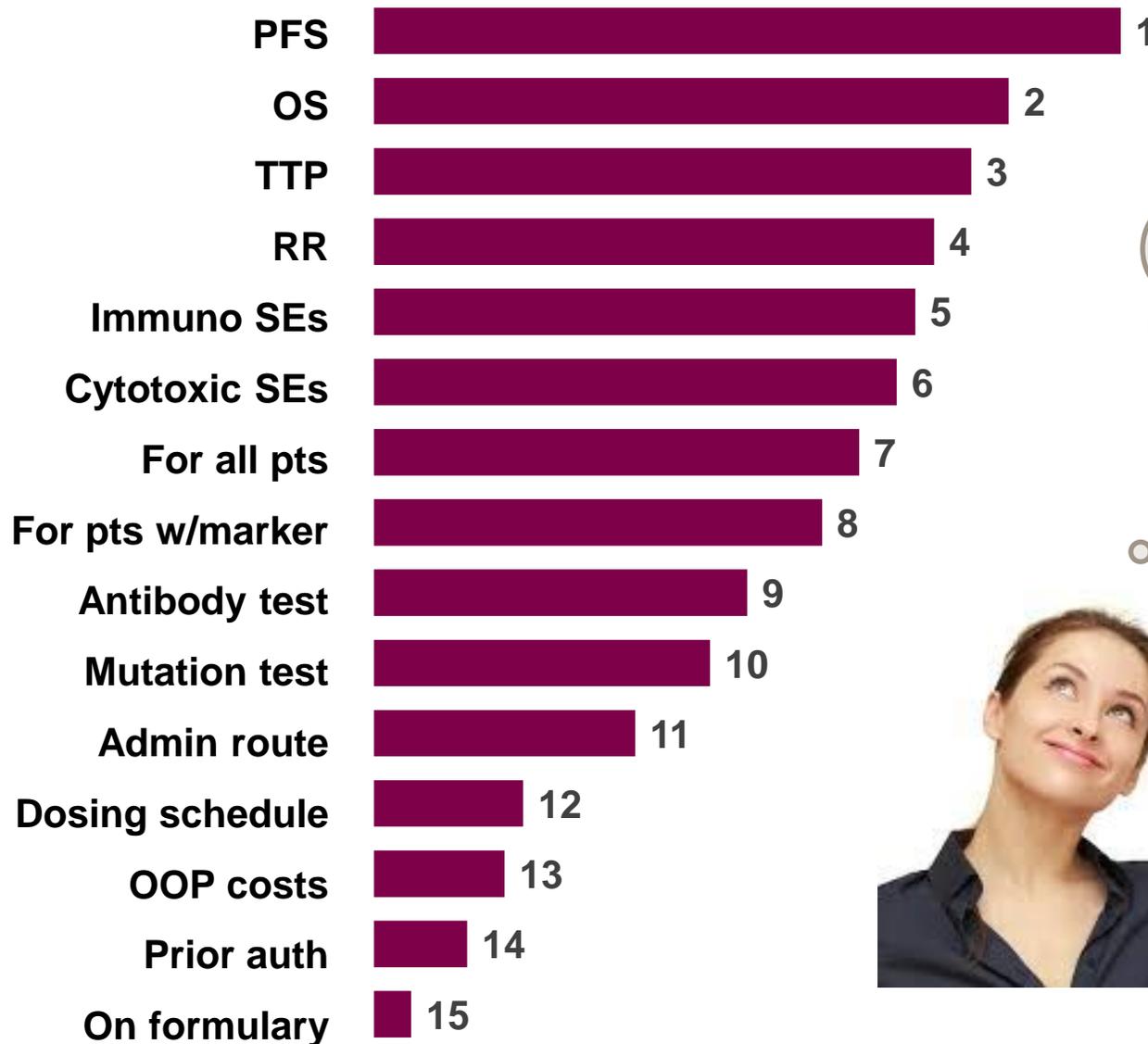


# Attribute order is preserved faithfully ... but distance is not

## Underlying Utilities vs. MaxDiff Utilities



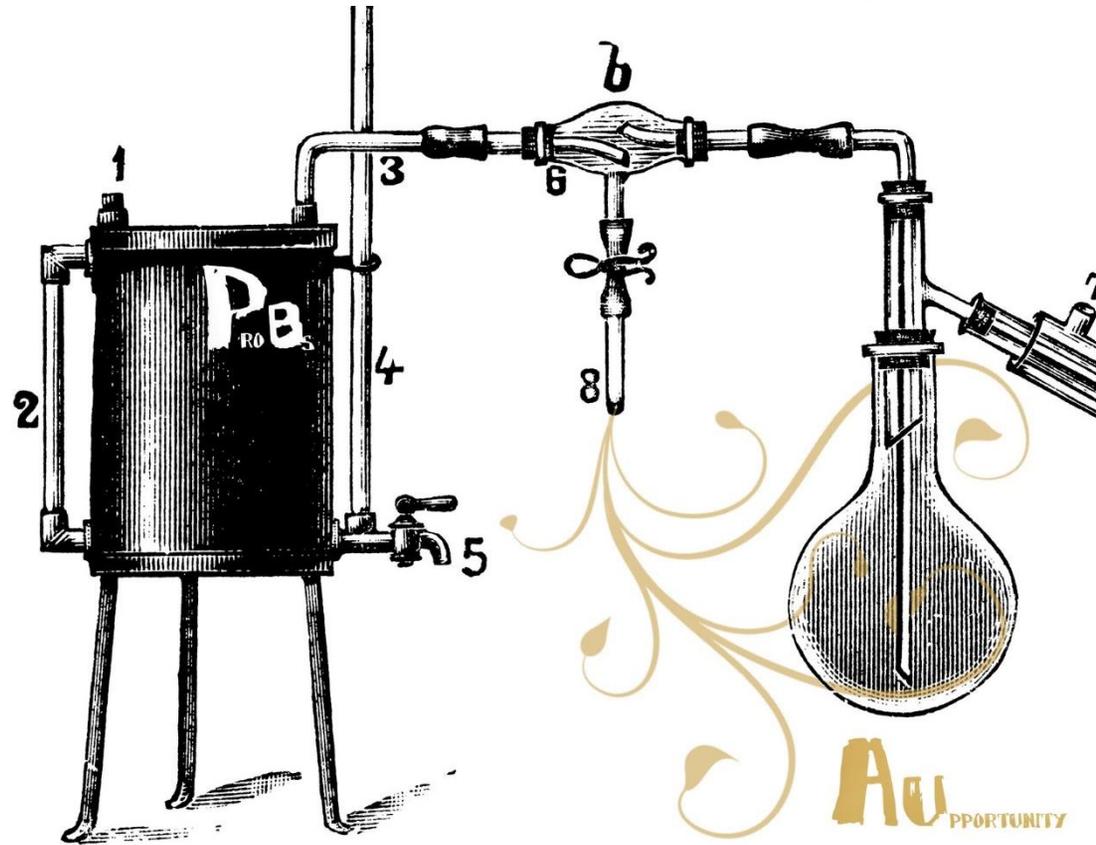
# What are the implications for interpreting **MaxDiff** results?



PFS is the most important thing here!



# MaxDiff models transform inputs in ways that lend them more luster than they deserve



## Two very different “close calls” on preference ...



vs.



44%

56%

Marginal preference  
**High** substitutability



vs.



46%

48%

Highly polarized  
**Low** substitutability

**MaxDiff** utilities cannot distinguish these scenarios

To summarize, **MaxDiff** does some things reliably



It does a good job of ranking many items on a dimension of evaluation



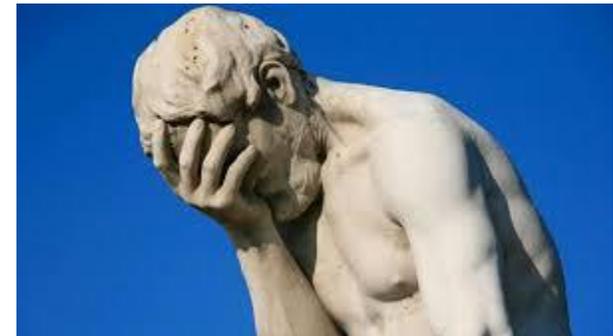
It can provide greater differentiation among attributes than ratings alone



But it CANNOT generate valid estimates of the relative intensity of preferences among items

# Beware the sin of over-interpreting **MaxDiff** results!

- ... it forces choice even where no true preference exists**
- ... in exchange for avoiding ties, it deprives us of real insight about intensity/distance and valence – a very high price to pay**
- ... it diverts attention from other tools and strategies that help train a sharper lens on closely tied attributes**



# Questions?



**Stephan Benzekri**

*Senior Group Director*

[sbenzekri@naxionthinking.com](mailto:sbenzekri@naxionthinking.com)

215.496.6835



**Debbie Kossman, Ph.D.**

*Senior Vice President*

[dkossman@naxionthinking.com](mailto:dkossman@naxionthinking.com)

215.496.6846

