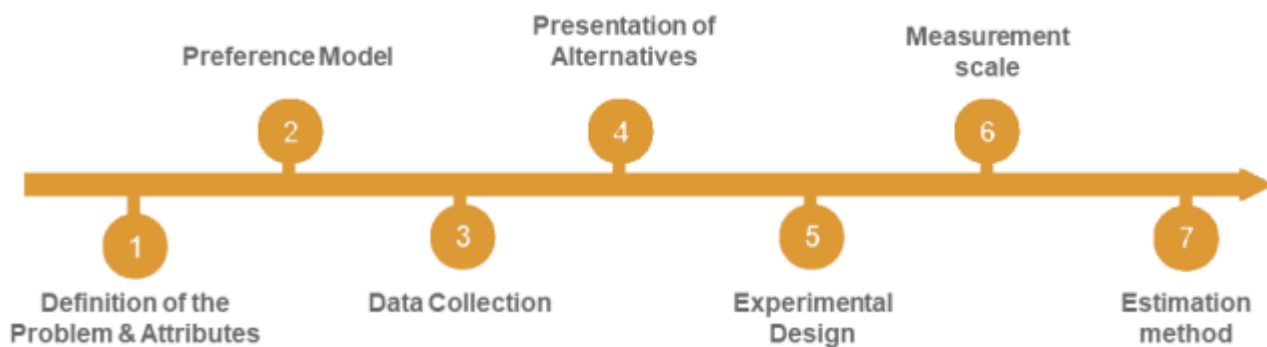


A conjoint analysis step by step guide.

Constructing a conjoint analysis is not as difficult, as it might seem. As we described in one of the previous articles, there are some things that need to be considered when constructing it. It is important for the conjoint analysis, that the steps fit to each other like pieces of one puzzle. As much as you cannot build a house if you do not choose the right materials that complement each other, a conjoint analysis will not be as effective if the steps do not complement each other. For this it might be necessary to go back to one of the earlier steps and adjust it. The whole design is an iterative process.



Step 1: The Problem & Attribute

At the very beginning of each conjoint analysis, you should define the problem and find attributes that you will want to collect. This is very important because the problem determines the purpose of the conjoint analysis and this will already limit you in certain ways. Whether you want to predict a market share or whether you want to understand your customers, for each case you will need different complementary components in order for it to work. When you want to understand a customer, you will most likely need to choose a part-worth model in step 2, because this one is the simplest one and enables you to present the preferences in plain figures. However, if you want to predict the market share, then a part worth model might not be able to generalize to special cases and you might need a mixed model.

After you have defined your problem well enough, it is crucial to select the right attributes

Economalytics is the analytics blog that shows you in plain and simple which methods are available and how you can use these methods to solve your problem. Check out www.economalytics.com for more!

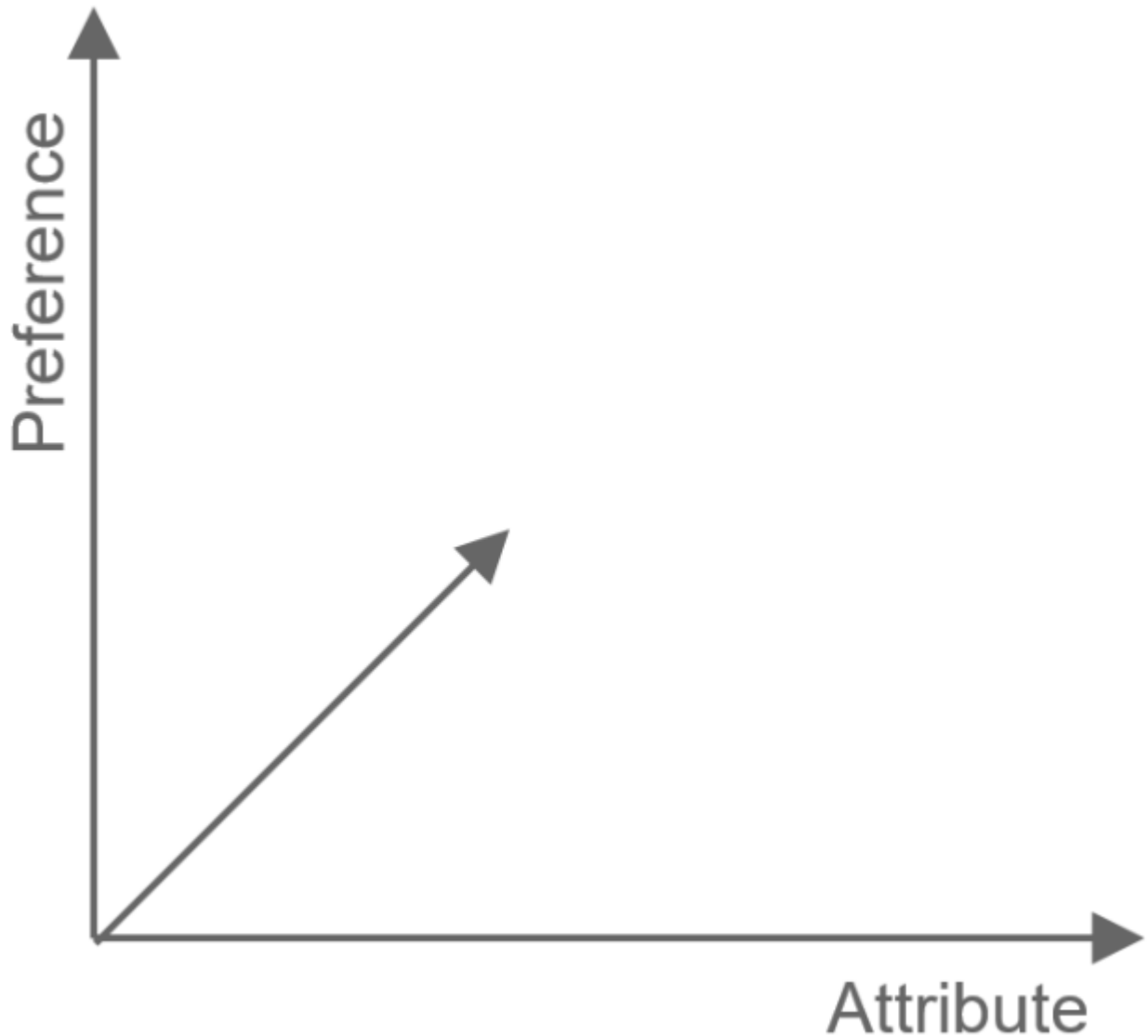
that matter for your purpose and try to limit yourself to only the really necessary attributes. It might be enough to have an domain expert to decide on the necessary attributes, but it might be even more beneficial to conduct interviews with consumers to identify relevant attributes from their perspective. For each attribute, you should decide what type it is (categorical, ordinal, and continuous) and what relationship you expect between utility and that attribute (linear, quadratic ...). Furthermore, you should look at the variables and see whether you expect any interactions and whether issues of *perceived correlation* might occur. Perceived correlation describes the phenomenon where the costumer expects a correlation between attributes when there is in fact none. For instance, a consumer will tend to assume that an expensive car must be better than a cheaper one, even though they might be identical. Perceived correlations can bias your conjoint analysis and render them less valid. Finally, you should spend thoughts on whether you attributes capture all the important variability in utility, e.g. that all relevant variables are included that determine the purchase decision.

Step 2: The Preference Model

When it comes to modelling preferences, you generally have the choice between 5 basic models that can be applied. This step is about deciding how you want to model the consumers preferences, e.g. how does the consumer actually make decisions in his head. The idea is to choose a model that resembles his internal way of decision making as much as possible, e.g. copy his internal way of thinking about decisions. The closer your model is to the actual way the consumer makes decision, the better your results will be. While we want to select the model that comes closest to the way the consumer things, there are also other considerations. Some models are more simple to understand and easier to present, others offer clear direction for action and others better complement other steps for instance. Here a short overview of some models that can be applied:

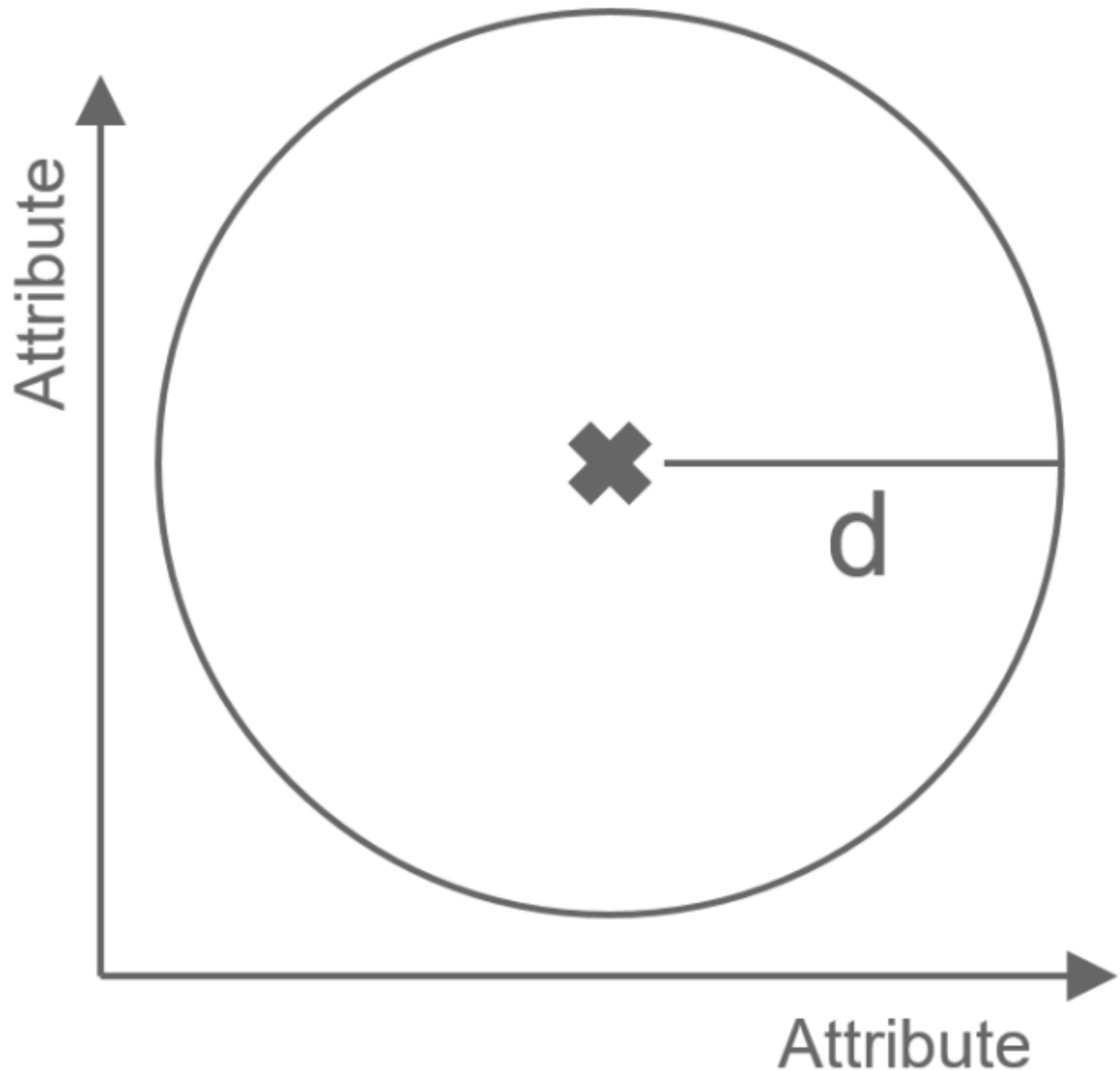
1. Vector model

A basic simple model that is well-suited for many purposes. However, the basic level does not allow for more complex relationships.



2. Ideal-Point model

The ideal-point model assumes that there is only one ideal preference each other combination of attributes can be rated by the distance from this attribute. The further away from the combination of the attribute, the less preferred. Depending on the assumption, it is also possible to use different distance measures such as quadratic distances or Manhattan distance depending on the assumptions. Ideal-point models offer the advantage that customers' preference can be literally plotted on a map and it can be easily presented.

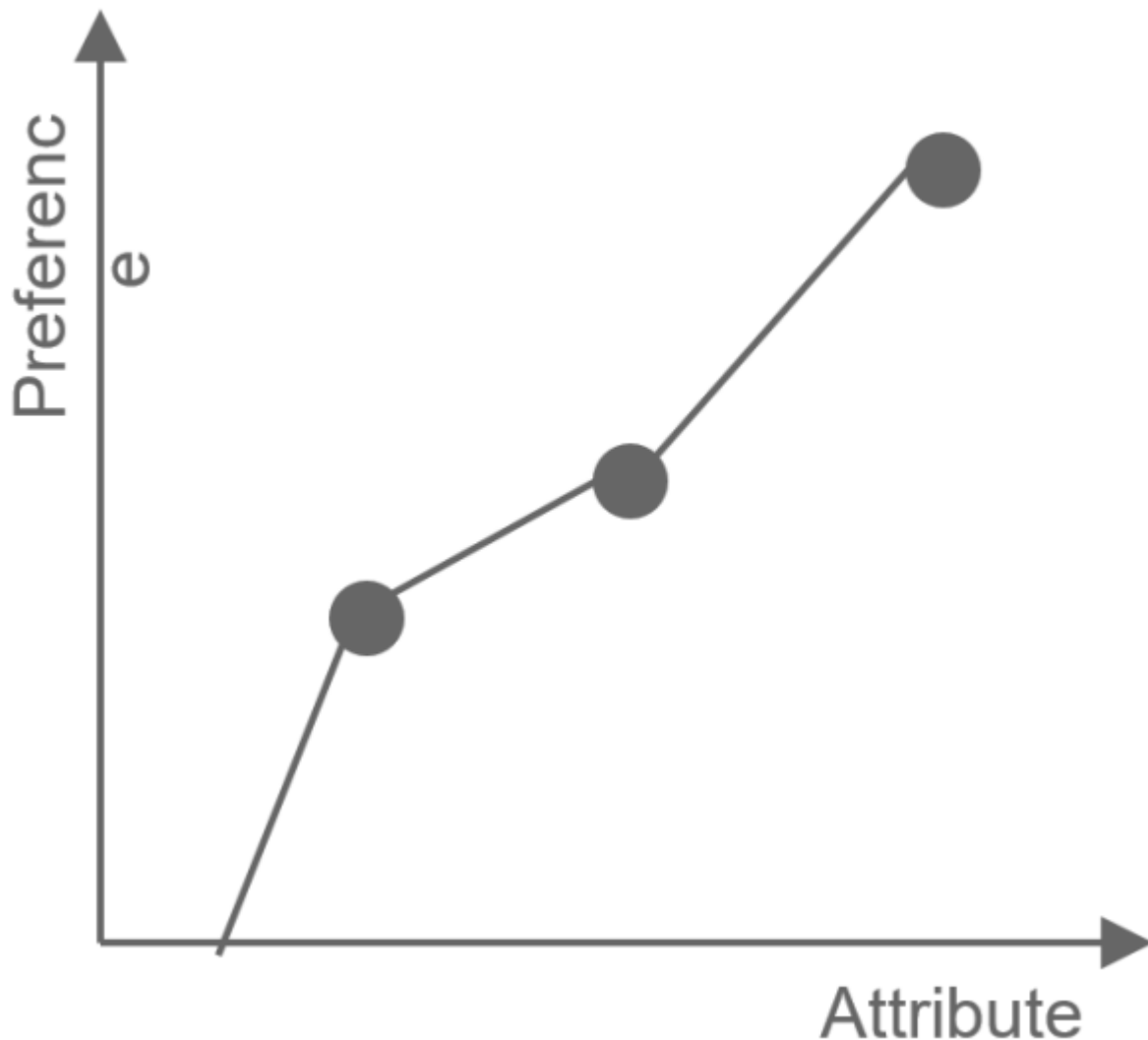


3. Part-worth function model

The part-worth model follows the idea of *marginal utility* from economics. It means it does not give an absolute value for the utility of an option, but rather assumes a reference alternative. All other alternatives are evaluated relative to the reference value, e.g. how much the utility would change if we changed the specific attributes of the reference alternatives to other values. All estimates in a part-worth model are on an *interval scale*. This makes it very difficult to produce predictions based on part-worth models. However, part-worth models offer several other advantages. They can be presented graphically, they can serve to calculate relative attribute importance,

Economalytics is the analytics blog that shows you in plain and simple which methods are available and how you can use these methods to solve your problem. Check out www.economalytics.com for more!

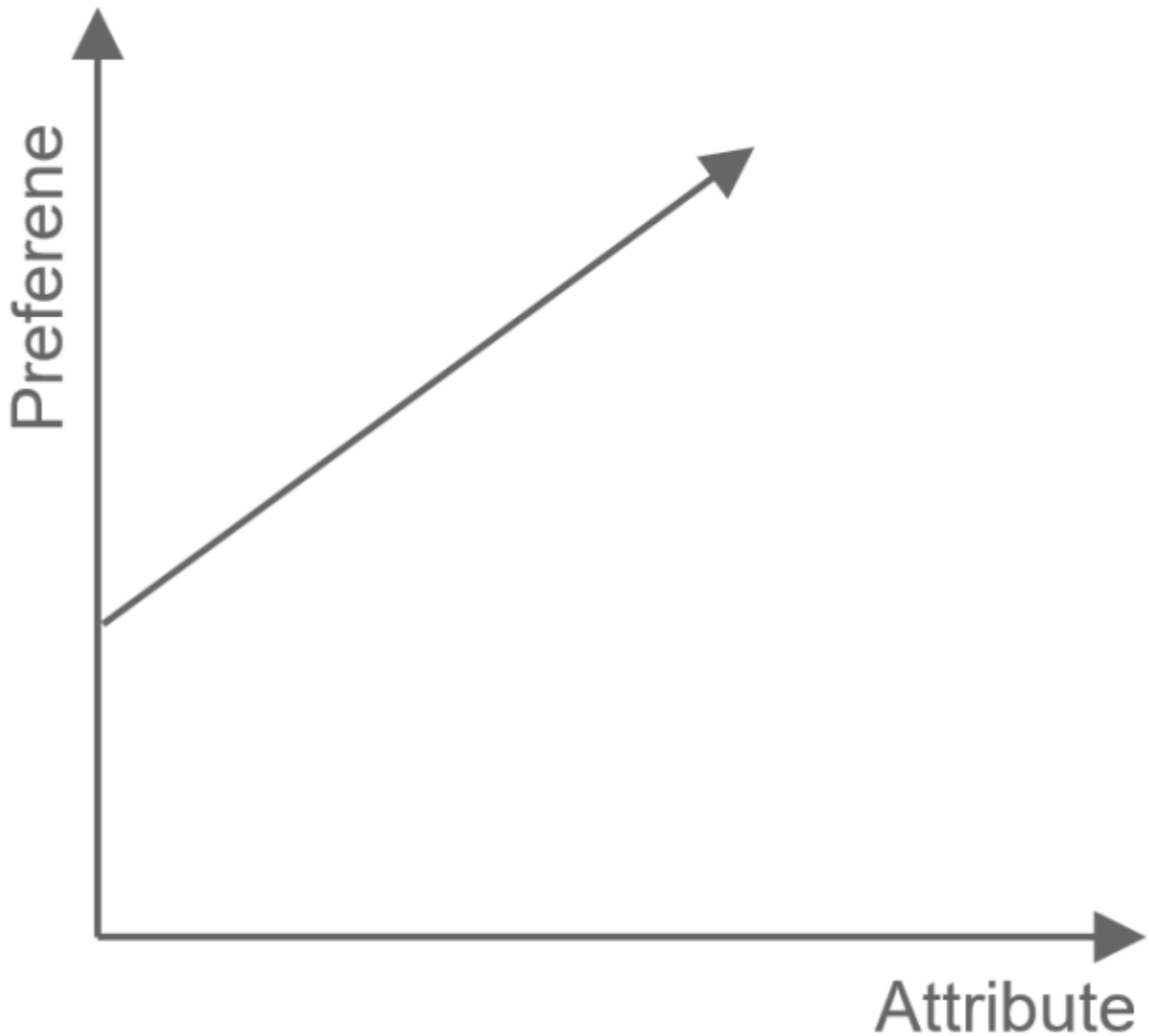
they are compatible with fractional factorial designs, and they offer clear suggestions on how to improve the product.



4. Mixed model

Mixed models offer the full range of possibilities in modeling and equal any type of regression model. It is possible to include interactions, factors, continuous variables, any type of transformation (logarithmic, quadratic, exponential etc.). This complexity enables to model more complex attribute-utility relationship and makes it easier to produce predictions. However, it has also drawbacks. One has to know what level of complexity is appropriate and how to use it, otherwise, it is recommended to stick to more simple methods. Furthermore, depending on the level of complexity, it might be more difficult to combine it with other components later. It might be difficult to understand the consumers given the complexity and it might be more difficult to

present the results to others.



5. Custom model

The custom model described special cases where an own model has been developed, the before mentioned models have been adapted or combined. If there is already knowledge on how consumers make decision, what mental rules they apply and how they approach the decision, than a custom model might be a good method. For instance, you know that the customer makes decisions in two steps, first he has k.o.-criteria and eliminates all option that do fulfill the criteria and then chooses the alternative that maximizes utility. Another example is *satisficing*, where your customer chooses the first alternative that fulfills the minimum criteria instead of going through all options and choosing the ideal alternative. The custom model enable you to adapt

the model to every specific instance, however, their compatibility with all the other steps is more difficult. It is definitely necessary that custom models are applied by real decision psychology experts, because the models still need to be scientifically valid

Step 1: Remove all options that do not fulfill x.

Step 2: Use a part-worth model for the remaining alternatives

Step 3: The Data Collection

The data collection step deals with how we obtain the data from the customers. Here you will have three general options each with advantages and disadvantages.

1. Trade-off

In trade-off presentation, you show the customer for each pair of attributes each combination of the levels of the two attributes. The customer then orders the combination from most preferred to least preferred or rates them. The advantage of trade-off analysis is that you obtain an accurate in-depth picture of how the attributes contribute to the overall utilities, interactions and non-linearities can be spotted. However, it will likely result in choice overload (imagine you have 4 attributes each

Economalytics is the analytics blog that shows you in plain and simple which methods are available and how you can use these methods to solve your problem. Check out www.economalytics.com for more!

with 3 levels, your customer will need rate 54 combinations). Furthermore, it is not a realistic scenario of how decisions are really made.

Noise Kitchen	High	Medium	Low
Modern	5	4	1
Average	8	6	2
Old	9	7	3

2. Concept evaluation

In concept evaluation, a whole alternative with all attributes is presented at a time and the consumer has to evaluate or order the alternatives according to its preference. This method is more realistic but might result in information overload. If a person is exposed to several options at the same time, it might be difficult to take all pieces of information into consideration.

Price	\$ 400
Kitchen	Old
Noise	Low
Distance	30min
Bedrooms	2

3. Choice making

The choice making option resembles the concept evaluation, but the difference is that usually one two or more options are presented to the customer and the customer has to make a choice, e.g. which one he would indeed buy. The clear advantage of this scenario is that choice data is collected first and it is more realistic.

Apartment 1		Apartment 2	
Price	\$ 100	Price	\$ 400
Kitchen	Old	Kitchen	Old
Noise	High	Noise	Low
Distance	10min	Distance	30min
Bedrooms	2	Bedrooms	2

Other considerations that need to be made here at this point is, for instance, how do you want to collect data (asking the participants personally, online questionnaire, give them two products that they can test and decide then ...) and how can you make your scenario as realistic as possible?

Step 4: Presentation of Alternatives

Realistic in this sense means that the scenario you create resembles the environment that your actual customers face, when they make purchasing decisions. An alternative approach to realistic as possible is that you want to make it for your customer as easy as possible to understand the alternative so he can estimate his score as precisely as possible. For the presentation of the alternative, you also have again three options:

1. Overview

An overview has the advantage that it presents all the important data in a short and concise way. However, often this way is not realistic and there is a danger that values without text might be misinterpreted.

Price	\$ 400
Kitchen	Old
Noise	Low
Distance	30min
Bedrooms	2

2. Paragraph

A paragraph description gives the opportunity to add a richer description and avoid misunderstanding. It also offers the possibility to also test different ways of phrasing. However, it might also easily lead to information overload and not every consumer will take time to read through everything.

The flat has **two bedrooms**, an **old kitchen**. It is not far but also not close from the university and walking to the university takes **30min**. The neighbourhood is generally **quiet** and the price per month is **\$ 400**.

3. **Picture or video**

The picture or a video is the option that will reduce choice overload. However, it does not necessarily provide any hard facts and it might be difficult to show an example if you are trying out a scenario and the product does not exist yet. Furthermore, this option can be considered more expensive than the previous two ones.



Finally, you should spend thoughts on how the user will process the information that you give him the way you present the scenario to him. Generally, you want to make the presentation of the information as realistic as possible, but at the same time, you want to

Economalytics is the analytics blog that shows you in plain and simple which methods are available and how you can use these methods to solve your problem. Check out www.economalytics.com for more!

avoid *information overload*. Information overload might bias the results for an individual because he simply does not know how to estimate the utility because it is too much information to form an opinion on. Also if you order the categories in a certain order, the ones that appear first might become more important because the customer does not really look for all, uses this one for constructing a cut-off rule (*satisficing*), or uses this one as an anchor to evaluate the following ones (*halo effect*). The way you present the attributes might also have an effect on *perceived correlation*.

Step 5: The Experimental Design

Finally, depending on what data model you choose, you will also need to specify an experimental design. This step should not be underestimated. This is especially true if you chose, for instance, to represent the preferences as a part-worth model or mixed model. The art of designing an experiment is something that little people will be familiar with and therefore I recommend you to stick to a full factorial or maximum fractional factorial design if you do not have any particular experience or no person with a Ph.D. in experimental sciences at hand. You can find an explanation including and step by step guide on how to create your own design on R for of full factorial design and for a fractional factorial design in one of the articles on this blog.

There are three main considerations to be made here. First, you should look at the list of attributes you prepared at step 1) and ask whether you expect to have any significant interaction effects. If it is certain that you will have several significant interaction effects between the attributes, then a fractional factorial design will produce biased results and cannot be used. In this case, either you change your attribute that you want to consider or you go for a full factorial design. If you expect that there are no significant interactions, you can choose between fractional factorial design and a full factorial design.

Secondly, you will need to think about whether you want to include any transformations (a logged variable, a quadratic variable etc.) if you decided to go for the mixed model solution. If this is the case, then you might need to do a higher number of experimental runs and maybe you will need to reduce the number of attributes. Be aware that the inclusion of such variables generally makes the interpretation of the results more difficult.

Third, you need to consider the size of the experiment. The longer the experimental design and the more runs it has, the more tiring it could be to the customer. Imagine you would need to determine a rating for 100 products that only differ slightly. I can bet that at some

point you would not take it serious any more, make mistakes or start to type just something in order to get done with the rating. Here comes the clear advantage of a fractional factorial design compared to a full factorial design. If you have 4 attributes with 3 levels, you might quickly end up with 3^4 81 runs in total, while fractional factorial design, as the name already says, enables you to reduce the size and run only a fraction of the total amount of runs. This is only possible if there are no significant interactions between the attributes.

Step 6: Measurement Scale

Finally, the measurement scale also matters depending on what method you have chosen. To give you a concrete example, if the goal of the conjoint analysis is about understanding the consumer and you chose to work with a part-worth model, then the ideal measurement scale would be categorical or maximum ordinal. This is generally also the case if you decide to work with a fractional factorial design. In this case, it is also necessary to use ordinal variables and only in exceptional cases, you can use continuous variables if you expect the attributes to be perfectly linear for instance.

On the other side, if your goal is to predict a future market preference share, then ordinal scales might not be enough, because you would face problems with predicting the utility for alternatives with attributes that go beyond your chosen attribute-levels. For instance, you want to predict the maximum utility for a laptop that has a total RAM of 32GB (such laptops exists), but when constructing the utility function, you only included the levels 4G, 8GB and 16GB for the attribute RAM. Then you will not be able to predict the actual utility score for any laptop that does not have any of these three levels for RAM, like the laptop with 32GB RAM.

Step 7: Estimation Method

There is a wide set of methods that can be applied to the estimation method which will ultimately depend on the model for the preferences you chose because you cannot estimate all models with all the methods. The method at the end will derive the importances and scores for the individual attributes for a specific person or a map of his or her preferences depending on the model chosen earlier.

Besides Prefmap, Linmap, Johnson's tradeoff algorithm, Monanova, probit, and logistic

Economalytics is the analytics blog that shows you in plain and simple which methods are available and how you can use these methods to solve your problem. Check out www.economalytics.com for more!

regression, the most practicable and known one will be a linear regression. Linear regression can be used to estimate a part-worth model as well as a mixed model. In the next articles, I will show you how to use a linear regression model to estimate a part-worth function as well as a mixed model.

Conclusion

Here it becomes evident now why conjoint analysis is a framework, e.g. a set of methods, as it exists in many different variations as well combinations depending on the specific situation at hand, the goal of the analysis and the available attributes. At the end of your self-designed conjoint analysis, you should spend thoughts on how you can measure the *validity* and *reliability* of your analysis, which I will not go into detail at this point as it might be too much for this article. Furthermore, you should always also consider the weaknesses of your own design. Every conjoint analysis will have weaknesses and the goal of constructing a conjoint analysis is not to eliminate all weaknesses, but rather to choose a conjoint analysis that makes its weaknesses irrelevant for your purpose and situation.

Finally, I want to hint out another possible problem that you will need to consider and my recommendation. You will need to think, whether you want to one preference model for a whole set of people, or whether you want to estimate one preference model for each person. A conjoint analysis is usually designed to estimate one preference model for each person, rather trying to produce the preference model for an “average” person and there are good reasons for this. An average person might only make sense if the population to which you want to generalize your model to is very homogeneous in their needs and utilities, which however is rarely the case. If this is not the case, you should calculate a preference model for each person individually, because as shown in the picture, the utility function for an average person might not represent the reality at all. Furthermore, if you have the utility functions for many people, you can start to apply clustering methods based on utilities to understand consumers in a different and greater way.

**Share this:**

- [Click to print \(Opens in new window\)](#)
- [Click to share on Facebook \(Opens in new window\)](#)
- [Click to share on LinkedIn \(Opens in new window\)](#)
- [Click to share on Twitter \(Opens in new window\)](#)
- [Click to share on WhatsApp \(Opens in new window\)](#)