

# Script: An Introduction to Factorial Survey Experiments (FSE), Part I

NCRM Online learning resources

Tamara Gutfleisch, Mannheim Centre for European Social Research (MZES), University of Mannheim

Hello everyone, my name is Tamara. I am a sociologist at the Mannheim Centre for European Social Research (MZES) in Germany. Before my current position, I was a PhD student at the University of Luxembourg, where I was involved in conducting a large-scale factorial survey experiment (FSE) on recruiters' hiring intentions that was the basis of my dissertation. Therefore, I am excited to talk about FSEs in this online course and hopefully get you excited about this method. FSEs are known under various terms that are used interchangeably such as factorial surveys or factorial survey analysis. Other terms often used are multifactorial survey experiment and multifactorial vignette experiment. I will mostly use the term factorial survey experiment or factorial survey in this course. This course is meant to be introductory, and will provide you with a first set of tools to develop your own research design based on this method.

The course is structured in three parts. The first part, which is covered in this presentation, gives a general introduction into factorial surveys (what they are and what they can do) The second part is a little bit more technical and discusses the different steps in designing and conducting a factorial survey. The last part is more hands-on and will show you how to perform some analyses and how to interpret the results.

At the end of this course, you will know what FSEs are and which type of research questions may be answered using this method. You will have an understanding of the decisions a researcher needs to make at different steps of the research process, and get an idea of how to analyse the data obtained from FSEs. As I said, my objective is to provide you with a starting point for building your own research design based on FSEs.

Before we begin, I want to draw your attention to some helpful reading. The book by Auspurg and Hinz provides a detailed introduction to factorial surveys and this course mostly relies on this source. I provide a further reading list in the supplementary material, which discusses the topics we address in this course and as well as topics that we cannot discuss in detail in this setting.

Factorial survey experiments are about surveys and experiments. It is therefore helpful to first look at surveys and experiments separately and discuss what each can do.

Surveys are widely applied and a relatively cost-effective method to collect information on a number of socially or politically relevant issues, such as political preferences, demographic information of individuals, attitudes, and many more. Many large-scale social surveys are applied in representative population samples, which allows to generalise the results to the whole population. Therefore, the results from representative surveys exhibit relatively high external validity. However, some confounders of the variables of interest might not be observed, which limits the possibility to establish causal relationships between variables. Lastly, respondents might give socially desirable answers if asked about sensitive topics (e.g., number of sexual partners).

In contrast, an experiment comprises at least two groups (a treatment group and a control group), whereby each respondent is randomly assigned to one of these groups. Lab experiments allow to control for unobserved influences on the treatment effect, since they are typically conducted in controlled environments. The results from experiments therefore have high internal validity when carefully designed, meaning that the researcher is able to measure what they intended to measure. Thus, the results can be interpreted in a causal way. However, they are typically based on selective

samples, such as students, which are not necessarily representative for the target population. Also, many lab experiments confront participants with fictitious situations. Therefore, the results obtained from lab experiments may have low external validity.

Factorial survey experiments combine the advantages of survey and experimental research to create a new method. The main idea behind factorial surveys is to present survey respondents with stimuli that resemble real-world evaluations. Specifically, the method makes use of hypothetical descriptions of persons, objects or situations, called vignettes, which are evaluated by respondents in a standard survey mode. For example, a vignette might describe job candidates, who are applying for a given job. The vignettes are experimentally manipulated by the researcher and vary in the values of multiple attributes. In the example of job candidates, such attributes might be the gender, educational background, and the previous work experience of applicants. The corresponding levels of these attributes might be male/female, university degree/no university degree, one/two/or three years of work experience. In the context of factorial surveys, the attributes are referred to as the dimensions of the vignettes, and the values of these dimensions are referred to as the levels. Depending on the design, single respondents may evaluate only one vignette, referred to as between-subjects design or several vignettes in sequential order, referred to as within-subjects design or mixed designs. In within-subjects designs, each respondent evaluates all of the existing vignettes. In mixed designs, only groups of respondents evaluate the same vignettes. In this course, we will focus on the latter type of factorial surveys, which are most often used in the social sciences. No matter the design, however, the vignettes are randomly assigned to respondents. We will talk about the main advantages and disadvantages of FSEs later in this course. For now, let's focus on what FSEs actually do.

The main goal is to assess how individuals interpret, weigh, and act upon information when making judgements or forming decisions in given scenarios. Returning to the example of job applicants, a researcher might be interested in the trade-offs employers make between applicant characteristics when hiring new employees, and which applicant characteristics are most important. In general, factorial surveys allow measuring the influence of single dimensions on individuals' evaluations of vignettes. Thus, the dimensions are the independent variables and individuals' evaluations of vignettes are the dependent variable. Depending on the scenario described in the vignettes, the outcome of interest may be attitudes, behavioral intentions, or normative judgements. Another objective of factorial survey experiments is to analyse the influence of respondent characteristics on vignette evaluations. For example, a researcher might be interested in whether the effect of applicant gender on employers hiring intentions varies between male and female employers. Factorial surveys typically test theoretical models about the single, joint, or interactive effects of different dimensions on the outcome of interest. The method is less suited for exploratory analysis because relevant confounders of treatment effects need to be considered in the experimental design to obtain unbiased estimates of the effects of dimensions on vignette evaluations.

There are other types of survey experiments that share some similarities with factorial surveys. There are not always clearly separated. For example, conjoint-analysis and choice experiments typically show respondents two vignettes simultaneously and respondents are asked to make a choice between them rather than to rate single vignettes sequentially. Also, many types of survey experiments are not based on a multifactorial design, that is they only focus on one or two treatments). Such designs limit the possibility to study the complexity that is often underlying social judgments. We will not talk about these other types of survey experiments in this course.

Now, let us look at some empirical examples. The first example is from a study, which was interested in the role of two dimensions, ownership and control, in individuals' fairness perceptions of couple's saving arrangements. The vignettes described different savings arrangements of a married couple in text format. Text format is a very common format in research based on FSEs. In the example on the

right, the vignette first provides background information on the married couple. The part written in *italic* varied between the vignettes. Thus, another vignette might show a situation, in which both partners decide equally when and for what the whole savings are spent. Ownership and control over savings are the dimensions of the vignettes, which are the same across the vignettes. The combinations of the levels of these two dimensions (e.g., both partners make decisions or only one partner makes the decisions) varies between vignettes. For each vignette, respondents were asked to evaluate the fairness of the situation described in a given vignette. The factorial survey was fielded in a representative sample of the German population.

The second example shows a vignette from a study interested in employers' willingness to retain older workers who are eligible for early retirement. The vignettes were displayed in a tabular format, as you can see on the right. The underlying principle, however, is the same as in the previous example. Respondents were first given a short introduction to the task. The vignettes were displayed in a tabular format, as you can see on the right. The underlying principle, however, is the same as in the previous example. Respondents were first given a short introduction to the task. The dimensions were the characteristics of the worker, such as the vignette person's age and attitudes towards retirement. The vignettes varied in the levels of these dimensions (for example regarding age, the vignette varied between the values 65, 63, 61, and 59). Respondents were asked to rate their willingness to retain the respective worker in their company. In contrast to the previous example, this study was not based on a representative population sample. Instead, a sample of managers of used identified through a Dutch household survey.

The last example is from my own research. The vignettes described applicants and varied systematically in the characteristics of these applicants, such as the applicants' gender. Other characteristics, such as the years of labour market participation, were held constant across the vignettes. The vignettes were displayed in the form of tabular CVs. The information that varied between the vignettes was highlighted in bold. In some cases, deviating from standard text or tabular vignettes may be useful in providing a more realistic experience for respondents. As in the previous two examples, the respondents were first given a general introduction into the task and provided with some background information. Respondents were real recruiters from Luxembourg. The respondents were asked to evaluate how likely it is that they would consider the applicant for a given position in their company.

With the three examples, I wanted to give you an idea of the various research questions that might be answered using FSEs. Simply typing „factorial survey“ or „vignette experiment“ in Google Scholar gives you a glimpse of the many possibilities of application. Factorial surveys have gained popularity in empirical social research in the 1970s, mostly to study normative judgments. In 2009, Lisa Wallander published a review of factorial surveys in sociology, showing the popularity in criminology and many other sociological subfields. The number of FSEs in social science research is increasing. They have been extensively applied in research on employer preferences, and to study educational choices and biases in teachers' evaluations of students. Recent examples also include individuals' preferences for patient prioritisation during the Covid-19 pandemic.

What are key features that all of these designs have in common? First, factorial surveys are characterised by a multidimensional experimental design. Researchers may vary the levels of three, five or more dimensions simultaneously. This is different from direct questioning in standard surveys, where a researcher would ask respondents multiple questions sequentially. For example, „how important is the crime level in your choice of residential area?“, „how important is the number of playgrounds in your choice of residential area?“, and so on. It is also different from standard lab experiments, which typically focus on only one treatment group. Second, the levels and dimensions that potentially influence the outcome of interest are randomly assigned to the vignettes. Finally, the

vignettes are randomly assigned to the respondents. Respondents might be randomly assigned only one or several vignettes. As I said in the beginning, we will focus on designs that assign several vignettes to respondents. Such designs allow for a larger number of observations with fewer respondents.

Let us talk about the main advantages of factorial surveys. From a conceptual point of view, factorial surveys allow to answer a broad spectrum of research questions and the vignettes can easily be adapted to fit the specific research context. Most importantly, factorial surveys allow to identify the principles underlying social judgments and decisions. They also offer the possibility to study rare combinations of characteristics that are hardly observed in reality.

There are also many advantages from a methodological point of view. As the dimensions (i.e., the independent variables in your theoretical model) are randomly assigned to the vignettes, they are independent from each other. Thus, the effects of dimensions that often correlate in reality (e.g., educational level and income) can be disentangled.

Potential confounders of treatment effects can be held constant in the experimental design, which ensures unbiased estimates of these effects. Factorial surveys therefore allow a causal interpretation of the results regarding the effects of vignette dimensions on the outcome.

Moreover, since multiple attributes are varied simultaneously, factorial surveys are probably less prone to social desirability bias than direct questioning because it is more difficult for respondents to detect the researcher's intentions.

Finally, factorial surveys can easily be implemented in large and representative samples, which increases the external validity of results. Note, however, that with experimental data, representative samples are not necessary to produce results with high internal validity. Causal relationships can still be established in convenience samples.

Despite these advantages, there has been some discussion about the external validity of factorial surveys. For example, the risk of methodological effects known in survey research such as social desirability can be mitigated but not fully excluded.

Also, so-called 'validation studies' have compared the results of factorial surveys and other types of survey experiments to real behaviour, providing mixed evidence as to the external validity of these kind of experiments. However, when reading these studies, it should be kept in mind that these studies themselves vary in their research design, which might have affected their results. Some of the criticism might apply more or less to prior studies depending on the respective research design. In general, these issues might be more relevant when studying decisions or behavioural intentions compared to general attitudes.

What have we learned so far? Factorial surveys can be a powerful tool to study the principles underlying normative judgements, attitudes, and behavioural intentions.

They allow to causally interpret the results due to the experimental set up. Note, however, that the causal interpretation only refers to the vignette dimensions. Since there is no control over respondent characteristics, the regression coefficients of these variables cannot be interpreted causally.

Although representative samples are ideal to be able to generalise the results, it is possible to establish causal relationships and have meaningful results also in convenience samples.

In Part II of this course, we will go a little bit more into detail and talk about how to set up the experimental design and the survey.

Thank you for listening to this presentation. I hope I was able to spark some interest in factorial surveys. Have a nice day.