Sawtooth Software

The survey software of choice

An Introduction to Max Diff

Agenda

- Why choose MaxDiff?
- Designing a MaxDiff experiment
- SSI Web Demo
- Analyzing MaxDiff Data
- Additional Analyses
- Conclusions

SECTION 1

Why Choose Max Diff?

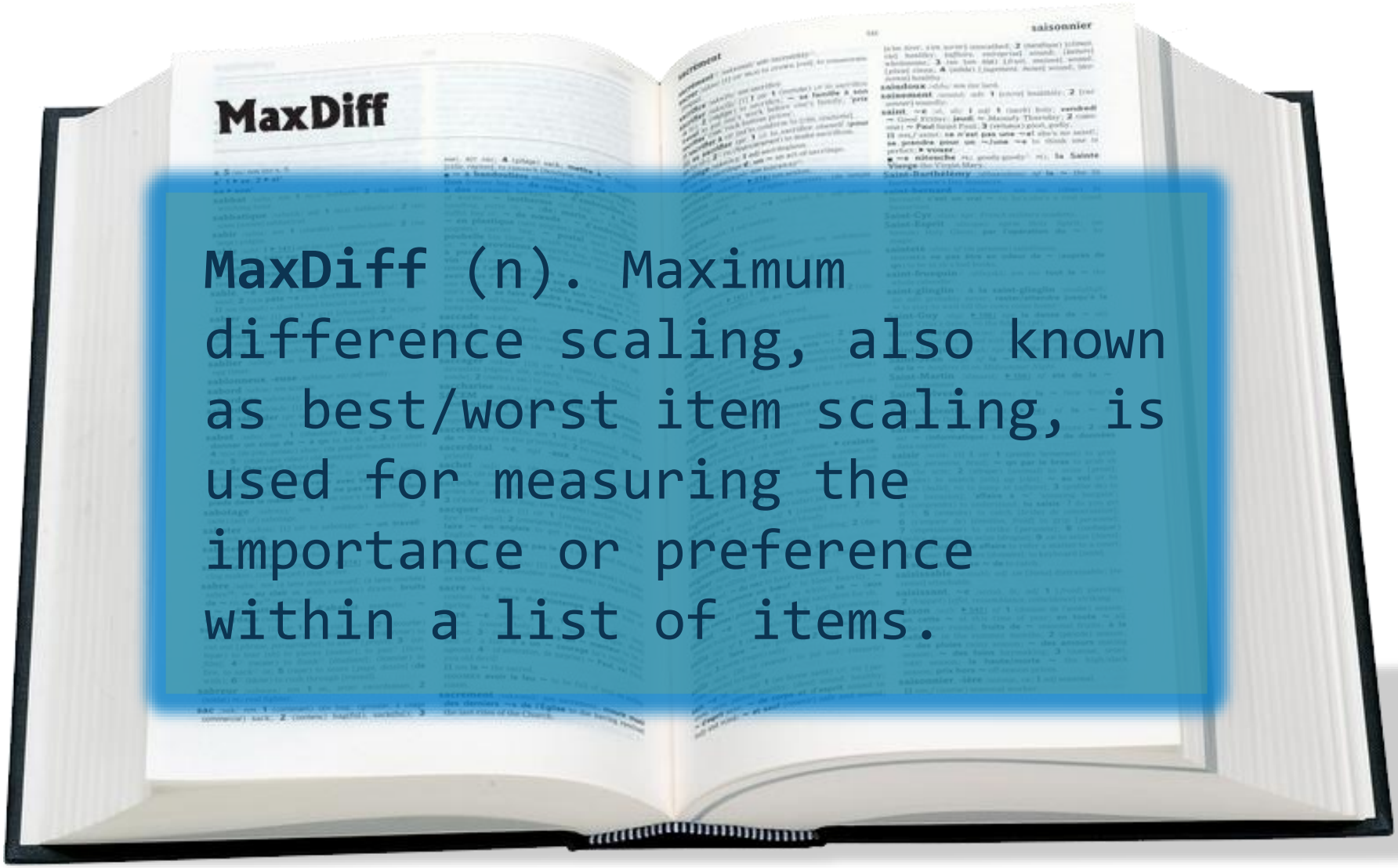
The Ubiquitous Monadic Rating Scale

How important are each of these factors when considering a vehicle for purchase?

	Not important 1	2	3	4	Extremely Important 5
It is luxurious	●	○	○	○	○
It offers a smooth ride	○	○	○	●	○
I get good value for the money	○	●	○	○	○
It features the latest high-tech gadgets	○	○	●	○	○

Possible Solutions

- Ranking exercises
  - Becomes impractical with >7 items
  - Ordinal scale results
- Constant-Sum Allocation
  - Becomes impractical with >7 items
  - Allocation may not be independent
  - Making answers to sum to a particular value is difficult


MaxDiff
MaxDiff (n). Maximum difference scaling, also known as best/worst item scaling, is used for measuring the importance or preference within a list of items.

Quick Overview

- Roughly comparable to a One-Attribute, multi-level CBC
- Respondents typically shown between 2-6 items at a time, asked to indicate which is best and/or which is worst
- Task is repeated many times, showing a different set of items in each task
- Resulting model provides ratio-scaled scores, or utilities, for each item

Jordan J. Louviere, PhD

Inventor of the MaxDiff technique

1944—

Dr. Louviere is currently the Executive Director of the *Centre for the Study of Choice* at the University of Technology in Sydney, Australia

MaxDiff Tutorial Example

Imagine you were making the decision to reenlist in the Navy today.
Which of the following four factors would make you most and least likely to want to reenlist?

(1 of 20)

Makes me **Most** want to reenlist		Makes me **Least** want to reenlist
○	$50 per month pay increase when out at sea	●
○	3-year reenlistment obligation	○
●	Increased reenlistment bonus ($5,000)	○
○	Live in 2-person barracks when in port	○

0% 100%

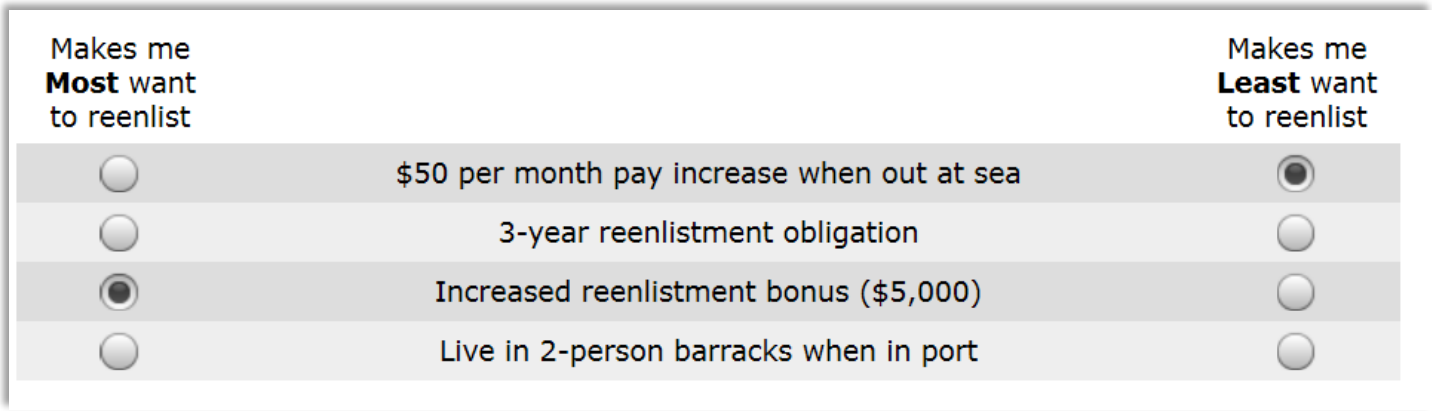
MaxDiff Tutorial Example

Imagine you were making the decision to reenlist in the Navy today.
Which of the following four factors would make you most and least likely to want to reenlist?

(5 of 20)

Makes me **Most** want to reenlist		Makes me **Least** want to reenlist
○	Three-quarters of on-job time using your skills and training	○
○	5-year reenlistment obligation	●
●	Choice between living in 2-person barracks or 4-person barracks when in port	○
○	Both location and duty guarantees for next assignment	○

0% 100%



Makes me **Most** want to reenlist		Makes me **Least** want to reenlist
○	$50 per month pay increase when out at sea	●
○	3-year reenlistment obligation	○
●	Increased reenlistment bonus ($5,000)	○
○	Live in 2-person barracks when in port	○

- We then know:
  - Reenlistment Bonus>$50/month increase
  - Reenlistment Bonus>3-yr reenlistment obligation
  - Reenlistment Bonus>2-person barracks
- And…
  - 3-yr reenlistment obligation>$50/month increase
  - 2-person barracks>$50/month increase
- From two “clicks” we learn about 5 of the 6 possible paired comparisons!
  - 3-yr reenlistment obligation ?? 2-person barracks

SECTION 2

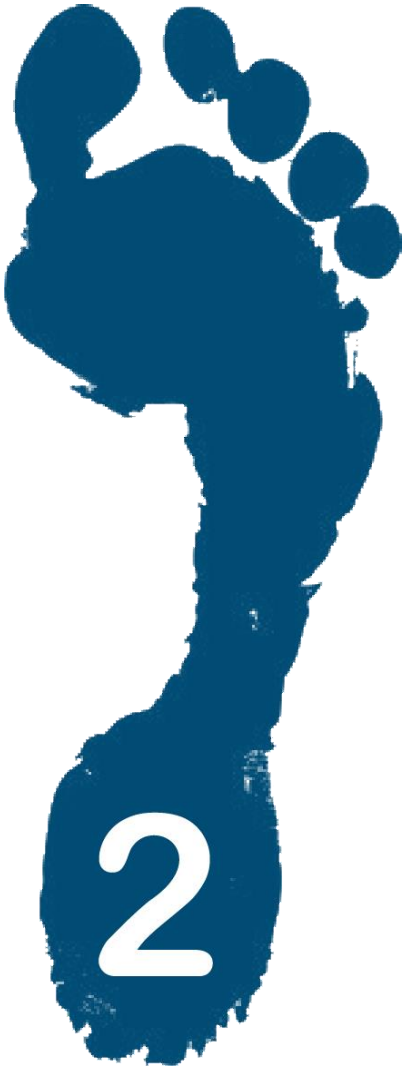
Designing MaxDiff Experiments

STEP 1

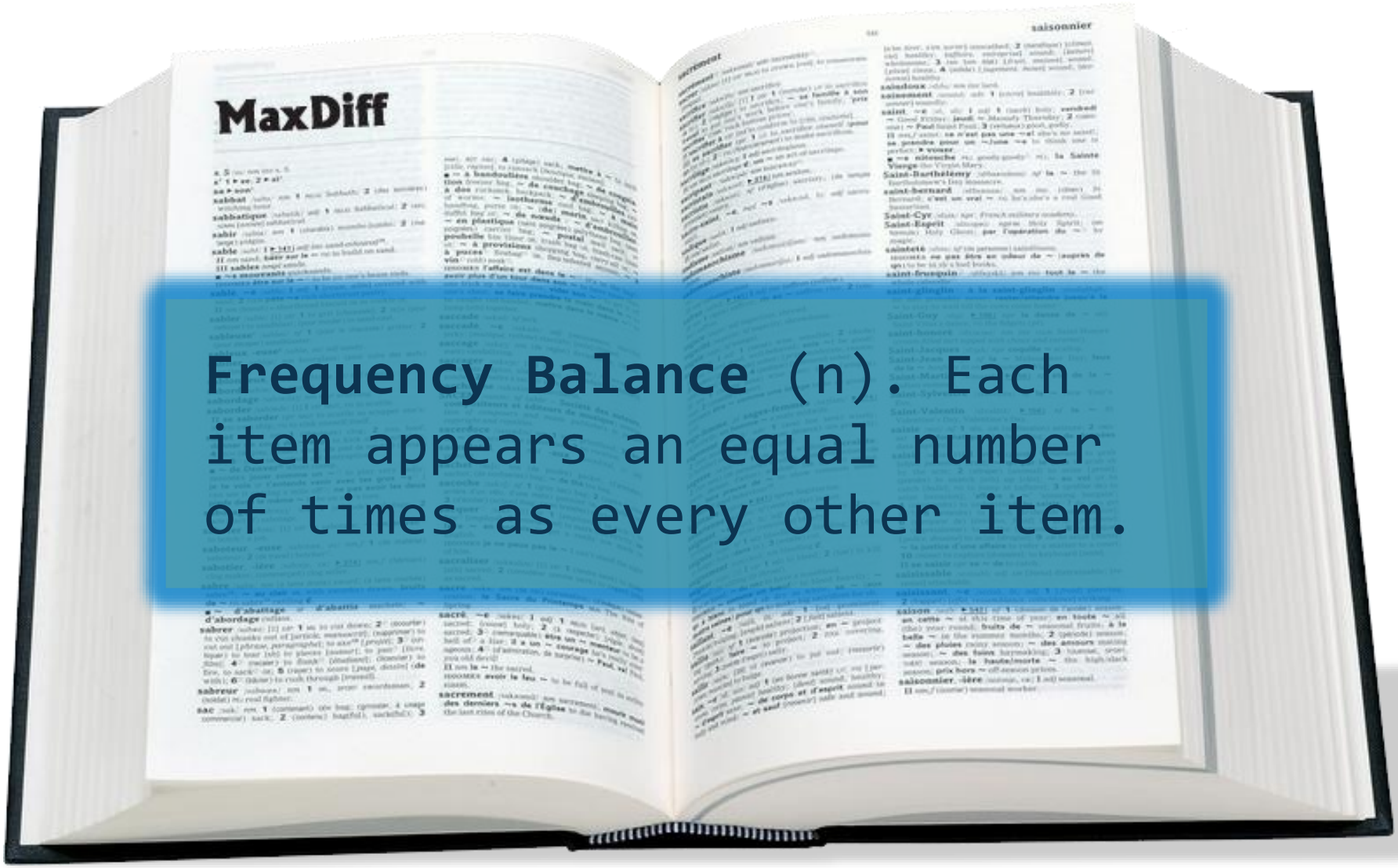
Develop attribute list and possible prohibitions

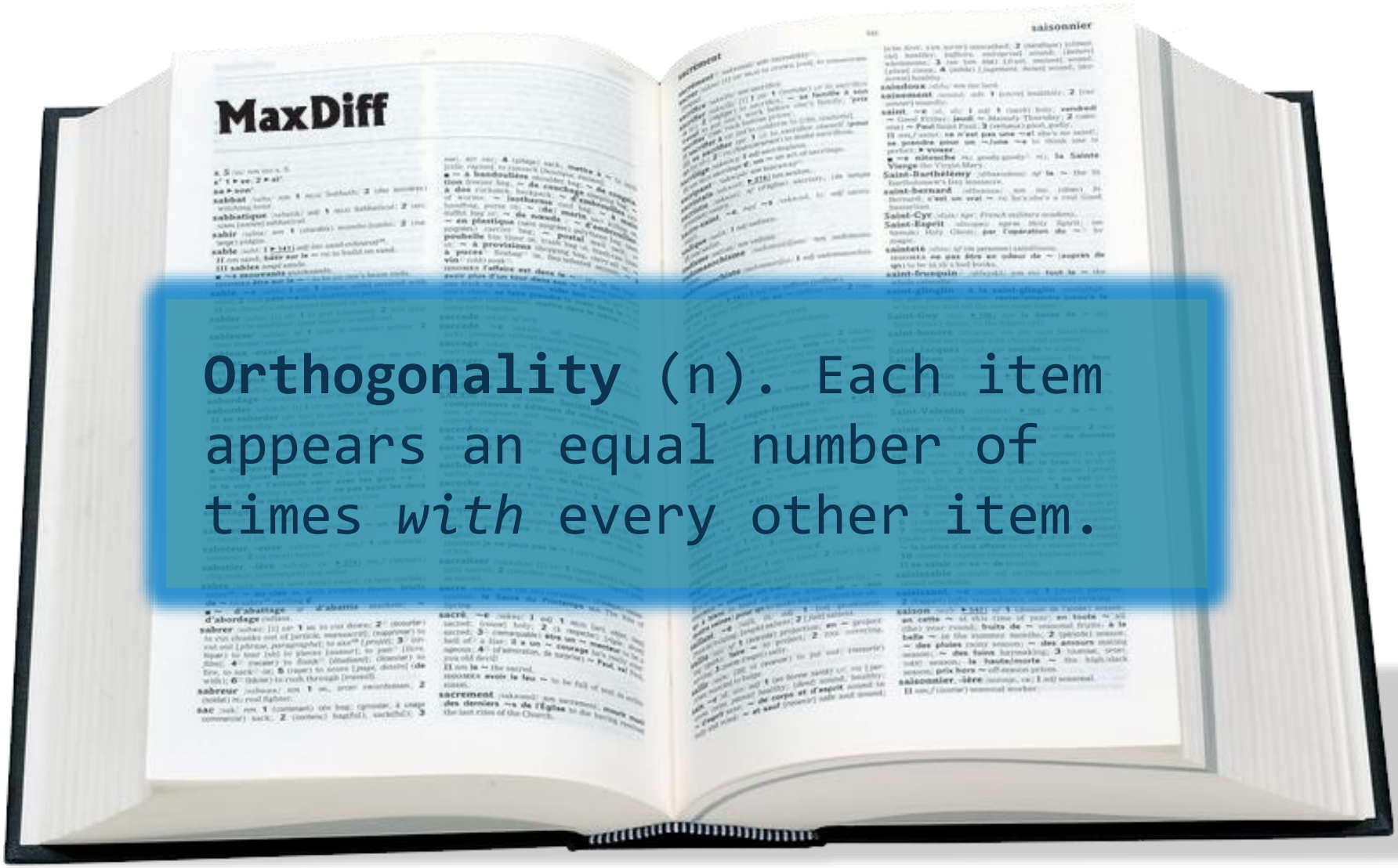
STEP 2

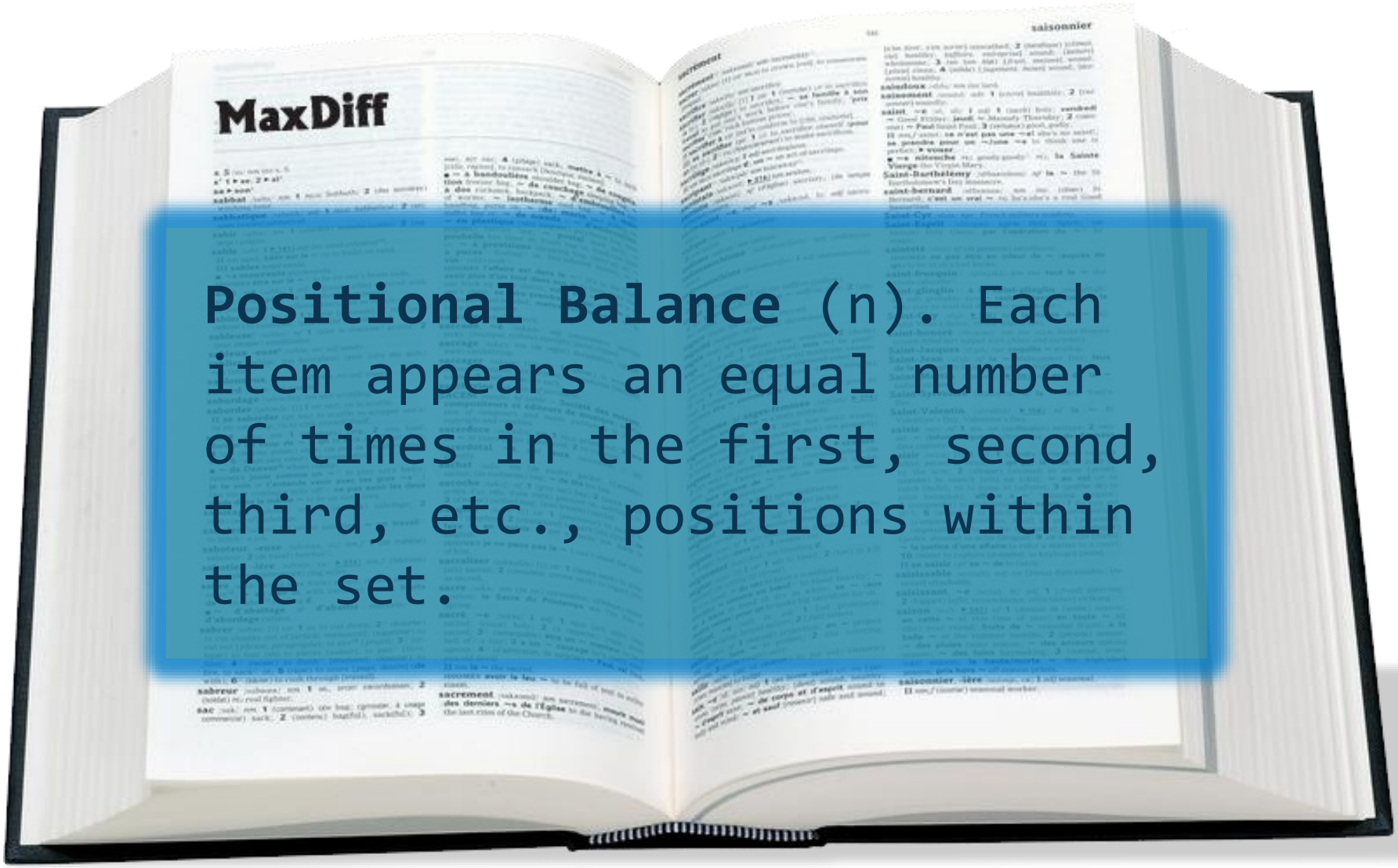
Choose the number of items, sets, and versions

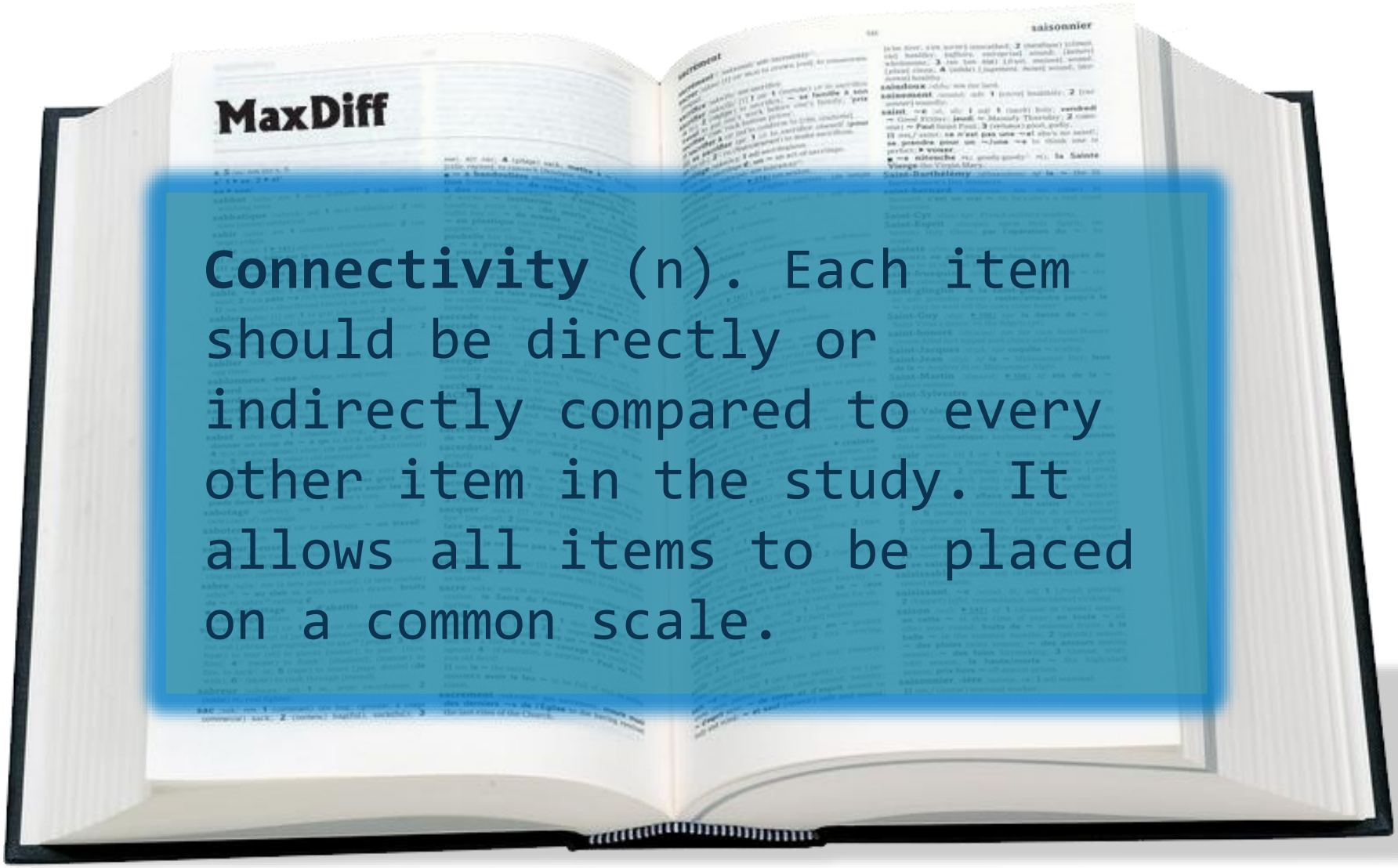


STEP 3
Generate design




MaxDiff
Orthogonality (n). Each item appears an equal number of times *with* every other item.


MaxDiff
Positional Balance (n). Each item appears an equal number of times in the first, second, third, etc., positions within the set.


MaxDiff
Connectivity (n). Each item should be directly or indirectly compared to every other item in the study. It allows all items to be placed on a common scale.

Develop attribute list and prohibitions

- Clean eating areas (floors, tables, and chairs)
- Clean bathrooms
- Has health food items on the menu
- Typical wait time is about 5 minutes in line
- Typical wait time is about 15 minutes in line
- Prices are a good value
- Your order is always completed correctly
- Has a play area for children
- Food tastes wonderful
- Restaurant gives generously to charities

Choose items, sets, and versions

Number of items:

Number of sets:

Number of versions:

Number of Items Per Set

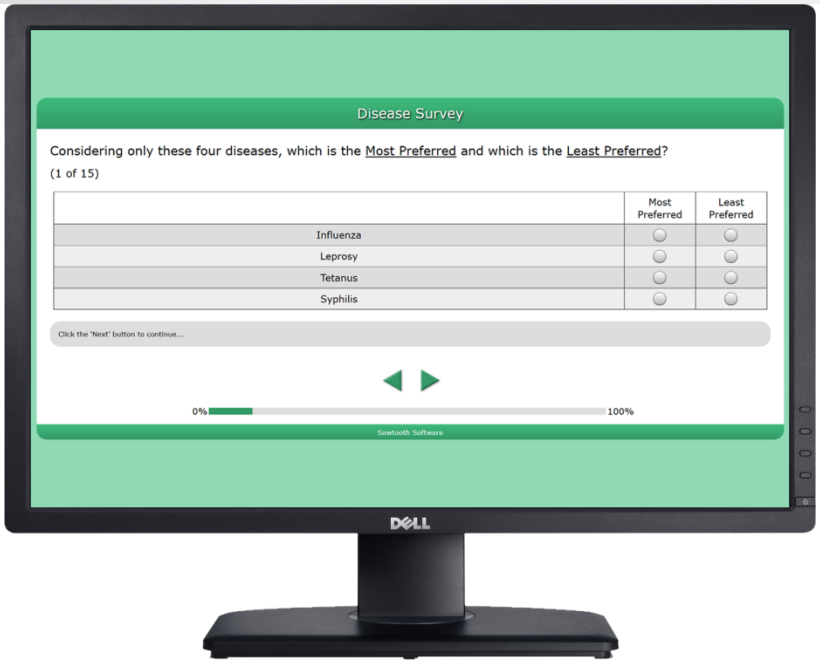
- 4 to 5 items per set
- Don’t show more than half of the total items in a set
- More than 5 items is detrimental

Number of Sets per Respondent

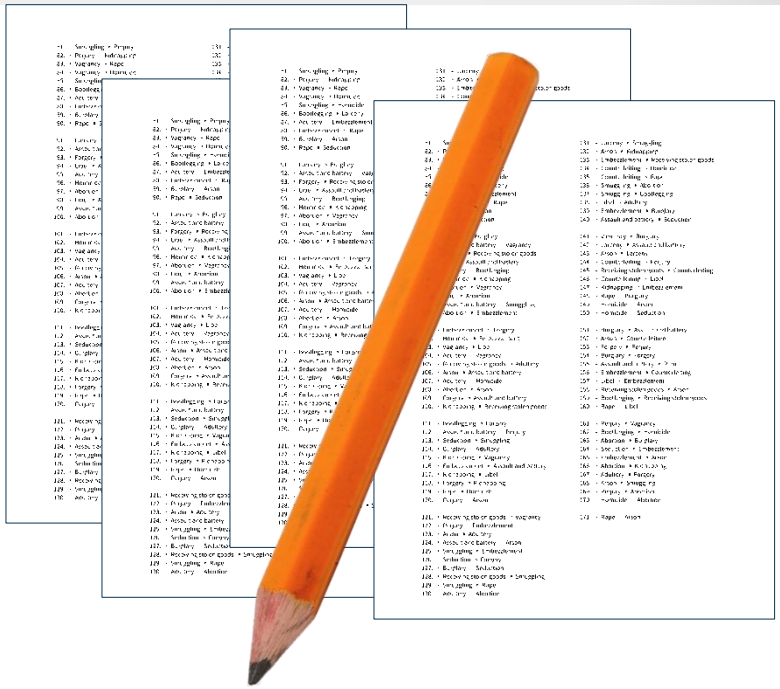
$$\frac{\text{\# tasks X \# items per task}}{\text{Total \# items}} > 3$$

$$\frac{?? \times 4}{10} > 3$$

Number of Versions/Blocks



300



4

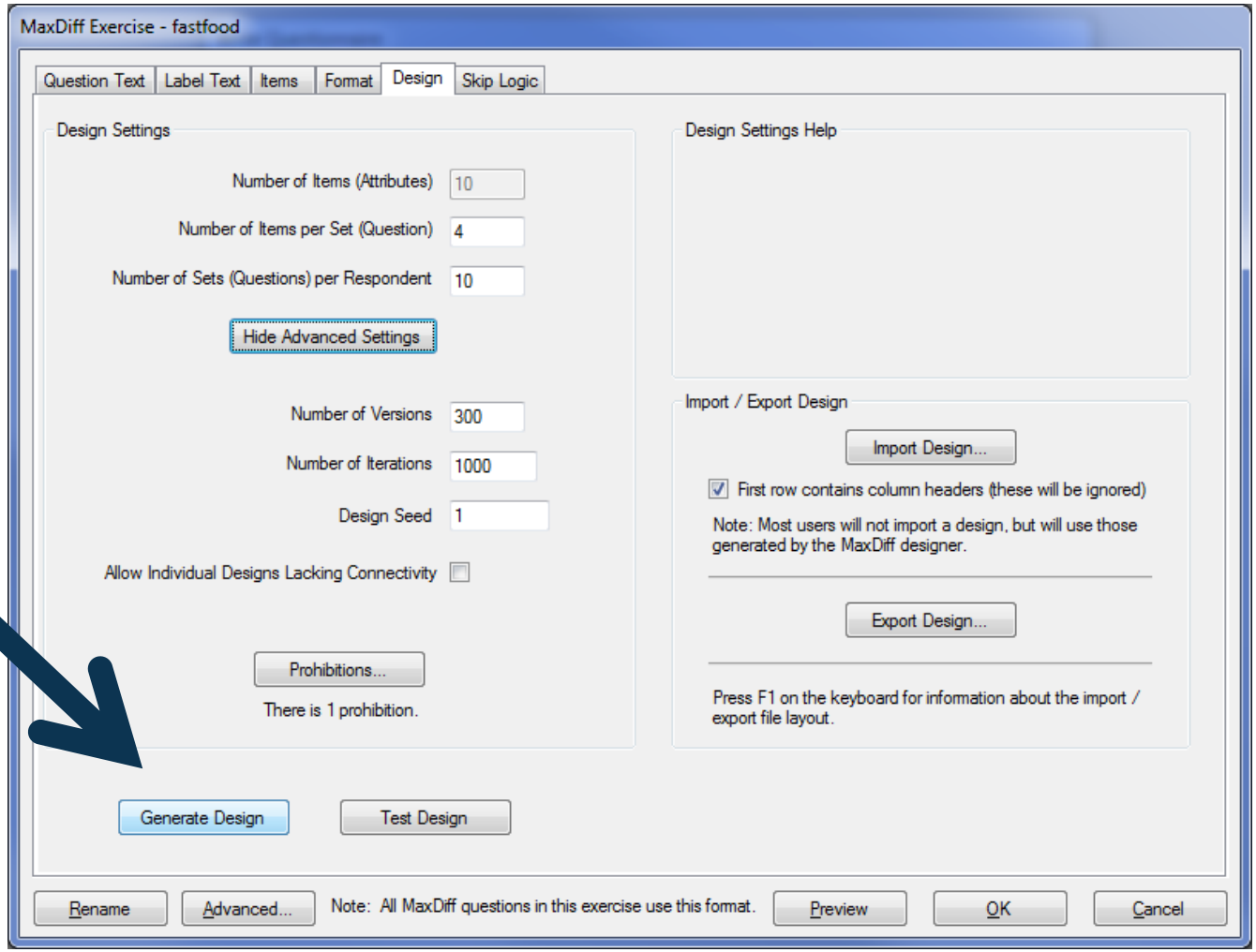
Choose items, sets, and versions

Number of items per set: **4**

Number of sets: **10**

Number of versions: **300**

Generate design



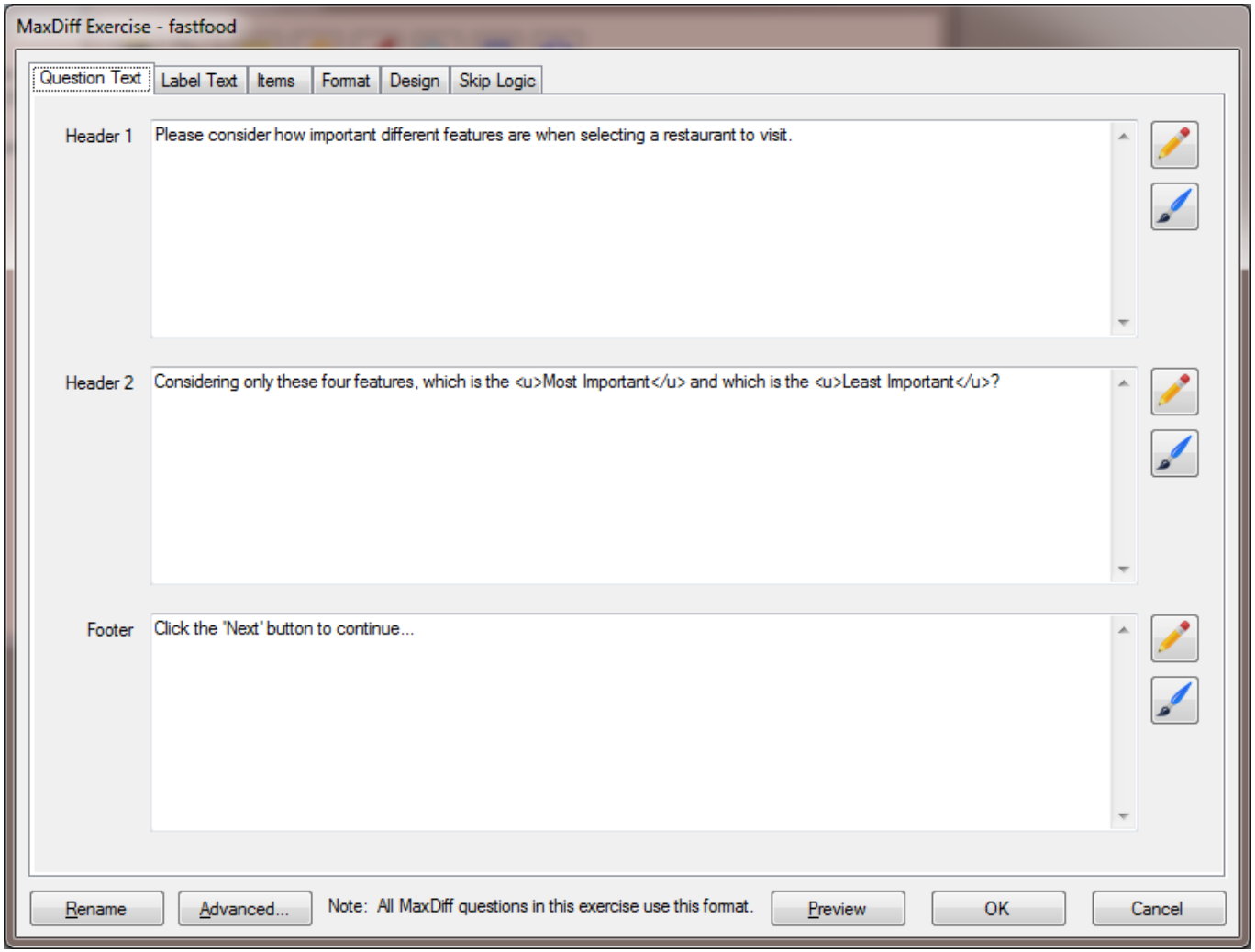
How the Design Program Works

- Design algorithm similar to CBC. Algorithm "builds up" the tasks, seeking one- and two-way frequency balance. Then, it "swaps" item positions on cards for positional balance.
- The design process is repeated 1000 separate times, and the replication that demonstrates the best one-way balance is selected.
  - If multiple designs have the same degree of one-way balance, then we select among those designs based on the best two-way balance.
  - If multiple designs have the same degree of one-way and two-way balance, then we select among those designs based on the best positional balance.

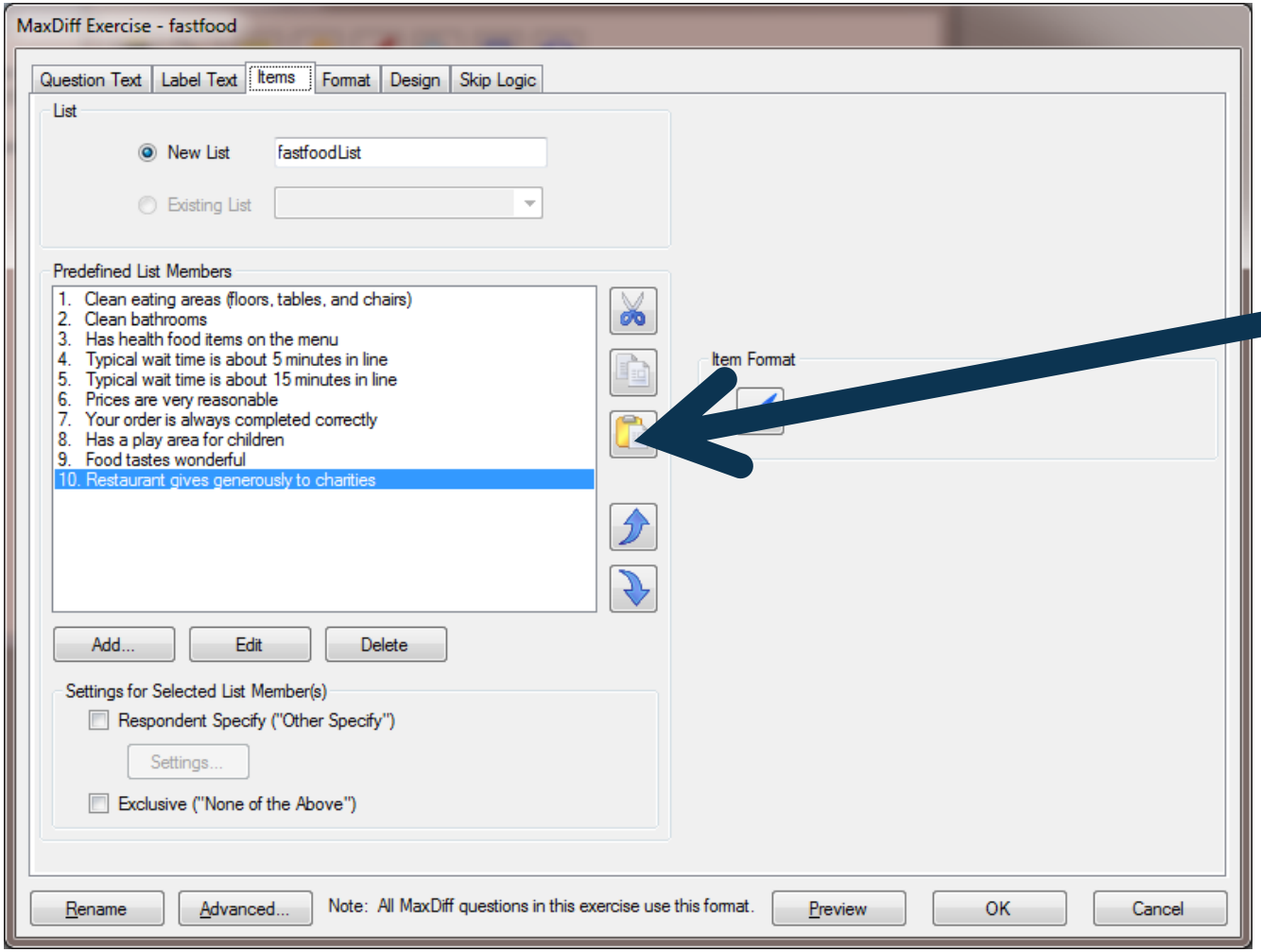
SECTION 3

SSI Web Demo

1. Enter question text

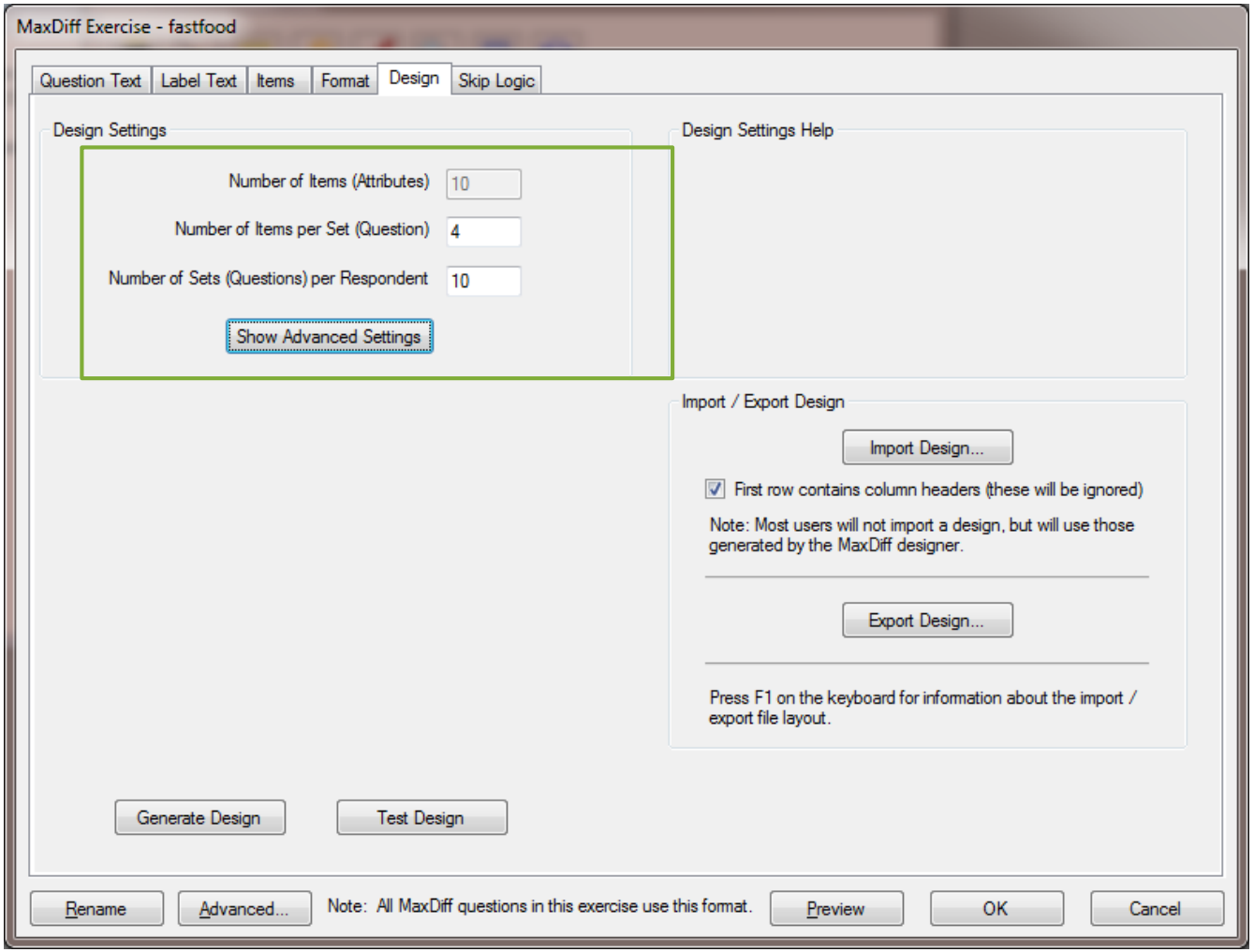


2. Paste items from document

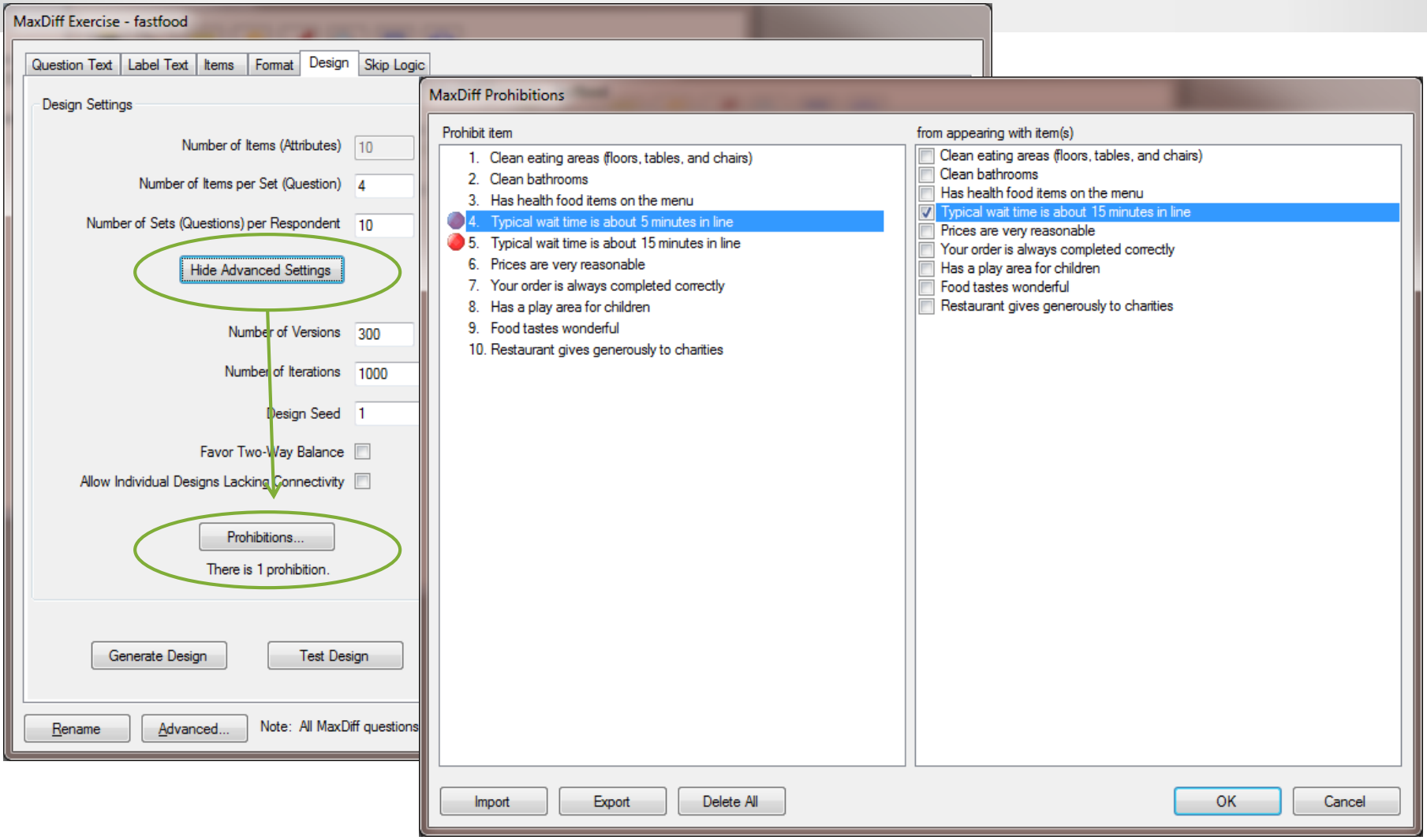


Click this button to paste items

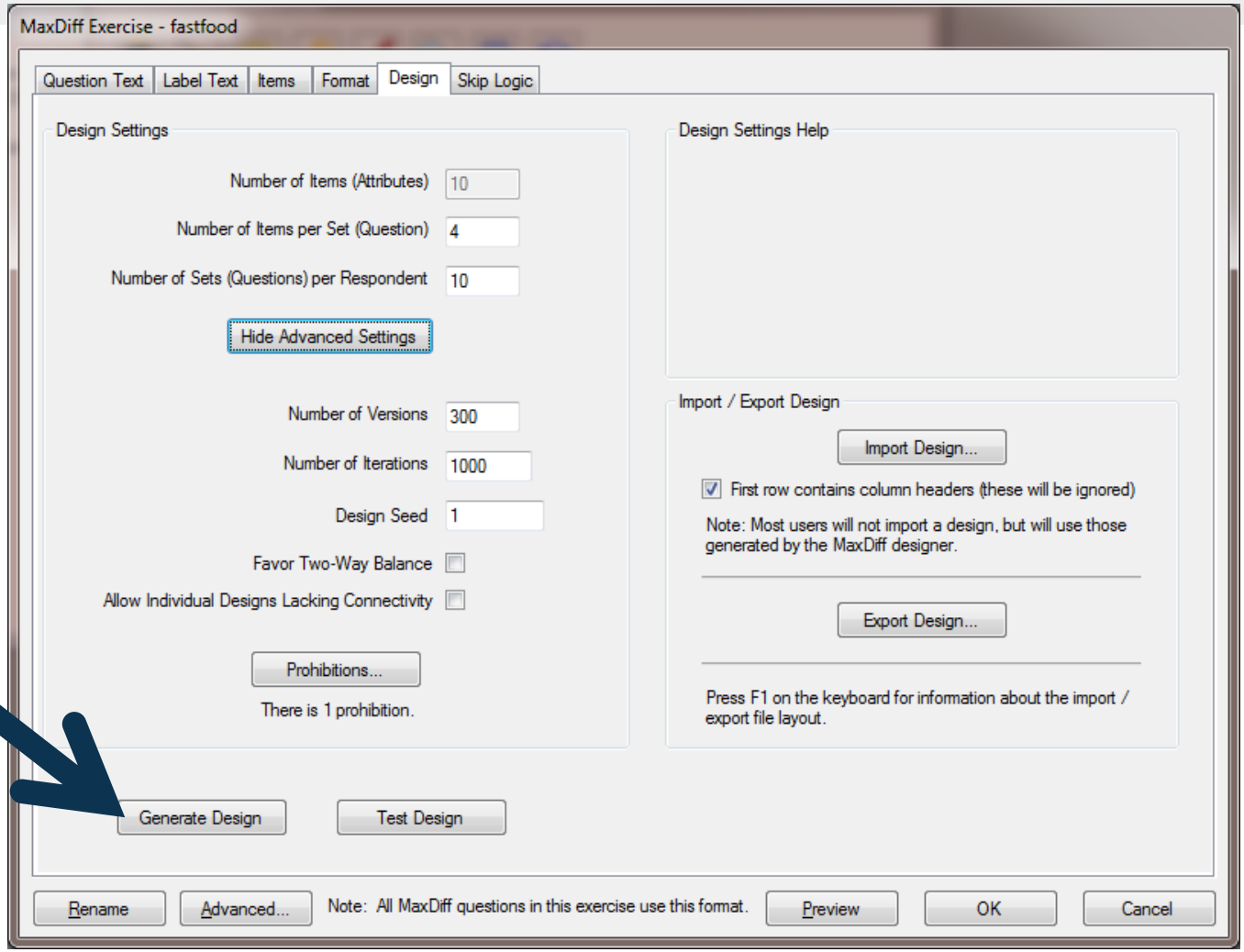
3. Type in your design numbers



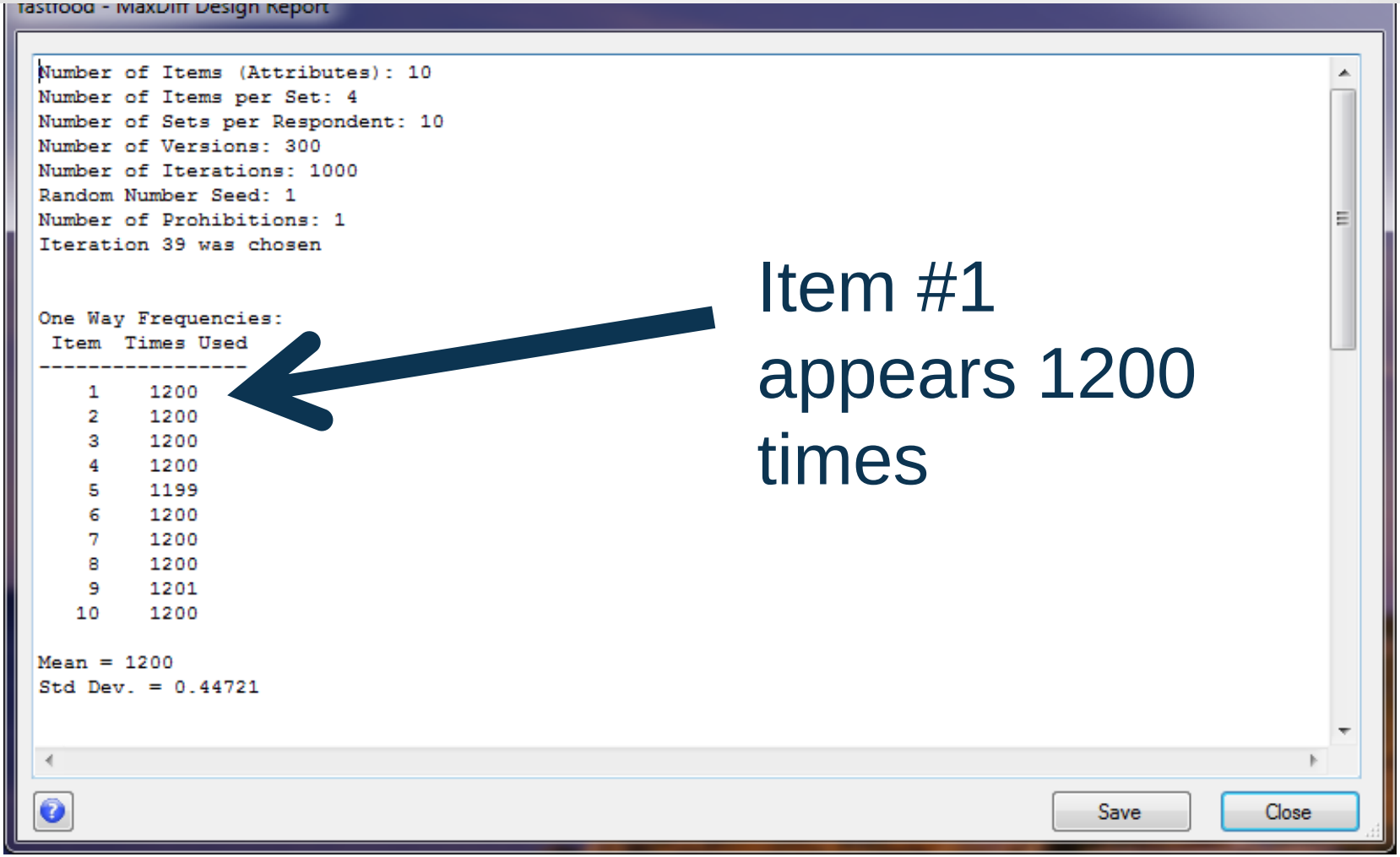
4. Specify the prohibitions



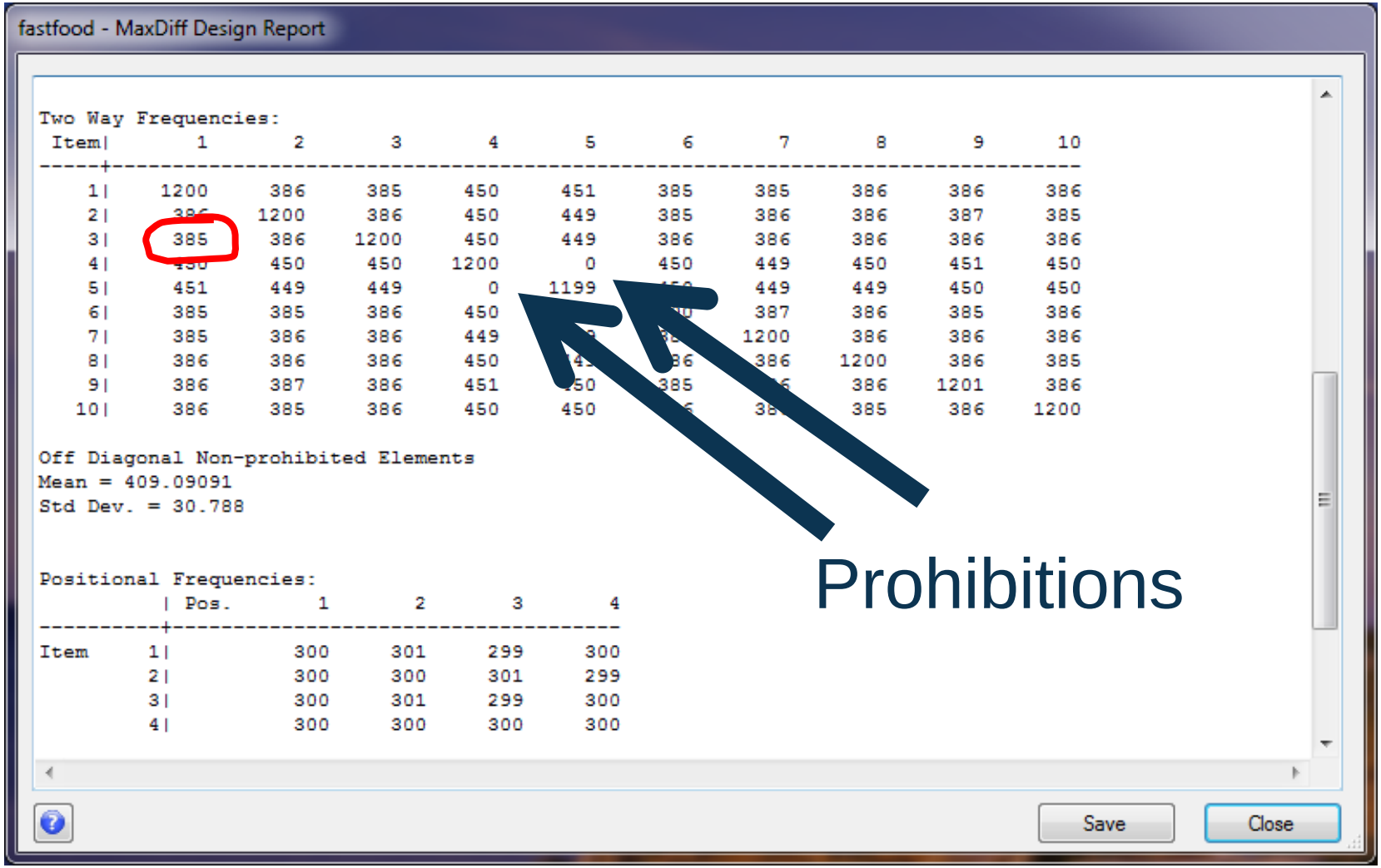
5. Click “generate design”

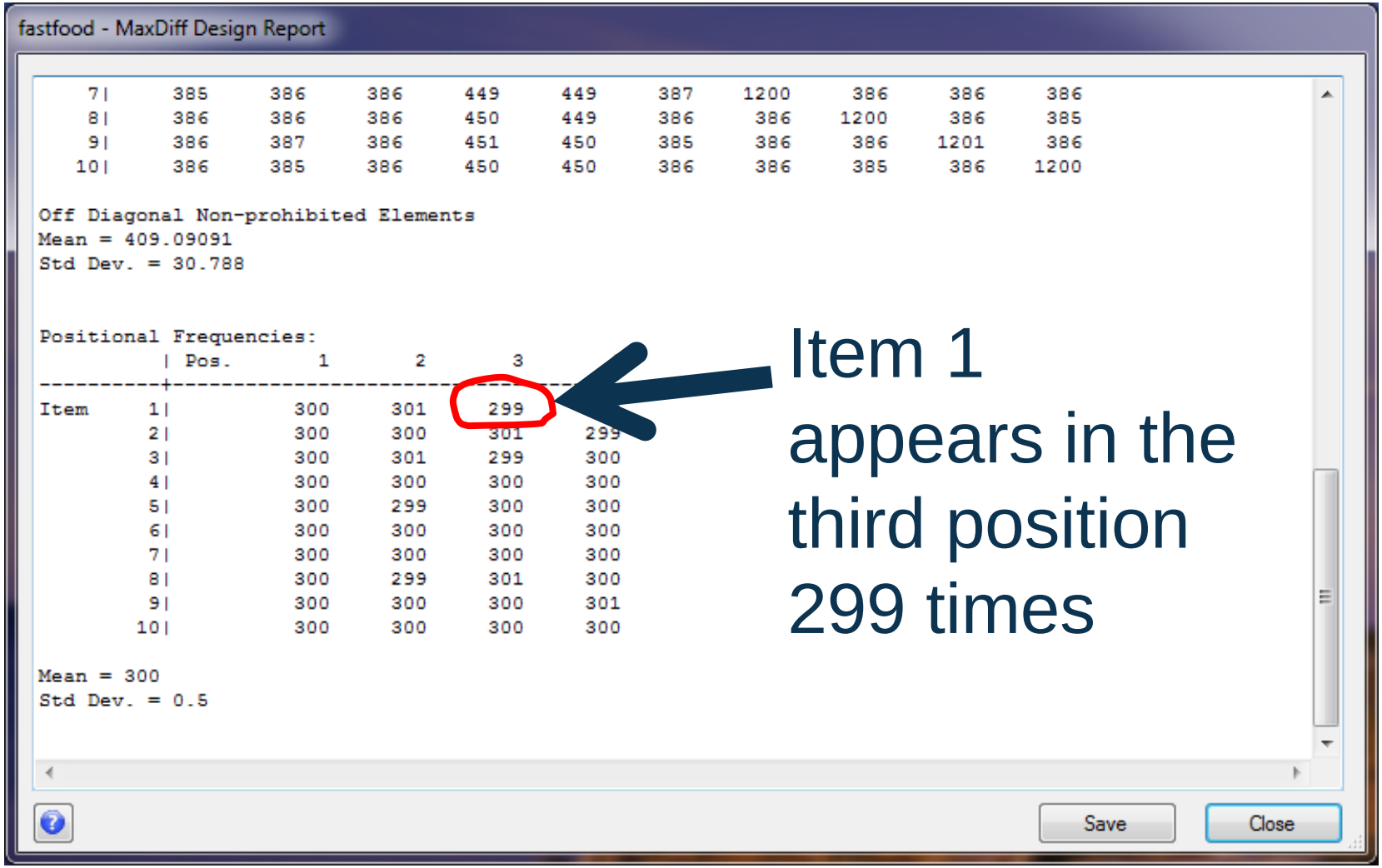


6. Review the design report



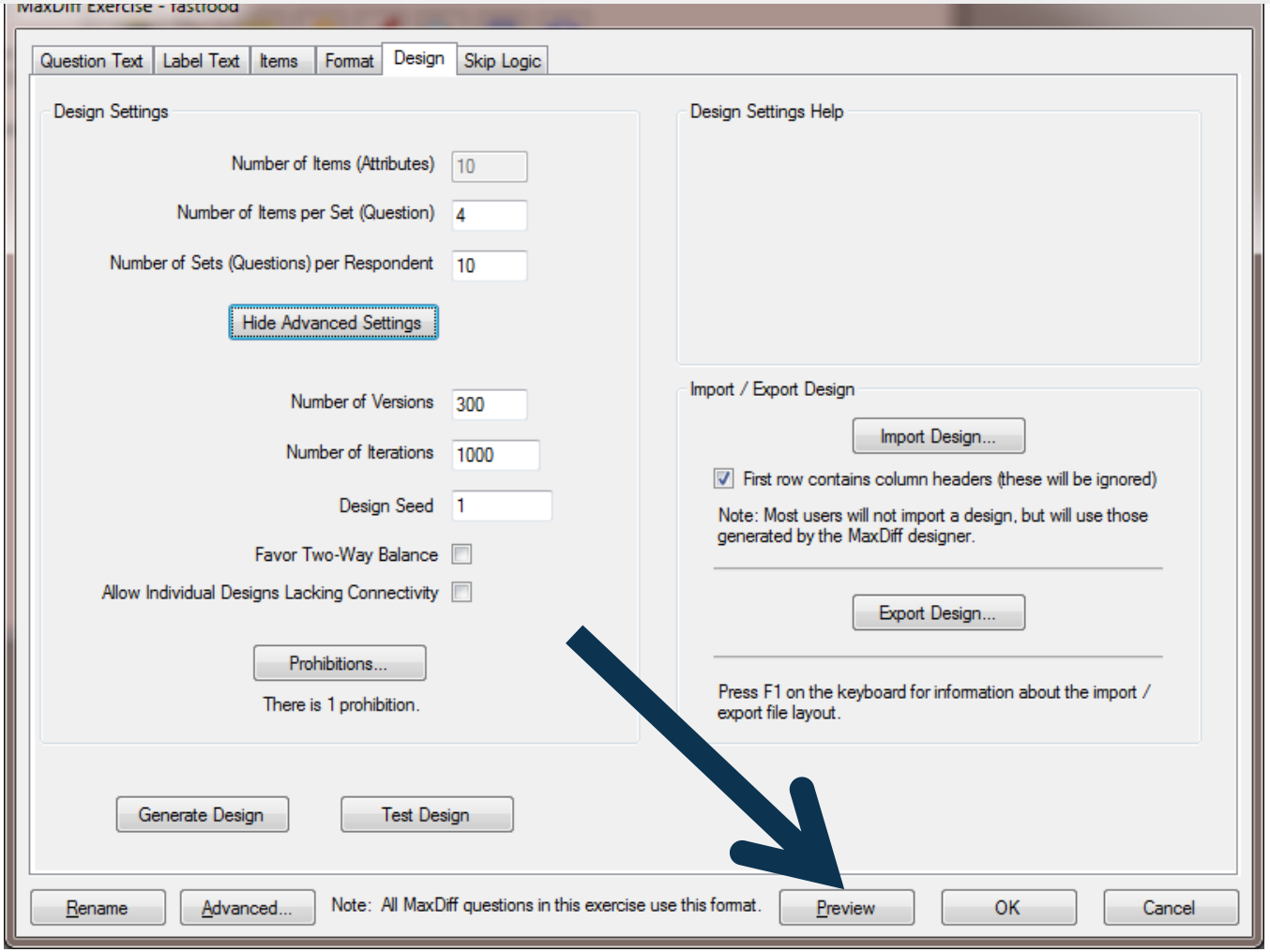
Item #1 appears 1200 times

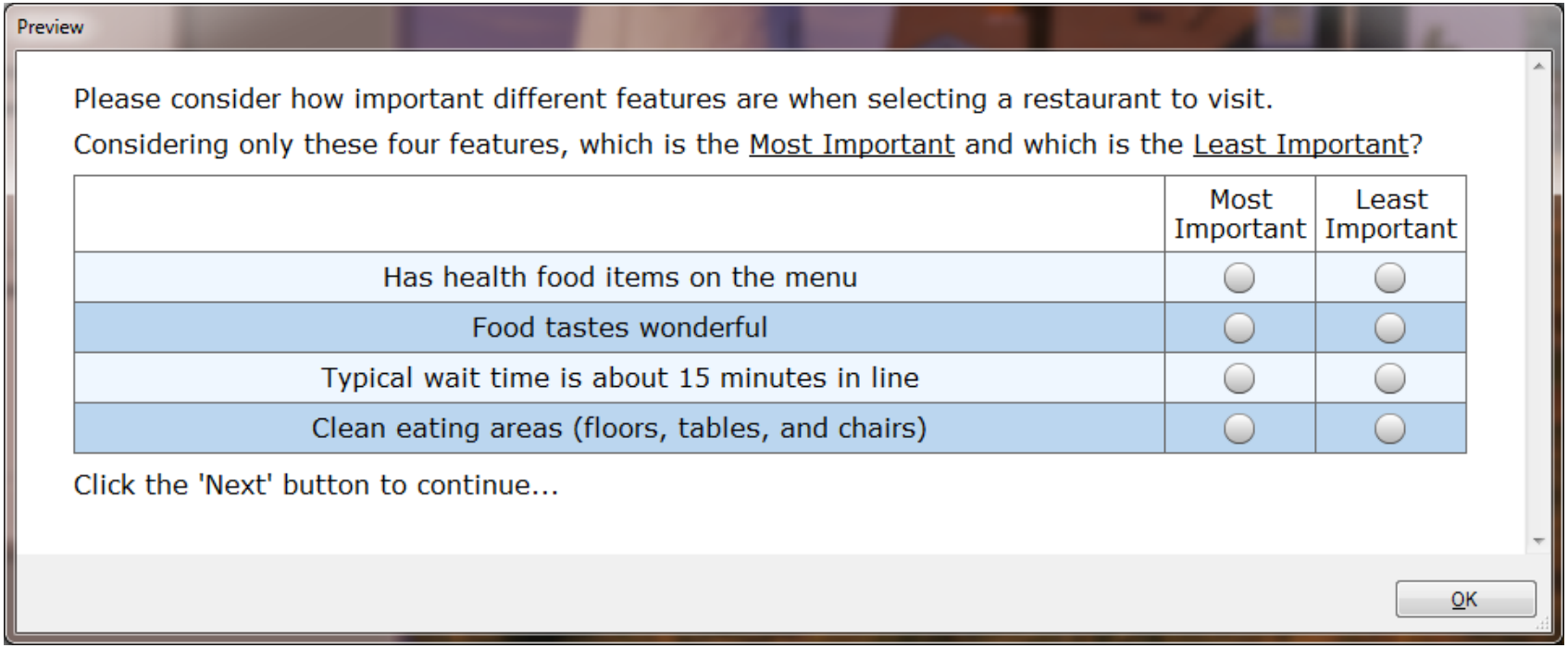

fastfood - MaxDiff Design Report
Two Way Frequencies:
Item| 1 2 3 4 5 6 7 8 9 10
1| 1200 386 385 450 451 385 385 386 386 386
2| 386 1200 386 450 449 385 386 386 387 385
3| 385 386 1200 450 449 386 386 386 386 386
4| 450 450 450 1200 0 450 449 450 451 450
5| 451 449 449 0 1199 449 449 450 450
6| 385 385 386 450 387 386 385 386
7| 385 386 386 449 1200 386 386 386
8| 386 386 386 450 386 386 1200 386 385
9| 386 387 386 451 385 386 1201 386
10| 386 385 386 450 450 385 386 1200
Off Diagonal Non-prohibited Elements
Mean = 409.09091
Std Dev. = 30.788
Positional Frequencies:
| Pos. 1 2 3 4
Item 1| 300 301 299 300
2| 300 300 301 299
3| 300 301 299 300
4| 300 300 300 300
Prohibitions
Save
Close



Item 1 appears in the third position 299 times

7. Click “Preview”




Preview
Please consider how important different features are when selecting a restaurant to visit.
Considering only these four features, which is the Most Important and which is the Least Important?
Most Important
Least Important
Has health food items on the menu
Food tastes wonderful
Typical wait time is about 15 minutes in line
Clean eating areas (floors, tables, and chairs)
Click the 'Next' button to continue...
OK

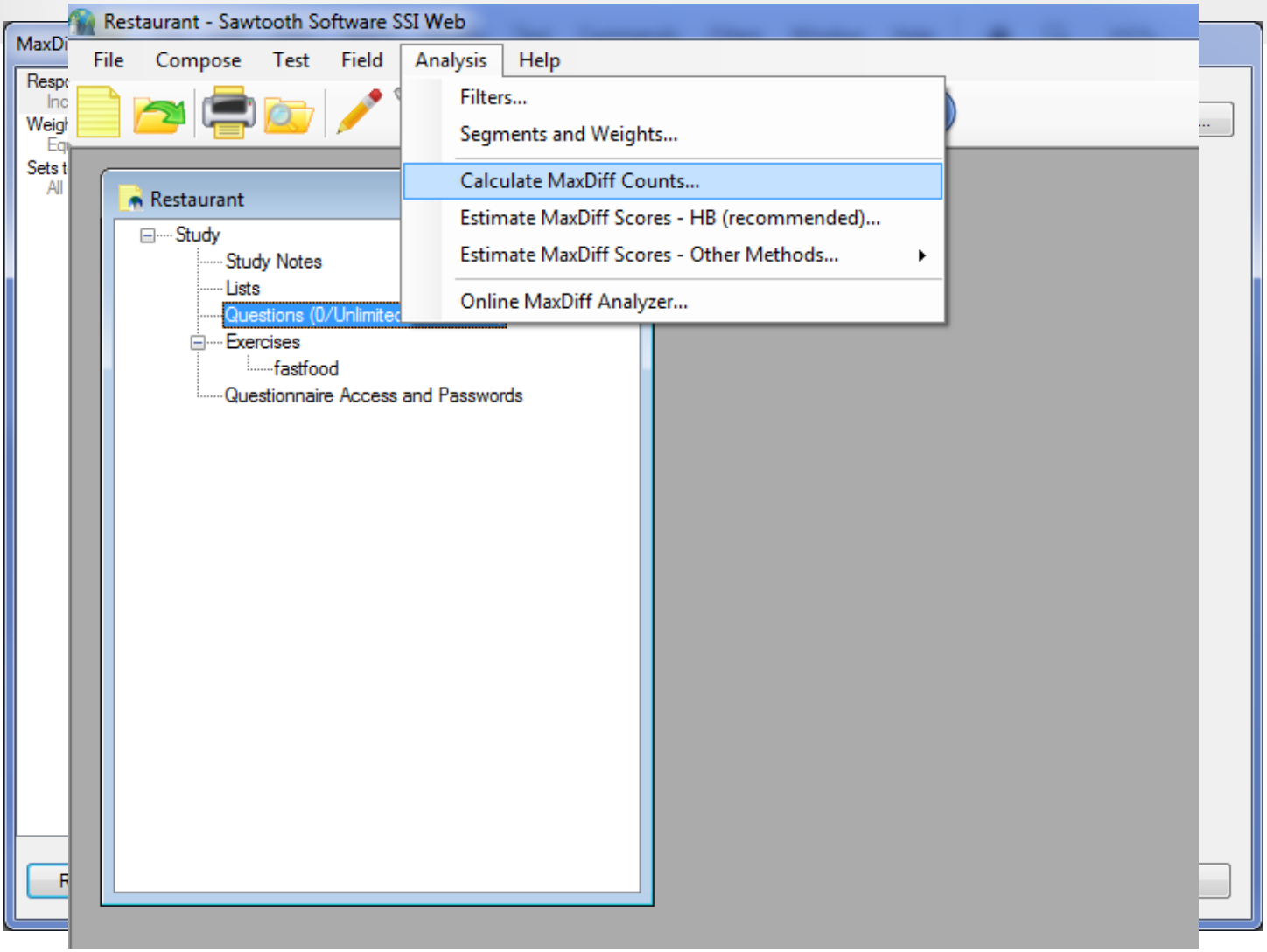
SECTION 4

Analyzing MaxDiff Data

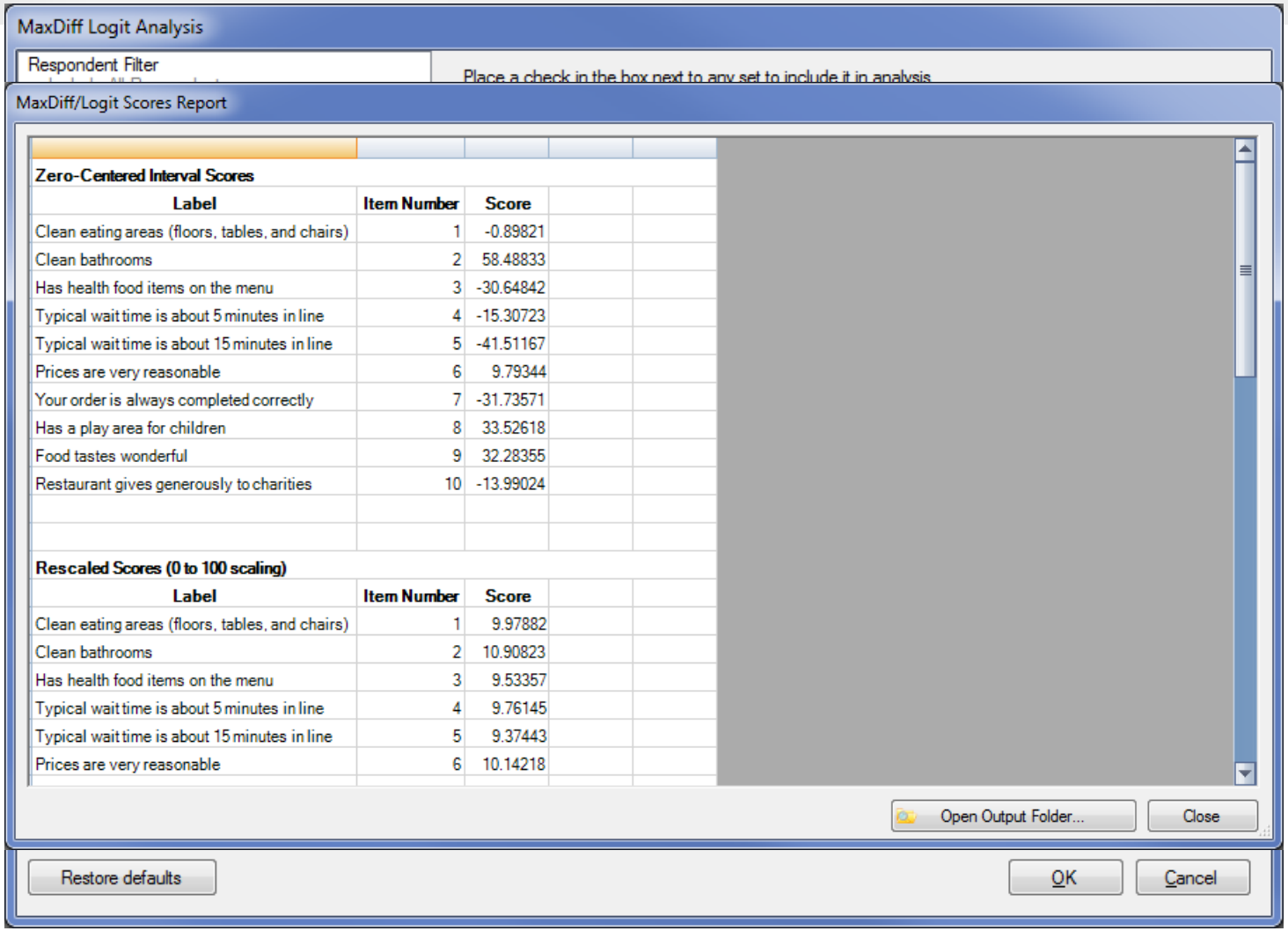
Methods of Analysis

- Counting
- Aggregate Logit
- Latent Class
- Hierarchical Bayes

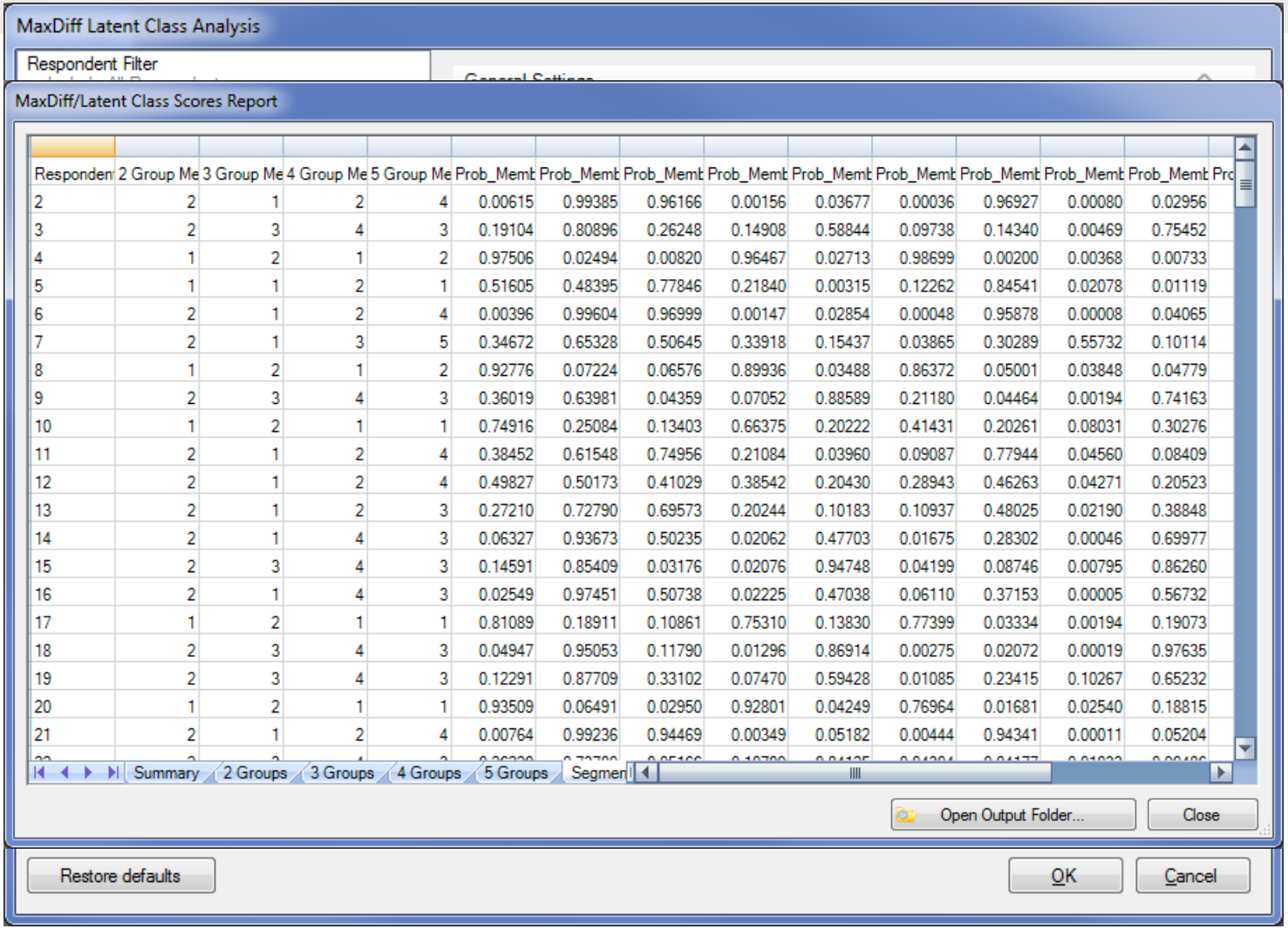
Counting


Restaurant - Sawtooth Software SSI Web
File
Compose
Test
Field
Analysis
Help
Filters...
Segments and Weights...
Calculate MaxDiff Counts...
Estimate MaxDiff Scores - HB (recommended)...
Estimate MaxDiff Scores - Other Methods...
Online MaxDiff Analyzer...
Restaurant
Study
Study Notes
Lists
Questions (0/Unlimited
Exercises
fastfood
Questionnaire Access and Passwords

Aggregate Logit

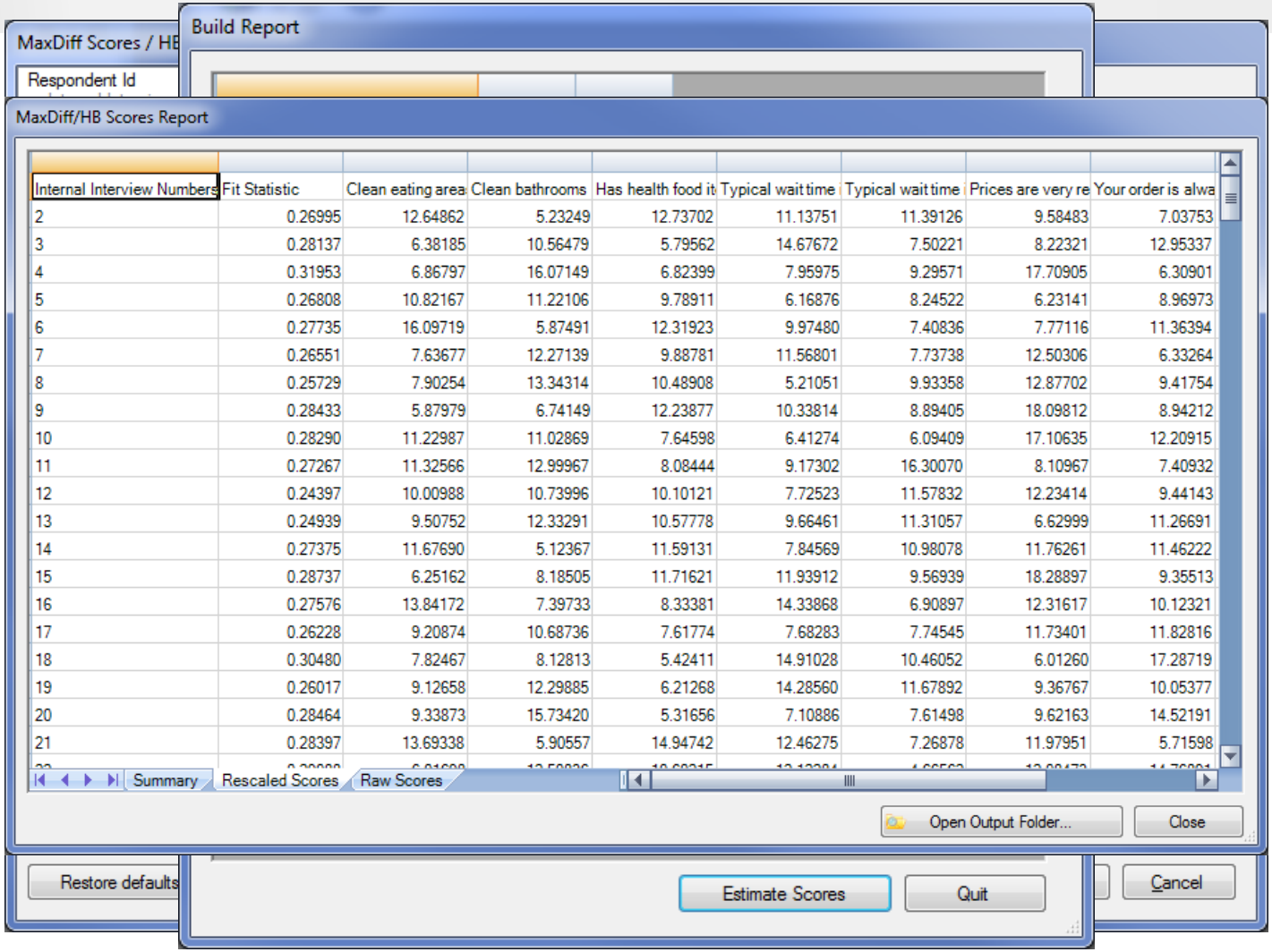


Latent Class



Respondent	2 Group Me	3 Group Me	4 Group Me	5 Group Me	Prob_Memb	Prob_Memb	Prob_Memb	Prob_Memb	Prob_Memb	Prob_Memb	Prob_Memb	Prob_Memb	Prob_Memb	Pro
2	2	1	2	4	0.00615	0.99385	0.96166	0.00156	0.03677	0.00036	0.96927	0.00080	0.02956	
3	2	3	4	3	0.19104	0.80896	0.26248	0.14908	0.58844	0.09738	0.14340	0.00469	0.75452	
4	1	2	1	2	0.97506	0.02494	0.00820	0.96467	0.02713	0.98699	0.00200	0.00368	0.00733	
5	1	1	2	1	0.51605	0.48395	0.77846	0.21840	0.00315	0.12262	0.84541	0.02078	0.01119	
6	2	1	2	4	0.00396	0.99604	0.96999	0.00147	0.02854	0.00048	0.95878	0.00008	0.04065	
7	2	1	3	5	0.34672	0.65328	0.50645	0.33918	0.15437	0.03865	0.30289	0.55732	0.10114	
8	1	2	1	2	0.92776	0.07224	0.06576	0.89936	0.03488	0.86372	0.05001	0.03848	0.04779	
9	2	3	4	3	0.36019	0.63981	0.04359	0.07052	0.88589	0.21180	0.04464	0.00194	0.74163	
10	1	2	1	1	0.74916	0.25084	0.13403	0.66375	0.20222	0.41431	0.20261	0.08031	0.30276	
11	2	1	2	4	0.38452	0.61548	0.74956	0.21084	0.03960	0.09087	0.77944	0.04560	0.08409	
12	2	1	2	4	0.49827	0.50173	0.41029	0.38542	0.20430	0.28943	0.46263	0.04271	0.20523	
13	2	1	2	3	0.27210	0.72790	0.69573	0.20244	0.10183	0.10937	0.48025	0.02190	0.38848	
14	2	1	4	3	0.06327	0.93673	0.50235	0.02062	0.47703	0.01675	0.28302	0.00046	0.69977	
15	2	3	4	3	0.14591	0.85409	0.03176	0.02076	0.94748	0.04199	0.08746	0.00795	0.86260	
16	2	1	4	3	0.02549	0.97451	0.50738	0.02225	0.47038	0.06110	0.37153	0.00005	0.56732	
17	1	2	1	1	0.81089	0.18911	0.10861	0.75310	0.13830	0.77399	0.03334	0.00194	0.19073	
18	2	3	4	3	0.04947	0.95053	0.11790	0.01296	0.86914	0.00275	0.02072	0.00019	0.97635	
19	2	3	4	3	0.12291	0.87709	0.33102	0.07470	0.59428	0.01085	0.23415	0.10267	0.65232	
20	1	2	1	1	0.93509	0.06491	0.02950	0.92801	0.04249	0.76964	0.01681	0.02540	0.18815	
21	2	1	2	4	0.00764	0.99236	0.94469	0.00349	0.05182	0.00444	0.94341	0.00011	0.05204	

Hierarchical Bayes



Internal Interview Numbers	Fit Statistic	Clean eating area	Clean bathrooms	Has health food it	Typical wait time	Typical wait time	Prices are very re	Your order is alwa
2	0.26995	12.64862	5.23249	12.73702	11.13751	11.39126	9.58483	7.03753
3	0.28137	6.38185	10.56479	5.79562	14.67672	7.50221	8.22321	12.95337
4	0.31953	6.86797	16.07149	6.82399	7.95975	9.29571	17.70905	6.30901
5	0.26808	10.82167	11.22106	9.78911	6.16876	8.24522	6.23141	8.96973
6	0.27735	16.09719	5.87491	12.31923	9.97480	7.40836	7.77116	11.36394
7	0.26551	7.63677	12.27139	9.88781	11.56801	7.73738	12.50306	6.33264
8	0.25729	7.90254	13.34314	10.48908	5.21051	9.93358	12.87702	9.41754
9	0.28433	5.87979	6.74149	12.23877	10.33814	8.89405	18.09812	8.94212
10	0.28290	11.22987	11.02869	7.64598	6.41274	6.09409	17.10635	12.20915
11	0.27267	11.32566	12.99967	8.08444	9.17302	16.30070	8.10967	7.40932
12	0.24397	10.00988	10.73996	10.10121	7.72523	11.57832	12.23414	9.44143
13	0.24939	9.50752	12.33291	10.57778	9.66461	11.31057	6.62999	11.26691
14	0.27375	11.67690	5.12367	11.59131	7.84569	10.98078	11.76261	11.46222
15	0.28737	6.25162	8.18505	11.71621	11.93912	9.56939	18.28897	9.35513
16	0.27576	13.84172	7.39733	8.33381	14.33868	6.90897	12.31617	10.12321
17	0.26228	9.20874	10.68736	7.61774	7.68283	7.74545	11.73401	11.82816
18	0.30480	7.82467	8.12813	5.42411	14.91028	10.46052	6.01260	17.28719
19	0.26017	9.12658	12.29885	6.21268	14.28560	11.67892	9.36767	10.05377
20	0.28464	9.33873	15.73420	5.31656	7.10886	7.61498	9.62163	14.52191
21	0.28397	13.69338	5.90557	14.94742	12.46275	7.26878	11.97951	5.71598

SECTION 5

Additional Analyses

Segmentation

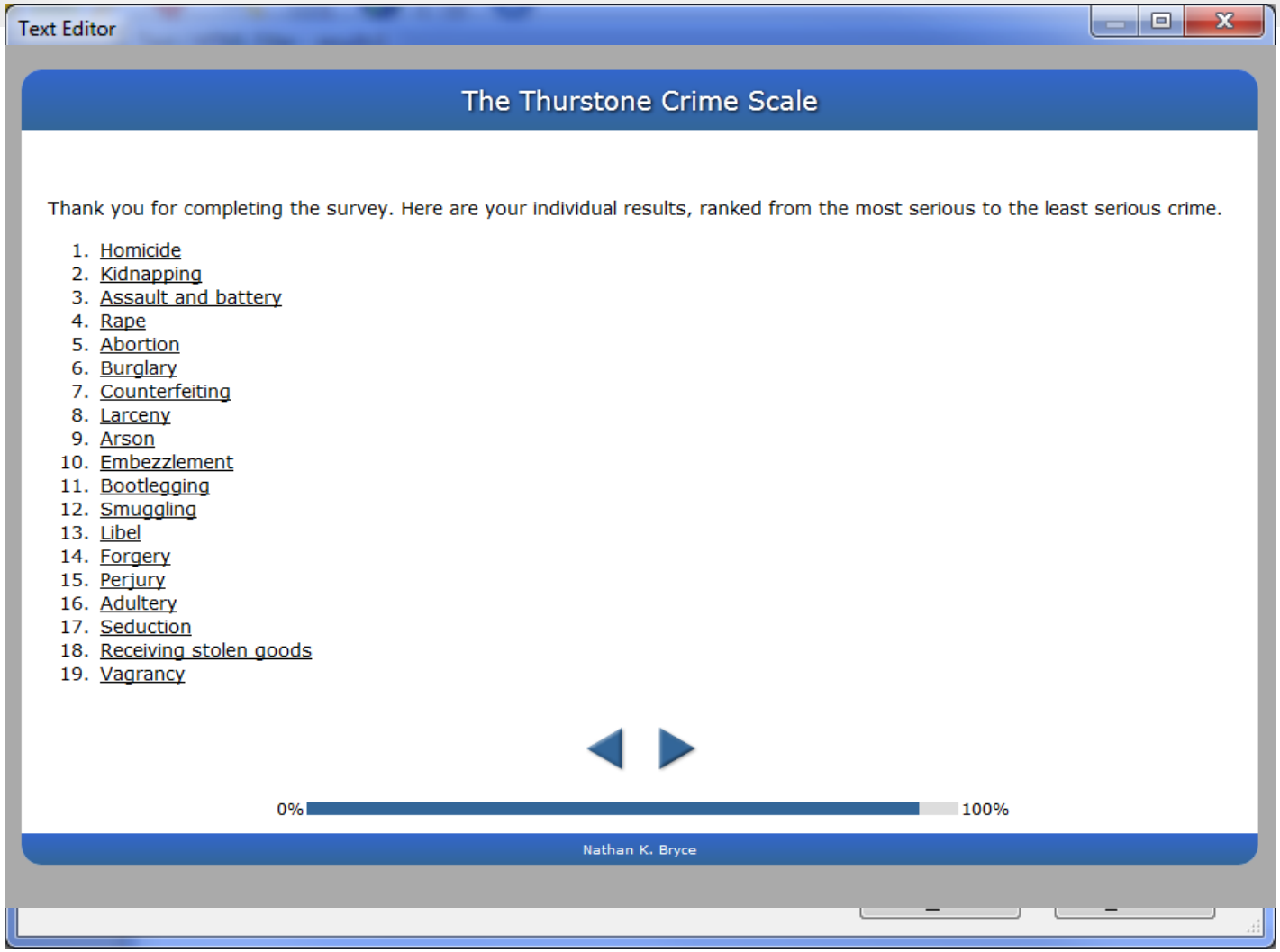
Strategic business decisions rely on segmenting the market and reaching the target.

Analyzing MaxDiff scores with our Latent Class or Convergent Cluster Ensemble Analysis software results in:

- Greater between-item and between-respondent discrimination
- Greater predictive accuracy

Than either monadic ratings or paired comparisons differentiated solutions.

“On The Fly” Scores



MaxDiff Analyzer

1. Display average values for the sample, or by segment
2. Conduct simulations, projecting “market choices”
3. Optimize portfolios of items to “reach” respondents, via TURF analysis
4. Automatically rescale scores, like probability or 0-100 scales, with 95% confidence intervals

Simulator

- **First Choice Rule**. Each respondent "casts a vote" for the item that has the highest score within the items included in the simulation set.
- **Share of Preference (Logit) Rule**. Respondents are allowed to split their votes across the items included in the simulation set.

Total Unduplicated Reach and Frequency (TURF) Analysis

Finds the items that will reach the maximum number of people

Anchoring

1

Ask to identify acceptable items from entire list (**Lattery's Direct Approach**)

2

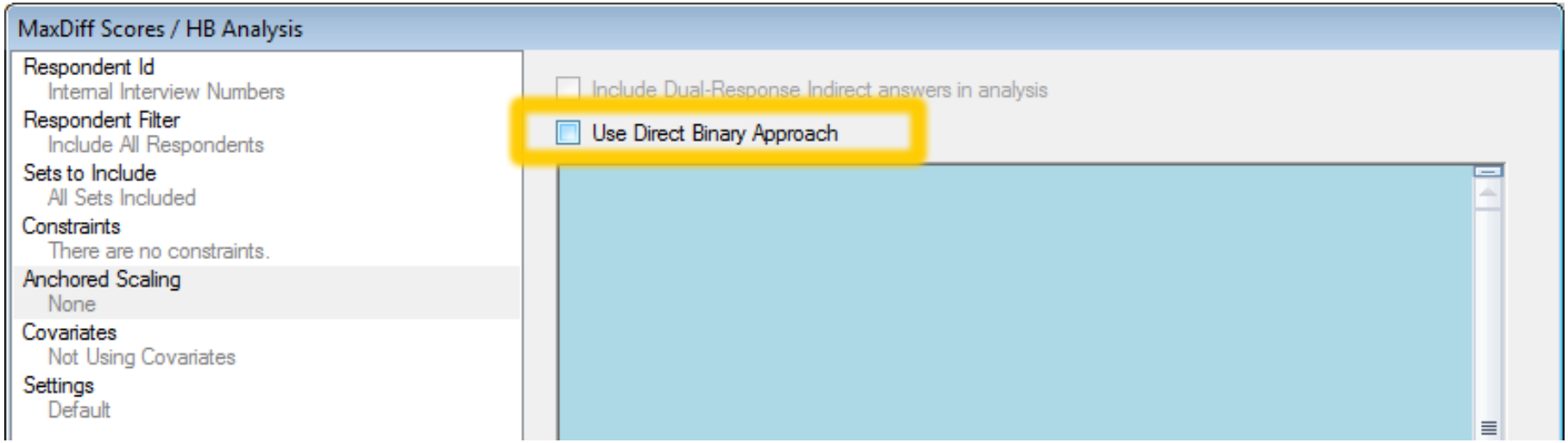
Ask to indicate whether all items in a set are all good, all bad, or some good and some bad (**Louviere's Indirect Approach**)

Direct Approach

This one is kind of long, but remember: For Science!

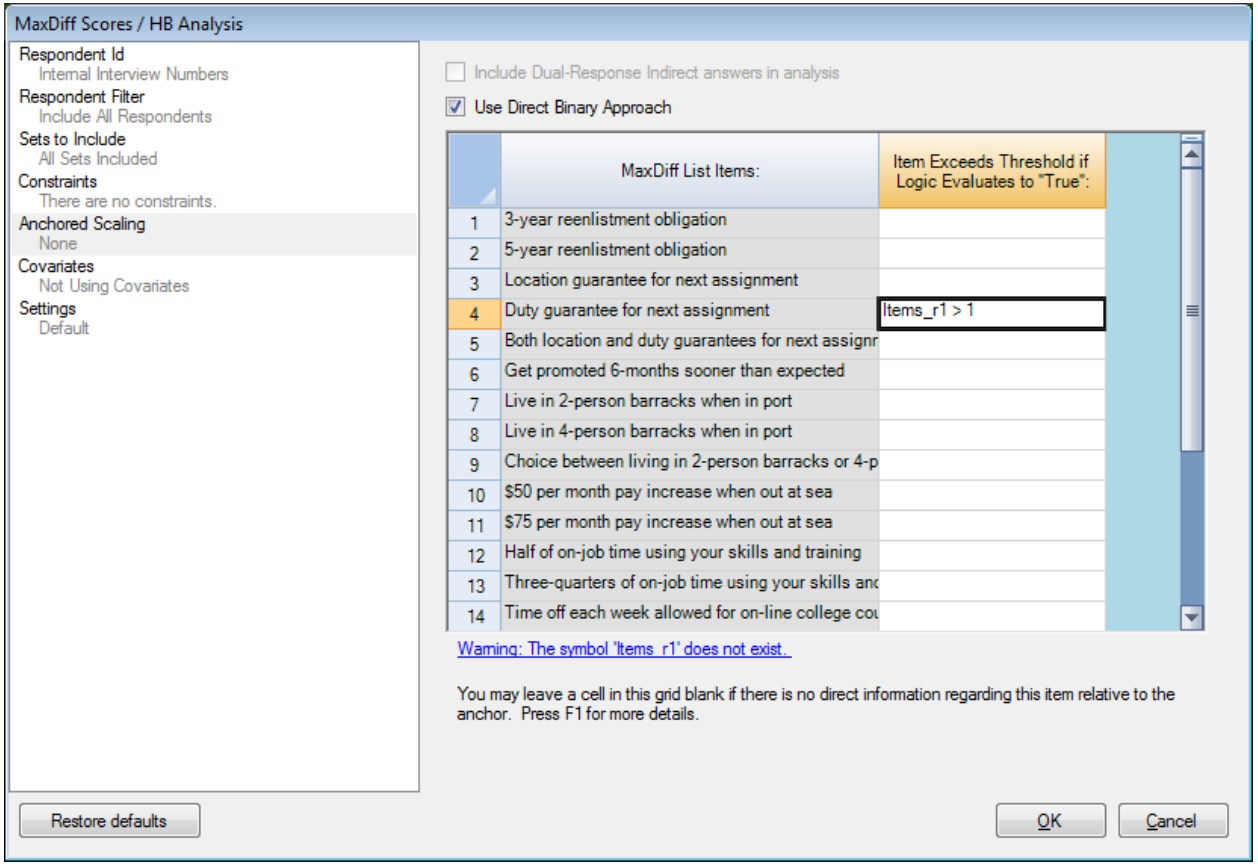
Please consider each pie below. If only that pie was available, would you take a slice?

	I would definitely take a slice	I probably would take a slice	Undecided	I probably would **not** take a slice	I would definitely **not** take a slice
Apple Crumb	○	○	○	○	○
Banana Cream	○	○	○	○	○
Berry Patch	○	○	○	○	○
Blackberry Bavarian Bliss	○	○	○	○	○
Blueberry Sour Cream	○	○	○	○	○
Caramel Pecan Silk Supreme	○	○	○	○	○
Cherry	○	○	○	○	○

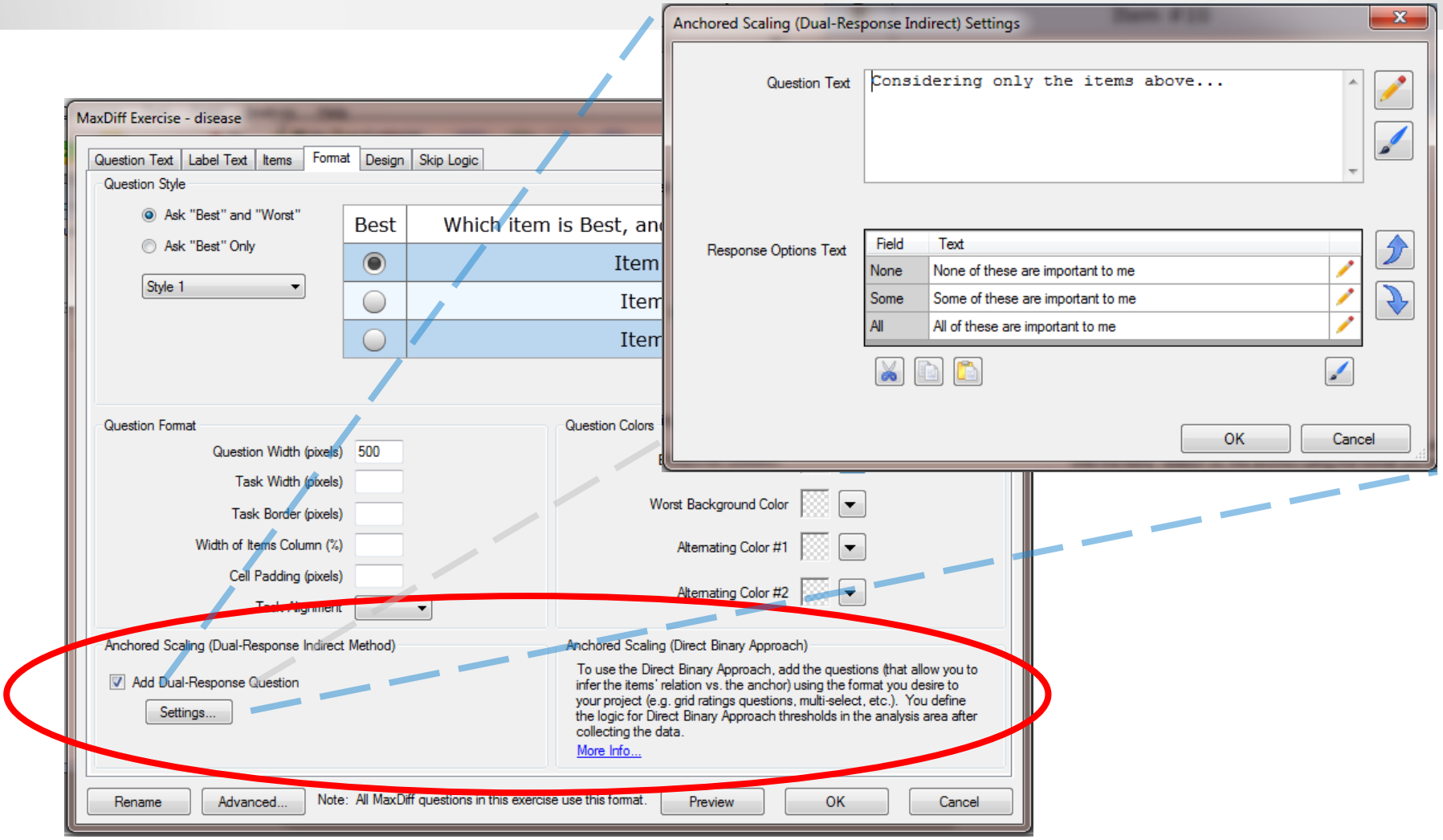


After you've collected the data, during score estimation, you simply indicate that you wish to use the **Direct Binary Anchoring** method.

You then specify the logic that references the scaling questions and identifies which values on the scale indicate that an item has exceeded the anchor threshold.



Indirect Approach



Considering only these four diseases, which is the Most Preferred and which is the Least Preferred?

(1 of 20)

Most Preferred		Least Preferred
○	Gum disease	○
○	Broken leg	○
○	Heart attack	○
○	Brain aneurysm	○

Considering only the items above...

- ○ None of these diseases are acceptable to me
- ○ Some of diseases are acceptable to me
- ○ All of these diseases are acceptable to me

Click the 'Next' button to continue...

In celebration of Pi Day, Sawtooth Software is going to pick up some pies. Hooray!

If Becky presented you with just the following 4 pie options, of which would you most like a slice?

Most Desirable		Least Desirable
○	**Triple Chocolate Cream**	○
○	**French Silk**	○
○	**Strawberry Rhubarb**	○
○	**Chocolate Caramel Delight**	○

If Becky went ahead and bought all 4 of these pies I would...

- ○ Not take a slice of any of them
- ○ Happily take a slice of any of them (or all of them!)
- ○ Probably take a slice of at least one

Anchoring Issues

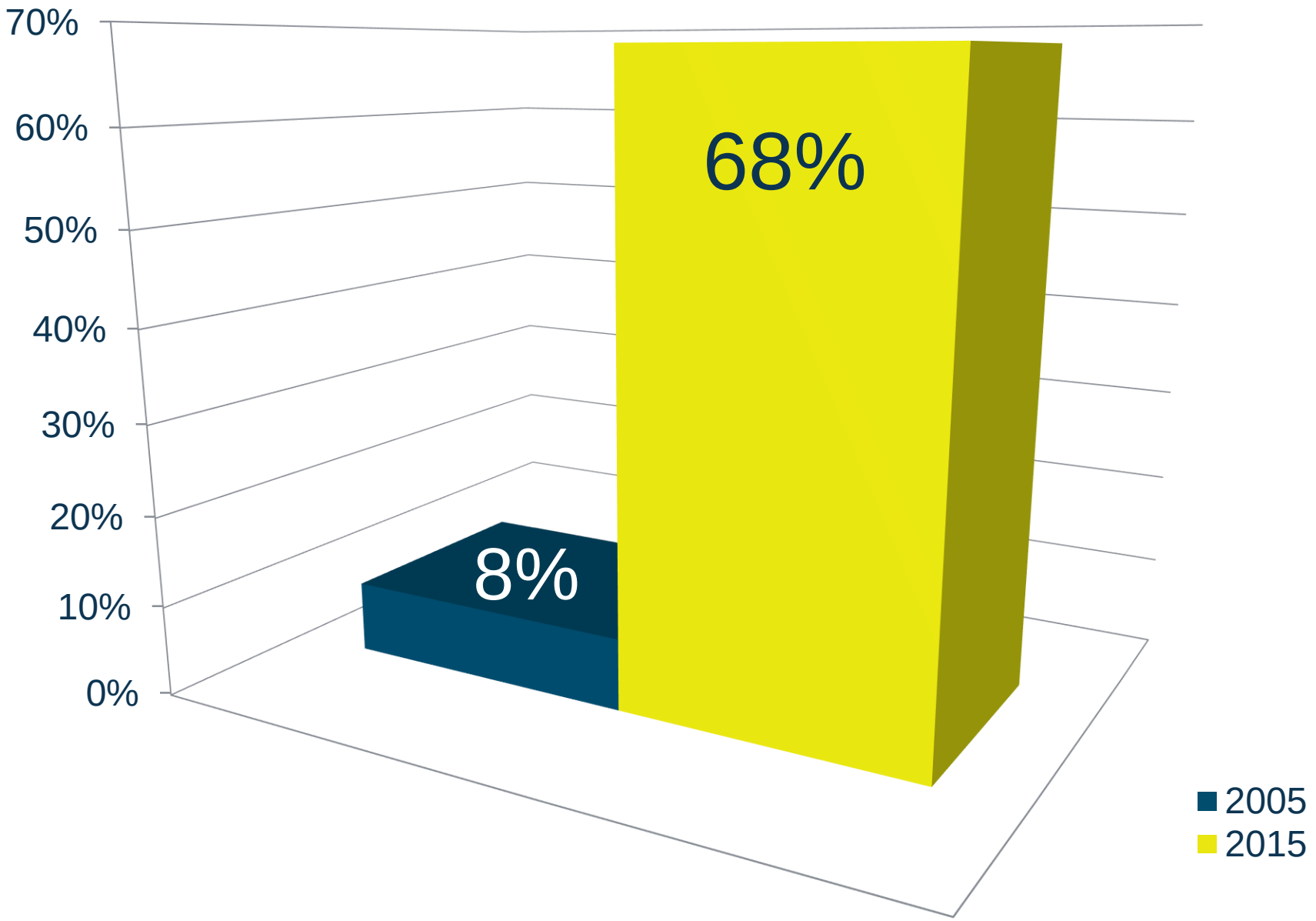
- Questionnaire is now longer
- Re-introduces scale bias
  - ✓ Still not using scale, but propensity to mark things as acceptable or not is different
  - ✓ Clustering may find similarities based primarily on use of cutoff rather than actual differences in which items are preferred

Express, Sparse and Constructed Lists

- Express – Creates different subsets of the larger number of items and asks a given respondent only about that subset, using HB analysis to fill in the blanks for the missing items.
- Sparse – Each respondent sees all items in the study, but fewer than the recommended number of times (<3, perhaps just once).
- Constructed list – Base the items on previous answers in a study, but all respondents must see the same number of items in total.

SECTION 6

Conclusions



Sawtooth Software

The survey software of choice

Megan Peitz – *Ingenuity Ambassador*

megan@sawtoothsoftware.com

Brian McEwan – *Director of Client Services*

brian@sawtoothsoftware.com

www.sawtoothsoftware.com

+1 801 477 4700

SECTION 7

References

Chrzan, Keith and Michael Patterson, “Testing for the Optimal Number of Attributes in MaxDiff Questions,” Sawtooth Software Conference Proceedings.

Chrzan, Keith (2004), “The Options Pricing Model: An Application of Best-Worst Measurement” White Paper available at www.sawtoothsoftware.com.

Cohen, Steven H. (2003), “Maximum Difference Scaling: Improved Measures of Importance and Preference for Segmentation” White Paper available at www.sawtoothsoftware.com.

David, H. A. (1963a). *The Method of Paired Comparisons*. Griffin, London. (2nd ed., 1988.)

Luce, R.D. (1959). *Individual Choice Behaviours: A Theoretical Analysis.* New York: J. Wiley.

Thurstone, L.L. (1927). A law of comparative judgement. *Psychological Review,* 34, 278–286.

Thurstone, L.L. (1959). *The Measurement of Values*. Chicago: The University of Chicago Press.

Orme, Bryan K. (2005), “Accuracy of HB Estimation in MaxDiff Experiments,” White Paper available at www.sawtoothsoftware.com.