How politics becomes news and news becomes politics. A comparative experimental study of the politics-media relationship

Helfer, L.

Chapter 2

Methodological challenges

The constant give-and-take taking place between politicians and journalists makes it challenging to empirically separate and study, how and when politics influences the media and vice versa. These challenges associated with the complexity of the interdependent politics-media relationship need to be addressed to be able to empirically study how politicians and journalists influence each other.

The study here focuses on the selection moment as underlined before. Focusing on one specific moment in this interaction, has the advantage that it allows to identify the crucial variables that affect these actors' decision-making at that very moment. Rich case studies describing the whole process of influence (e.g. Melenhorst, 2015; Wolfsfeld, 2004) often find it hard to generalize their findings beyond the case(s) studied. Zooming in on one specific moment of influence, however, allows studying potentially more generalizable mechanisms of influence and, more importantly, makes an empirical study more feasible.

However, even if researchers focus on one specific moment in the Politics-Media Wheel, there is still a considerable complexity due to the large number of potential sources of influence. Therefore, sender, message and receiver characteristics are differentiated as categories of influence on how these actors select messages. Furthermore, the sources of influence are systematized into individual-level, organizational-level and finally political system influences affecting these actors, from their individual experience on the job to the country's electoral system (see subsection 1.2.2). This systematization has important consequences for the research design. First, it means that a method is required that allows (statistically) separating these sources of influence which are often confounded in reality. A design should have maximal control over the contextual variables shaping the selection studied. Second, to study effects on the country level, the study needs to be carried out in more than one political system. An approach is required that can be applied across countries and still return reliable results.

In sum, even if one specific moment of influence is chosen, the decisions journalists and politicians have to make are of high complexity. Multivariate experimental research

designs model (part of) the complexity of such decision-making in their design (Hainmueller et al., 2014). They manipulate several variables at the same time leading to more realistic conditions, while at the same time giving the researcher maximal control over the context within which the study is carried out. The factorial survey experiment, a variant, is chosen for the studies in this book. Experimental designs are not common in the study of the politics-media relationship. Yet, they are particularly apt to establish a causal link between an (independent) event and the behavior of the relevant actors in this constant interaction. Because how else can we be sure that a politician's particular behavior is a consequence of something she or he read in the media and not from another source of information (Davis, 2007, p. 182)? Isolating cause and effect in political reporting is equally challenging and experimental approaches can thus be particularly insightful; "[experiments] guide theoretical development by providing a means for pinpointing the effects of institutional rules, preference configurations, and other contextual factors whose impact can be difficult to gauge using other forms of inference." (Druckman et al., 2006, p. 627)

The next section first introduces the factorial survey experiment. This experimental approach is not yet widely known in communication or political science but it has a lot of potential to contribute to the field beyond the questions addressed in this book. Following this introduction, the political and media systems of Switzerland and the Netherlands are introduced with a particular focus on the political system characteristics expected to affect the politics-media relationship. In particular the differences between the two countries with regards to the distribution of political power between parties and the electoral systems are discussed. Because the same study is carried out in two countries, it contributes to filling the void of truly comparable comparative studies in the field of political communication (Esser et al., 2012, p. 140).

2.1 Introducing the factorial survey experiment

Experimental methods have become very popular in political science in the past decennia (Druckman et al., 2006). Since 1990, the number of articles published referring to experimental methods has increased tremendously. A simple search in the Worldwide Political Science Abstracts database shows that in 2011 a record 481 peer-reviewed articles using the word "experiment" in their abstract were published. 20 years earlier there had only been 8 such articles. At the same time there are methodological books being published on experimental methods in political science (see for example Druckman et al. (2011); Morton and Williams (2010)) and journals (re-)established such as the *Journal of Experimental Political Science*. Despite the popularity of the method, experimental studies

with politicians and journalists as subjects are rather rare.[1] Many authors have pointed to the difficulty of recruiting those respondents for experimental research. For instance, when presenting results from a large number of interviews with journalists and politicians in Westminster, Davis (2007, p. 185) observed that "it is unlikely that cooperation can be gained for experimental, focus group, ethnographic, or extensive survey research." He concludes that "the most realistic methods [of data collection] involve interviews, content analysis, participant observation, and use of other survey data." Also Kepplinger (2007, p. 5), calling for a "theory of media effects on decision makers" writes about the desirability of experimental research to further the understanding of effects on elites, yet mentions that "it is nearly impossible to set up experimental studies. In such studies, one would have to show subjects news reports about themselves or about something in their field of activity." Others also pointed to the difficulty of recruiting these elites into laboratory settings (Hanitzsch and Engesser, 2014).

However, even if those elites could be recruited, one of the main obstacles remains their limited numbers. Elites are by definition a small population when compared with the general public. Yet experimental designs comparing treatment and control groups in a traditional between-respondent design require a rather large number of participants. That is usually not feasible in elite research. In politics, for example, although sizes of parliaments vary between countries, there are usually no more than a couple of hundred seats available at the national level. As a consequence, the number of participants in studies with politicians and journalists is often small. Overall, the small populations combined with the difficulty of gaining access to these actors results in relatively small numbers of respondents.

This means that experimental approaches used for studies in general population samples have only limited applicability. Other more innovative approaches are better suited to conduct an experiment in the small and inaccessible populations of journalists and politicians. The factorial survey experiment provides such an avenue. It is more commonly applied in sociology to measure normative rules and attitudes (for a review see Wallander, 2009). A number of studies have also applied this design to study intended behavior, which is particularly interesting in the context of this book. Abraham and colleagues (2010) studied the probability of accepting a job offer in dual-earner partnerships, in particular the influences of the partner. Another study focused on the propensity to move into a neighborhood (Shlay, 1986). To study the politics-media relationship in this book, the main focus in on the intended behavior of journalists and politicians. What party press releases do journalists select for reporting? And similarly, what media coverage can trigger politicians to react?

[1] Experimental studies with journalists were for instance carried out by Patterson and Donsbach (1996), Hudson (1992) or Zhong and Newhagen (2009). Examples of experimental studies with politicians are Fatas Neugebauer and Tamborero (2007) with Spanish politicians or Linde and Vis (2015) with Dutch politicians. Other studies relate more to real-world experimental designs in the US context, for instance Protess and colleagues (1987) on how media, the public and politics influence each other or Clinton and Enamorado (2014) on how FOX news affects members of the Congress.

2.1. Introducing the factorial survey experiment

The factorial survey experiment is a multivariate or factorial experimental approach combining a between-and within-respondent design. The multivariate design means that several variables are manipulated simultaneously, similarly to conjoint experiments (see Hainmueller et al. 2014 on the application of conjoint experiments in political science). Such a design means that the influence of several variables can be tested in one study instead of multiple independent ones, and it also means that interaction effects between variables of interested can be estimated. When politicians react to media coverage, the influence of their own background, for instance their tenure, might be moderated by the content of the news report they read; not all politicians react to the same media coverage. In a multivariate design, the *relative* influence of variables can furthermore be gauged allowing to draw up a hierarchy of influences on the phenomenon studied. This allows putting the many variables that influence how journalists and politicians select messages in perspective.

Center piece of the factorial survey experiment are so called "vignettes". Those are descriptions of objects, situations or persons on which a respondent is asked to pass judgment. To study how politicians react to media content, they are for example shown a media report and asked, whether they would take political action based on the report. The basic assumption of the factorial survey experiment is that people's underlying judgments guide their behavior and that this can be captured by letting them give their judgments on those vignettes (Jasso, 2006). To study what factors influence these judgments, certain aspects of these descriptions are varied. Politicians for example might be more likely to react to coverage of an issue their party regularly reacts to than others. Those manipulated characteristics are referred to as "dimensions", in the experimental literature commonly referred to as factors or also variables. These dimensions or variables in turn can take on several values, for instance the issue crime or asylum seekers. These values are referred to as "levels" in the factorial survey terminology.

It is confusing for some that although referred to as "vignettes", the factorial survey should be clearly distinguished from so-called "vignette studies". Vignette studies also use short descriptions of situations or persons. In contrast to the factorial survey experiment, they however commonly use a between-respondent design with one or two manipulated factors instead of the multivariate design of factorial survey experiments. Next to the within-respondent aspect of the design, sampling of experimental conditions to create statistically efficient designs is a key component of the factorial survey experiment I will elaborate in more detail below. These characteristics distinguish the factorial survey experiment from other vignette or scenario studies which are occasionally used in political science (e.g. Hopkins and King, 2010) and journalism studies (e.g. Kepplinger et al., 1991). Because the factorial survey approach is not yet widely known in the field, the next section introduces the method in detail with the steps required from design to analyses.

2.1.1 The factorial survey from design to analyses

Designing a factorial survey experiment has its challenges. As with other experimental designs, the preparatory phase is crucial as once data is collected, the design cannot be adjusted anymore. Particularly the sampling of experimental conditions, the randomizations required for data collection and finally the statistical procedures for data analyses can add to the complexity of the design. Although a number of methodological articles have been published related to the factorial survey experiment (e.g. Sauer et al., 2011; Dülmer, 2007), only recently the first handbook that describes the method in detail was published by Auspurg and Hinz (2015). The section below introduces readers to this relatively new method in political and communication science research. The steps required from the design to the analysis of the obtained data are presented in an overview in Figure 2.1.

Define variables and their values Experimental approaches can be particularly fruitful when researchers have an idea of the factors that influence the phenomenon studied. The first step is thus to get a clear idea of the dependent and independent variables that should be investigated. For the studies here the dependent variable is the likelihood that either journalists or politicians would act based on the information they receive, either a party press release for journalists or a news report for politicians. The main question then is, which variables to include in the design. Theoretical considerations and past research guide these decisions and the relevant chapters describe the choice of variables and values listed in Table 2.1 in detail. Here, discussion focuses on the methodological considerations which are as important when drawing up a factorial survey.

First, because of the multivariate design where several variables are included at the same time variables cannot be considered in isolation. The choice of one variable potentially affects the effects that can be measured on others. Studies show for example that if there is a scandal politicians have to react publically and possibly take political action (Protess et al., 1987; Cook et al., 1983). If a variable with such strong effects as whether or not information concerns a scandal is included in the experimental design, chances are high that the effects of other variables in the design will be overpowered. In more methodological terms, such a variable runs the risk of being used as a heuristic by respondents. One way to fine-tune the influence of such a variable is to increase its number of values. While this might be challenging in case of political scandals as there are no "half scandals", in a study with journalists the standing of the political actor could be manipulated. Studies of how often political actors are mentioned in the media often compare prime ministers with ordinary members of parliament and conclude that the minister is mentioned in the media more often. In this case, one could think about increasing the values on this variable by including party leaders as a group of actors with

a political standing that lies between a prime minister and an ordinary politician. By including a more fine-grained scale, the contrast between dimensions can be reduced and chances that these are used as a heuristic reduced.

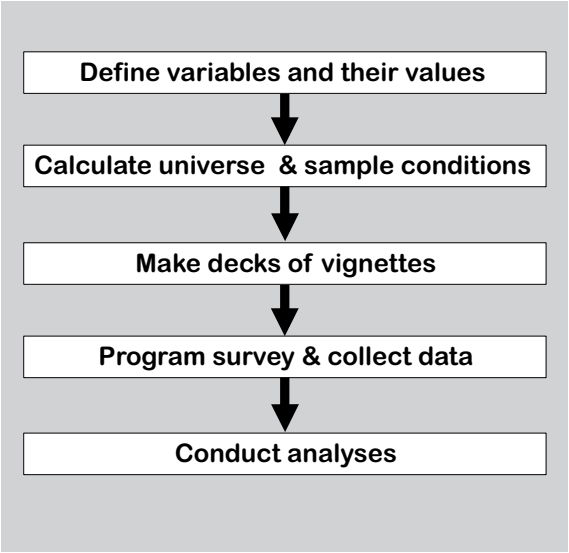
Table 2.1: Experimentally manipulated variables and values of the studies of the selection of messages by journalists and politicians

Selection by journalists	
Variable	Values
Politician's political power	Party leader – ordinary MP
Party's political power	Government – opposition party
Conflict	Government criticism – none
Unexpectedness	Party's not owned – owned issue
Magnitude of political action	Law proposal – question
Selection by politicians	
Reputation outlet	Quality – popular
Negativity	Negative – positive development
Potential for conflict	Responsibility to politics – not [a]
Investigative reporting	Investigative journalism – government report
Party issue ownership	Party's owned – not owned issue

Note. [a]Conflict was originally operationalized with the following four values: responsibility to national politics, responsibility to European Union, responsibility to real world developments, no responsibility mentioned. These were later collapsed into the mentioned two categories.

Another challenge when choosing the variables and according values is the external validity of the design. While multivariate designs often make the descriptions more realistic (Hainmueller et al., 2014) and at the same time allow the researcher to include cases that might not be common in reality such as a government party member criticizing government, some combinations of variables and values might not be realistic anymore. In the study of how journalists select party messages for reporting for example, the influence of the political position of the actor sending the report and his or her party are tested. If all political positions, from the ordinary member of parliament to the (prime) minister, would be included, some impossible combinations of experimental factors would occur; opposition parties do not have ministers in government. In such situations, researchers can either choose to exclude these illogical cases and subsequently account for this situation both when sampling and analyzing the data (see below). Another strategy is to adjust the research design to make sure there are no illogical cases, the approach chosen for the studies presented in the following chapters. Instead of ministers, the influence of a political actors' position on the selection of news by journalists was studied by contrasting ordinary politicians with their party leaders. This is also relevant from a theoretical perspective because it allows studying the influence of political power on a more fine grained level to see whether there are still differences between political positions.

Figure 2.1: Overview of steps in designing a factorial survey experiment



Calculate universe and sample conditions Once the number of levels for each of the dimensions is defined, the "vignette universe" is calculated. It consists of all possible combinations of experimental stimuli and forms the basis for the sampling of vignettes. The sampling of experimental conditions is one key aspect that distinguishes the factorial survey from other experimental designs more common in the social sciences. Usually, all possible combinations of experimental stimuli are included. However, in many cases, such a full factorial design is not statistically efficient because not all possible interaction effects between the included variables need to be estimated. They do not make sense from a theoretical perspective. Sampling of experimental conditions furthermore has the advantage of decreasing the number of conditions that need to be tested without losing precision in the estimation. The smaller number of respondents required can be particularly important when dealing with small elite populations such as journalists and politicians where resources are scarce. In agriculture, medical research or manufacturing fractional experimental designs are much more common (Gunst and Mason, 2009). For factorial surveys, sampling of experimental conditions is guided by considerations about the number of respondents and the number of vignettes each respondent should evaluate.

While many authors rely on random sampling of conditions in the existing factorial survey literature (Wallander, 2009), more advanced sampling procedures perform much better (Dülmer, 2015). In factorial survey research, strategic sampling is based on a number of theoretically pre-defined criteria, making it a form of stratified sampling. The researcher first determines which interaction effects between the tested variables make sense from a theoretical perspective to ensure that with sampling of conditions,

the researcher has maximal control over the information that is lost. Auspurg and Hinz (2015, p. 24f) provide an elaborate account of how sampling of conditions is related to introducing correlation in the design and there is a whole body of literature on fractional factorial experimental designs (e.g. Dülmer 2007; 2015 on factorial survey experiments). Many different sampling strategies are possible depending on the goal of the research project (e.g. Gunst and Mason, 2009, for some examples). For the studies in this book, a half fraction factorial sample is chosen. It means that half of the cases of the full factorial design, all possible combinations, are sampled. The resulting sample of vignettes is both orthogonal and balanced. Balance means that each stimulus is represented an equal amount of times. In an orthogonal design, each of the combinations of values of a variable is measured the same number of times. These two criteria ensure that the correlation between variables is low leading to standard errors that are not inflated by the experimental design. The smaller standard errors in turn increase the likelihood of obtaining significant results, also when measuring small effects. Algorithms can be helpful to identify the relevant sample of vignettes; the program SAS includes an algorithm for example.

Overall, the strategic sampling of experimental conditions is a key aspect in the factorial survey research which can be particularly challenging. Particularly because fractional factorial designs are not (yet) commonly used in experimental research in the social sciences. The examples here, however, show that it can be a fruitful strategy when resources are scarce, for example due to the small respondent population. Such sampling strategies might become more widely used the more researchers are challenged by increasingly scarce resources. At the same time, some researchers call for adopting more factorial designs in the field to study decision making in a more realistic setting (Hainmueller et al., 2014).

Make decks As noted in the beginning, the factorial survey applies a within-respondent design, where multiple vignettes are shown to each respondent. At the same time, not each of the respondents receives all vignettes that were sampled, a feature of a between-respondent design. The factorial survey thus uses both characteristics of within-respondent and between-respondent designs. So-called "decks" of multiple vignettes are thus shown to each respondent.

To make decks of vignettes, the vignettes that were sampled in the previous step need to be distributed across decks. The number of vignettes presented to each respondent depends on the length of the survey and the difficulty of the task for the group of respondents. If respondents are judging very short vignettes tapping into something they do on a daily basis, more vignettes might be presented. Journalists and politicians for example have to evaluate information on a daily basis, evaluating press releases or news reports as the studies here do is therefore a relatively "easy" task. It is important to avoid fatigue effects as this causes respondents to judge vignettes only on a limited

number of variables (Sauer et al., 2011). In the studies here, political journalists are asked to rate six or seven fictional party press releases while politicians evaluate four news reports. This is well within the recommended 10 vignettes per respondent to avoid fatigue effects (Auspurg and Hinz, 2015, p. 122).

As when sampling vignettes, the distribution of the vignette samples into decks should also not be done randomly. All values of variables should be distributed orthogonally and balanced across the decks. This means that all variables and values are represented with equal frequency within a deck of vignettes making the design stronger because respondents are less likely to use a specific variable as a heuristic. If journalists were shown three press releases from one party and only one from another party, this latter single press release is likely rated differently because it is seen as "special". In a balanced design however, a journalist receives two press releases from one party and two from another one. For the studies here, decks of vignettes were balanced on every dimension.[2] Ideally, decks are balanced on every variable included in the experiment.

There are a number of mathematical constraints as to the designs in which this is possible. In the politician study for example, each variable had either two or four values (see Table 2.1). To be orthogonal and balanced, decks need to be a multiple of the number of values for each variable. The smallest number of vignettes is chosen for the politician study, four vignettes per deck. A related consideration can be the expected number of respondents. Each vignette in the sample needs to be evaluated by several respondents to discern respondent from vignette effects. Generally, the more heterogeneous a group of respondents, the more often each of the vignettes should be judged. As a rule of thumb, five judgments should usually suffice for fairly homogeneous groups of respondents. However, more judgments are of course better, particularly if one expects that some groups of respondents will react differently to some manipulated dimensions. In the study of the media's influence on politics for example, opposition parties will react more to negative reporting than government parties research shows. In the final analyses, cross-level interaction effects need to be included between the party of the respondent and the manipulated content of the report. Consequently, a sufficient number of government and opposition politicians need to have evaluated the reports. The bigger the decks, the more evaluations on vignettes are gathered per respondent. For the politician study, a response rate of around 30% is set as a goal as this comes close to what other studies have managed to achieve. Four vignettes per respondent should be sufficient to estimate results.

In sum, the size of the vignette decks is determined firstly by the number of variables and according values chosen. At the same time, the choice is also guided by considering the number of respondents and fatigue effects which might occur depending on the vari-

[2] The exception is the journalist study for situations where respondents receive seven party press releases. Most respondents evaluate 6 press releases in a balanced and orthogonal design.

ables included in the design. A clear idea of the expected response rate can be helpful at this stage. Within decks, the presentation of vignettes is again randomized to control for order effects.

Program survey and collect data So far, the presentation focused on how the vignettes are drawn up balancing theoretical and statistical considerations when designing a factorial survey. Often, one has to move back and forth between the steps described so far. Next to the experimental conditions, the factorial survey experiment usually also includes questions tapping into the background of respondents. (Control) variables are measured that are expected to influence the phenomenon studied. To be able to determine whether a journalists' political orientation influences their selection of political news, information on their political orientation is needed. Similarly, to study whether politicians differ with regards to their reactions to media coverage, data on their interests or tenure need to be collected. That the factorial survey experiment explicitly models differences between respondents in their reaction to the experimental conditions is particularly important for the studies presented here. As elaborated when presenting the overall research design (see section 1.3), journalists and politicians selecting messages are expected to be influenced by factors such as their own background and their media outlet or their political party. In factorial survey experiments, variation between respondents is expected and modeled into the design (Auspurg and Hinz, 2015, p. 88).

In a factorial survey, each respondent receives a slightly different survey. First, respondents only receive one deck of vignettes, assigned randomly from all the decks that were created in the previous step. Similarly, within decks of vignettes, the order is randomized to avoid order effects. Their strength mainly depends on the complexity of vignettes and the number of dependent variables research shows (Auspurg and Jäckle, 2012). Although it is possible, including these randomizations in a paper and pencil survey is a lot of work, and not only because a different paper survey needs to be printed for every respondent. The randomizations and differences between surveys also need to be taken into account when the data are entered into the system. This process is susceptible to mistakes if not carried out carefully.

Online survey tools make this somewhat easier and the tool Qualtrics is chosen for the studies here. Next to including the randomization and a wide possible range of other survey questions, Qualtrics allows sending personalized survey links to respondents to keep track of which ones need to receive a reminder. Additionally, it is easy to collect data on a tablet computer important for data collection in the Swiss parliamentary buildings for the study of the selection by politicians. The methodological sections of the relevant chapters elaborate how data is collected for the studies.[3]

[3] A manual on how to program a factorial survey in Qualtrics is available on my website at http://luziahelfer.wordpress.com/manuals

Conduct analyses Once data are collected, the responses on the vignettes need to be connected to the manipulations included in each of the vignettes again to be able to analyze the results. There are a number of different strategies that can be used for analyzing the data. First, the isolated effects of the experimental manipulations can be estimated. Because each respondent evaluates multiple stimuli in the within-respondent design, observations are not independent. Multilevel regression models can account for this clustering of observations per respondent, allowing for a correct estimation to ensure that the null-hypotheses is not incorrectly refuted. T-tests or other approaches often used to analyze experimental data are not suitable for factorial survey experiments.

Next, in a factorial survey experiment the evaluations of the vignettes depend on the background of respondents. Journalists from media outlets that are published weekly for example might be less inclined to react to a party press release than those working for a daily newspaper (Abbott and Brassfield, 1989). To model these differences between respondents, such respondent variables can be included in the multilevel regression models on a second level. Those background characteristics can either be obtained in the survey following the experimental stimuli, or from other independent sources. For the journalist survey, most of these variables are obtained in the survey while politicians' party membership or their field of specialization are from official parliamentary record. If respondents are again distributed into subgroups, for example government and opposition parties or classified in specific media organizations, more levels can be included in the regression models.

Finally, the multilevel approach also allows the estimation of interaction effects between the manipulated variables and the respondent's background. These interaction effects are particularly interesting because they show whether some respondents are affected differently by some content. Senior politicians for example react differently to media content than junior ones. It also allows studying whether opposition party politicians are indeed more likely to react to negative coverage while their government party colleagues react to positive coverage (Thesen, 2012). Overall, including interaction effects help study the conditionality of an independent variable's influence on the phenomenon at hand even further.

The parallel comparative design (see section 1.3) where the exact same experiments are carried out in two countries provide the possibility for additional analyses. While the above described analyses strategies allow to compare the influence of specific variables, analyzing the Intraclass Correlation (ICC) coefficient can provide additional insights. It is a measure for comparison across contexts. Like other correlation measures, it ranges from 0 to 1 and estimates the proportion of variance of the dependent variable due to variation among respondents. In the words of Auspurg and Hinz (2015, p. 89), "this coefficient states how much of the variance of the outcome(s) is a reflection of different respondents evaluating the vignettes." Higher values indicate that respondents are very

similar in their evaluation of the vignettes, while lower values mean that they are less in agreement on how to evaluate the stimuli. In the concluding section of this book where the studies are compared on a more general level (see section 5.1), the ICC serves as a measure to compare between journalists and politicians and also between countries. A one-way ANOVA model was used to calculate the ICC on a baseline model only including the evaluation as the dependent variable and controlling for the cluster (respondents) with the loneway command in Stata 13.

2.1.2 Strengths and weaknesses

Generally, a method's strengths and weaknesses depend on the phenomena being investigated: one method might be well-suited for answering one research question, but not for studying another phenomenon. Internal and external validity are the "gold standards" according to which research designs are evaluated. It usually involves trade-offs: maximizing internal validity in experimental designs often involves accepting trade-offs in external validity. For many research questions in the social sciences, the factorial survey experiment allows maximizing both. It combines maintaining control over several variables of an experiment with the possibilities of obtaining a representative sample of respondents in survey research, something that is often not possible for experimental studies (Atzmüller and Steiner, 2010). Additionally, because of the multivariate design, relatively realistic experimental conditions can be created. Before elaborating on the advantages of a factorial survey experiment however, it is important to also point out the challenges of applying such a design.

The factorial survey experiment can be challenging to set up due to the complexity of the multivariate design. Researchers require solid methodological knowledge of experimental designs and statistics. Particularly sampling the vignettes and programming the survey can be challenging (see subsection 2.1.1). Also when analyzing the data, the complexity of the within-and between-respondent design should be taken into account.

Some researchers might find it challenging that sampling of experimental conditions is not (yet) common in political science or communication science. However, there is a body of literature on sampling of experimental conditions from other fields such as psychology and even more in the natural sciences. Moreover, methodological studies on the factorial survey method are relatively scarce (for exceptions see for example Dülmer, 2015; Sauer et al., 2011). However, the recently published handbook on the method provides a concise set of recommendations on how to design a factorial survey (Auspurg and Hinz, 2015). Nevertheless, because of its use of scenarios the factorial survey is not a completely new method in the field of course and multivariate experimental designs are not new either. Even more complex multivariate designs for which only a small fraction of all experimental conditions are sampled would be able to make full use of the potential of the factorial survey experiment.

The factorial survey experiment has a number of advantages. One of the most important aspects is that, as emphasized earlier, a statistically efficient experimental design paired with the mixture of the within-and between-respondent design can significantly reduce the resources associated with conducting an experiment. This does not only relate to costs due to the length of surveys or the number of respondents, but also to the burden for each respondent.

Another advantages of the factorial survey already mentioned is the high internal validity. Internal validity refers to the extent to which the researcher can be sure that the inferences drawn are indeed caused by the variables she or he assumes. Do politicians indeed react to the media content they have seen, or did they obtain this information from colleagues or other sources? Because experimental design allow maximal control over the context, many researchers consider them superior to other methods, such as surveys (McDermott, 2009). Experimental designs often, however, score lower when it comes to the general applicability of the findings, the ecological validity. Subjects have to be placed in laboratory settings or have to answer in a survey and do not behave as if no one was watching them. This can lead respondents to give answers that are more socially desirable impairing the external validity of the studies. For example, journalists are not likely to say that they have a preference for some parties over others because impartiality and objectivity are the most important norms of journalistic reporting (Brown, 2011). Because of the multivariate design of factorial survey experiments, the risk of respondents showing social desirability bias is reduced compared to conventional survey items (Gaines et al., 2007; Alexander and Becker, 1978, p.95). Instead of asking journalists whether they prefer some parties over others to measure the effects of a journalists' political orientation, respondents are asked to rate a number of party press releases. The researchers' goal of measuring differences between parties is obscured because other variables are manipulated at the same time.

External and internal validity always have to be seen as a balance. In many experimental designs maximizing internal validity means making trade-offs with regard to external validity. So which is more important? "Perhaps the best way to conceptualize the balance between internal and external validity in experimental design is to think about them in a two-step temporal sequence. Internal validity comes first, both sequentially and practically" (McDermott, 2009). Factorial surveys have the advantage that these trade-offs are often relatively limited. Because several variables are manipulated simultaneously, factorial survey experiments allow constructing more realistic experimental conditions. Journalists and politicians for example are exposed to a constant stream of complex information and have to decide what to do with it. This complexity of reality can be modeled in multivariate designs (Hainmueller et al., 2014). Additionally, because data is collected through a survey, usually more respondents are willing to participate than in a laboratory setting. Particularly, elites like journalists and politicians, who are

not likely to agree to participate in experiments, can be reached more easily with a survey. This means that experimental research with these elites becomes more feasibly, and findings more generalizable because of the higher number of respondents.

While factorial surveys have clear advantages and score high on external validity because of the multivariate design, they ultimately only measure intended and not actual behavior. This is a drawback and should always be discussed by researchers. In an ideal world, researchers would be able to follow politicians or political journalists around, record all the information these actors receive and their behavior. Without considerable resources and complete cooperation of respondents, such designs are usually not possible. There are some cases where journalists collected information they received and allowed researchers to analyze the data: so called input-output gate keeping studies starting with Whites' (1950) seminal study, or more recently Gant and Dimmick (2000). Although more rare, similar studies are conducted with politicians as subjects. Orton and colleagues (2000) followed two British members of parliament for a duration of four weeks to study their information seeking behavior. Although such studies indeed measure actual behavior of those actors, their generalizability is lower because they are limited to only few cases. While in factorial surveys only intended behavior can be assessed, they do provide the advantage that data can be collected from a larger group of respondents relatively easily thus leading to considerably more generalizable findings if the project is carried out with care. This combination of factors makes the factorial survey experiment particularly suitable to study how journalists and politicians select messages, the question this book sets out to investigate.

2.2 Alike but different: Switzerland and the Netherlands

Journalists and politicians do not operate in a vacuum. They political and media system within which they are embedded provides the boundaries for their actions, for example through certain norms and values. The comparative nature of the studies presented in this book provide the unique opportunity to study how such differences affect the behavior of these actors and the politics-media relationship more broadly. Although Switzerland[4] and the Netherlands share many characteristics, particularly with regards to their media systems, there are a number of important differences that likely influence how politicians and journalists react to each other. Two macro level political system characteristics that differ between these two countries are particularly interesting with regards to selection. First, the distribution of political power between parties affects how journalists select news coverage and also politicians' reactions to news coverage. Journalists follow the "trail of power" (Bennett, 1996; Gans, 1979) and the fact that there is no real opposition in the Swiss parliament affect how journalists select news. Second, it is particularly intriguing to investigate, how politicians as strategic actors react to different kind of news content due to differences in the electoral systems of the two countries. While there is a body of literature on legislative behavior, it has seldom been applied to study reactions of politicians to news coverage.

With this comparative approach, the studies can make important contributions to our understanding of the nature of the politics-media relationship in multiparty systems more generally. Donsbach and Patterson (2004, p. 253) already noted that despite being particularly interesting, "explanatory comparative analyses, however, are virtually nonexistent, even though the field is of high scientific interest, especially where theories of news selection are concerned. How much of the explained variance can be attributed

[4] For Switzerland, this book focuses on the German speaking population of political journalists and politicians. There are four official languages spoken in Switzerland. German is the first language of the biggest part of the population (64.5 %) and has the biggest group of representatives in parliament. This is followed by the French (22.6%) and Italian (8.3%) speaking population. Only a small minority speaks Rhaeto-Romanic (0.5%). To draw general inferences, this study required a sufficiently large population of politicians. Consequently, I chose to focus on the German speaking members of the Swiss Lower House and the respective journalists. While in politics there is no apparent segregation between the language regions with regards to the political parties that can be voted on or the electoral system which might affect politicians' behavior, there are important differences with regards to the media systems which is more fragmented. Each language region has its own broadcasters and the newspapers often have a local focus although the major ones are read across language regions. The choice was made to focus on the biggest language group, the German speaking members of parliament and accordingly journalists of German language outlets. As underlined before, politicians and journalists are hard to reach elite populations and chosing for the biggest group makes it more likely that results can at least be generalized to this particular group and subsequently compared with the results obtained in the other country, the Netherlands. Because the electoral system expected to affect the way politicians react to media coverage is the same across language regions in Switzerland, results are expected to generalizable to members of parliament from other language regions in Switzerland (see subsection 4.4.3).

– technically speaking – to general patterns of human behavior and how much to specific circumstances, is an interesting question." To be able to isolate how system level macro variables affect the complex interaction between politics and the media, focusing the studies on two countries instead of multiple ones has a number of advantages. Next to making a comparative study with hard to reach elite actors feasible, this "strategy of paired comparison" (Tarrow, 2010) can be particularly insightful. Both characteristics of the media and the political system are likely to influence the interdependent politics-media relationship. Including more diverse countries can have the advantage of many degrees of freedom, allowing the researcher to statistically test for the influence of very specific political context variables on the phenomenon studied (Tarrow, 2010, p. 239). However, such strategies are often not fruitful when it comes to investigating more complex relationships. Despite increasing degrees of freedom with more countries, also the number of confounding factors on which those countries differ increases. This can make it very difficult to draw valid causal inferences, particularly when studying the politics-media relationship. By opting for a two-country comparison, the context and its potential effects on the phenomenon studied can be better controlled.

Additionally, construct equivalence of the experimental stimuli is more likely if only a limited number of countries are chosen (Landman, 2008, p. 69). Particularly if an experimental approach is applied in a cross-country setting, drawing up experimental conditions that can be applied across contexts is challenging. For the present studies, it is particularly important that news reports and party press releases are perceived as realistic by respondents. Relatedly, respondents need to have the same understanding of the meaning of these messages. This so called construct equivalence is the minimal basis for drawing valid comparisons in comparative research (Wirth and Kolb, 2004, p. 88). Including more countries would likely mean that part of the validity of the research would have to be sacrificed (Livingstone, 2003, p. 488). For example, to study effects of party issue ownership, isolating effects of the issue itself and the party is important. If across countries, different parties would be associated with different issues each time, such a comparison would soon become obsolete because more issues would need to be included.

In sum, comparing two relatively similar countries allows an in-depth study of how the political system affects how journalists and politicians select messages instead of an investigation of many country level variables. The number of confounding variables is limited by selecting cases that are as similar as possible on a number of key characteristics and thus "capture diversity within a common framework" (Livingstone, 2003, p. 487). Comparing two cases is "an intermediate step between a single-case study, which suggests a general relationship, and a multicase analysis that tests or refines a theory" (Tarrow, 2010, p. 245). For studying how politics and media influence each other, this approach can provide particularly fruitful insights. In the next sections, the media and political

systems of Switzerland and the Netherlands are briefly presented. Both countries are part of the democratic corporatist models (Hallin and Mancini, 2004) with strong traditions as consociational democracies with weak ties between media outlets and politics. The key differences expected to influence the selection moments studied here, in particular the distribution of political power and the voting system, are discussed in more detail.

2.2.1 Media systems compared

Related to the media system and the political reporting styles, Switzerland and the Netherlands are highly similar. First, the characteristics of the media system, and in particular its relationship with the political system, in both countries are strongly alike. According to the seminal work of Hallin and Mancini (2004, p. 143), Switzerland and the Netherlands both belong to the 'democratic corporatist' type. The printed press has a high circulation with former partisan ties that have however subsided in the past years (Bakker and Scholten, 2014). Newspapers cannot be attributed to a particular political party anymore. Like other democratic corporatist countries, the freedom of the press is highly valued and seen as an integral part of the democratic institutions. In the past years, like all media the printed press market is more and more exposed to a strong competition for readership, particularly in the fragmented Swiss system where most regions have their own publication. However, a number of newspapers are read across the country, one popular (*Der Blick*) and one considered more a quality newspaper (*Neue Zürcher Zeitung*). Whether politicians are affected differently by either one of these newspapers is tested in the study of the selection by politicians (see chapter 4). Next to their considerable readership, they are chosen is because they compare best with national Dutch newspapers.

The Dutch print market is also divided into a strong national press and newspapers with a regional focus. In urban areas, national newspapers are more important, while in other areas regional newspapers are more important (Bakker and Scholten, 2014, p. 22). The present study focuses on national titles, of which the popular newspaper *De Telegraaf* has the highest circulation. There are also a number of daily newspapers that are considered more quality outlets (e.g. *NRC Handelsblad*). With its focus on comparison between countries, two newspapers were chosen that are similar to the Swiss ones mentioned before. They are expected to play a comparable role for politicians when they consume the news; they are widely read by the public and have a dedicated staff for political reporting.

With regards to broadcasting, Switzerland and the Netherlands have strong public service broadcasters. Its position is however much stronger in Switzerland than in the Netherlands. In Switzerland, it is the undisputed market leader in current affairs across language regions. The situation presents itself differently in the Netherlands. Since the market was opened up in 1990, the Dutch broadcasting market is more competitive. Since

its foundation, the private channel RTL has been able to build a reputation with its main evening news broadcast RTL Nieuws (market share 24,1 %) aired half an hour before the main evening broadcast of the public service provider NOS (market share 34,2 %). In a re-analysis of the Hallin and Mancini classification, the relatively weak influence of public broadcasters in the Netherlands led scholars to compare the Dutch media system to that of countries such as the US (Brüggemann et al., 2014). Nevertheless, the competition has not (yet) led to a substantively different type of coverage of Dutch politics, coverage has not become dominated by criteria of sensationalism and negativity so often ascribed to the liberal media system in the UK and the US (Hallin and Mancini, 2004).

In his study of political journalists and their reporting in several European countries, Van Dalen (2010) shows that there are structural differences and divides countries into those with a more pragmatic and sacredotal culture. In another comparative study Esser and Umbricht (2013) propose a similar distinction. They distinguish between a US model of rational news analysis, an Italian model that is based on polarized reporting and a Germanic model of disseminating news with views. Both Switzerland and the Netherlands are part of a more pragmatic journalistic news culture and belong to a Germanic model in their reporting styles, in line with the characteristics associated to a democratic corporatist country (Hallin and Mancini, 2004). Although no longer recognizable nowadays, both the Swiss and Dutch printed presses have parted from their partisan affiliations relatively late (Blum, 2005; Pleijter et al., 2012). Generally, the style of political reporting in the media is very similar in both countries. In the Netherlands, the media are now looser from their former "pillarized" origins and more commercially oriented (Van der Eijk, 2000, p. 312). Since the late 1960s, reporting has become more critical of the political elite (Brants and Van Kempen, 2013; Brants and Van Praag, 2006). Similar developments have taken place in Switzerland, although commercialization trends are probably less pronounced than in the Netherlands. Research on the media coverage of politics in campaign environments of both countries indicates that although conflict and horserace coverage is on the rise, a substantial part of the reporting still focuses on actual issue positions (Kleinnijenhuis et al., 2007; Hänggli and Kriesi, 2010). Journalists' reporting on politics is guided by the goal of providing analysis and interpretation (Van Dalen and Van Aelst, 2012, p. 520 for the Netherlands).

In sum, although the Swiss and Dutch media systems are clearly not identical, they are comparable on many aspects related to political coverage. The degree of government intervention and the reporting styles of politics are similar. As a consequence, differences in journalists' selection of news are much more likely explained by differences in the political system, as the next section shows.

2.2.2 Political systems compared

There are a few studies that directly compare politics and political traditions of Switzerland and the Netherlands. Daalder (1971) for example identifies common factors but also differences in how the two countries developed historically, some of which are still visible today. More recently, Schenkel (2000) compared policy making on climate issues in the two countries. He shows how these countries, despite shared traditions and many similarities, take different approaches to climate policy and its formation. With regards to the politics-media relationship studied here, these contributions first underline the usefulness of such a comparative approach; the countries share many similarities but also have important differences, some of which will affect the relationship between politics and the media. I will first briefly mention the common characteristics of the two countries before discussing the differences expected to affect how journalists and politicians select messages. Table 2.2 gives an overview of several important aspects related to the political systems of the two countries.

Table 2.2: Comparison of Swiss and Dutch political systems

Characteristic	Switzerland	Netherlands
Number of chambers	2	2
Number of voting districts	26	1
Electoral system (Lower House)	proportional	proportional
Seats Lower House (total)	200 (246)	150 (225)
Turnout at national elections (year)	49% (2011)	75% (2012)
Composition of executive	"Zauberformel"	Coalition parties

Switzerland and the Netherlands are countries with a strong tradition as consensus democracies. They belong to the democratic-corporatist democracies in the widely used classification by Hallin and Mancini (2004). Countries in this group have a history of early democratization and organized pluralism with a strong welfare state and independent media. In both Switzerland and the Netherlands, there is a bicameral system in parliament where multiple parties are represented. Moreover, the executive is supported by the majority of the elected parliament, albeit in different ways. In comparison to other countries such as the US, Switzerland and the Netherlands have a balance of power between the executive and the legislative branch. With their tradition as consociational democracies, politicians in both countries search for the best compromise and focus on inclusion, in order to obtain the largest majority possible in support of legislation. Both Switzerland and the Netherlands rank among the top 10 most corporatist countries in the world (Siaroff, 1999, p. 198). There is a close relationship between interest groups and the government, which is largely based on cooperation and influence of support. In Switzerland, for instance, interest groups are already part of policy making process in the pre-parliamentary phase (Scarini, 2006,

p. 496). In the Netherlands, advisory expert comittees are also an integral part of the policy making process (Andeweg and Irwin, 2014, p. 172).

While the two countries share many similarities, there a few important differences which affect the politics-media relationship. First, the electoral system and the level of control it gives parties over who gets elected. Although the legislative branch is elected through a system of proportional representation in both countries,[5] in Switzerland the high level of federalization paired with the open list system provide a different motivational structure for national-level politicians to react to media coverage. The allocation of seats in the Dutch Tweede Kamer is based on the d'Hondt method, which results in a very high degree of proportionalism by treating the whole of the Netherlands as one single electoral district (Andeweg and Irwin, 2014, p. 98). Political parties have major influence over who gets elected because of the semi-closed list system; although possible, preference votes hardly affect the outcome of Dutch elections. In contrast, in Switzerland preference votes have a big influence on election results in the open-list proportional system. Switzerland is a highly federalized country with a total of 26 cantons which serve as electoral districts within which the seats of both Chambers are allocated. These cantons have far reaching powers because according to the constitution, all functions that are not explicitly attributed to national institutions remain with them (Vatter, 2008, p. 82). For politicians, this means that they likely have different reasons to react to media coverage. While in the Netherlands they may want to make sure to toe the party line and not upset the leadership, Swiss politicians will focus on cultivating a vote at the local level and will not be as concerned with the division of labor within their party for instance. I will elaborate on this aspect in more detail in the relevant chapter later in this book (see subsection 4.2.3).

A second important difference between the Swiss and Dutch political system is expected to affect how journalists select political news. It relates to the distribution of political power between political parties in parliament, in particular to the presence of a (strong) opposition. Although in both countries executive government consists of several different parties, in Switzerland there is no "real" opposition. In Switzerland the government (*Bundesrat*) is elected by parliament on the basis of a 'magic formula' (*Zauberformel*). It ensures that all major parties are represented largely according to their electoral strength. Once elected for a term of four years, the Council cannot be dissolved. If a member decides to step down, a new one is chosen. As a consequence, parliamentary elections do not affect the composition of the executive directly. This is in contrast to the Netherlands, where the government is formed on the basis of a coalition of parties, which in turn appoint their ministers. If the coalition falls apart,

[5] The Lower Houses of Switzerland and the Netherlands are both elected through proportional representation within voting districts. Because the study in this book was conducted with representatives of the Lower Houses only, differences in the electoral system of the Upper House are not relevant in this context.

parliament can be dissolved by the government. This difference in the distribution of political power also has some consequences for policy making.

An important consequence of how power is allocated between parties is the say of each of these parties over legislation. In the Netherlands, the political parties in government usually hold a majority in parliament, which means they can determine legislation (Andeweg and Irwin, 2014). In recent years, the coalition tends to be a minimal winning coalition, and always leaves several parties in opposition. In Switzerland, however, the four biggest parties in parliament are represented in government.[6] These parties hold different positions on most issues which means they hardly ever vote unanimously. "Coalitions" are formed on the spot, depending on the issue at stake (Linder et al., 2006). Parties might agree on some issues and thus work together, while on others they are actually opponents. Hence, power is fairly evenly distributed among the major parties (Kriesi and Trechsel, 2008). Consequently, while Dutch government parties have significantly more political power than those in opposition, no clear distinction can be made in terms of political power between the main parties represented in the Swiss government. This affects journalists' selection of political messages because they want to report what is consequential and important. As I will elaborate in more detail later in this book (see chapter 3), political actors with more power are more present in news coverage. Applied to the studies here this means that while in the Netherlands coalition parties have a distinct advantage to have their messages selected, such a mechanism is not present in the Swiss case. There, because no party has significantly more power than another, journalists weigh other message aspects more for their selection.

In sum, as traditional consensus democracies with a multi-party system, Switzerland and the Netherlands share many characteristics. They also differ in several aspects, such as electoral systems. Or the distribution of political power in parliament. These two aspects in particular are expected to have an influence on the relationship between politics and media studied here. In the relevant empirical chapters, I will formulate hypotheses relating to the selection mechanisms of politicians and journalists.

[6] In 2003, a change has taken place when political party BDP entered the government, despite being a small party due to its split from the SVP. The subsequent shift towards a more polarized system, although still without a substantive opposition in the Swiss parliament, has led some authors to conclude that Switzerland has moved from the extreme case of consensus democracy more towards other systems that fall in this category (Vatter, 2008, p. 11).