Managing a Successful CBC Project
- from the Conjoint Experts

Agenda

▶ Proposal

- Uncover Business Objectives & Listen for Trigger Words

▶ Sample & Survey Design

- The Conjoint Exercise
  - Attributes & Levels
  - Design

▶ Analysis

▶ From One Consultant to Another

- Mistakes to Avoid
- Things to Consider

DURING THE PROPOSAL PHASE...

With any proposal, there are "Big Picture Issues" to address…

▸ What is the goal of this research? What are the key business questions to answer?

▸ Who is the target for this research? What sample size is available?

  ▪ Will we conduct analysis for the overall market or deep dive into segments?

▸ What is the budget and timing for deliverables?

In those conversations, there are certain trigger words that imply conjoint analysis is a good fit

▸ Optimization

▸ Pricing

▸ Preference

▸ Trade-off

▸ Simulate

▸ "Market share"

▸ Competition

▸ Cannibalization

Assuming conjoint analysis is the right approach, you may also want to ask…

▸ How do consumers make purchases in this space?

The answer to this question might help you determine which conjoint method to use.

For the sake of this webinar, we'll assume that Choice-Based Conjoint, or CBC, is the right method.

▸ A 2015 customer survey reports that CBC makes up about 80% of our users' conjoint-type research projects, with the remaining 20% split across ACBC, CVA, ACA, and MBC respectively.

▸ See the appendix for tips on which conjoint method to use and check out our Interactive Advisor online.

CBC Example

Tourism Example

If these were your only choices for vacation packages, which would you choose?
Or would you choose to not go on vacation?

# of tasks

1/10

Levels

Attributes

Concept

	San Francisco, CA	Washington, DC	Las Vegas, NV	None: I would prefer not to go on vacation rather than choose any of these.
Destination:	San Francisco, CA	Washington, DC	Las Vegas, NV	
Number of Nights:	5 nights	3 nights	7 nights	
Accomodation:	Luxury (5 star)	Upscale (3 star)	Deluxe (4 star)	
Hotel Type:	Boutique (with distinct style)	Resort (usually with spa, golf, etc.)	Business	
Car Rental:	Full-Size	None included	Compact	
Price (per person):	$1,380	$810	$1,500	

SURVEY DESIGN

Attributes & Levels

▶ How many attributes are necessary to answer the business objectives?

▶ Levels –

- Graphical or text representation?

- Qualitative attributes – Make sure levels are well-defined and easy to understand

- Quantitative attributes - Make sure levels list accommodates the entire range of prices. You can always interpolate, but should never extrapolate!

▶ Prohibitions? Alt-Specific? Conditional Pricing/Display?

Sample Size

▸ NTC/L >= 500 rule for quick sanity checks

- N= number of respondents

- T= number of tasks per respondent

- C= number of concepts per task

- L= largest number of levels for any one attribute in your study (or if you are wanting to examine interactions, L is the product of the number of levels between two attributes)

▸ Build your CBC and use the Advanced test
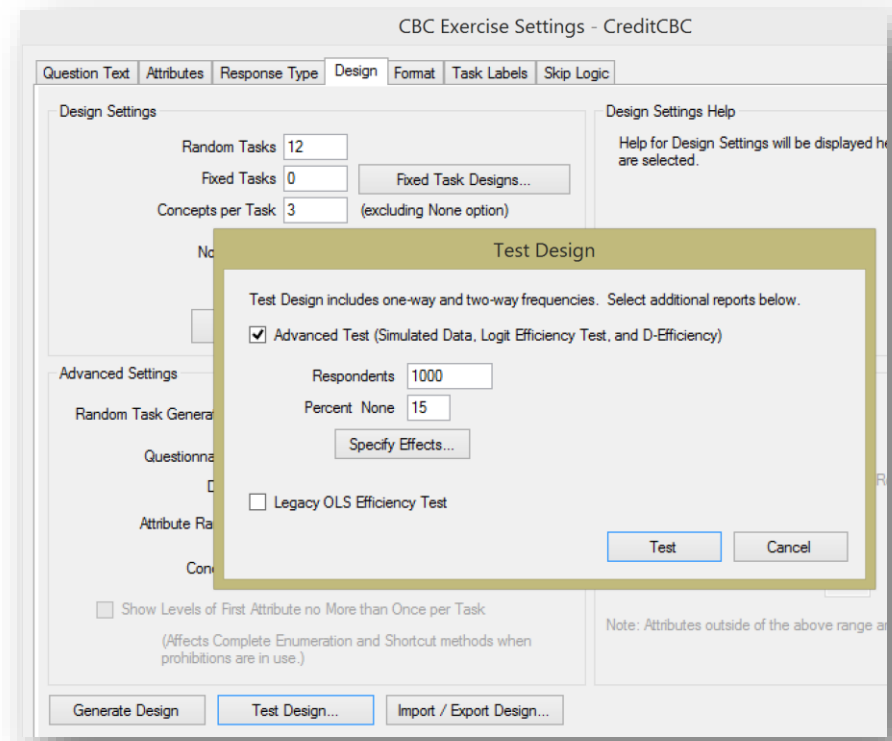
Base Case

▸ A base case typically reflects a current (or future) market scenario: your brand vs. the relevant competition. If there is no relevant competition, or your conjoint study was designed to model only your product, the base case may be a single product, reflecting a likely configuration.

▸ What is the base case scenario and the range of simulations from base case?

CBC LAYOUT

CBC Layout

▸ Do buyers consider just a few alternatives in the consideration set or many alternatives?

  ▪ Fewer concepts per task or more, depending on response

▸ Is it a low involvement purchase? Is there variety seeking? Are there many alternatives to consider?

  ▪ Fast moving consumer goods (FMCG)

  ▪ Typically more concepts per task

  ▪ Shelf-set CBC exercise

CBC Layout

▶ One None alternative? Multiple Opt-Out alternatives? Dual-Response None?

▶ Number of choice tasks?

- How much resolution do we need at the individual-level?
- What kind of overall sample size are we working with?

▶ What % of respondents will be using mobile phones to answer? Do we want to prohibit this?

CBC Design

▸ Overlap (level repeating across concepts within the same task) or minimal overlap?

▸ Worry about dominating concepts or not?

- Should we create corner prohibitions?

- Should we just massage the .CSV design file to root out the worst offenders in terms of potential domination? If doing this, no need to use more than 30 versions…you just cause more work for yourself.

Fixed Tasks (Holdout Tasks)

▶ Is there a base case that the client is quite interested in and will be the basis for the majority of the simulations we'll be doing?

- If so, maybe add it as a fixed task and include it in utility estimation?

▶ Is the client wary of conjoint and wants proof that it can predict held out tasks?

▶ Will we need holdouts to compare models (e.g. check for significant interaction effects)?

- Use of CBC/HB Model Explorer lets you get the benefits of holdouts without actually specifying fixed holdouts (due to its jackknifing across "random" tasks). You can use Model Explorer to investigate proper priors and interaction terms that practically matter.

- We need more than just a handful of fixed holdouts for within-sample model validation, and there are ways to design your study that allow us to get enough of them.

Covariates?

▸ Do we want to include some questions about past brand usage, budget threshold, motivations, or BYO questions to include as covariates in HB estimation?

▸ We could use these as covariates in HB estimation to bring out more heterogeneity in the data.

Example of Prohibitions

Cruise Line Example

Imagine you were in a position to take a cruise with your spouse, significant other, or friend in the next 2 years. If these were your only options, which would you choose?

1/21

Destination:	Alaska (sailing out of Seattle, WA)	Mexican Riviera (sailing out of Los Angeles, CA)	Western Caribbean (sailing out of Tampa, FL)	Mediterranean (sailing out of Barcelona, Spain)
Cruise Line:	Disney	Norwegian	Royal Caribbean	Princess
Number of Days:	7 days	10 days	7 days	8 days
Stateroom:	Ocean view stateroom, porthole window	Inside stateroom (no windows)	Balcony stateroom, sliding door to private balcony	Ocean view stateroom, porthole window
Ship Amenities/Age:	Fewer amenities, older ship	More amenities, newer ship	Fewer amenities, older ship	More amenities, newer ship
Price (per person):	$875 / person	$1,000 / person	$1,400 / person	$800 / person
	○	○	○	○

⬅ ➡

Prohibitions…

	Norwegian	Disney	Royal Caribbean	Princess	Holland Am	Carnival
Mexican Riv	X	X				
E Carrib	X			X	X	
W Carrib	X			X	X	
Alaska		X	X			X
Norway		X	X			X
Medit			X			X

▸ If we believe it is likely that there are significant interaction effects and if the client insists that these combinations are prohibited in reality (i.e. Disney cruise going to Alaska), what should we do?

Options:

▸ No prohibitions: Main effects only estimation
  ▪ (6-1)+(6-1)=10 parameters between attributes 1 & 2

▸ No prohibitions: Main effects + interactions
  ▪ (6-1)+(6-1)+(6-1)(6-1)=35 parameters between attributes 1 & 2

▸ Prohibitions: Main effects only estimation
  ▪ (6-1)+(6-1)=10 parameters between attributes 1 & 2

▸ Collapse into 1 attribute to accomplish prohibitions: Main effects only estimation
  ▪ (20-1)=19 parameters

Conclusion

▸ If we think that there's a good possibility of interactions between attributes 1 and 2, then it makes sense to do the prohibitions and collapse into a single factor – destination x cruise line.

	Norwegian	Disney	Royal Caribbean	Princess	Holland Am	Carnival
Mexican Riv	X	X	L7	L10	L14	L18
E Carrib	X	L4	L8	X	X	L19
W Carrib	X	L5	L9	X	X	L20
Alaska	L1	X	X	L11	L15	X
Norway	L2	X	X	L12	L16	X
Medit	L3	L6	X	L13	L17	X

Client Response

"But it just doesn't make sense to show certain cruise companies sailing to destinations that they don't serve. Many of the respondents will recognize that these just don't exist and I worry they will lose faith in the realism of the interview."

"I don't want the survey to look stupid/unrealistic to my boss."

"Wouldn't it help if we just had respondents see and respond only to the combinations of cruise line & destination that are real?"

What About Alt-Spec Effects?

▸ Make Attribute 1, Destination, the "Primary Attribute"

Mexican Riv	E Carrib	W Carrib	Alaska	Norway	Medit
Royal Carrib	Disney	Disney	Norwegian	Norwegian	Norwegian
Princess	Royal Carrib	Royal Carrib	Princess	Princess	Disney
Holland Am	Carnival	Carnival	Holland Am	Holland Am	Princess
Carnival					Holland Am.

▸ Notes:

- There are still 19 parameters to estimate
- But, this design will force level balance for the primary attribute
- If you were to make cruise line the primary attribute, then you force level balance for it

PRETEST

Internal Testing

▶ Take it first yourself. Play the part of a real respondent; don't just randomly click.

▶ Have other researchers around you, and your client, take the survey.

▶ Natural tendency is for researchers/clients to <u>underestimate</u> how much information respondents can manage well per CBC screen.

Sample Size for testing

▸ N=50 is usually enough. Check time to complete survey; time to complete CBC tasks, counts, aggregate logit, and HB (RLH & Percent Certainty).

DATA CLEANING

One Possible Approach…

▸ Remove speeders with HB's RLH output

- With four alternatives per task, the null RLH is 0.25 (i.e. random guess would say we'd get it right ¼ times, or 25%).

▸ Recommend more than one bad data quality check

- Straight-lining
- Total time to complete survey
- Different answers to same question

▸ Depending on the category/topic, I would assume ~10% bad data

- If the client expects you to report on n=200, you should budget for n=220 completes

ANALYSIS

Counts

▸ Examine counts results for different sample sizes (ex. n=200, n=50). Chi-Square stats should go up and differences between best and worst levels should increase as sample size decreases.

▸ Look for reversals in counts.

Logit

▸ Look for reversals in utilities.

▸ Make sure standard errors are all in line. Rule of thumb: 0.05 or less for main effects; 0.10 or less for interaction terms.

▸ Look at linearity of price. Convert to probabilities by taking the antilog. Look at elasticity (about unitary elastic -1.0).

Latent Class

▸ Look for reversals of utilities within segments…if so, then consider constraining and re-running.

▸ Examine the scree plot of BIC and find where the plot levels off for a recommendation of number of segments.

▸ Compare across segments using the rescaled utilities.

Latent Class

▶ What's your goal?

  ▪ Managerial interpretability for strategy?

  ▪ Capturing heterogeneity and improving predictability of your model?

▶ If interpretability, then maybe 5 or 6 segments would be about the right choice.

▶ If predictive ability, then maybe 20 or 30 segments…maybe ensembles (multiple solutions that you combine to make a single prediction).

HB

▶ HB, look at Percent Certainty (pseudo R-squared).

- Based on 25 commercial and methodological CBC studies (prior var=1, DF=5), I find average Percent Certainty of ~68%, min~57% and max~77%.

▶ Look at RLH at the individual level across respondents.

- If respondents had to choose among 4 concepts, expected RLH given chance is 1/4 or 25%.

▶ If you have overnight to run it, use Model Explorer to check if default priors work well…and for significant interaction terms.

▶ If you have additional time, think about valuable covariates.

- Covariates are dummy-coded (K-1 coding).
- Don't try to use more than ~10 coded columns in your covariates independent variable matrix or things get really slow.

Beyond the Defaults

▸ Most commonly changed settings:

- Constraints

- Covariates

- Omitting tasks as warm-ups

- Adding interaction terms

- Prior variance, Prior D.F.

    - Use the Model Explorer

    - Use meta analysis recommendations from Orme and Williams 2016 paper "*What are the Optimal HB Priors Settings for CBC and MaxDiff Studies*"

- Skip factor for printing (makes runs with lots of parameters go faster)

HB Stuff I Never Change

▶ Acceptance Rate

▶ Custom Prior Alpha

▶ Code Variables Using Dummy coding

And don't forget the Simulator!

▶ The simulator is the most powerful output

- It is a "choice laboratory" for testing alternatives
- Results are expressed in terms that make sense to managers
- Helps to answer strategic questions:
  - At what price will people switch to a competitor?
  - Can we modify our product to reduce cost while maintaining share?
  - Should we launch a high-end product or a budget model (or both)?
  - Will the new product cannibalize our own sales?

▶ Examining utilities and importances only gets you so far

- Average utilities cannot tell the whole story
- Fallacy of Division

Preferred Color?

▸ Consider the following utilities:

	Blue	Red	Yellow
Respondent #1	50	40	10
Respondent #2	0	65	75
Respondent #3	40	30	20
Average	**30**	**45**	**35**

▸ Red has the highest <u>average</u> preference

▸ But, does any <u>one</u> respondent prefer red?

Chosen Color?

▸ Each respondent's preferred color:

	Blue	**Red**	**Yellow**	**"Choice"**
Respondent #1	50	40	10	**Blue**
Respondent #2	0	65	75	**Yellow**
Respondent #3	40	30	20	**Blue**
Average	**30**	**45**	**35**	

▸ Blue "chosen" twice, Yellow once

Choose the appropriate simulation method

First Choice Rule ("Maximum Utility")	Logit Probability Rule ("Share of Preference")	Randomized First Choice
• Pro - Simple and easy to understand – assumes respondent chooses the product with the highest utility • Con – Learn which product is preferred but don't capture anything about relative preferences • Works well with large sample sizes and in a "winner take all" model (e.g. automobiles, refrigerators)	• Pro – Probability of purchase; Buyers never purchase with 100% certainty and this method takes that into account (i.e. error in model, out-of-stock, variety seeking category) • Con – Independence of Irrelevant Alternatives or IIA (Red-Bus/Blue-Bus Problem)	• Pro – Splits shares but reflects more accurate substitution effects for similar products; is tunable in terms of scale and product similarity • Con – If a correction for similarity is applied to price, your demand curve can be distorted. (Solution - Turn off correlated error for Price attribute); If simulating many products (20+) some shares can become so small the random component introduced by RFC can make those shares imprecise. • Note – If using Excel simulator, RFC plug-in does not work on a Mac

Additional considerations

▸ Think about tuning down the scale factor to better fit market share information or "holdout choices".

 ▪ As $\lambda \rightarrow \infty$, shares become First Choice (best alternative gets 100% share)

 ▪ As $\lambda \rightarrow 0$, shares flatten to become equal

▸ Since a conjoint simulator assumes equal distribution, awareness, etc., use even more advanced techniques to correct for this.

 ▪ Check out the new Choice Simulator offered by Sawtooth Software, available to all subscription users

 ▪ Want to know more about Market Simulations? Check out the webinar series here

FROM ONE CONSULTANT TO ANOTHER...

Avoid these common mistakes…

▶ Don't tell your client that the results will "directly tells us how many people will buy this product".

- Conjoint gives us preference share, not necessarily MARKET share and preference share is only partially indicative of real market results.

▶ Conjoint analysis does not assess how good or bad a feature is, but rather how preferred one feature is to another .

- Just like in MaxDiff, the feature with the highest part-worth, doesn't necessarily mean it's the best. It could be, as one of my client's would say, *the smartest kid in summer school* – or the *best of the worst*.

▶ Just like any quant survey, weigh the pros and cons of increasing sample.

- Yes, you will minimize your confidence interval and standard errors by increasing your sample size, but it could come at exorbitant sample costs.

And consider this…

▸ Conjoint analysis is a great technique for pricing research, but be mindful that there are typically many more factors that could impact price than allotted for in the exercise or model.

▸ Build a simulator! If you only look at the averages, or each feature independent of the other features, you may be missing out on the bigger picture.

▸ Be careful when reporting importances – they're relative!

▪ Including (or omitting) a very popular or unpopular level on one attribute will alter the "importance" of every other attribute!

Want to learn more?

▶ <u>Singapore, April 19-21</u>

  ▪ Track 1: CBC/MaxDiff/ACBC hands-on workshop

  ▪ Track 2: Advanced concepts track

▶ <u>Huntington Beach, CA May 22-24</u>

  ▪ 3 day choice modeling workshop

▶ <u>Park City, UT July 17-21</u>

  ▪ 3 day choice modeling workshop

  ▪ Becoming an Expert in Conjoint Analysis Seminar NEW

  ▪ Menu-Based Choice (MBC) Workshop

QUESTIONS?

Keith Chrzan
SVP, Sawtooth Analytics
keith@sawtoothsoftware.com

Megan Peitz
Ingenuity Ambassador
megan@sawtoothsoftware.com

www.sawtoothsoftware.com
+1 801 477 4700
@sawtoothsoft

APPENDIX

Which conjoint method should you use?

Method	MaxDiff	CVA	ACA	CBC	ACBC	MBC
Sample Size	Small	Small	Small	Large	Small	Very Large
Attributes	None	Up to 6-7	Up to 30	Up to 6-7*	Any	Any
Levels	30-40, but up to 500	Up to 4-5	Up to 15, usually <6	Up to 15, usually <6	Any	Any
Pricing	N/A	Yes, but limited	Not recommended	Yes	Yes (Summed)	Yes
Complexity	No	No	No	Yes	Some	Yes
Fielding	Paper or Computer	Paper or Computer	Computer	Paper or Computer	Computer	Paper or Computer
Typical Use Cases	• List of brands • Positioning Statements • Specific product concepts • Flavors, etc.	• Measure purchase likelihood or other discrete scale elicitation	• Measure purchase likelihood	• Competitive scenarios where choice is among multiple alternatives • Pricing studies • Alternative Specific designs • Chip allocation studies • Shelf-facing studies	• Pricing studies • Finding the best product • Respondent determines which attributes/levels are shown	• Multi-part decisions • Complex models • Bundling • Mixed designs