

Welcome to

The National Center for Database Marketing

NCDM
2009

Produced by:

DIRECT
Direct Marketing Association

DMA[®]
Direct Marketing Association

**SEIZE THE
DATA**

TIME IS MONEY. MONETIZE
YOUR CUSTOMER DATA NOW.

Multivariable Testing: **Statistics and Strategies to Achieve Long-term Benefits**

Chris Law, Principal, NNE Marketing

Fred Holland, Vice President, LucidView

December 9, 2009

Testing has always been important

- “Almost any question can be answered, cheaply, quickly and finally, by a test campaign.”
 - Claude Hopkins, Scientific Advertising, 1923
- “Testing is still the best way to find true breakthroughs.”
 - Bob Stone, Successful Direct Marketing Methods, 2008
- Perhaps one time in fifty a guess may be right. But fifty times in fifty an actual test tells you what to do and avoid.
 - Claude Hopkins, My Life in Advertising, 1927

is now more efficient

Testing ~~has always been important~~

Many questions

- “~~Almost any question~~ can be answered, cheaply, quickly and finally, by a test campaign.”

^
multivariable

What is Multivariable Testing?

- A vast array of test designs and statistical techniques developed to test variables more efficiently
- Change many variables at once—in an organized way—so you can separate out the impact of each
 - Number of test recipes (test versions)
 - Each with a unique combination of all elements
 - Analyzing all recipes together to quantify effects
- Test design = collection of related test recipes



Marketing and Multivariable Testing

- Testing is paramount to success
- Test, test again, and then test again
- Become intimately involved in your data
 - Gain a crystal clear understanding of -
 - all that has occurred in the past
 - use that learning to develop future testing strategies.
- Everything is imminently testable:
 - Creative, Offer, Message, Ask, Audience, etc..
- Apply same vigilance and methodical approach in all facets of the program

4 Key Benefits → Faster, Deeper Insights and Larger ROI

Benefit	Scientific Multivariable Testing	Split-run Testing
Efficiency	2 - 35 variables in one test	1 variable = 1 test
Speed and Sample size	Constant sample size (no matter how many variables)	Increasing sample size
Depth of insights	Accurate, robust, and comparative main effects and interactions	One main effect (conditional, with no interactions)
Flexibility	Wide range of test designs	One choice

“New” techniques are based on 75 years of statistical research

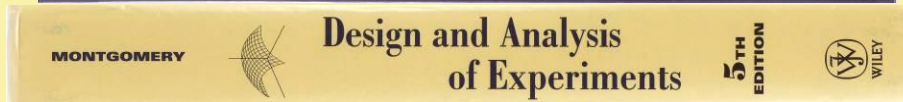
year

2007



2000

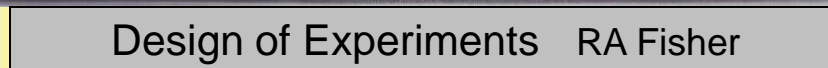
1997



1978



1935



Scientific articles on marketing applications

**Multivariable tests create a “stable framework”
for uncovering insights within marketing programs...**



... Even when the marketplace gets messy



Testing remains art & science

Good Ideas



Well executed



With the right test design



Solid statistics and clear data



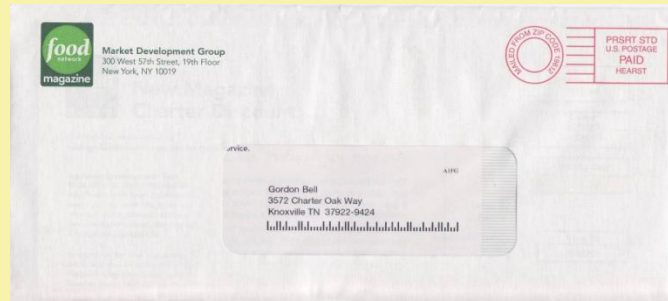
And reasonable interpretation of results

Good Ideas – Marketing Meets Science

- Objectives and Goal Setting
- Brainstorming w/o Barriers
- Criteria Evaluation
 - Costs
 - Risk / Reward
- Prioritization of Ideas
- Test Plan Creation

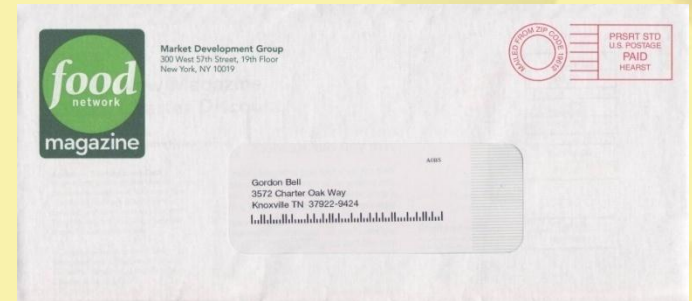
A simple split-run test: logo on envelope (presented by Hearst at DMA09)

Control



n = 50,000

Large Logo



n = 50,000

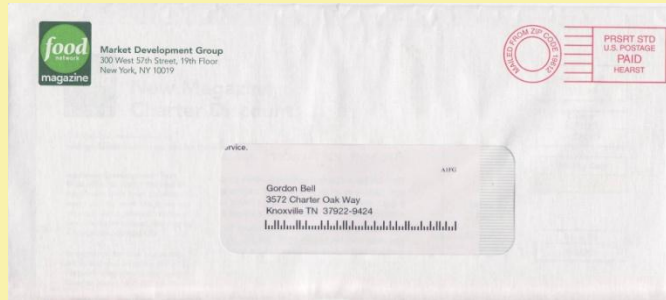
Red Logo



n = 50,000

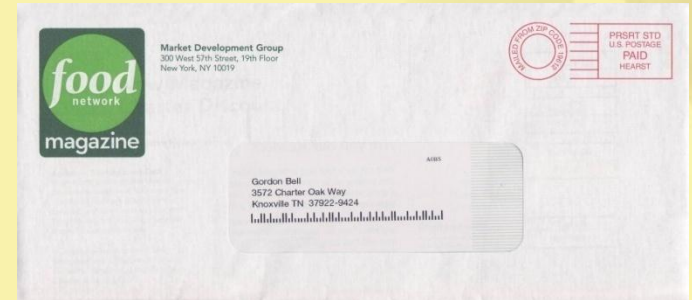
A simple multivariable test

Control



~~n = 50,000~~ 25,000

Large Logo



~~n = 50,000~~ 25,000

Red Logo



~~n = 50,000~~ 25,000



~~n = 50,000~~ 25,000

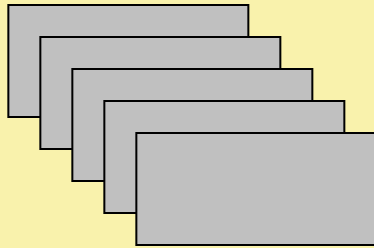
(3) Analyze main effects + the interaction

(2) Cut sample size by 1/3rd

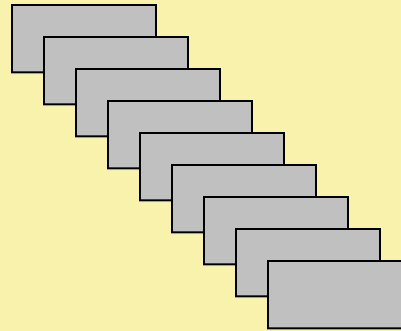
(1) Add one test cell

Larger tests: Split-run vs. Multivariable

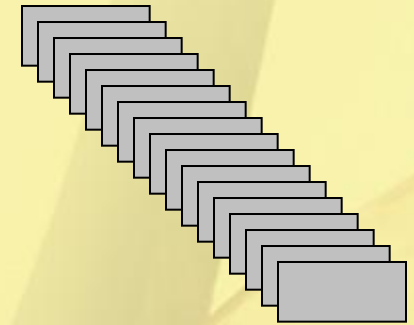
4 test elements
(+ control)



8 elements



16 elements



Split-run

n = 250,000

n = 450,000
or 2 campaigns

n = 850,000
or 4 campaigns

Multivariable

n = 100,000
1 campaign

n = 100,000
1 campaign

n = 100,000
1 campaign

A simple split-run test

<u>Test Element</u>	<u>(-) Control</u>	<u>(+) New Idea</u>
A: Logo size	Small	Large
B: Logo color	Green	Red

Split-run → Multivariable test matrix

	Logo color	Logo size		
Recipe	A	B	AB	Response
1	-	-	+	1.40%
2	+	-	-	1.10%
3	-	+	-	1.60%
4	+	+	+	1.50%

Calculating Effects

- Main effect = average of all plus recipes – average of all minus recipes
(because the design is orthogonal)

$$\text{Effect (A)} = \bar{y}_{A+} - \bar{y}_{A-}$$

$$\text{Effect (B)} = \bar{y}_{B+} - \bar{y}_{B-}$$

Average response of all “+” recipes

Average response of all “-” recipes

Calculating Effects

- Main effect = average of all plus recipes – average of all minus recipes
- Interaction effect = how the main effect changes depending upon the setting of another test element

Main effect of A when B is at the “+” level Main effect of A when B is at the “-” level

Effect (AB)

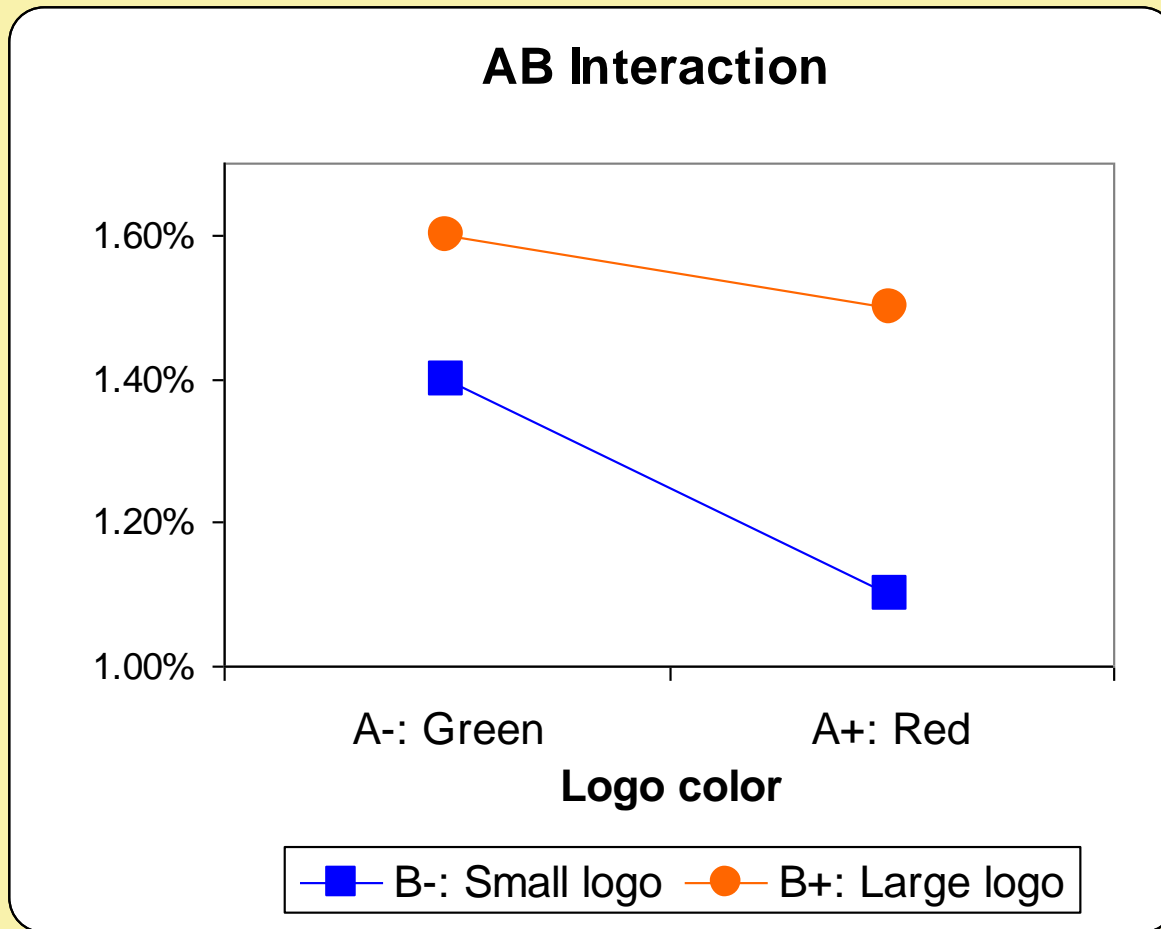
$$= \frac{1}{2} [(\bar{y}_{A+} - \bar{y}_{A-})_{B+} - (\bar{y}_{A+} - \bar{y}_{A-})_{B-}]$$

$$= \bar{y}_{AB+} - \bar{y}_{AB-}$$

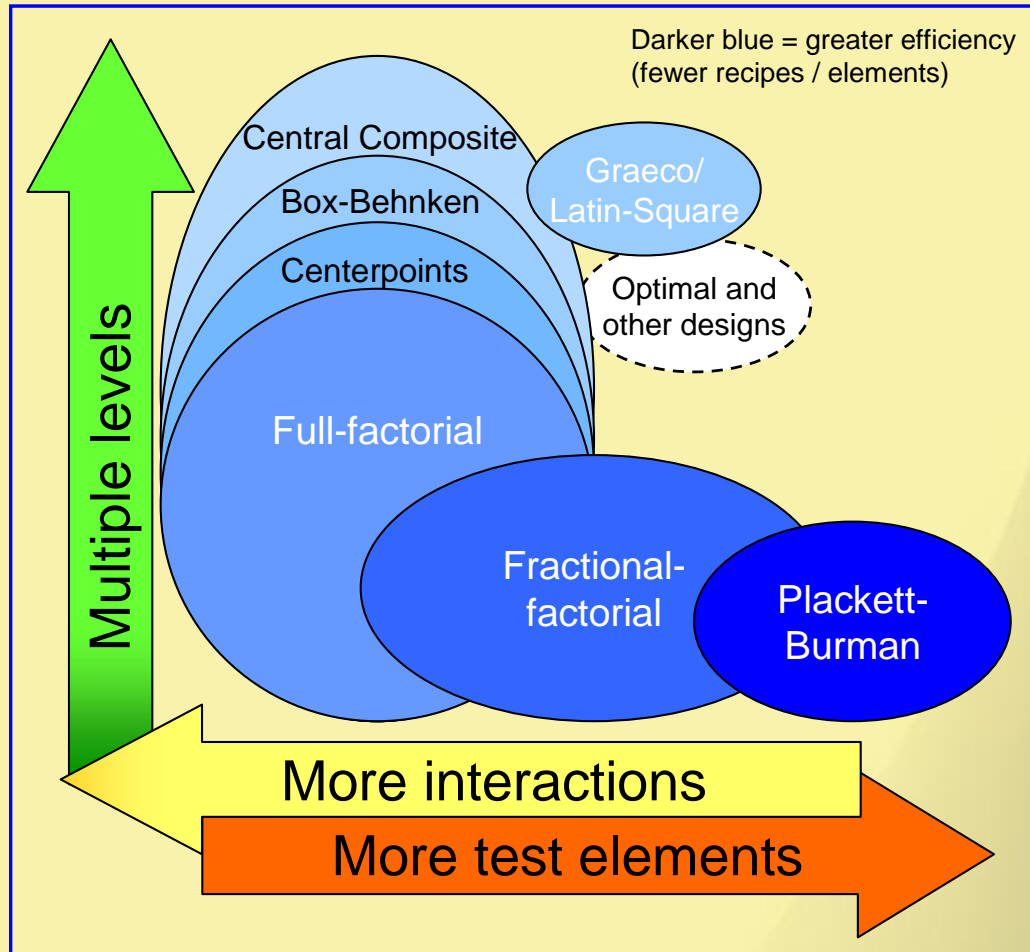
Or... just calculate avg(+)-avg(-) using the +/- signs in the AB column

Size-Color Interaction

Incremental change in one effect caused by another variable



Test Design Options



Important Principles of Multivariable Testing

1. Orthogonal Test Designs
2. Balanced Test Designs
3. Controlled Variation
 - Avoid making changes outside of the test
 - Maintain consistent execution of test elements
 - Block out known sources of variation
4. Replication
5. Randomization (of test units and recipes)
6. Measurement precision
7. Process stability

Case Study: Retail Test

	<u>Test Elements</u>	<u>Control (- level)</u>	<u>New Idea (+ level)</u>
A	Circular cover	Control	More promotional
B	Advertised discounts	Dramatic	Fewer, less dramatic discount
C	In-store signage	Current	Additional sale endcap displays
D	Copies of circular in store	Yes	No
E	Paper quality	Current	Higher quality paper
F	Products advertised in circular	Fewer	More
G	Hype end date	No	Yes, highlight sale end date
H	Product sequence in circular	Current	New sequence
J	Upsell tags in-store	No	Yes, "for only \$X more..."
K	Extra items on sale in-store	No	Yes
L	Circular market penetration	Current	Less

= circular advertising test elements
 = in-store test elements

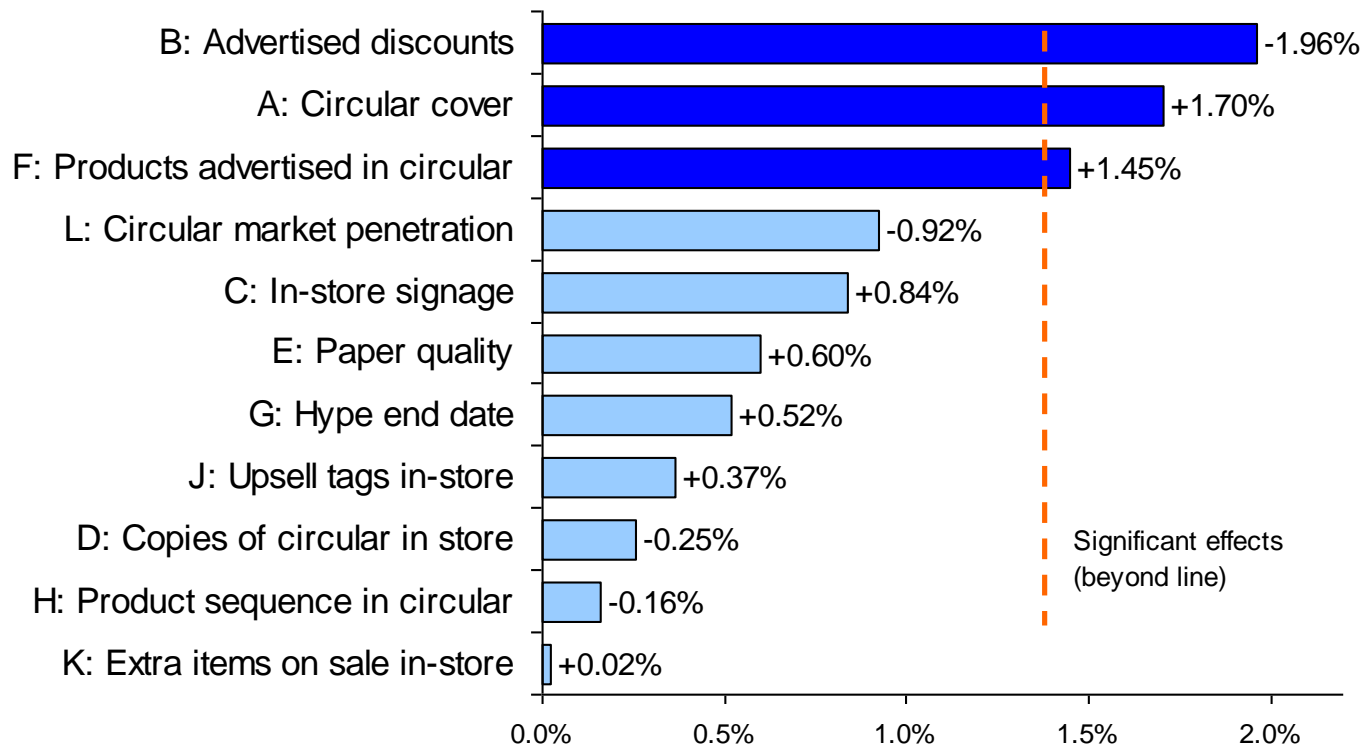
Retail Test Design

Recipe	Circular cover Advertised discounts In-store signage Copies of circular in store Paper quality Products advertised in circular Hype end date Product sequence in circular Upsell tags in-store Extra items on sale in-store Circular market penetration											Sales (4-week average)			Basket Size (4-week average)		
	A	B	C	D	E	F	G	H	J	K	L	Market 1	Market 2	Average	Market 1	Market 2	Average
1	+	+	-	+	+	-	+	-	-	-	+	-0.47%	1.94%	0.73%	25.61	24.99	\$25.30
2	+	-	-	-	+	+	+	-	+	+	-	7.46%	3.95%	5.71%	26.82	25.68	\$26.25
3	+	+	-	+	-	-	-	+	+	+	-	-0.20%	1.74%	0.77%	25.56	24.85	\$25.20
4	-	+	-	-	-	+	+	+	-	+	+	-1.83%	1.83%	0.00%	26.44	25.75	\$26.10
5	-	+	+	-	+	-	-	-	+	+	+	-0.56%	0.54%	-0.01%	26.96	26.28	\$26.62
6	-	+	+	+	-	+	+	-	+	-	-	1.23%	2.77%	2.00%	27.15	26.58	\$26.86
7	-	-	-	+	+	+	-	+	+	-	+	0.67%	3.59%	2.13%	25.48	25.33	\$25.40
8	-	-	+	+	+	-	+	+	-	+	-	1.83%	3.40%	2.61%	24.73	23.84	\$24.28
9	+	+	+	-	+	+	-	+	-	-	-	4.22%	2.81%	3.51%	27.91	27.34	\$27.63
10	+	-	+	+	-	+	-	-	-	+	+	2.92%	4.85%	3.88%	25.42	25.15	\$25.28
11	+	-	+	-	-	-	+	+	+	-	+	3.89%	2.90%	3.40%	27.14	26.18	\$26.66
12	-	-	-	-	-	-	-	-	-	-	-	1.65%	0.45%	1.05%	26.29	25.70	\$26.00

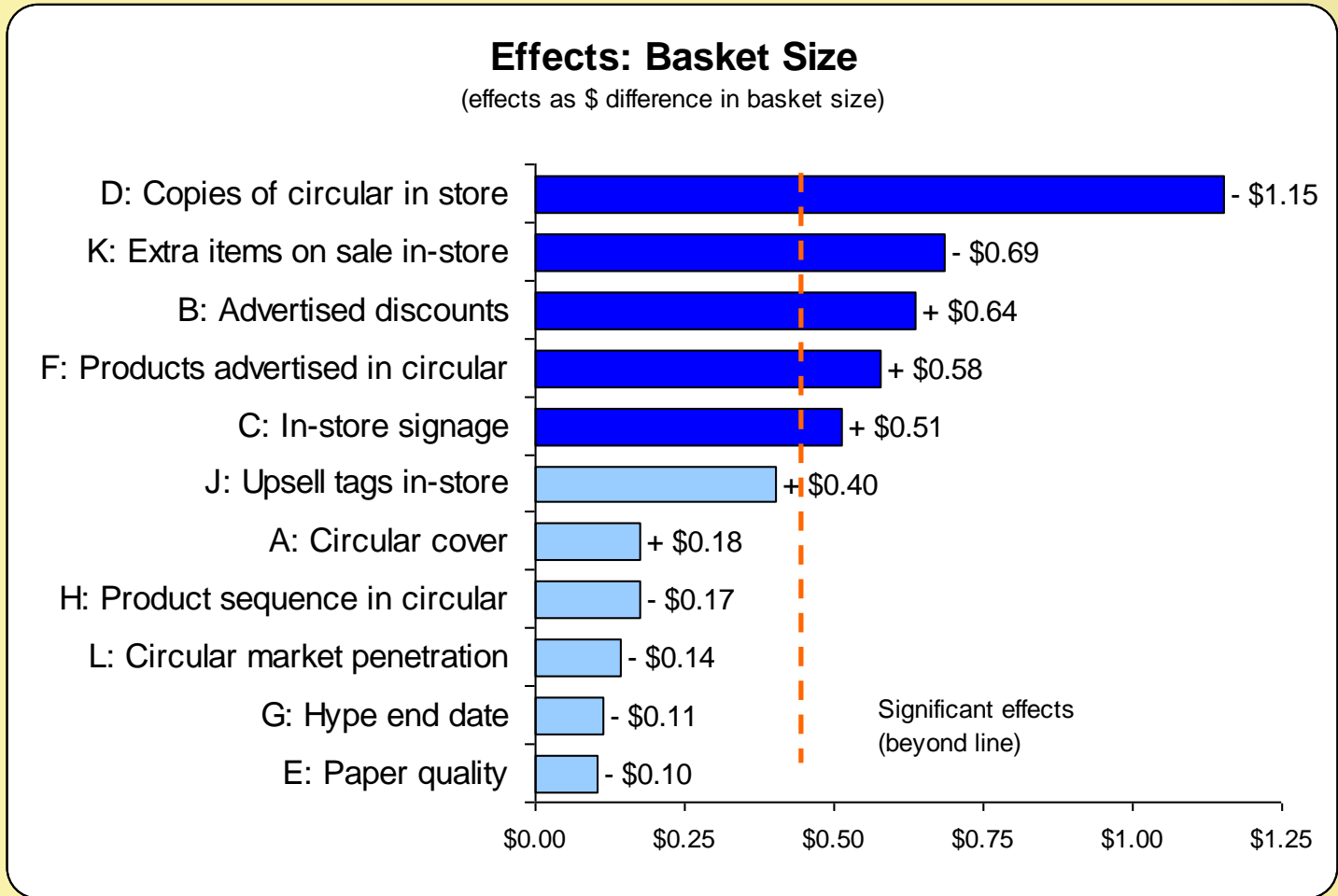
Retail Test Results - Sales

Effects: Sales

(effects as difference in % change of test period versus baseline)



Retail Test Results – Basket size



Calculating Predicted Improvement

- Create a linear equation with all significant effects and interactions

Predicted average sales/basket size

$$\hat{\bar{y}} = \bar{\bar{y}} + \frac{1}{2} (E(A) \cdot A + E(B) \cdot B + E(AB) \cdot A \cdot B + \text{all other significant effects})$$

Overall average across all test recipes

Each significant effect x optimal level, (+1) or (-1)

Retail Test

Predicted Improvement

$$\hat{\bar{y}} = \bar{\bar{y}} + \frac{1}{2} (E(A) \cdot A + E(B) \cdot B + E(AB) \cdot A \cdot B + \text{all other significant effects})$$

1) The predicted increase in **Sales**

$$= 0.0215 + \frac{1}{2} [E(B)(-1) + E(A)(+1) + E(F)(+1)]$$

$$= 0.0215 + \frac{1}{2} [0.0196 + 0.0170 + 0.0145]$$

$$= 0.0215 + 0.02555$$

$$= \mathbf{4.705\%} \rightarrow \text{the predicted increase in sales versus baseline period}$$

Notice that the “control” sales increase is 1.05% and the overall average during the test period is 2.15%

Retail Test

Predicted Improvement

$$\hat{\bar{y}} = \bar{\bar{y}} + \frac{1}{2} (E(A) \cdot A + E(B) \cdot B + E(AB) \cdot A \cdot B + \text{all other significant effects})$$

2) The predicted increase in **Basket Size**

$$\begin{aligned} &= \$25.97 + \frac{1}{2} [E(D)(-1) + E(K)(-1) + E(B)(+1) + E(F)(+1) + E(C)(+1)] \\ &= \$25.97 + \frac{1}{2} [1.15 + 0.69 + 0.64 + 0.58 + 0.51] \\ &= \$25.97 + \$1.79 \\ &= \mathbf{\$27.76} \quad \rightarrow \text{predicted actual basket size} \end{aligned}$$

Notice that the “control” basket size is \$26.00 (about the same as the overall average), so the optimal recipe increases basket size 6.8%

Case Study: Contact Strategy Test

- Objectives and Goal Setting
 - Determining the Optimal Contact Strategy
 - Improving short and long term revenue and customer retention
- Brainstorming w/o Barriers
 - Open discussion regarding all possible contact elements
- Criteria Evaluation
 - Costs – what are the costs associated with each element
 - Risk / Reward – Are the ideas big and bold, is there too much risk associated with testing
- Prioritization of Ideas
 - Culling the list to a manageable test plan
- Test Plan Creation

Case Study: Contact Strategy Test

<u>Test Element</u>	<u>Control (-)</u>	<u>New Idea (+)</u>
A Schedule of mailings	Monthly	Every two weeks
B Early mailing	None	Yes, add early drop
C Expiration mailing	None	Yes, week of contract expiration
D Late mailing	None	Yes, add later drop
E Outer envelope	Constant	Change every drop
F Insert	No	Yes
G Expiration e-mail	Yes	No

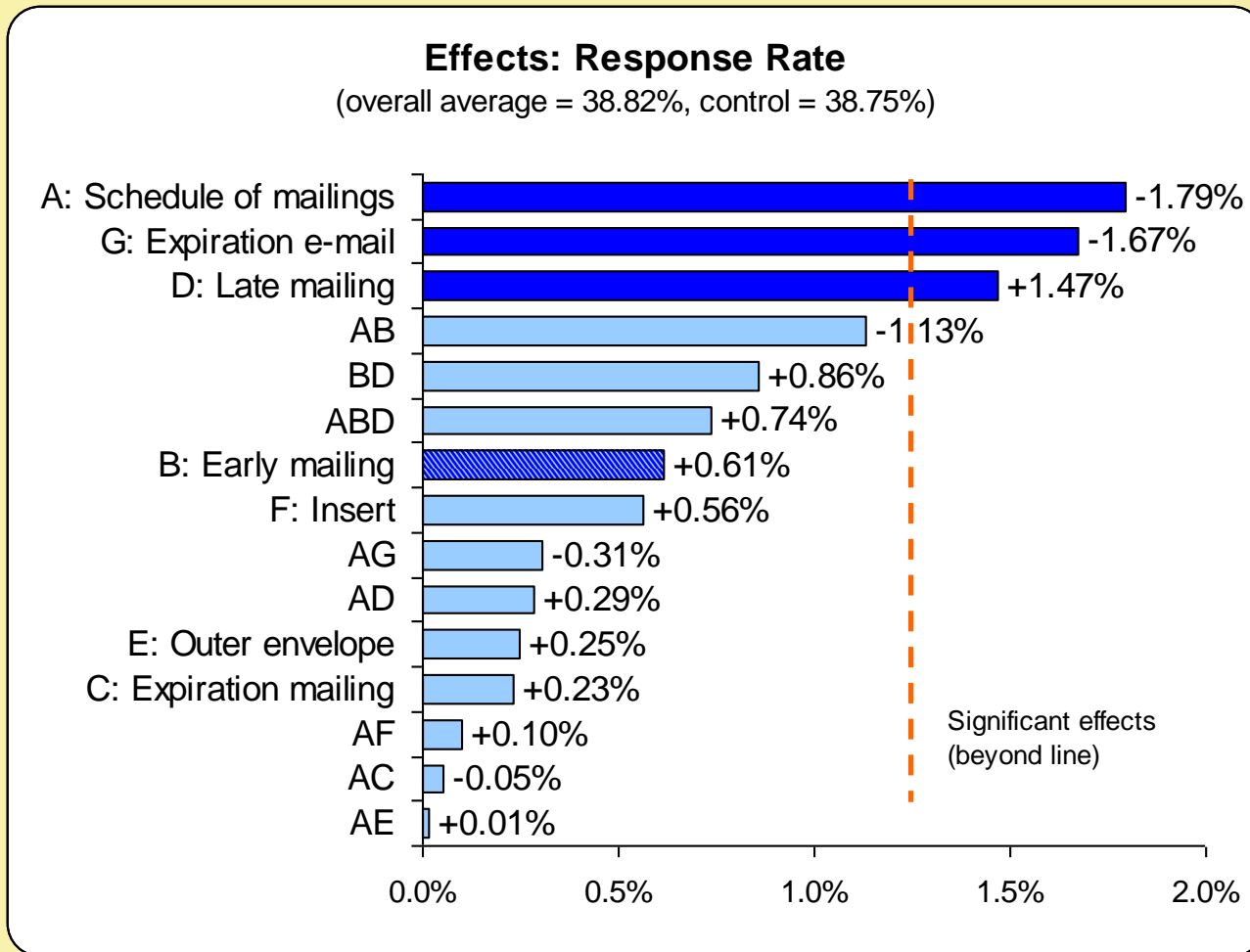
Contact Strategy Test Design

Recipe	Mailing schedule	Early mailing	Expiration mailing	Late mailing	Outer envelope	Insert	Expiration e-mail	Response Rate				
	A	B	C	D	E	F	G	Group 1	Group 2	Group 3	Group 4	Total
1	-	-	-	-	-	-	-	36.1%	40.1%	38.2%	39.6%	38.4%
2	+	-	-	-	+	-	+	35.9%	33.7%	41.6%	36.2%	37.0%
3	-	+	-	-	+	+	-	45.7%	38.5%	38.7%	40.4%	40.9%
4	+	+	-	-	-	+	+	35.5%	37.4%	36.6%	31.0%	35.1%
5	-	-	+	-	+	+	+	38.5%	36.7%	39.9%	37.4%	38.2%
6	+	-	+	-	-	+	-	36.9%	42.2%	40.0%	38.4%	39.3%
7	-	+	+	-	-	-	+	38.8%	40.3%	39.5%	37.5%	39.1%
8	+	+	+	-	+	-	-	39.2%	34.1%	38.8%	35.4%	37.0%
9	-	-	-	+	-	+	+	36.1%	43.3%	34.1%	41.2%	38.5%
10	+	-	-	+	+	+	-	38.7%	39.4%	38.4%	42.0%	39.6%
11	-	+	-	+	+	-	+	41.3%	45.4%	37.2%	37.3%	40.2%
12	+	+	-	+	-	-	-	39.7%	38.6%	40.2%	40.6%	39.8%
13	-	-	+	+	+	-	-	39.7%	40.5%	40.3%	39.8%	40.1%
14	+	-	+	+	-	-	+	37.3%	37.6%	36.6%	36.0%	36.9%
15	-	+	+	+	-	+	-	44.5%	41.8%	39.1%	43.4%	42.1%
16	+	+	+	+	+	+	+	36.9%	38.4%	40.9%	39.4%	38.9%

... and what you see beyond the basic matrix

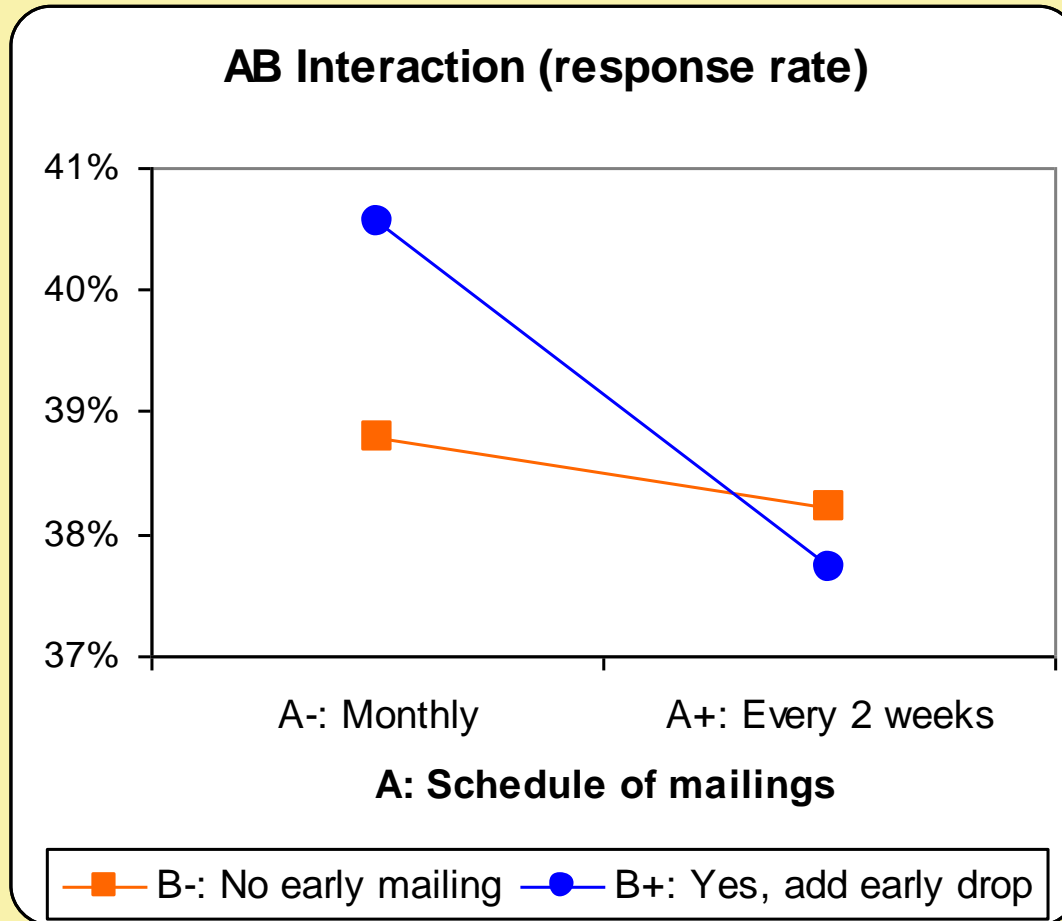
	Mailing schedule														
	Early mailing	Expiration mailing	Late mailing		Outer envelope					Expiration e-mail					
Recipe	A	B	C	D	AB	AC	AD	BC	BD	CD	F ABC	ABD	G ACD	F BCD	ABCD
1	-	-	-	-	+	+	+	+	+	+	-	-	-	-	+
2	+	-	-	-	-	-	-	+	+	+	+	+	+	-	-
3	-	+	-	-	-	+	+	-	-	+	+	+	-	+	-
4	+	+	-	-	+	-	-	-	-	+	-	-	+	+	+
5	-	-	+	-	+	-	+	-	+	-	+	-	+	+	-
6	+	-	+	-	-	+	-	-	+	-	-	+	-	+	+
7	-	+	+	-	-	-	+	+	-	-	-	+	+	-	+
8	+	+	+	-	+	+	-	+	-	-	+	-	-	-	-
9	-	-	-	+	+	+	-	+	-	-	-	+	+	+	-
10	+	-	-	+	-	-	+	+	-	-	+	-	-	+	+
11	-	+	-	+	-	+	-	-	+	-	+	-	+	-	+
12	+	+	-	+	+	-	+	-	+	-	-	+	-	-	-
13	-	-	+	+	+	-	-	-	-	+	+	+	-	-	+
14	+	-	+	+	-	+	+	-	-	+	-	-	+	-	-
15	-	+	+	+	-	-	-	+	+	+	-	-	-	+	-
16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

Contact Strategy Test Results



Contact Strategy Test Results

Interaction of Mailing Schedule and Early Mailing



Contact Strategy Test:

Predicted Improvement

Predicted response rate = 42.16%
(8.8% improvement over the control)

Case Study: CPG Test

Central Composite Design

- Objectives and Goal Setting
 - Determining the optimal price point
 - Increasing sales volume
- Brainstorming w/o Barriers
 - Open discussion regarding all possible pricing options
- Criteria Evaluation
 - Costs – what are the costs associated with each element
 - Risk / Reward – Are the ideas big and bold, is there too much risk associated with testing
- Prioritization of Ideas
 - Culling the list to a manageable test plan
- Test Plan Creation

Case Study: CPG Test

Central Composite Design

<u>Test Elements</u>	<u>(-) Level 1</u>	<u>(0) Centerpoint</u>	<u>(+) Level 2</u>
A Number of package sizes	3	4	5
B Spread among price points	Tight	Normal	Wide
C Pricing strategy	Discount	Competitive	Premium
D Shelf arrangement (left to right)	Small to large	Random	Large to small

Central Composite Designs

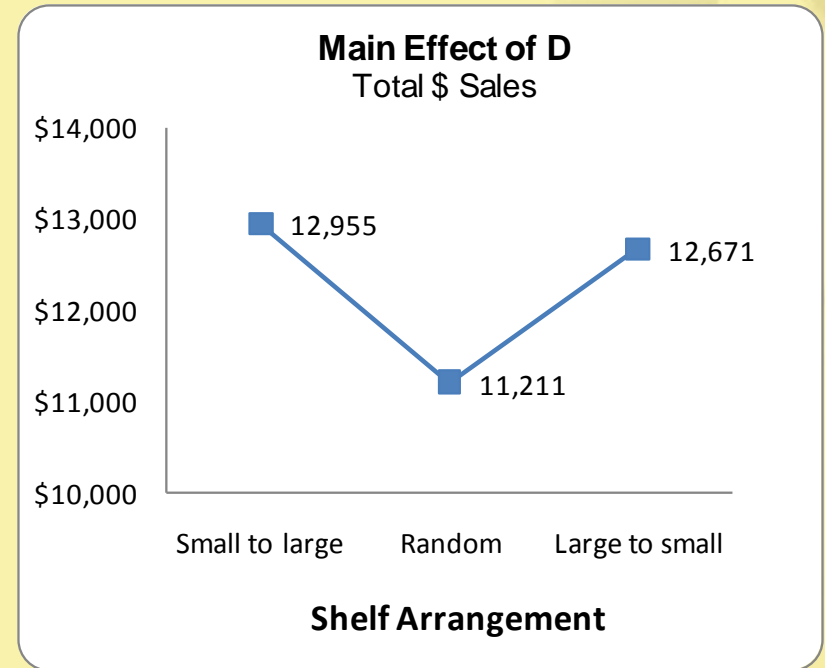
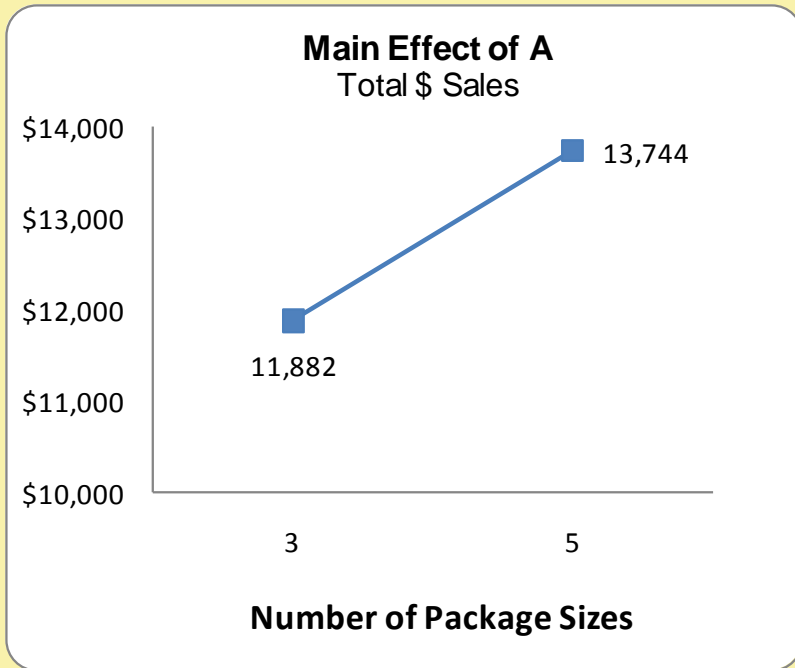
- Full factorial with centerpoints plus additional axial recipes
- Greater accuracy of curvature analysis
- More levels of each element are tested

Test Design and Data

Recipe	Number of package sizes				Number of Stores	Total Unit Sales	Average Purchase Price	Total Dollar Sales
	A	B	C	D				
1	-	-	-	-	20	4639	2.06	9,539.85
2	+	-	-	-	20	4764	3.15	15,001.96
3	-	+	-	-	20	4309	3.38	14,552.22
4	+	+	-	-	20	3955	3.04	12,029.21
5	-	-	+	-	20	3600	3.62	13,042.72
6	+	-	+	-	20	3881	3.69	14,308.76
7	-	+	+	-	20	3120	3.92	12,225.60
8	+	+	+	-	20	3321	3.90	12,938.39
9	-	-	-	+	20	3516	3.24	11,407.01
10	+	-	-	+	20	4130	3.13	12,936.51
11	-	+	-	+	20	4108	3.07	12,622.43
12	+	+	-	+	20	3956	3.96	15,646.64
13	-	-	+	+	20	3652	3.48	12,725.72
14	+	-	+	+	20	3829	3.96	15,160.48
15	-	+	+	+	20	2585	3.46	8,940.76
16	+	+	+	+	20	3144	3.79	11,931.07
17	-	0	0	0	20	2454	3.28	8,060.62
18	+	0	0	0	20	3356	3.47	11,652.03
19	0	-	0	0	20	3470	3.17	10,995.87
20	0	+	0	0	20	2989	4.24	12,686.17
21	0	0	-	0	20	3786	3.19	12,075.20
22	0	0	+	0	20	2763	3.92	10,821.59
23	0	0	0	-	20	3721	3.61	13,421.36
24	0	0	0	+	20	3629	3.68	13,345.96
25	0	0	0	0	60	9606	3.81	36,551.90

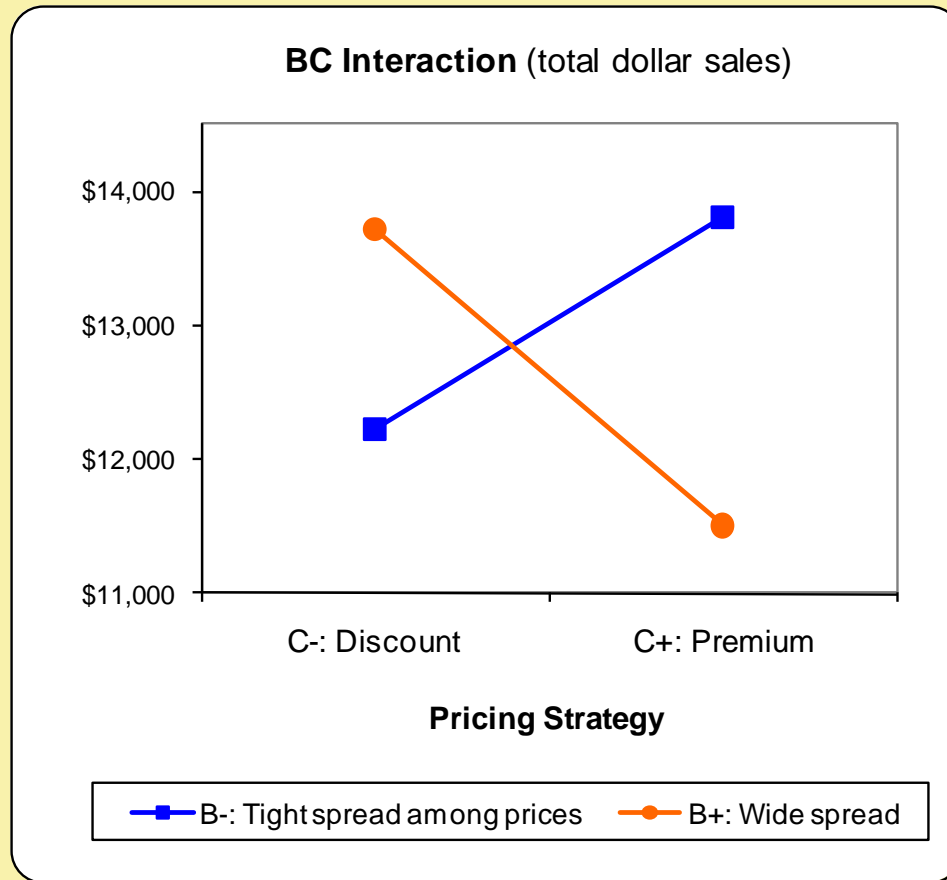
CPG Test Results

Main Effects



CPG Test Results

Interaction of Price Spread and Pricing Strategy



CPG Test Results

Optimal Settings

Optimal Combination for Total \$ Sales

- A+ 5 package sizes
- B- Tight price point spread
- C+ Premium pricing strategy
- D- Not random, either L to R or R to L

Case Study: Non-profit Renewals Test

- Objectives and Goal Setting
 - Increasing response and average gift within the campaign
- Brainstorming w/o Barriers
 - Open discussion regarding all possible items that might increase response or average gift or the combination
- Criteria Evaluation
 - Costs – what are the costs associated with each element
 - Risk / Reward – Are the ideas big and bold, is there too much risk associated with testing
- Prioritization of Ideas
 - Culling the list to a manageable test plan
- Test Plan Creation

Case Study:

Non-profit Renewals Test

<u>Test Elements</u>	<u>(-) Control</u>	<u>(+) New idea</u>
A Campaign appeal	Stand-alone appeal	Campaign branded appeal
B OE postage	NP indicia	Live NP stamp
C Insert	No insert	New insert
D Voucher	No voucher	Premium voucher
E Gift array	Control array	New gift array
F Reply envelope	No handwriting on RE	Handwritten reminder on RE
G Response channels	Mail-in response only	Multiple response options: mail-in, phone & e-mail
H Salutation	"Dear Friend" salutation	"Dear First Name" salutation
J Copy message	Control message	New message
K Package type	# 10 envelope	6 x 9 Envelope

Test Design and Data

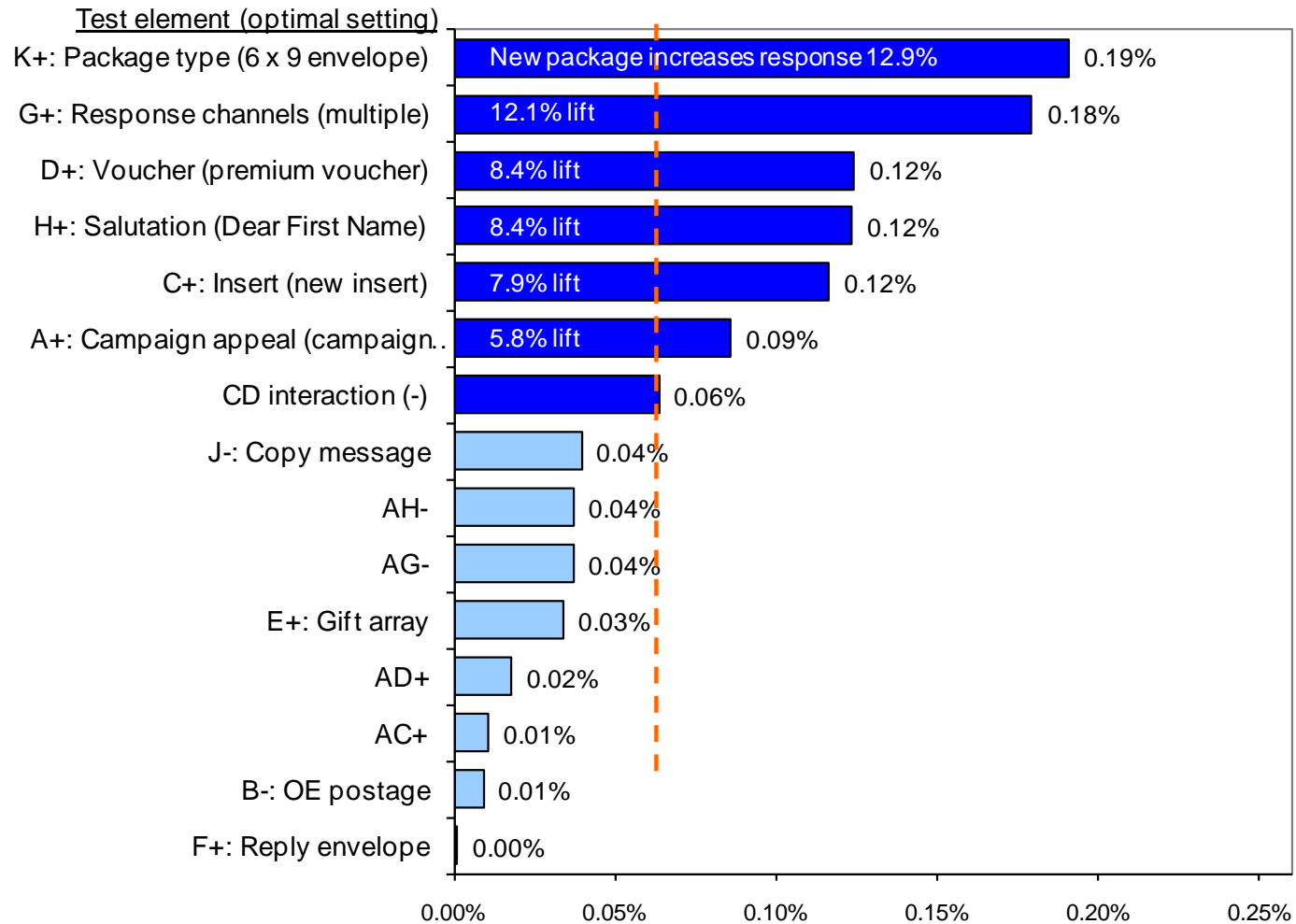
RENEWALS

Recipe	Campaign appeal											AC BG DF EH	AD BH CF EG	AF BE CD GH JK	AG BC DE FH	AH BD CE FG	Mailed	Responses	Response rate	Online responses	Resp rate (mail+online)
	A	B	C	D	E	F	G	H	J	K											
1	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	31,250	463	1.48%	0	1.48%	
2	+	-	-	-	-	+	+	+	+	+	-	-	+	+	+	31,250	585	1.87%	45	2.02%	
3	-	+	-	-	+	-	+	+	+	+	+	+	+	-	-	31,250	574	1.84%	62	2.04%	
4	+	+	-	-	+	+	-	-	-	-	-	-	+	-	-	31,250	519	1.66%	0	1.66%	
5	-	-	+	-	+	+	+	-	+	-	-	+	-	-	+	31,250	532	1.70%	57	1.88%	
6	+	-	+	-	+	-	-	+	-	+	+	-	-	-	+	31,250	660	2.11%	1	2.12%	
7	-	+	+	-	-	+	-	+	-	+	-	+	-	+	-	31,250	628	2.01%	2	2.02%	
8	+	+	+	-	-	-	+	-	+	-	+	-	-	+	-	31,250	564	1.80%	29	1.90%	
9	-	-	-	+	+	+	-	+	+	-	+	-	-	+	-	31,250	564	1.80%	1	1.81%	
10	+	-	-	+	+	-	+	-	-	+	-	+	-	+	-	31,250	650	2.08%	40	2.21%	
11	-	+	-	+	-	+	+	-	-	+	+	-	-	-	+	31,250	592	1.89%	49	2.05%	
12	+	+	-	+	-	-	-	+	+	-	-	+	-	+	-	31,250	585	1.87%	2	1.88%	
13	-	-	+	+	-	-	+	+	-	-	-	-	+	-	-	31,250	602	1.93%	52	2.09%	
14	+	-	+	+	-	+	-	-	+	+	+	+	+	-	-	31,250	636	2.04%	0	2.04%	
15	-	+	+	+	+	-	-	-	+	+	-	-	+	+	+	31,250	589	1.88%	2	1.89%	
16	+	+	+	+	+	+	+	+	+	-	-	+	+	+	+	31,250	639	2.04%	29	2.14%	

Effects: Response Rate

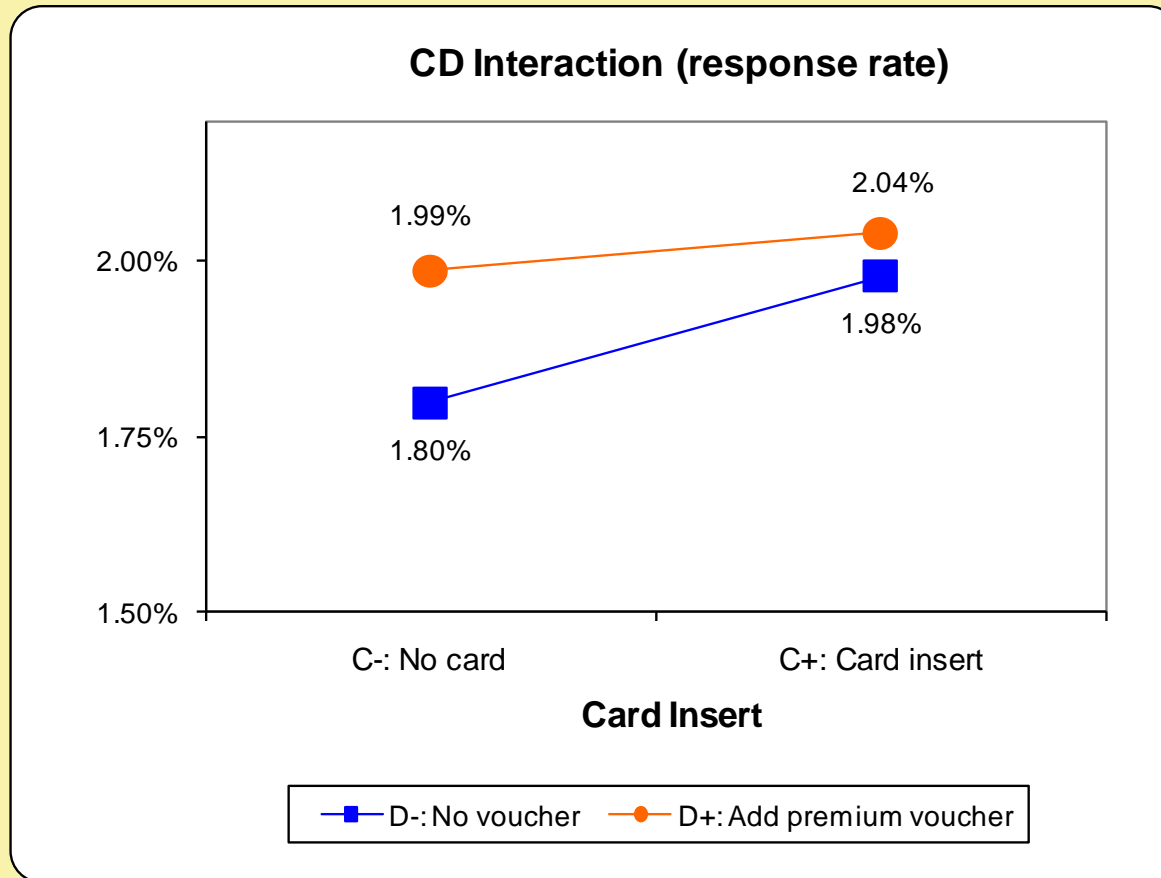
Response Rate (all channels): Non-profit Renewals campaign

control = 1.48%, optimal = 2.33% (57.4% increase)



Non-profit Renewals

Interaction of Card Insert and Premium Voucher



Non-profit Renewals Test

Predicted Improvement

Predicted response rate = **2.33%**
(57.4% improvement over the control)

Important Principles of Multivariable Testing

1. Orthogonal Test Designs
2. Balanced Test Designs
3. Controlled Variation
 - Avoid making changes outside of the test
 - Maintain consistent execution of test elements
 - Block out known sources of variation
4. Replication
5. Randomization (of test units and recipes)
6. Measurement precision
7. Process stability

Thank you attending our session.

Can we answer any questions?