Proceedings of the Sawtooth Software Conference

April 2021

ENHANCE CONJOINT WITH A BEHAVIORAL FRAMEWORK

PETER KURZ
STEFAN BINNER
BMS - MARKETING RESEARCH + STRATEGY

BEHAVIORAL FRAMEWORK

As shoppers process information and act on it, they are not simple stimulus-response robots. Creating a behavioral framework prior to answering choice tasks therefore helps respondents select from choice tasks as if they were in a real purchase situation. If price and assortment changes are the focus of the research, it is particularly important to understand shopper perceptions of prices and values. Again, a behavioral framework is useful for interpreting consumer decisions, as simulated by the results of the choice model, in the appropriate context.

To create such a behavioral framework, prior to each conjoint exercise, we apply nine standardized, binary "Behavioral Calibration Questions" regarding each respondent's individual shopping behavior for the focal category. Based on principles from behavioral economics, these questions help consumers recall their usual buying habits. "Behavioral Calibration Questions" are also used to describe the context of consumer choices, including how purchase decisions are made within a specific category, as they reveal typical patterns of buying habits, purchase repertoires, and brand value perceptions, as well as price knowledge.

BEHAVIORAL CALIBRATION QUESTIONS

We use the derived contextual information about each respondent's individual disposition towards brand and price knowledge (or lack thereof), past behavior, and perceptions within the category in our analysis. Retrieving a prior shopping situation and their individual dispositions helps consumers to make decisions in the following choice experiment. Currently, the set contains nine "Behavioral Calibration Questions" (semantic differentials) with respect to buying habits along three dimensions: brand, price, and innovation.

"Behavioral Calibration Questions" are used in our research context for several purposes:

- to establish a behavioral framework before respondents answer the choice tasks,
- as covariates in the hierarchical Bayes estimation process, and
- as segmentation/filter variables in the choice simulator.

Furthermore, we store the responses to the "Behavioral Calibration Questions" in a benchmark database to anchor further conjoint studies in the different product categories.

OUR STANDARD BEHAVIORAL CALIBRATION QUESTIONS

In each of our conjoint questionnaires, we combine the binary questions with our nine semantic differentials and ask respondents which of two statements (left or right) is more related to their last shopping trip.

We would like to learn a few things about you and your general thoughts, feelings, and opinions when it comes to home upkeep, construction adhesives.
Please read each pair of statements. For each pair, please indicate whether you agree with the statement on the left or the statement on the right more, and how much more.
If both statements describe your opinion well, choose the one that best describes you. If neither seems to describe you well, choose the one that comes the closest.
Select one response for each.

	Agree Left	Agree Right	
I think that brands differ a lot	○	○	I think that all brands are more or less the same
I always know exactly what brand I'm going to buy before I enter the shop	○	○	I decide what brand I'm going to buy when I'm standing in front of the shelf
I always buy the brand I bought last time	○	○	I switch between different brands
I compare prices very carefully before I make a choice	○	○	To be honest, I compare prices only superficially
I always search for special offers first	○	○	Special offers are not the first thing I look out for
I always know the price of the products I buy	○	○	I never really know what products cost
I'm always interested in new products	○	○	I prefer to stick to what I know
I think that products in this category need to be improved	○	○	I'm completely satisfied with the products as they are
I find it easy to make the right choice for me	○	○	I find it very difficult to make the right choice for me

Example from R&D study in US (2020, context: construction adhesives)[1]

Of the nine semantic differentials, three are related to the "Role of Price," three to the "Role of Brand," and three to the "Role of Innovation." This approach allows respondents to recall past behavior when buying a product in this category.

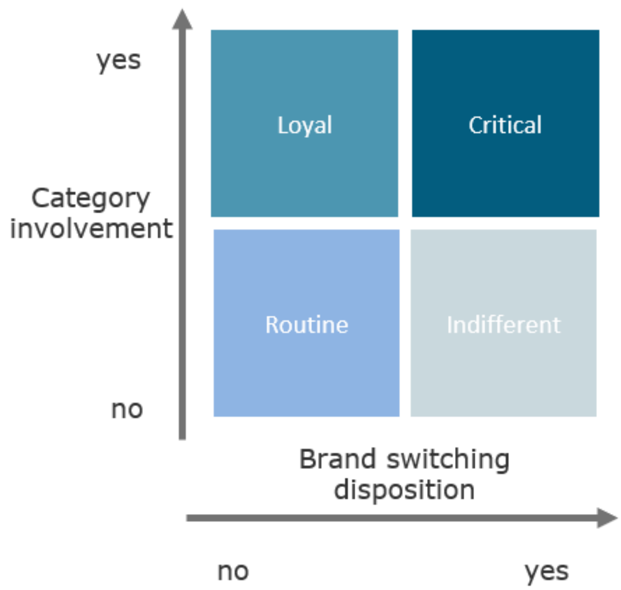
Over the years, we have adapted sets of nine semantic differentials to different product categories, as not all statements behave similarly in distinct shopping situations or categories. For example, when buying a new car, virtually no respondents would answer, "I never really know what the car I buy would cost." In this situation, one needs a question such as "I never really know what the competitive brands cost; I more or less compare only within my preferred brand." Such adaptations for each category are necessary to produce a valid framing of respondents from different target groups.

[1] We are uncertain of the origin of these questions; we first encountered them in a segmentation approach from Research International in 2008. In this approach, the questions were asked as scale questions and used to derive consumer segments.

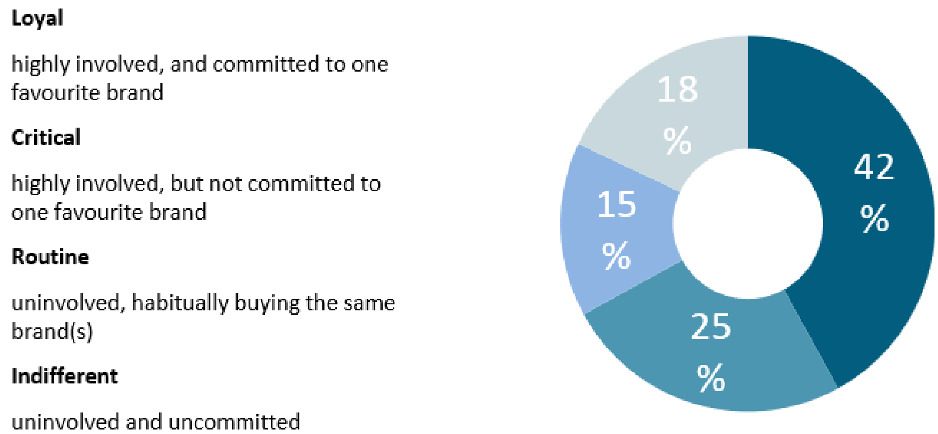
	FMCG	Media (Pay TV, VoD)	Mobiltelefone	PC_Notebook	Fernseher	Auto	Mobiltarife	Telekomtarife	Stromtarife	Girokonto	Spar-Anlagen	Krankenkasse	KFZ-Versicherung	Airlines	Pauschalreisen
1 - Markendifferenzierung	Die Marken unterscheiden sich stark voneinander	Die Anbieter unterscheiden sich stark voneinander	Die Marken unterscheiden sich stark voneinander	Die Marken unterscheiden sich stark voneinander	Die Marken unterscheiden sich stark voneinander	Die Automarken unterscheiden sich stark voneinander	Die Anbieter unterscheiden sich stark voneinander	Die Anbieter unterscheiden sich stark voneinander	Die Anbieter unterscheiden sich stark voneinander	Die Banken unterscheiden sich stark voneinander	Die Anbieter unterscheiden sich stark voneinander	Die Krankenkassen unterscheiden sich stark voneinander	Die Versicherungen unterscheiden sich stark voneinander	Die Airlines unterscheiden sich stark voneinander	Die Veranstalter unterscheiden sich stark voneinander
	Alle Marken sind mehr oder weniger gleich	Alle Anbieter sind mehr oder weniger gleich	Alle Marken sind mehr oder weniger gleich	Alle Marken sind mehr oder weniger gleich	Alle Marken sind mehr oder weniger gleich	Alle Automarken sind mehr oder weniger gleich	Alle Anbieter sind mehr oder weniger gleich	Alle Anbieter sind mehr oder weniger gleich	Alle Anbieter sind mehr oder weniger gleich	Alle Banken sind mehr oder weniger gleich	Alle Anbieter sind mehr oder weniger gleich	Alle Krankenkassen sind mehr oder weniger gleich	Alle Versicherungen sind mehr oder weniger gleich	Alle Airlines sind mehr oder weniger gleich	Alle Reiseveranstalter sind mehr oder weniger gleich
2 - Wichtigkeit der Marke bei der Kaufentscheidung	Ich weiß jedes Mal genau, welche Marke ich kaufen werde	Ich weiß jetzt schon, welchen Anbieter ich wähle	Ich weiß jetzt schon, welche Marke ich kaufen werde	Ich weiß jetzt schon, welche Marke ich kaufen werde	Ich weiß jetzt schon, welche Marke ich kaufen werde	Ich weiß jetzt schon, welche Marke ich kaufen werde	Ich weiß jetzt schon, welchen Anbieter ich wähle	Ich weiß jetzt schon, welchen Anbieter ich wähle	Ich weiß jetzt schon, welchen Anbieter ich wähle	Ich weiß jetzt schon, welche Bank ich wähle	Ich weiß jetzt schon, welchen Anbieter ich wähle	Ich weiß jetzt schon, welche Krankenkasse ich wähle	Ich weiß jetzt schon, welche Versicherung ich wähle	Ich weiß immer genau, welche Airline ich buchen werde	Ich weiß genau, welchen Veranstalter ich wählen werde
	Ich entscheide mich erst beim Kauf für eine Marke	Ich entscheide mich erst bei Abschluss für den Anbieter	Ich entscheide mich erst beim Kauf für eine Marke	Ich entscheide mich erst beim Kauf für eine Marke	Ich entscheide mich erst beim Kauf für eine Marke	Ich entscheide mich erst unmittelbar vor dem Kauf	Ich entscheide mich erst bei Abschluss für den Anbieter	Ich entscheide mich erst bei Abschluss für den Anbieter	Ich entscheide mich erst bei Abschluss für den Anbieter	Ich entscheide mich erst kurz vor Abschluss für d. Bank	Ich entscheide erst kurz vor Abschluß über den Anbieter	Ich entscheide mich erst kurz vor Abschluss	Ich entscheide mich erst kurz vor Abschluss	Ich entscheide mich erst bei Buchung für eine Airline	Ich entscheide mich erst bei Buchung für einen Anbieter
3 - Markenloyalität / Autopilot	Ich kaufe immer die gleiche Marke	Ich wähle immer den gleichen Anbieter	Ich kaufe immer die gleiche Marke	Ich kaufe immer die gleiche Marke	Ich kaufe immer die gleiche Marke	Ich wähle immer die gleiche Automarke	Ich wähle immer den gleichen Anbieter	Ich wähle immer den gleichen Anbieter	Ich wähle immer den gleichen Anbieter	Ich wähle immer die gleiche Bank	Ich wähle immer den gleichen Anbieter	Ich wähle immer die gleiche Krankenkasse	Ich wähle immer die gleiche Versicherung	Ich buche fast immer die gleiche Airline	Ich buche fast immer den gleichen Reiseveranstalter
	Ich wechsle zwischen verschiedenen Marken	Ich wechsle zwischen verschiedenen Anbietern	Ich wechsle zwischen verschiedenen Marken	Ich wechsle zwischen verschiedenen Marken	Ich wechsle zwischen verschiedenen Marken	Ich wechsle zwischen verschiedenen Automarken	Ich wechsle zwischen verschiedenen Anbietern	Ich wechsle zwischen verschiedenen Anbietern	Ich wechsle zwischen verschiedenen Anbietern	Ich wechsle zwischen verschiedenen Banken	Ich wechsle zwischen verschiedenen Anbietern	Ich wechsle zwischen verschiedenen Krankenkassen	Ich wechsle zwischen verschiedenen Versicherungen	Ich wechsle zwischen verschiedenen Airlines	Ich wechsle zwischen verschiedenen Reiseveranstaltern
4 - Preisvergleiche	Ich vergleiche Preise sehr genau, bevor ich kaufe	Ich vergleiche Preise sehr genau, bevor ich abschließe	Ich vergleiche Preise sehr genau, bevor ich kaufe	Ich vergleiche Preise sehr genau, bevor ich kaufe	Ich vergleiche Preise sehr genau, bevor ich kaufe	Ich vergleiche Preise sehr genau, bevor ich kaufe	Ich vergleiche Preise sehr genau, bevor ich abschließe	Ich vergleiche Preise sehr genau, bevor ich abschließe	Ich vergleiche Preise sehr genau, bevor ich abschließe	Ich vergleiche Konditionen sehr genau vor Abschluss	Ich vergleiche Konditionen sehr genau vor Abschluss	Ich vergleiche Kosten sehr genau vor Abschluss	Ich vergleiche Kosten sehr genau vor Abschluss	Ich vergleiche Preise sehr genau, bevor ich buche	Ich vergleiche Preise sehr genau, bevor ich buche
	Ehrlich gesagt vergleiche ich Preise eher oberflächlich	Ehrlich gesagt vergleiche ich Preise eher oberflächlich	Ehrlich gesagt vergleiche ich Preise eher oberflächlich	Ehrlich gesagt vergleiche ich Preise eher oberflächlich	Ehrlich gesagt vergleiche ich Preise eher oberflächlich	Ehrlich gesagt vergleiche ich Preise eher oberflächlich	Ehrlich gesagt vergleiche ich Preise eher oberflächlich	Ehrlich gesagt vergleiche ich Preise eher oberflächlich	Ehrlich gesagt vergleiche ich Preise eher oberflächlich	Ich vergleiche Konditionen eher oberflächlich	Ich vergleiche Konditionen eher oberflächlich	Ehrlich gesagt vergleiche ich Kosten eher oberflächlich	Ehrlich gesagt vergleiche ich Kosten eher oberflächlich	Ehrlich gesagt vergleiche ich Preise eher oberflächlich	Ehrlich gesagt vergleiche ich Preise eher oberflächlich
5 - Sonderangebote	Ich suche immer zuerst nach Sonderangeboten	Ich suche immer zuerst nach Sonderangeboten	Ich suche immer zuerst nach Sonderangeboten	Ich suche immer zuerst nach Sonderangeboten	Ich suche immer zuerst nach Sonderangeboten	Ich suche immer zuerst nach Sonderangeboten	Ich suche immer zuerst nach Sonderangeboten	Ich suche immer zuerst nach Sonderangeboten	Ich suche immer zuerst nach Sonderangeboten	Ich suche immer zuerst nach Sonderkonditionen	Ich suche immer zuerst nach Sonderkonditionen	Ich suche immer zuerst nach Sonderkonditionen	Ich suche immer zuerst nach Sonderkonditionen	Ich suche immer zuerst nach Sonderangeboten	Ich suche immer zuerst nach Sonderangeboten
	Sonderangebote sind nicht das Erste, wonach ich suche	Sonderangebote sind nicht das Erste, wonach ich suche	Sonderangebote sind nicht das Erste, wonach ich suche	Sonderangebote sind nicht das Erste, wonach ich suche	Sonderangebote sind nicht das Erste, wonach ich suche	Sonderangebote sind nicht das Erste, wonach ich suche	Sonderangebote sind nicht das Erste, wonach ich suche	Sonderangebote sind nicht das Erste, wonach ich suche	Sonderangebote sind nicht das Erste, wonach ich suche	Sonderkonditionen sind nicht das Erste wonach ich suche	Sonderkonditionen sind nicht das Erste wonach ich suche	Sonderkonditionen sind nicht das Erste wonach ich suche	Sonderkonditionen sind nicht das Erste wonach ich suche	Sonderangebote sind nicht das Erste, wonach ich suche	Sonderangebote sind nicht das Erste, wonach ich suche
6 - Preiskenntnis	Ich kenne genau die Preise der Produkte, die ich kaufe	Ich kenne die Preise der für mich interessanten Angebote	Ich kenne die Preise der für mich interessanten Modelle	Ich kenne die Preise der für mich interessanten Modelle	Ich kenne die Preise der für mich interessanten Modelle	Ich kenne die Preise der für mich interessanten Modelle	Ich kenne die Preise der für mich interessanten Tarife	Ich kenne die Preise der für mich interessanten Tarife	Ich kenne die Preise der für mich interessanten Tarife	Ich kenne genau die Konditionen meines Girokontos	Ich kenne genau die Konditionen meiner Anlagen	Ich kenne genau die Kosten meiner Krankenkasse	Ich kenne genau die Kosten meiner Versicherung	Ich kenne mich bei Flugpreisen gut aus	Ich kenne mich bei Reisepreisen gut aus
	Ich weiß nie genau, was die Produkte kosten	Ich weiß nie genau, was die einzelnen Angebote kosten	Ich weiß nie genau, was die einzelnen Modelle kosten	Ich weiß nie genau, was die einzelnen Modelle kosten	Ich weiß nie genau, was die einzelnen Modelle kosten	Ich weiß nie genau, was die einzelnen Modelle kosten	Ich weiß nie genau, was die einzelnen Tarife kosten	Ich weiß nie genau, was die einzelnen Tarife kosten	Ich weiß nie genau, was die einzelnen Tarife kosten	Ich weiß nie genau die Konditionen meines Girokontos	Ich weiß nie genau die Konditionen meiner Anlagen	Ich weiß nie genau, was meine Krankenkasse kostet	Ich weiß nie genau, was meine Versicherung kostet	Ich weiß nie genau, was die Flüge kosten	Ich weiß nie genau, was die Reisen kosten
7 - Interesse an neuen Innovationen	Ich bin immer an neuen Produkten interessiert	Ich bin immer an neuen Angeboten interessiert	Ich bin immer an neuen Mobiltelefonen interessiert	Ich bin immer an neuen Computern interessiert	Ich bin immer an neuen Fernsehern interessiert	Ich bin immer an neuen Autos interessiert	Ich bin immer an neuen Tarifen interessiert	Ich bin immer an neuen Tarifen interessiert	Ich bin immer an neuen Tarifen interessiert	Ich bin immer an neuen Girokonto-Angeboten interessiert	Ich bin immer an neuen Anlage-Angeboten interessiert	Ich bin immer an neuen Angeboten interessiert	Ich bin immer an neuen Angeboten interessiert	Ich bin immer an neuen Flug-Angeboten interessiert	Ich bin immer an neuen Pauschalangeboten interessiert
	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne	Ich bleibe lieber bei dem, was ich schon kenne
8 - Zufriedenheit mit Status Quo	Ich glaube, die Produkte müssen verbessert werden	Ich denke, das Angebot muss noch verbessert werden	Ich denke, das Angebot muss noch verbessert werden	Ich denke, das Angebot muss noch verbessert werden	Ich denke, das Angebot muss noch verbessert werden	Ich denke, das Angebot muss noch verbessert werden	Ich denke, das Angebot muss noch verbessert werden	Ich denke, das Angebot muss noch verbessert werden	Ich denke, das Angebot muss noch verbessert werden	Ich denke, das Angebot muss noch verbessert werden	Ich denke, das Angebot muss noch verbessert werden	Ich denke, das Angebot muss noch verbessert werden	Ich denke, das Angebot muss noch verbessert werden	Ich denke, das Angebot muss noch verbessert werden	Ich denke, das Angebot muss noch verbessert werden
	Ich bin voll zufrieden mit den Produkten, wie sie sind	Ich bin voll zufrieden mit den Angeboten, wie sie sind	Ich bin voll zufrieden mit Mobiltelefonen, wie sie sind	Ich bin voll zufrieden mit den Computern, wie sie sind	Ich bin voll zufrieden mit den Fernsehern, wie sie sind	Ich bin voll zufrieden mit den Autos, wie sie sind	Ich bin voll zufrieden mit den Tarifen, wie sie sind	Ich bin voll zufrieden mit den Tarifen, wie sie sind	Ich bin voll zufrieden mit den Tarifen, wie sie sind	Ich bin voll zufrieden mit den Girokonten, wie sie sind	Ich bin voll zufrieden mit den Anlagen, wie sie sind	Ich bin voll zufrieden mit Krankenkassen, wie sie sind	Ich bin voll zufrieden mit KFZ-Tarifen wie sie sind	Ich bin voll zufrieden mit dem Flugangebot, wie es ist	Ich bin voll zufrieden mit dem Reiseangebot, wie es ist
9 - Sicherheit der Wahl vs. Ambivalenz	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht	Die richtige Wahl für mich zu treffen, fällt mir leicht
	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer	Die richtige Wahl für mich zu treffen, fällt mir schwer

BEHAVIORAL CALIBRATION QUESTIONS

The first insight we derive from these nine questions is the identification of four respondent segments. Two semantic differentials, "brands differ a lot" and "always buy the same brand," can be used to classify consumers according to "Brand Loyalty" and "Category Involvement," thereby providing useful insights about the product category in general. Quantifying these different buyer segments is useful for identifying the best-performing strategies for products under investigation in the choice model.



This classification mostly refers to consumers' attitudes towards brands. A consumer classified as "Indifferent" is not necessarily indifferent to other attributes. Segment names should not be taken too literally, as classifications represent only a rough outline of consumers' personalities. For instance, a "Loyal" consumer may actually have a relevant set of two or three brands. What makes her a "Loyal" consumer is her self-perception as someone who sticks to her brand(s) (as opposed to consumers who are indifferent to brand), and her belief that the difference between her brand(s) and others really matters.



The figure above shows an example of the distribution of the four consumer types within the "laundry detergent" category. We see this as an initial blueprint for each category to begin interpreting the results of our choice models. Based on the benchmark from past studies, the client can easily determine how target consumers think about this category.

Experience with Behavioral Calibration Questions

Asking the nine "Behavioral Calibration Questions" before our choice exercise helps respondents to recall their behavior during their last shopping trip in a specific category. Therefore, we assume that the nine questions improve their decision-making process in the subsequent choice exercise, supporting a realistic answering behavior comparable to real shopping situations. Therefore, this approach helps generate more realistic data. Using the derived shopper classifications as segmentation variables in the choice simulator provides deeper insights into respondents' preference structure. Based on our findings from numerous conjoint exercises, we learned that answering the nine questions results in better "Share of Choice" estimates as compared with conjoint exercises performed without the calibration questions. Furthermore, part-worth estimates, which include the "Behavioral Calibration Questions" as covariates, further improve share predictions against holdout samples (ensembles with the questions and other covariates offer marginal improvement in results).

Empirical Validation of Behavioral Calibration Questions

For validation purposes, we conducted nine empirical R&D studies over the last two years, in which we asked 50% of respondents the nine "Behavioral Calibration Questions" prior to answering the choice model, whereas the other 50% answered the choice model without being exposed to the semantic differentials prior to the choice tasks.

We addressed the following hypotheses in this paper:

- The framing offered by the "Behavioral Calibration Questions" results in improved answering behavior among our respondents, leading to part-worth estimates that are more stable and valid.
- Adding the answers to the "Behavioral Calibration Questions" as covariates in the HB estimation further improves the part-worth estimates.
- Using the questions as filter/segment variables in the choice simulator provides additional insights in the data as "Shares of Choice"; elasticities differ according to the derived segments based on roles of brand, price, and innovation.

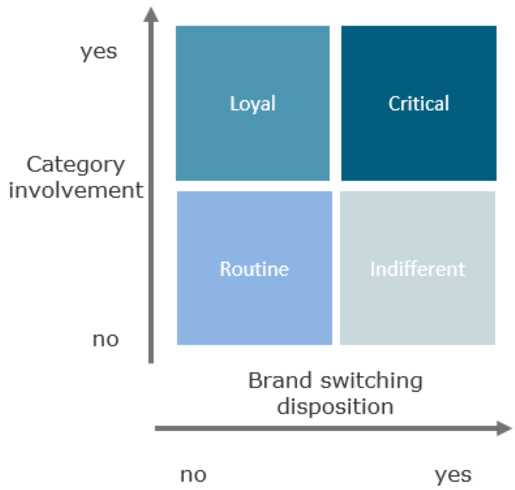
All studies were conducted with respondents recruited from online access panels in 2019 and 2020 and the samples were split as outlined above (i.e., Behavioral Calibration Questions shown or not). The studies varied in terms of categories, number of attributes, number of levels, number of concepts, and number of tasks. Sample sizes depended on the number of parameters to be estimated and varied between 250 and 1,000 respondents.

Only one study ("super glue") differed slightly from the others, as we conducted 4 sample splits to create an opportunity to validate the estimation samples with separate validation samples. (For the two estimation samples, n=500 interviews, and n=250 interviews for the two validation samples.) These four split cells enable cross-validation of the part-worth estimates derived from asking or not asking the "Behavior Calibration Questions" and including or excluding them from the hierarchical Bayes estimation.

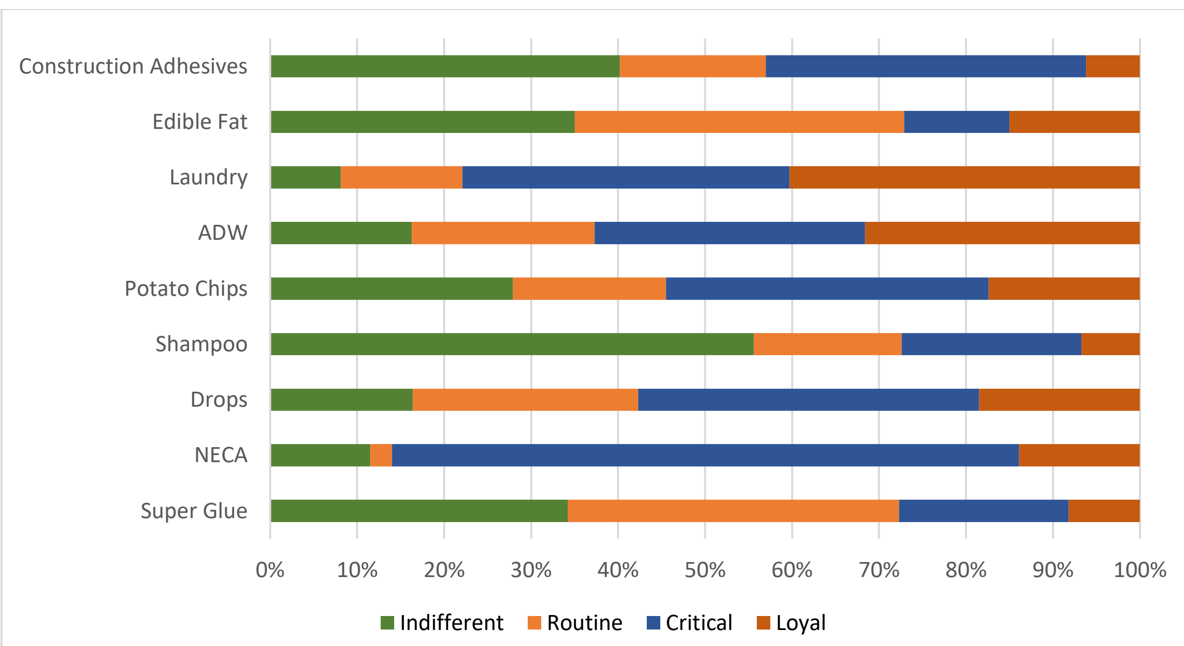
Project	N=	Attributes	Levels	Tasks/Concept per Task	Model Specifics	Covariates
Detergent ADW	1006	6	10+2*2+2*3+6	12/8+None	502/504	Socio-demographic, Purchase Behavior
Construction adhesives	510	30	6	15/4+None	250/260	Socio-demographic Purchase Behavior
Drops	1030	3	47+3+47*3	15/12+None	530/500	Socio-demographic, Purchase Behavior
Edible Fat	2030	12	12+6*2+3*2+2*5	15/6+None	1030/1000	Socio-demographic, Purchase Behavior
None Electric Air freshener	500	5	7+2*9+2+7	15/5+None	250/250	Socio-demographic, Purchase Behavior
Hair Shampoo	1016	46	96+ 40*2 +3*5	15/12+None	509/507	Socio-demographic, Purchase Behavior
Potato Chips	800	38	45+2+2+30*5	15/5+None	400/400	Socio-demographic, Purchase Behavior
Laundry Detergent	980	16	96 + 15*5	15/12+None	580/400	Socio-demographic, Purchase Behavior
Super Glue	1500	23	22+ 22*5	15/12+None	500/500/250/250	Socio-demographic, Purchase Behavior

Buying Habits and Involvement

The four segments derived from the "Behavioral Calibration Questions" have real potential to differentiate between categories and to identify promising strategies. For example, a significant proportion of "indifferent" consumers may have a larger effect on strategies for new product development, compared with a large share of "critical" or "loyal" consumers.



A comparison of the nine empirical studies shows that the different product categories have different compositions in terms of the four consumer segments. For example, in the "super glue" category, the study identifies an equal number of "Indifferent" and "Routine" consumers, whereas only a small proportion are "Loyal." In contrast, in the category "laundry," "Loyal" customers are by far the largest group, followed by the "Critical," "Indifferent," and "Routine" consumers. Furthermore, the "non-electric air freshener" (NECA) category has by far the highest share of "Critical" consumers. In the context of introducing new, innovative products, this category seems to offer many more opportunities as compared with the "super glue" category.



Behavioral Roles

The three behavioral roles represent a second possible usage of the nine calibration questions to understand the behavior of the respondents during shopping trips in a particular category. We found these roles helpful for interpreting the "Share of Choice" from simulations. They allow deeper insights into the reasons respondents behave differently in their choices.

The three roles derived from the "Behavioral Calibration Questions" are:

- Role of Price
- Role of Brand
- Role of Innovation

Each is represented by three semantic differentials that ask (in a binary manner) whether the left or the right statement better corresponds to the respondent's last shopping trip in that category.

The following results, based on nine empirical studies, demonstrate the diversity of behavior within the different categories.

Role of Price

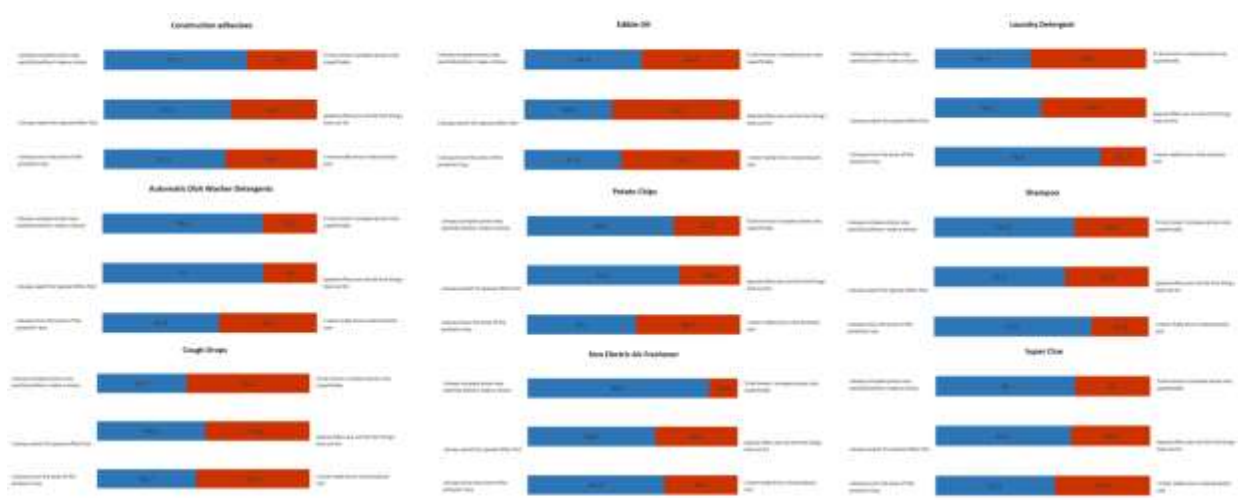
The "Role of Price" (RoP) is based on the following three semantic differentials:

"I compare prices very carefully before I make a choice"
vs
"To be honest, I compare prices only superficially"

"I always search for special offers first"
vs
"Special offers are not the first thing I look out for"

"I always know the price of the products I buy"
vs
"I never really know what products cost"

The following table shows how differently consumers behave when buying within these nine categories:



The number of consumers who always compare prices carefully varies between 41.9% (“cough drops”) and 86.4% (“NECA”). In the “edible oil” category, 40.8% are looking for special offers first, compared with 75% in the “automatic dish washer detergent” (“ADW”) segment. Price knowledge varies between 42% (“edible oil”) and 78% (“laundry detergent”).

Such differences in consumer behavior are useful for interpreting results from choice models. For example, in the “cough drops” category, a price increase is more likely to be accepted, given that 58% of the consumers do not compare prices. In contrast, only 15% do not compare prices in the “NECA” category, so price increases could have a much higher impact on preference shares.

ROLE OF BRAND

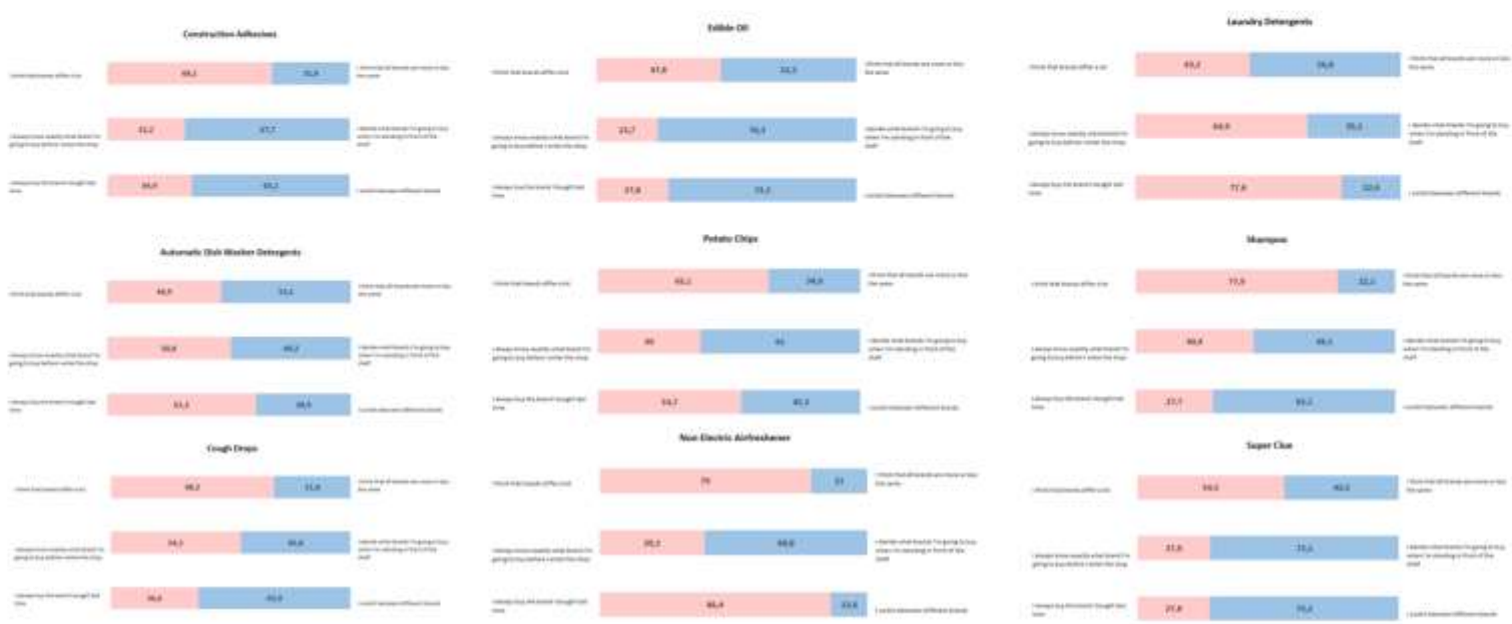
The “Role of Brand” (RoB) is represented by following differentials:

“I always buy the brand I bought last time”
vs
“I switch between different brands”

“I think brands differ a lot”
vs
“I think that brands are more or less the same”

“I always buy the brand I bought last time”
vs
“I switch between different brands”

These three differentials provide insights into the RoB, thus deepening understanding of consumers’ behavior in this regard. This approach provides further insight when interpreting simulations based on choice models.



Again, there are significant differences between the segments: The brand-switching attitude varies between 21% ("NECA") and 56.8% ("laundry detergents"), representing a significant difference when a company aims to "introduce a new brand" into a category. Because 72% of "super glue" customers think that all brands are more or less the same, compared with only 14% of "NECA" customers, it seems that having a strong brand has more equity in the "NECA" category as compared with "super glue."

Role of Innovation

For "Role of Innovation" (RoI) the following three differentials are used:

"I'm always interested in new products"
vs
"I prefer to stick with what I know"

"I think products in this category need to be improved"
vs
"I'm completely satisfied with the products as they are"

"I find it easy to make the right choice for me"
vs
"I find it difficult to make the right choice for me"

With these differentials, we can derive insights about the opportunities for new products in the different categories.

"Satisfaction with current products" ranges from 16.6% ("NECA") to 82.2% ("Potato Chips"), representing a large difference in suppliers' opportunity to develop new products. Another example: 39% find it "easy to make the right choice" in the dishwasher detergent category, compared with 92% for "NECA." This may indicate the need for differentiation, such as by developing and clearly communicating specific USPs for different products.

Initial Conclusions

Results from the nine "Behavioral Calibration Questions" indicate that they have a potential to differentiate between respondents' buying habits. Regarding our hypothesis, behavior during the most recent shopping trip (within a category) influences the answering behavior in the choice exercise. Considering this, these questions should help respondents to recall their decisions during their last shopping trip more effectively; therefore, responses to the following conjoint tasks should be much easier and clearer to them. Bivariate analysis of the "Behavioral Calibration Questions" suggests that our hypothesis may be correct and that it is worthwhile to invest the additional interview time to improve the answering behavior of respondents on the choice task.

Enhance Conjoint

To explore how the calibration questions enhance the conjoint exercise that follow, we consider three different mechanisms:

- Simply asking the questions helps respondents to recall their most recent shopping trip, which results in more reliable answers.
- Using these questions as covariates improves the Bayesian estimation of the part-worth utilities and results in better "Share of Choice" estimations, better hit rates, and less error.
- The three roles may provide further insight when using them as segmentation variables in the choice simulator.

All nine empirical studies were analyzed using the same settings to avoid methodological bias. We used Sawtooth Software CBC/HB (190,000 burn-in-draws, write out 1,000 draws by using every tenth draw). For SoC simulation, we used the average over these 1,000 draws, as well as the Sawtooth Software default settings for prior variance and degrees of freedom (1.0/5), with an acceptance rate of 30%. For the comparisons, we used three different estimations for the sample split cells with "Behavioral Calibration Questions":

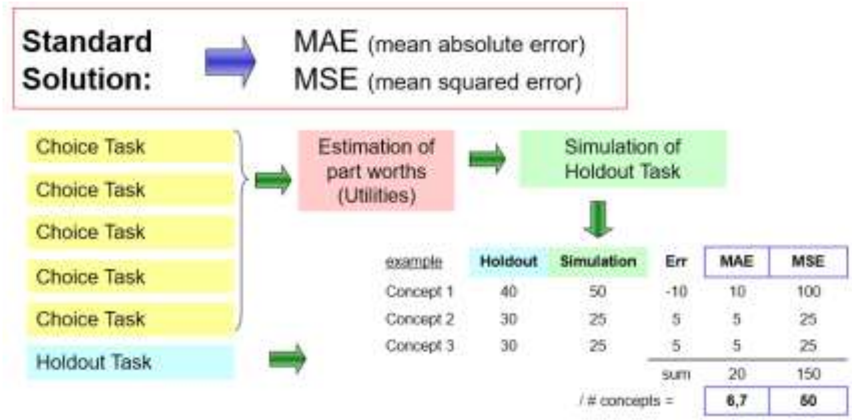
- Standard HB estimation,
- HB with the nine binary questions as covariates, and
- Ensemble of nine estimation runs with one of the questions used as a covariate in each run.

One of the great achievements of machine learning is certainly the use of ensembles. An ensemble approach generates multiple diverse models, include HB estimations with different covariates as in this study. First, we can make predictions with each of the specific HB models individually. Due to the different covariates, these models are diverse in the sense that each provides different predictions and has its own unique strengths and weaknesses. For the ensemble approach, we take the nine different models and blend the SoC predictions to reduce bias from the individual models, thereby generating more robust and accurate predictions.

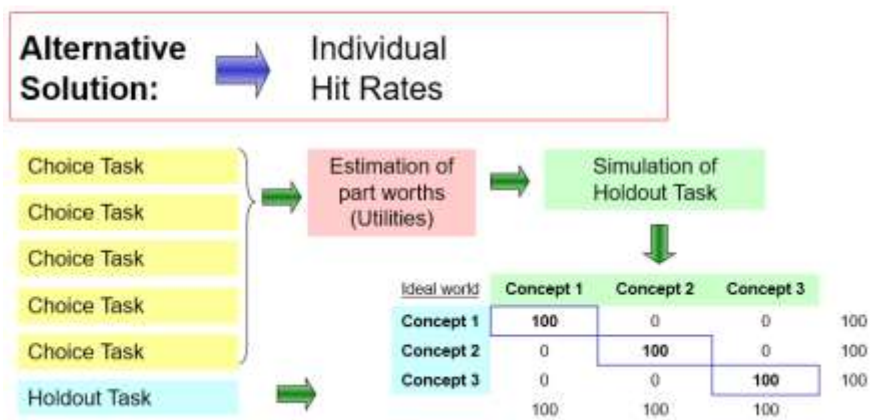
SHORT REMINDER: HOW WE MEASURE THE VALIDITY OF CONJOINT STUDIES

Before describing the results of the different approaches, we would like to review how the measures of validity are computed.

The standard approach is to use one predefined choice task not used for the estimation process as a "holdout task." This task is then simulated and the MAE (mean absolute error) or the MSE (mean square error) for the whole sample is calculated taking into account the number of concepts in the task.



The alternative approach is to individually simulate the "holdout task" for each respondent and match it with his or her actual answer to this task during the interview.



If real market data (e.g., market shares) are available, the root mean squared error (RMSE) is used.

EMPIRICAL RESULTS

The nine studies analyzed confirm our hypotheses: the "Behavioral Calibration Questions" are effective, and hit rates can be significantly increased by asking these questions up front, even if they are not used in the HB estimation. Using the "Behavioral Calibration Questions" as covariates in the HB estimation further improves hit rates. Finally, an ensemble of part-worth utilities from nine estimations based on one calibration question as covariate in each run results in slightly higher hit rates compared with a single HB run that use all nine questions as covariates. Within a category, the more specific the three different roles of our behavioral calibration questions are, the more the hit rates can be improved by using this additional information in the estimation. For example, in the "NECA" category, where numerous consumers compare prices and search for new innovative products, the hit rate could be improved from 43.5% to 53.5%.

		Behavioral Calibration Questions			
Hitrate in %	Chance-Rate	not shown	shown	used as covariate	Ensemble
ADW	11,11	36,50	41,60	41,90	43,20
Construction adhesives	20,00	53,90	55,30	55,60	57,10
Cough Drops	7,69	32,40	39,10	40,20	41,90
Edible oil	14,29	41,20	49,30	51,10	53,20
NECA	16,67	43,50	52,40	52,80	53,50
Hair Shampoo	7,69	30,90	32,10	33,00	33,40
Potato Chips	16,67	47,10	52,40	52,60	52,90
Laundry Detergent	7,69	31,80	36,20	37,20	37,80
Super Glue	7,69	34,20	38,70	39,10	39,60

Out-of-sample calculations were done by splitting the samples into estimation and validation samples (80% and 20%, respectively).

Only the "super glue" study design consisted of four sample splits, such that the possibility of using validation samples and estimation samples was built into the design. In this study, the separate validation samples each had 250 respondents answering or not answering the Behavioral Calibration Questions, whereas the estimation samples had 500 respondents each.

As we could not calculate part-worth estimates for out-of-sample tests, and therefore could not simulate preference shares in the traditional way, we used logCounts (described in Johnson, Orme, Pinnell 2006). The conclusion is roughly the same as that for the above-mentioned hit rates: almost all studies have better RMSE values when "Behavioral Calibration Questions" were implemented. Only the "Shampoo" study seemed to not benefit from use of the "Behavioral Calibration Questions," but the framing did not harm the results.

	within-sample				out-of-sample			
RMSE	not shown	shown	used as covariate	Ensemble	not shown	shown	used as cova	Ensemble
ADW	2,12	2,01	1,97	1,96	2,67	2,48	2,31	2,26
Construction adhesives	1,74	1,69	1,66	1,61	2,19	1,98	1,89	1,84
Cough Drops	2,51	2,43	2,41	2,36	3,21	3,17	2,94	2,89
edible oil	2,45	2,42	2,40	2,39	3,39	3,25	3,11	3,06
NECA	2,72	2,62	2,58	2,57	3,94	3,37	2,89	2,81
Hair Shampoo	3,21	3,23	3,22	3,20	4,63	4,65	4,71	4,61
Potato Chips	2,16	2,05	2,01	1,96	3,12	2,93	2,73	2,67
Laundry Detergent	2,38	2,19	2,14	1,99	2,99	2,74	2,54	2,44
Super Glue	1,84	1,79	1,67	1,66	3,87	2,56	2,17	2,06

Only two of our nine studies have reliable, "real" market shares information and can therefore be compared against them. In the "super glue" case, we estimated separate models for the validation and estimation splits and compared them with the RMSE measure.

The results support the same deductions as the above comparisons: simply asking the “Behavioral Calibration Questions” improves the predictions. The inclusion of these questions as covariates or in an ensemble approach further improves the “Share of Choice” simulations.

	Share of Choice - Market Shares			
RMSE	not shown	shown	used as covariate	Ensemble
Construction adhesives	5,68	5,21	5,19	4,96
NECA	10,23	9,63	9,60	9,38
Super Glue	8,36	7,56	7,47	7,18

Use Behavioral Calibration as Segmentation

Our third approach in using the “Behavioral Calibration Questions” is based on the three “Roles.” For each, we calculated a filter variable based on the three questions to derive specific “Share of Choice” values for the splits.

The following example shows different price elasticities for one SKU in the “edible oil” study. Again, it is clear that asking the “Behavioral Calibration Questions” results in different elasticities:

The different elasticities correspond with our expectations regarding the role of price, in that price-sensitive buyers with brand-switching behavior (i.e., respondents who switch to a different brand when price increases) have higher elasticities:

Innovation seekers are less price sensitive. Simply exposing respondents to the "Behavioral Calibration Questions" results in different elasticities ("edible oil" study):

For a more detailed inspection of these effects, we calculated the arc-elasticities of demand for the different segments. The differences between the segments provide detailed insights into the influence of consumer behavior on price and can be leveraged for more insightful recommendations to clients.

ARC - Elasticities of Demand					
	€	2,49-2,99	2,99-3,49	3,49-3,99	3,99-4,49
Role of Brand	yes	-1,00	-1,44	-2,14	-1,00
	no	-0,94	-1,19	-0,61	-0,53
	all	-0,94	-1,22	-0,78	-0,58
Role of Price	yes	-1,57	0,00	-0,88	-1,16
	no	-0,89	-1,33	-0,77	-0,53
	all	-0,94	-1,22	-0,78	-0,58
Role of Innovation	yes	-1,29	0,00	-0,52	0,00
	no	-0,89	-1,41	-0,83	-0,69
	all	-0,94	-1,22	-0,78	-0,58
Behavioral Calibration	shown	-0,94	-1,22	-0,78	-0,58
	not shown	-0,26	-0,61	-0,49	-0,42

FINDINGS

Based on our nine empirical studies, we can conclude that the "Behavioral Calibration Questions" represent a useful extension to DCM exercises. Our findings suggest that all three hypotheses may be verified. The "Behavioral Calibration Questions" help the respondents to recall their most recent shopping trip in a particular category and thereby positively influence answering behavior in the ensuing conjoint model. The data-generation process comes closer to representing a real shopping trip. Using the questions as covariates can also help improve the estimation results, rendering them more meaningful for simulations. The use of nine different estimations based on the "Behavioral Calibration Questions" in an ensemble approach slightly improves the results and always performs slightly better than a single estimation with nine covariates. Due to the modest improvement, one should decide if the additional effort required by this approach is justified. Using the "Behavioral Calibration Questions" as filter variables provides more detailed insight into the data structure and helps to improve recommendations for clients.

Consequently, it seems that further investing in these additional questions is worthwhile to improve our conjoint models.

FUTURE RESEARCH

The nine "Behavioral Calibration Questions" are a good starting point for further developments. To take advantage of such a framing exercise, the "Behavioral Calibration Questions" could be extended to include more than the three roles. For instance, three additional semantic differentials about the importance of features could be added to generate a fourth role. More specific wording for different categories should also be developed and validated.

From a more methodological point of view, a next step could be the use of the questions as an input for a Bayesian variable selection model to improve part-worth estimates.

Benchmarks should be established by building a database of Category Behavioral Calibration results to position tested concepts in commercial studies.

Peter Kurz

Stefan Binner

References

Allenby, G.M.; Rossi, P.E. (2006): Hierarchical Bayes Models, in: Grover, R.; Vriens, M. (Eds.): *The Handbook of Marketing Research: Uses, Misuses, and Future Advances*, 418–440, SAGE Publications Inc., Thousand Oaks.

Binner, S. (2006): Do Individual Hit Rates Matter at All? (Design & Innovations Conference München 2006).

Hein, M.; Kurz, P.; Steiner, W. (2013): Limits for Parameter Estimation in Choice-Based Conjoint Analysis: A Simulation Study, European Conference on Data Analysis 2013.

Hein, M., Kurz, P., Steiner, W. (2019): Analyzing the Capabilities of the HB Logit Model for Choice-Based Conjoint Analysis: A Simulation Study. In: Journal of Business Economics (forthcoming).

Johnson, R., Orme, B., Pinnell, J. (2006): Simulating Market Preference with Build Your Own Data. Proceedings of the 2006 Sawtooth Software Conference.

Kurz, P; Binner, S. (2011): Added Value through Covariates in HB Modeling?, Proceedings of the 2011 Sawtooth Software Conference.

Liakhovitski, D.; Shmulyian, F. (2011): Covariates in Discrete Choice Models: Are They Worth the Trouble? ART Form Presentation.

Research International. (2010): Using the landscape questions in pricing research (promotional material) RI Hamburg.

Sentis, K. and Li, L. (2001): One Size Fits All or Custom Tailored: Which HB Fits Better? Proceedings of the 2001 Sawtooth Software Conference.

Sentis, K.; Geller, V. (2011): The Impact of Covariates on HB Estimates, Proceedings of the 2011 Sawtooth Software Conference.