UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Expanding the methodological toolbox: Factorial surveys in journalism research

Otto, L.P.; Glogger, Isabella

Citation for published version (APA):
Otto, L. P., & Glogger, I. (2020). Expanding the methodological toolbox: Factorial surveys in journalism research. *Journalism Studies*, *21*(7), 947-965. https://doi.org/10.1080/1461670X.2020.1745663

Expanding the Methodological Toolbox: Factorial Surveys in Journalism Research

Lukas P. Otto & Isabella Glogger

Routledge
Taylor & Francis Group

Expanding the Methodological Toolbox: Factorial Surveys in Journalism Research

Lukas P. Otto [a] and Isabella Glogger[b]

[a]Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Amsterdam, The Netherlands; [b]School of Journalism and Mass Communication, TU Dortmund University Dortmund, Germany

ABSTRACT

Experimental designs to examine attitudes and behavior are crucial to make causal inferences. However, studies that assess attitudes and behavior of journalists are still dominated by correlational designs, such as used in surveys with journalists. Elaborating on historical and practical reasons for that, we argue in this paper why journalism scholars may benefit from including a certain experimental approach to their toolbox: the factorial survey experiment. Using data from a factorial survey with German newspaper journalists, we illustrate the application of factorial surveys from their conceptualization to the data analysis. Suggestions for further fields of application are made.

Ever since journalism became an empirical discipline within communication research, scholars are interested in how journalists work: how they *decide* which news to select, how to frame certain events, or which stories to publish (Hanitzsch and Wahl-Jorgensen 2009). Journalistic decision-making, thinking, and behavior are the focus of journalism research; consequentially, some of the most important theories in the field describe journalistic thinking and behavior, be it agenda-setting, framing, or gatekeeping (see for an overview, Wahl-Jorgensen and Hanitzsch 2009). Despite the importance of the investigation of journalistic perceptions, decision making, and behavior, journalism scholars scarcely make use of a methodological design that is especially suited to describe and predict these kinds of processes, to draw causal inferences and, in turn, to test theories – experimental designs. Among these experimental designs, the factorial survey experiment is particularly unknown, not only in journalism studies, but in communication research in general (Reineck et al. 2017).

Factorial survey experiments or factorial surveys (FS) are widely applied in other social sciences such as sociology (Wallander 2009). They are particularly suitable to test models, theories, and hypotheses in journalism studies since they: allow for testing *multidimensional* situational influences, come with high power, and can include higher-level variables on the person, organization, or even country level. They combine the advantages of experiments, which have been rather rarely applied so far to study journalistic behavior, with the

widely used survey method in this field (Hanitzsch and Wahl-Jorgensen 2009; Steiner and Atzmüller 2006). FS, in contrast to the "traditional" survey experiment, can investigate complex situations and, at the same time, carry the strength of causal inferences (Auspurg and Hinz 2015a). Given these advantages, this paper is concerned with the application of FS in journalism research, to be precise with studying attitudes and behavioral intentions of journalists. We argue – based on a review of literature on FS in communication science and an empirical example of an FS with newspaper journalists – why journalism scholars should add FS to their methodological portfolio.

To draw this conclusion, we will briefly discuss historical but also practical reasons why journalism scholars might so far have refrained from using experimental designs when studying journalistic attitudes and behavior. After which, we show that some of these challenges journalism studies face can be solved by using *factorial survey experiments*. Finally, we introduce this design and illustrate its application in a study on hard and soft news perception of print journalists in Germany.

The Method Repertoire in Journalism Studies

When describing the field of journalism studies, many research questions, models, and theories focus on journalistic decision making, such as, news values, agenda setting, and the question of how researchers can explain and predict journalistic working processes. To explain and predict these processes, scholars make use of the whole (social science) method repertoire. Scholars use content analyses, for example, to infer from the journalistic product to journalistic selection processes (e.g., news values (Harcup and O'Neill 2001)). Qualitative interviews with journalists are especially suitable to gain in-depth knowledge about the work in the newsroom since these very complex situations are hardly measurable in surveys (Dupagne and Garrison 2006) and the amount and quality of data accompanying digital journalism make it possible and necessary to apply big data and automated methods to the field (Boumans and Trilling 2016).[1]

However, the usage of research methods and designs in a discipline is not always a deliberate and logical decision, but rather reflects historical developments and self-conceptions of a field. The sociological turn in the 1970s and 1980s has brought about an extensive repertoire of qualitative approaches, reflecting critical standpoints and theories in journalism (Hanitzsch and Wahl-Jorgensen 2009). The global comparative turn in the 1990s and technical developments of the decades after made it easier to establish comparative large scale surveys and comparing journalistic norms and cultures across countries (Hanusch and Hanitzsch 2017).

Yet, when investigating and describing the methods used in journalism research, it becomes obvious that designs other than *experimental designs* are dominant. For example, when searching the abstracts in the database "Communication and Mass Media Complete", and entering "journalism" in combination with "content analysis" there are four times more hits than for "journalism" and "experiment"; nine times more hits for "journalism" and "survey" as well as for "journalism" and "interview". [2] Of course, this can only be a subtle hint for an unequal distribution of methods in journalism studies. Yet when comparing the same methods to the neighboring discipline of political communication, the method distribution is less skewed towards non-experimental designs.[3]

This trend and composition of methods in the field are confirmed by systematic reviews and content analyses of published journalism research in the last decades. Covering the period from 2000 to 2007, Cushion (2008) was not able to find any experimental approaches in the journals "Journalism Studies" and "Journalism: Theory, Practice and Criticism". A recent content analysis of published research in the realm of journalism studies, covering a period from 2000 to 2015 and a variety of communication journals, showed that only around 2% of articles used experiments (Hanusch and Vos 2019). In contrast, covering the whole field of mass communication in the 1980s and 1990s, Kamhawi and Weaver (2003) state that 13% of all studies apply experimental designs.

These are notable observations since experimental research has not only become one of the most intensively used approaches in communication research as a whole (Matthes et al. 2015) but also carries strengths that cannot be provided by other methods. Experiments are (a) the research design to establish *causal* relationships, and (b) the most suitable for testing *theories* (Thorson, Wicks, and Leshner 2012). Explaining the investigation of causal inferences goes well beyond the scope of this paper, but establishing a causal relationship between two or more variables usually needs experimental designs (Leshner 2012). A consequence of this notion is that theories, representing a description of the relationship between a set of variables, are to be tested within experimental designs.

Concerning this, experimental designs should be one tool in the toolbox of journalism scholars. Yet, the question of why they are less prevalent than in other disciplines remains. From our point of view, there are at least five reasons for this observation. First, there is the historical development of journalism studies, which, in contrast to other sub-disciplines of communication science, is characterized by qualitative approaches. This should of course not be criticism, but a strength of the sub-discipline, especially in a time where qualitative research is scarce and scholars and journal-editors call for high-quality qualitative approaches (Haven and van Grootel 2019). While the historical development and self-conception of journalism studies close the gap for qualitative approaches, it might explain why experimental designs are still less prominent in the field. Of course, this is not to say that – until now – there are no experiments with journalists.[4] There are, of course, some remarkable exceptions. For example, Patterson and Donsbach (1996) applied a quasi-experimental design to identify partisan bias in news selection and journalistic decision processes. In this study, they even used short vignettes, similar to what this paper presents, however, without making use of a factorial design. More recently, Mothes (2016) also applies a quasi-experimental "vignette design" to investigate how confirmation bias and objectivity values shape perceptions of news for journalists and laypersons. Additionally, in a field experiment using a vignette within a mail to journalists, Graves, Nyhan, and Reifler (2016) examine the usage of fact-checking in journalism. Furthermore, in applying a within- and between-subject approach, McGregor and Molyneux (2018) investigated the influence of tweets vs. headlines on the perceived newsworthiness of a news item. These seminal examples show that experiments can contribute to the field and are certainly able to answer long-standing questions about journalistic bias, as well as more recent developments of digital journalism.

Second, researchers might argue that journalism scholars are often interested not only in *situational* influences on journalistic thinking and behavior, but (maybe even more) in variables on the person-level, e.g., education or attitudes of journalists, the organizational level, e.g., market-orientation of a medium, or the country-level, e.g., the media system in a country. In fact, experimental designs are most suitable when it comes to the interplay of

situational, personal, and higher-order factors on the organizational- or even country-level. After all, experimental designs are the only research designs that can precisely distinguish these influences since some of these variables can be manipulated and the interplay between different levels of variables can be determined, as we will see in this paper.

Third, journalism scholars might claim that journalistic processes are too complex to explain in a simple experiment that manipulates a few factors without addressing the complex situations journalists face. After all, "traditional" (between-subjects) experimental designs are limited in terms of factors they can manipulate and are, therefore, not able to carry the characteristics of very complex situations.

Fourth, a very practical reason to refrain from using experimental designs is the necessity to use journalists as participants when aiming to investigate their behavior. Journalists are – depending on the definition of the population, the country, and the circumstances – not a large group and not easy to access for journalism scholars. If you now imagine a classical between-subjects experimental design, dividing your very valuable, busy, and small sample into groups of (at least) two, this seems to be a very high effort for such a high-value sample.

Fifth, the external validity of experiments is a major concern for scholars (Hanasono 2017). While other disciplines like psychology often use simple stimuli, which allow for a highly controllable experimental design, communication scholars are interested in the effects of media content which is – by nature – more complex, and its reception and production is embedded in an as equally as complex social setting. However, real-world media stimuli decrease the internal validity of an experiment. A solution to this problem is using media stimuli that are less complexed and only designed for the experiment – an approach that faces the criticism of being highly artificial (Klimmt and Weber 2013).

In sum, many of the aspects that keep journalism scholars from conducting experimental research with journalists can be faced by using a certain experimental design – the factorial survey experiment. We now describe the method and design, explain strengths and weaknesses for journalism research, and show an application of the research design.

Factorial Surveys

Factorial surveys (FS) belong to the family of multifactorial experiments which all aim at assessing participants' attitudes, preferences, and behavioral intentions (Auspurg and Hinz 2015b). To reach this aim, multifactorial experiments make use of short descriptions of objects, situations, or individuals which or who are composed of a set of characteristics (also called *dimensions*). These dimensions can have various *values*. Next to factorial surveys, conjoint analyses and choice experiments are the most prominent representatives of the multifactorial experiment family (Auspurg and Hinz 2015b).[5]

FS, also known as vignettes studies, have been introduced as early as in the 1950s. Peter Rossi was confronted with the problem of describing the relationship between social status and socio-demographic variables for his dissertation. His advisor, no one less than Paul Larzarsfeld, suggested "to create 'vignettes' describing fictitious families whose essential status characteristics … would be described in thumbnail sketches" (Rossi 1979, 178). These thumbnail sketches or vignettes are the centerpiece of FS and represent fictitious descriptions of objects, individuals, or situations. In vignettes, characteristics of these objects, individuals, or situations are systematically varied.

In FS, participants get presented with the vignettes and asked to evaluate them, stating their attitudes or judgments towards the described object in the vignettes (Auspurg and Hinz 2015a). The systematic variation of the characteristics of the objects or situations allows analyzing the impact of each characteristic individually on the participants' evaluation of the vignettes (Jasso 2006). FS are incorporated in a survey that allows, furthermore, assessing the influence of characteristics of the participants on this evaluation.

FS are described as a mix of experiments and surveys, combining the best of both in one approach (Steiner and Atzmüller 2006). Like experiments, FS use a manipulation of an independent variable which allows making statements about the causal relationship between the independent and dependent variables (Auspurg and Hinz 2015a). Due to randomization, the influences of confounding variables get eliminated and the independent variables manipulated in the vignettes are not correlated with respondent characteristics (Alexander and Becker 1978). In contrast to experiments, participants in FS get presented with several stimuli (Auspurg and Hinz 2015a). Like surveys, FS make use of questionnaire items to assess, for example, characteristics of the participants, such as journalistic political leaning (Helfer and van Aelst 2015). Contrary to surveys, FS "force respondents to make judgements based on trade-offs" (Auspurg and Hinz 2015a, 11). For example, rather than stating which factors impact publication decisions and probably raking all of them equally important, participants in a FS are forced to decide based on the given factors simultaneously. Hence, the combination of the experimental design of a FS incorporated in a survey makes it possible to observe decision-making processes in a more direct way (Nisic and Auspurg 2009).

Communication scholars, for example, have used FS in the realm of political communication. Henn, Dohle, and Vowe (2013) were interested in the very concept of political communication and assessed what communication scholars and students understand by it. Others focused on the effects of news on politicians (Helfer 2016). Kruikemeier and Lecheler (2018) analyzed how news consumers evaluate digital sources news, while Engelmann and Wendelin (2015) concentrated on recipients and looked at the influence of news values and the amount of user comments on online news selection. In the realm of studies with journalists, which is the focus of the paper at hand, Helfer and van Aelst (2015) were interested in news values and examined how they influenced how political journalists perceived the newsworthiness of party press releases. Vos (2016) used vignettes to determine whether or not journalists would mention certain members of the Belgium parliament in a news item. Glogger and Otto (2019) analyzed the journalistic understanding of hard and soft news. Lilienthal et al. (2014) applied a FS to determine which characteristics of tweets influence journalists' and bloggers' trust and intention for interaction with the tweeter.

In sum, the broad range of applications shows that FS are not limited to one field of communication studies, but seem to be suitable for various research purposes. Why FS are – from our point of view – particularly suited to conduct research on journalists is subject to the following.

Factorial Surveys for Journalism Research

As outlined above, there are historical and practical reasons why we only find few experiments on journalists to examine their attitudes, intentions, and behaviors. How can a FS help to overcome these problems?

First, FS allow for assessing *situational* as well as *individual* and *structural* influences on journalistic thinking and behavior (Reineck et al. 2017). While the systematic variation of the characteristics of the variable of interest enables researchers to elicit situation-based reaction on the vignettes, the survey part of FS can address: (1) person-level variables, such as education, (2) organizational-level variables, such as market-orientation of the medium the journalists work for, and (3) country-level variables, such as the media system in a country. This is of particular value when we think of the various sources of influences on media content (e.g., Shoemaker and Reese 1996). When considering appropriate sample sizes on the different levels (Maas and Hox 2005), FS are well suited to analyze the influences stemming from all these levels in one study *simultaneously*.

Second, we emphasized that journalistic processes are sometimes too complex to be explained in a simple experiment by varying only few variables of interest. FS address this shortcoming of traditional experiments since they vary more than one factor within one stimulus. But not only journalistic products are too complex to be explained when varying only one factor per stimulus. This also holds true for the choices and judgments journalists make. Above, we have argued that FS allow to assess influences on, for example, news selection from different hierarchical levels. FS also enable us to include various factors and to vary them within one vignette, that is, the situational influence in such studies. Consider a study about selection decisions in the context of the news value theory. Journalists will not base their selection decision on *one* news value but on the simultaneous influence of several news values. In FS, we can measure the impact of these different news values on a multidimensional choice at the same time, detangling the effects of the individual news values on the decision statistically (Nisic and Auspurg 2009).

Third, and in line with our second rationale of varying different variables within one treatment, FS are regarded to be a resource-friendly method (Auspurg and Hinz 2015a). In times of shrinking response rates (Sheehan 2001) and aiming at journalists as a population which is often described as hard to access (e.g., Engelmann 2012) as well as to be under great time pressure at the workplace (Reinardy 2011), this is an advantage of FS over traditional experiments. Since FS participants are presented with more than one experimental stimulus, rather small samples compared to multifactorial experiments are sufficient to reach a satisfactory power.[6]

Fourth, we have pointed out why external validity is a concern of communication scholars and why real-world media content, or the best approximation to it, is required to resemble real-world media consumption and production. Of course, this is to the expansion of internal validity (Hanasono 2017). Since in FS more than one factor is manipulated, the stimuli used in FS are a closer reflection of real media content which is complex and multidimensional.

Conducting a Factorial Survey Experiment

To show the application of FS, we use a dataset which has been gained with German local newspaper journalists (Glogger and Otto 2019). We will use this example to illustrate the necessary steps for an FS: (1) experimental design and vignette design, (2) vignette allocation, (3) measures, and (4) analysis.

Background Information on the Used Dataset: The Hard and Soft News Example

The concept of hard and soft news (HSN) is regarded as one of the key concepts in journalism research (Esser, Strömbäck, and de Vreese 2012). It has developed from a one-dimensional concept which only used, for example, the topic of a news item to distinguish hard from soft news (Adams 1964) to a multidimensional concept. These multidimensional concepts differentiate hard from soft news by referring – next to the topic – also to presentational dimensions (e.g., Reinemann et al. 2012). In its most recent academic conceptualization, HSN is determined by five dimensions: the *topic* of a news item, how it is *framed*, the *relevance* of the topic stressed, as well as the degree of *emotionality*, and journalistic *opinion* expressed in the news (Reinemann et al. 2012). We conducted a FS to investigate whether these theoretically assumed dimensions reflect in journalists' thinking about news (Reinemann et al. 2012). Data that were so far not analyzed enable us, furthermore, to assess individual differences in journalists' understanding of HSN, assessing whether the perceived readers' interest in hard and soft news accounted for differences in the journalistic understanding of HSN.

We used the PR-database Zimpel, which had been used in earlier surveys with German journalists (Obermaier, Koch, and Riesmeyer 2015), to invite journalists to participate in our study. After manually selecting only e-mail addresses that belong to individual journalists, we randomly selected 1,500 addresses and sent out an email with the link to the online FS. In total, 149 journalists from 73 German newspapers participated in the study. 69% of the participants were male. The journalists were on average 48 years old (*SD* = 11.00) and had worked as journalists for on average 22 years (*SD* = 11.11). The sample comprised journalists from various levels of responsibility (chief editor: 4%, middle management: 31%; editor: 54%; trainee: 6%; others: 5%).

Experimental Design and Vignette Design

The *first step* of FS is to decide on the experimental design and to design the vignettes. This includes determining the dimensions of interest, hence the characteristics of a given object or situation, and the values of these dimensions.

FS are mostly "motivated by specific social theories" (Auspurg and Hinz 2015b, 308) which should guide researchers' decision-making throughout the entire process. In our example, we are interested in the concept of HSN, following a recent conceptualization by Reinemann et al. (2012) which offers five dimensions of HSN: topic, framing, relevance, opinion, and emotionality. Others, for example, used news value theory as their theoretical foundation and derived the dimensions of interest from the large corpus of literature in this area (e.g., Helfer 2016). Next, the values of the dimensions should be determined. Taking the topic dimension as an example, one could think of an infinite amount of values of this dimension – in other words: topics. Adams (1964), for example, provides a list of various topics and categorizes dichotomously them into hard (e.g., politics) and soft news topics (e.g., human interest). For our example, we decided for dichotomous dimensions, that is, every dimension had exactly two values. Table 1 summarizes the dimensions and values that we used in the FS.

Having determined the dimensions and values of interest, the *second step* in a FS is to convert the value codes into vignettes. First, we combine the numerical codes assigned to the values of the dimensions to unique identifiers for each vignette. Combining the

Table 1. Dimensions and values of HSN.

Dimensions	Values	Operationalizations
Topic	1	Hard topic: unemployment
	2	Soft topic: neglected dogs
Relevance	1	Societal consequences emphasized
	2	Individual consequences emphasized
Framing	1	Thematic framing
	2	Episodic framing
Opinion	1	Opinion-free reporting
	2	Opinionated reporting
Emotionality	1	Unemotional picture
	2	Emotional picture

dimensions and values illustrated in Table 1, we can think of an example vignette about unemployment (dimension topic, value 1), in which the relevance of this topic for an individual is stressed (dimension relevance, value 2), which is framed thematically (dimension framing, value 1), reported opinionatedly (dimension opinion, value 2) and emotionally (dimension emotionality, value 2). Second, the values for all dimensions must be operationalized. In our example, we wrote short text snippets that represented the HSN dimensions and combined them to the full vignette. Reineck et al. (2017) emphasize that vignettes used in communication studies should be as similar as possible to media content to enhance external validity.

Figure 1 shows the example vignette in its final form. The text snippets are marked as listed in Table 1. Since the topic of a news item is also a dimension of hard and soft news, we decided to operationalize the topic not only by using different headlines in the vignettes but also by adapting the text snippets we wrote for the different values of the

Figure 1. Exemplary vignette with answer scale (both translated).

dimensions for both topics. For example, while in vignettes about the presumably harder topic "unemployment", the thematically framed text snippet used statistics about unemployment figures. On the other hand, vignettes about the softer topic of neglected dogs referred to statistics about how many pets get abandoned every summer.

Having followed this approach so far, the next step is building the *vignette universe*, i.e., the totality of all possible combinations of the values and dimensions (Auspurg and Hinz 2015a). Given our five dimensions and two values for each dimension, our universe comprised $2 \times 2 \times 2 \times 2 \times 2 = 32$ vignettes. This shows how easily the size of the vignette universe can increase, given an exponential growth of the total amount of vignettes with adding dimensions and values. For example, Reinemann et al. (2012) suggested a polytomous operationalization of the HSN dimensions with four or three values. This could result in $4 \times 3 \times 3 \times 3 \times 3 = 324$ vignettes. Participants in a FS get asked to rate vignettes and to state, for example, their attitude toward the described situation – or as in our case to give a judgment whether a vignette represents a hard or soft news item.

Vignette Allocation

Vignette universes are rarely small enough for the participants to rate each vignette, and researchers need to allocate vignettes to participants as a *third step*. There are different strategies to allocate the vignettes to the participants. We illustrate how to allocate the entire vignette universe, and – if this is not possible – how to draw different types of vignette samples for allocation. Which strategies to choose, depend on the size of the universe. The first strategy is to allocate the entire universe similar to a classical within-subjects experimental design. However, we easily see the limitations of this strategy if we look at the average reading time of one vignette in our example. It took the participating journalists on average 28.93 s ($SD = 8.36$) to read one vignette or an average of 10.43 min to complete the study. Given a universe size of 32 vignettes, rating all vignettes would have resulted in around 15 min on average to finish evaluating only the vignettes – far too long for a study with journalists who regularly suffer from work stress. Auspurg and Hinz (2015a) warn that large universes can result in "fatigue, boredom, and unwanted methodological effects, such as response heuristics" (17).

To avoid this, vignette allocation comes into play, i.e., using samples of vignettes that are allocated to participants. Samples can be random or fractionalized (Reineck et al. 2017). Drawing a *random sample* is accompanied by severe disadvantages. The confounding structure of the study set-up renders interpretation of the main effects almost impossible since the main effects could be confounded with other main or interaction effects (Atzmüller and Steiner 2010). This also holds true for the commonly applied strategy of drawing an individual random sample for each participant without replacement (Su and Steiner 2018). Therefore, so-called fractionized samples are recommended.

Compared to complete factorial designs, in which all combinations of the dimension values are used, fractional factorial designs present only a fraction of the vignette universe, e.g., a half, to the participants. However, this fraction is drawn systematically, based on the aim to lose as little information as possible and to ensure an as equally as possible distribution of the effects of interest (Steiner and Atzmüller 2006). In comparison to random samples, fractional factorial designs show the advantage of only higher-order interaction effects to be aliased with the main effects. These higher-order interaction effects are often not of interest (Atzmüller and Steiner 2010). Compared to full factorial designs, fractions

are also regarded to be more economical since fewer participants are needed and participating individuals require less time to complete the study (Gunst and Mason 2009). As the complexity of fractional factorial designs increases with the number of dimensions and values, one can seek the help of statistical programs that offer this function to build a fractional sample (e.g., Wheeler 2014) and refer to literature that discusses the various sampling techniques for block-building in greater detail (Su and Steiner 2018).

Both strategies – fractional and random sampling – also become relevant when applying the third strategy to reduce the vignettes universe: set building (Auspurg and Hinz 2015a). Sets, also called blocks or decks, comprise a fraction of vignettes. Compared to the strategies elaborated above, participants are presented with different sets. In our example, the 32 vignettes were split into eight sets, comprising four vignettes each; hence, each journalist had to evaluate one set of four vignettes. Set building is advantageous over simple random or fractional samples since more vignettes in total can be assigned to the participants, increasing the efficiency of the design (Auspurg and Hinz 2015a).[7]

If we are aware of the timely and cognitive constraints of participants in FS, one must determine how many vignettes a participant can be exposed to. These constraints are determined, on the one hand, by the amount of the vignettes one participant has to rate. Auspurg, Hinz, and Liebig (2009) warn that FS can lead to learning effects and, consequently, to consistent responses when the participants evaluate more than one vignette. On the other hand, the complexity of a vignette increases the more dimensions and values there are. Fatigue leading to inconsistent response behavior or response heuristics might be the consequence. The authors suggest to manipulate around seven (+/− two) dimensions and to apply not more than ten vignettes per participant (Atzmüller and Steiner 2010; Auspurg and Hinz 2015a).

Measures

While having determined the independent variables right at the beginning of the research endeavor with deciding which dimensions and values to include, one also has to decide on the dependent variable of the FS in a *fourth step*. Which option researchers have here, is subject to the following section.

Whether FS are used to elicit behavioral intentions or attitudinal judgments – participants' reaction toward the vignette is the core variable in each study. They represent the dependent variable and can be assessed by using several types of response scales. Jasso (2006) distinguishes between dependent variables that define a total amount ("How much would you pay for a subscription of this newspaper?"), a probability ("How likely is it that you publish a text like this vignette?"), or a set of ordered categories which we used in our study. To examine what journalists understand by hard and soft news, we asked them if the news item in the form of the displayed vignette represented hard or soft news from their point of view on a 7-point-scale.[8] Figure 1 shows an exemplary vignette and the corresponding answering scale.

A core advantage of FS is that we can implement the experimental module of the vignettes in regular (online) surveys. That means, we are able to account for two sources when trying to explain variance in the dependent variable: (1) the dimensions that we systematically varied in the vignettes, and (2) the characteristics of the participants that we

measured with a questionnaire. This decomposes the sources of explanation into a situational factor, a personal factor, and further factors (e.g., organizational). Thus, the independent variable can stem from different levels with the most basic level 1, where the characteristics of the vignettes are located, to level 2 where we find the characteristics of the participating journalists. Further levels, for example, when journalists from different media organizations (level 3) or various countries (level 4) participate in the FS, are possible (Figure 2).

The example study comprises a two-level structure. On level 1, we measured the vignettes characteristics as our first independent variable. All hard and soft news dimensions were operationalized dichotomously (Table 1). Characteristics of journalists are modeled at level 2; we assumed the perceived interest of readers in hard and soft news might influence how journalists evaluate the vignettes. We assessed the perceived influence with a single item, asking the participants if their readers preferred rather hard (1) or rather soft (7) news ($m = 3.68$; $SD = 1.07$).[9]

Analysis

Finally, we illustrate the basic steps to analyze factorial survey data. Data gained in FS is distinct since, firstly, participants evaluate more than one vignette. In our example, the journalists were asked to evaluate four vignettes in total, which means that the data for the dependent variable was not independent but clustered in participants. Judgments stemming from one journalist are more likely to be similar than judgments from different journalists.[10] When analyzing data from FS, it is particularly important to account for this nested structure (Nezlek 2011). FS data are, secondly, special since the independent variables can stem from different levels.

To account for the nested data structure, one can use ordinary least squares (OLS) regression with cluster-robust-standard errors[11] (Auspurg and Hinz 2015a) – but only when just interested in the influence of level 1 predictors on the dependent variable. Most researchers, though, might be interested in additional explanatory variables stemming from further levels, such as characteristics of the participants. *Multilevel modeling* is appropriate since this approach meets the needs of both peculiarities of FS data. It does not only consider the nested data structure, but also allows phenomena to "be
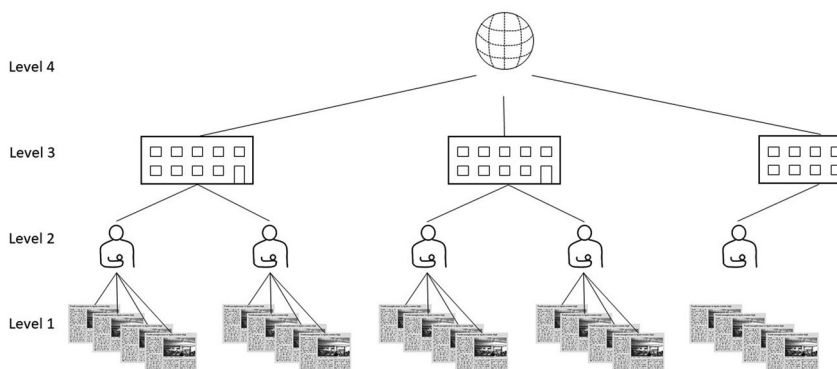

Figure 2. Hierarchical levels in a factorial survey.

examined simultaneously at different levels of analysis" (Nezlek 2011, 4) – in our case the influence of independent variables stemming from the HSN dimensions varied in the vignettes and characteristics of the journalists. Additionally, multilevel models enable researchers to analyze interactions between variables on different levels. In our example, this means that, by applying multilevel modeling, we are able to determine (1) the influence of the HSN dimensions, (2) the influence of the perceived readers' interest in hard and soft news as an individual characteristic of journalists, and (3) the interactions between the HSN dimensions and the individual characteristic on journalists' judgments of the vignettes. Table 2[12] shows the results of a step-by-step approach which adds variables from different levels to the model one after another.[13] [14]

First, the so-called null model contains no independent variables. It allows, however, to estimate the intraclass correlation (ICC) which states how much of the overall variation can be explained by the clustering (Nezlek 2011). Table 2 shows that 13% of the overall variance is accounted for the fact that journalists judged more than one vignette. Some researchers recommend using the ICC as an indicator for whether or not multilevel modeling is required in the first place (Bickel 2007), nevertheless, they warn that "even a very weak intraclass correlation can substantially deflate standard errors of regression coefficients" (Bickel 2007, 9).

Second, we determine the influence of the HSN dimensions on journalists' judgments on the vignettes. In other words, we are interested in the situational component of these

Table 2. Results of multilevel modeling.

	Model 0		Model 1		Model 2		Model 3	
Fixed effects								
Constant	3.51***	(0.09)	1.93***	(0.13)	1.93***	(0.13)	1.92***	(0.12)
HSN Dimensions (level 1)[a]								
Topic (unemployment)			1.15***	(0.12)	1.15***	(0.12)	1.14***	(0.12)
Relevance (societal)			0.21*	(0.09)	0.21*	(0.09)	0.21*	(0.09)
Framing (thematic)			1.43***	(0.10)	1.43***	(0.10)	1.43***	(0.10)
Opinion (opinion-free)			0.36**	(0.11)	0.36**	(0.11)	0.36**	(0.11)
Emotionality (neutral)			0.02	(0.11)	0.02	(0.11)	0.03	(0.11)
Journalistic characteristic (level 2)								
Perceived HSN interest					0.18*	(0.07)	0.20	(0.11)
Cross-level interactions								
HSN interest x topic							−0.35**	(0.13)
HSN interest x relevance							0.06	(0.09)
HSN interest x framing							−0.03	(0.09)
HSN interest x opinion							0.24*	(0.11)
HSN interest x emotionality							0.04	(0.12)
Random effects								
Residual variance								
HSN dimension (level 1)	2.73		1.56		1.56		1.49	
Journalist (level 2)	0.39		0.68		0.64		0.66	
HSN interest (freed slope)					0.004		0.004	
ICC level 1	0.13		0.30		0.29		0.31	

Notes: $n_{observation} = 596$, $n_{groups} = 149$; robust standard errors in parentheses. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.
Answer to the questions: "Please indicate whether the news item was rather hard news (1) or soft news (7) from your point of view" (7-point Likert-type scale).
HSN = hard and soft news; ICC = intraclass correlation;
[a]Reference categories in parentheses.

judgments based on the HSN characteristics varied in the vignettes. Model 1 in Table 2 reveals that four out of five HSN dimensions had a significant influence on whether the participating journalists regarded the presented vignettes as hard or soft. For example, vignettes about the presumably softer topic of abandoned pets were more likely to be evaluated as soft news than items about the harder topic of unemployment ($b = 1.15$, $p < .001$). The emotionality dimension, in contrast, did not influence significantly how the journalists judged the vignettes.

Third, next to the situational aspect, FS also allow considering higher level influences. Based on the assumption that hard and soft news are regarded as one way to attract the audience (Patterson 2000), we were interested in whether or not the perceived audience interest in hard and soft news influenced journalists' evaluations of the vignettes. This perceived interest represents a level-2-variable which is located at the individual level. In Model 2 (Table 2), we see that the perceived interest of the audience indeed influenced how journalists judged the vignettes. The more interested in soft news the journalists perceived their readers to be, the softer they rated the vignettes ($b = 0.18$, $p = .01$). This result, however, is of little empirical value since it only gives us a general idea about how the individual journalistic characteristic of perceived audience interest in HSN influenced the judgment of the vignettes. It does not, though, incorporate the situational component of the vignette characteristics. So-called cross-level-interactions allow – as a fourth step – to make assumptions about the interplay of participant and vignette characteristics.

We, therefore, included interaction terms of the perceived audience interest and the HSN dimensions of topic, relevance, framing, opinion, and emotionality to the model (Model 3, Table 2). Two cross-level-interactions yielded a significant result. Journalists who perceived their readers as preferring hard news were more likely to evaluate news items about unemployment as hard news than journalists writing for an audience with soft preferences ($b = -0.35$; $p = .006$). In contrast, we found a positive interaction effect between the opinion dimension and the perceived interest of the readers ($b = 0.24$; $p = .009$).

Discussion

Investigating journalists' attitudes, values, or behavior lies at the heart of journalism studies. This field is, to date, dominated by correlational designs (Weaver 2008). In this paper, we have discussed potential reasons why journalism scholars may have refrained from experiments with journalists and suggested FS – a research design that combines the advantages of experiments and surveys – as a solution to these problems. We believe that FS will – once added to the methodological toolbox of journalism scholars – advance our understanding of journalistic thinking and behavior.

We illustrated the advantages of the method by using data from a FS with German newspaper journalists (Glogger and Otto 2019). However, assessing journalists' understanding of concepts is only one possible field of application. We see further directions of applying FS in studies with and on journalists. First, the question of how and why journalists make certain selection decisions is a core question in journalism research (Hanitzsch and Wahl-Jorgensen 2009). FS can extend our understanding of these decision-making processes by using stimuli in which we manipulate various characteristics simultaneously.

Thereby, researchers can get closer to the multi-dimensional decisions that journalists face in their job (e.g., Helfer and van Aelst 2015). Since journalists do not only decide which news to publish or which picture to select in photojournalism, but also *how* to present news, we regard FS also suited for studies on the latter decisions. Fields of the application might be journalistic framing, journalistic objectivity, or sensationalism.

Second, we see a possible application of FS in studies that examine journalists' attitudes. Take for example ethical considerations on controversial journalistic research methods (Reineck et al. 2017; Sanders 2003). Under which conditions do journalists, for example, justify paying for information? While traditional surveys can face the problem of social desirability when assessing such moral issues, FS are less prone to this phenomenon (Auspurg and Hinz 2015b). To analyze journalistic attitudes like moral judgments, researchers can also create vignettes in which, for example, a journalist is portrayed in a situation in which he or she makes a questionable decision like bribing officials for information. FS as "dry runs" (Reineck et al. 2017, 103) allow participating journalists to judge fictive colleagues' behavior without being in the situation themselves.

Finally, comparative studies in journalism research are a field in which scholars can apply FS, following calls for more experimental designs (Hanitzsch 2009). As discussed earlier, FS enable us to assess influences on various levels simultaneously. In our example, we limited the influencing factors to two levels, only assessing situational and individual levels. However, given appropriate sampling strategies to gain enough participants on an organizational or even country level, FS are suitable to compare journalistic behaviors and attitudes across media organizations, types of media, and media systems. Such a detailed comparative judgment is not possible within a traditional survey as both the situational and personal influences are affected – and thus confounded – by the country context. Only cross-cultural experimental designs can overcome this shortcoming of a correlational design.

Despite these advantages, we cannot neglect that FS provide challenges to (communication) scholars and that they are not suited to answer every research question. Firstly, since vignettes should resemble real-world media content which is evaluated by the producers of media content – the journalists, the construction of the vignettes should be given thorough attention to prevent implausible value combinations (Auspurg, Hinz, and Liebig 2009). Pretesting of the vignettes with a journalistic sample is, therefore, advised. Secondly, FS, of course, share disadvantages with other quantitative approaches. Whenever scholars aim to gain in-depth insights, qualitative methods might be advantageous. Especially an academic discipline which studies a subject which is in constant change (e.g., Weaver and Willnat 2016), FS with their high degree of standardization do not give us the possibility to, for example, explain new journalistic phenomena in an exploratory way (Iorio 2014). Thirdly, traditional large-scale surveys might be advised when scholars are interested in assessing more than one journalistic attitude or behavior. Take as example studies that aim to describe the "profile of global journalists" (Weaver and Willnat 2012, 529) and do not aim at explaining journalistic behavior or decision making. Similarly, traditional experiments should be preferred over FS when one is interested in the effect of only very few independent variables since the construction of the vignettes comes with the challenges described above.

However, given the advantages of FS for studying journalistic behaviors and attitudes, we believe that FS can be an enrichment to journalism scholars' methodological toolbox.

When accounting for the – admittedly – challenges when planning, conducting, and analyzing such a study in an appropriate way, FS will be more than a "flash in the pan" (Reineck et al. 2017, 114).

Notes

1. It is beyond the scope of this paper to give a review of research designs used in journalism studies – we rather would like to show that experimental designs are not as prevalent as in neighboring sub-disciplines. Following this argument, we aim to introduce factorial surveys as a further possibility to consider when researching journalists.
2. A Boolean search in abstracts of the database "Communication & Mass Media Complete" on January 25, 2019, yielded 148 hits for "journalism" AND ("experiment" OR "experimental study"), 438 for "journalism" AND ("content analysis" OR "content analyses"), 901 for "journalism" AND "survey", and 1072 for "journalism" AND "interview".
3. When entering similar search strings for political communication, the method survey was most prevalent (more than 1500 hits) – experiment, interview and, content analysis where almost equal with 500–600 hits (abstract search).
4. We are, of course, aware of the fact that journalism research is not only concerned with studying journalists. However, investigating the work, behavior, attitudes, and decisions of journalists is at the core of the discipline and could benefit a lot from using FS designs, as we show in this paper.
5. One has to be aware, though, that in the literature the terms are sometimes used interchangeably. To prevent confusion, we follow the terminology and threefold differentiation by Auspurg and Hinz (2015b). Introducing conjoint analysis and multifactorial experiments would go beyond the scope of this paper as they differ largely in research tradition, data structure, and, most important, data analysis. For introductions to (conjoint) choice experiments, see, e.g., Knudsen and Johannesson (2018), to conjoint analyses, e.g., Green, Krieger, and Wind (2001).
6. Of course, low response rates in factorial surveys are still a problem concerning the nonresponse bias (e.g., Fowler 2013).
7. However, using only some vignette comes with statistical disadvantages. While in a complete vignette universe the "all of the dimensions and interactions between the dimensions are uncorrelated with each other" (Auspurg and Hinz 2015a, 16), fractions lack this desirable feature. Furthermore, some dimensions might be oversampled in a fraction. Consequently, the statistical strength of the FS is reduced.
8. The English terms of hard and soft news are familiar to German journalists since they are used in text-books in journalism school (e.g., Hooffacker and Meier 2017). Nevertheless, we made sure that our participants were acquainted with the terms by asking them before participating in the study if they knew the terms. Only n=3 neglected this and were excluded from further participation.
9. We group-mean-centered the independent level-2 variable. It is beyond the scope of the paper to discuss the advantages and disadvantages of the group- and grand-mean centering in multilevel modeling (e.g., Enders and Tofighi 2007).
10. As described above, journalists can also be clustered in news organizations or even countries. We refrained from clustering journalists also in newspapers since the n of these clusters would have been too small for an analysis (*min*=1, *max*=14; *m*=2) (Hox 2010).
11. We focus on statistical methods suited for a metric answer scale.
12. We report random intercept models with fixed effects for the predictors (level 1) and random effects for the residual variance of the levels (Levels 1 and 2). It is beyond our scope to discuss the possibilities in multilevel-modeling, but we refer the readers to appropriate literature, e.g., Bickel (2007); Nezlek (2011). Because the residuals of the full model were not normally distributed, we also decided for Huber/White/sandwich estimators using the xtmixed, vce(robust) command in Stata 14.0 Hoechle (2007)

13. It is beyond the scope of this paper to give a detailed description of how to conduct multilevel analyses. Readers new to this approach may want to follow the rule of thumb though that multi-level modeling "[is] just regression" (Bickel 2007, 1) when interpreting the results (Table 2) and to pay less attention to random effects which account for clustering of data. For a comprehensive description of multilevel analyses, see, for example, Raudenbush and Bryk (2002).
14. Within the sets, the vignettes were presented in random order. We, furthermore, checked for potential influences of the sets, i.e., the so-called deck or set effect (Auspurg and Hinz 2015a), by including the set numbers as independent variables to the model. Since neither the set numbers yielded significant results nor did we see a notable change of the results, we report the model without the set number.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

ORCID

Lukas P. Otto 🔟 http://orcid.org/0000-0002-4374-6924

References

Adams, J. B. 1964. "A Qualitative Analysis of Domestic and Foreign News on the AP TA Wire." *Gazette* 10 (4): 285–294.

Alexander, C. S., and H. J. Becker. 1978. "The Use of Vignettes in Survey Research." *Public Opinion Quarterly* 42 (1): 93–104.

Atzmüller, C., and P. Steiner. 2010. "Experimental Vignette Studies in Survey Research." *Methodology* 6 (3): 128–138. doi:10.1027/1614-2241/a000014.

Auspurg, K., and T. Hinz. 2015a. *Factorial Survey Experiments*. Thousand Oaks, CA: Sage.

Auspurg, K., and T. Hinz. 2015b. "Multifactorial Experiments in Surveys: Conjoint Analysis, Choice Experiments, and Factorial Surveys." In *Experimente in den Sozialwissenschaften*, edited by M. Keuschnigg and T. Wolbring, 291–315. Baden-Baden: Nomos.

Auspurg, K., T. Hinz, and S. Liebig. 2009. "Komplexität von Vignetten, Lerneffekte und Plausibilität im Faktoriellen Survey [Complexity of Vignettes, Learning and Plausibility]." *Methoden, Daten, Analysen* 3 (1): 59–96.

Bickel, R. 2007. *Multilevel Analysis for Applied Research: It's Just Regression!*. New York: Guilford Press.

Boumans, J. W., and D. Trilling. 2016. "Taking Stock of the Toolkit." *Digital Journalism* 4 (1): 8–23. doi:10.1080/21670811.2015.1096598.

Cushion, S. 2008. "Truly International? A Content Analysis of Journalism: Theory, Practice and Criticism and Journalism Studies." *Journalism Practice* 2 (2): 280–293.

Dupagne, M., and B. Garrison. 2006. "The Meaning and Influence of Convergence." *Journalism Studies* 7 (2): 237–255. doi:10.1080/14616700500533569.

Enders, C. K., and D. Tofighi. 2007. "Centering Predictor Variables in Cross-Sectional Multilevel Models: A New Look at an Old Issue." *Psychological Methods* 12 (2): 121.

Engelmann, I. 2012. *Alltagsrationalität im Journalismus: akteurs- und organisationsbezogene Einflussfaktoren der Nachrichtenauswahl [Day-to-day Rationality in Journalism: Person- and Organization-Related Factors Influencing the Selection of News]*. Konstanz: UVK.

Engelmann, I., and M. Wendelin. 2015. "Relevanzzuschreibung und Nachrichtenauswahl des Publikums im Internet [Audience's Relevance Attribution and News Selection in the Internet]." *Publizistik* 60 (2): 165–185.

Esser, F., J. Strömbäck, and C. de Vreese. 2012. "Reviewing key Concepts in Research on Political News Journalism: Conceptualizations, Operationalizations, and Propositions for Future Research." *Journalism* 13 (2): 139–143. doi:10.1177/1464884911427795.

Fowler, F. J. 2013. *Survey Research Methods*. Los Angeles, CA: Sage.

Glogger, I., and L. P. Otto. 2019. "Journalistic Views on Hard and Soft News: Cross-Validating a Popular Concept in a Factorial Survey." *Journalism & Mass Communication Quarterly. Advance Online Publication*, doi:10.1177/1077699018815890.

Graves, L., B. Nyhan, and J. Reifler. 2016. "Understanding Innovations in Journalistic Practice: A Field Experiment Examining Motivations for Fact-Checking." *Journal of Communication* 66 (1): 102–138.

Green, P. E., A. M. Krieger, and Y. Wind. 2001. "Thirty Years of Conjoint Analysis: Reflections and Prospects." *Interfaces* 31: 56–73.

Gunst, R. F., and R. L. Mason. 2009. "Fractional Factorial Design." *Wiley Interdisciplinary Reviews: Computational Statistics* 1 (2): 234–244.

Hanasono, L. 2017. "External Validity." In *The SAGE Encyclopedia of Communication Research Methods*, edited by M. Allen, 480–483. Thousand Oaks, CA: Sage.

Hanitzsch, T. 2009. "Comparative Journalism Studies." In *The Handbook of Journalism Studies*, edited by K. Wahl-Jorgensen and T. Hanitzsch, 413–427. London: Routledge.

Hanitzsch, T., and K. Wahl-Jorgensen. 2009. "Introduction: On why and how we Should do Journalism Studies." In *The Handbook of Journalism Studies*, edited by K. Wahl-Jorgensen and T. Hanitzsch, 3–16. London: Routledge.

Hanusch, F., and T. Hanitzsch. 2017. "Comparing Journalistic Cultures Across Nations." *Journalism Studies* 18 (5): 525–535. doi:10.1080/1461670X.2017.1280229.

Hanusch, F., and T. P. Vos. 2019. "Charting the Development of a Field: A Systematic Review of Comparative Studies of Journalism." *International Communication Gazette*, 1748048518822606.

Harcup, T., and D. O'Neill. 2001. "What is News? Galtung and Ruge Revisited." *Journalism Studies* 2 (2): 261–280.

Haven, T. L., and L. van Grootel. 2019. "Preregistering Qualitative Research." *Accountability in Research*, 1–16. doi:10.1080/08989621.2019.1580147.

Helfer, L. 2016. "Media Effects on Politicians: An Individual-Level Political Agenda-Setting Experiment." *The International Journal of Press/Politics* 21 (2): 233–252.

Helfer, L., and P. van Aelst. 2015. "What Makes Party Messages Fit for Reporting?: An Experimental Study of Journalistic News Selection." *Political Communication* 33 (1): 59–77. doi:10.1080/10584609.2014.969464.

Henn, P., M. Dohle, and G. Vowe. 2013. "Politische Kommunikation": Kern und Rand des Begriffsverständnisses in der Fachgemeinschaft ["Political Communication". Core and Periphery of the Community's Understanding of the Concept]." *Publizistik* 58 (4): 367–387.

Hoechle, D. 2007. "Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence." *Stata Journal* 7 (3): 281–312.

Hooffacker, G., and K. Meier. 2017. *La Roches Einführung in den praktischen Journalismus [La Roche's Introduction to Practical Journalism]*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Hox, J. J. 2010. *Multilevel Analysis: Techniques and Applications (2nd ed.). Quantitative Methodology Series*. New York: Routledge.

Iorio, S. H. 2014. *Qualitative Research in Journalism: Taking it to the Streets*. Mahwah, NJ: Routledge.

Jasso, G. 2006. "Factorial Survey Methods for Studying Beliefs and Judgments." *Sociological Methods & Research* 34 (3): 334–423. doi:10.1177/0049124105283121.

Kamhawi, R., and D. Weaver. 2003. "Mass Communication Research Trends From 1980 to 1999." *Journalism & Mass Communication Quarterly* 80 (1): 7–27.

Klimmt, C., and R. Weber. 2013. "Das Experiment in der Kommunikationswissenschaft [The Experiment in Communication Studies]." In *Handbuch standardisierte Erhebungsverfahren in der Kommunikationswissenschaft*, edited by W. Möhring and D. Schlütz, 125–144. Wiesbaden: Springer.

Knudsen, E., and M. P. Johannesson. 2018. "Beyond the Limits of Survey Experiments: How Conjoint Designs Advance Causal Inference in Political Communication Research." *Political Communication*, 1–13. doi:10.1080/10584609.2018.1493009.

Kruikemeier, S., and S. Lecheler. 2018. "News Consumer Perceptions of New Journalistic Sourcing Techniques." *Journalism Studies* 19 (5): 632–649. doi:10.1080/1461670X.2016.1192956.

Leshner, G. 2012. "The Basics of Experimental Research in Media Studies." In *The International Encyclopedia of Media Studies*, edited by A. N. Valdivia, Vol. 40, 236–254. Oxford: Wiley. doi:10.1002/9781444361506.wbiems181.

Lilienthal, V., S. Weichert, D. Reineck, A. Sehl, and S. Worm. 2014. *Digitaler Journalismus [Digital Journalism]*. Berlin: Vistas.

Maas, C. J. M., and J. J. Hox. 2005. "Sufficient Sample Sizes for Multilevel Modeling." *Methodology* 1 (3): 86–92.

Matthes, J., F. Marquart, B. Naderer, F. Arendt, D. Schmuck, and K. Adam. 2015. "Questionable Research Practices in Experimental Communication Research: A Systematic Analysis from 1980 to 2013." *Communication Methods and Measures* 9 (4): 193–207. doi:10.1080/19312458.2015.1096334.

McGregor, S. C., and L. Molyneux. 2018. "Twitter's Influence on News Judgment: An Experiment among Journalists." *Journalism*, 1464884918802975.

Mothes, C. 2016. "Biased Objectivity: An Experiment on Information Preferences of Journalists and Citizens." *Journalism & Mass Communication Quarterly. Advance Online Publication*, doi:10.1177/1077699016669106.

Nezlek, J. B. 2011. *Multilevel Modeling for Social and Personality Psychology*. London: Sage.

Nisic, N., and K. Auspurg. 2009. "Faktorieller Survey und klassische Bevölkerungsumfrage im Vergleich [Comparing Factorial Surveys and Traditional Surveys]." In *Klein Aber Fein!: Quantitative Empirische Sozialforschung mit kleinen Fallzahlen*, edited by P. Kriwy and C. Gross, 211–245. Wiesbaden: VS Verlag für Sozialwissenschaften.

Obermaier, M., T. Koch, and C. Riesmeyer. 2015. "Deep Impact?: How Journalists Perceive the Influence of Public Relations on Their News Coverage and Which Variables Determine This Impact." *Communication Research* 45 (7): 1031–1053. doi:10.1177/0093650215617505.

Patterson, T. E. 2000. *Doing Well and Doing Good: How Soft News and Critical Journalism are Shrinking the News Audience and Weakening Democracy-and What News Outlets can do About it*. Cambridge: Harvard University Press.

Patterson, T. E., and W. Donsbach. 1996. "News Decisions: Journalists as Partisan Actors." *Political Communication* 13 (4): 455–468.

Raudenbush, S. W., and A. S. Bryk. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, CA: Sage.

Reinardy, S. 2011. "Newspaper Journalism in Crisis: Burnout on the Rise, Eroding Young Journalists' Career Commitment." *Journalism* 12 (1): 33–50. doi:10.1177/1464884910385188.

Reineck, D., V. Lilienthal, A. Sehl, and S. Weichert. 2017. "Das faktorielle Survey. Methodische Grundsätze, Anwendungen und Perspektiven einer innovativen Methode für die Kommunikationswissenschaft [Factorial Survey. Methodology, Usage, and Perspective of an Innovative Method in Communication Science]." *M&K Medien & Kommunikationswissenschaft* 65 (1): 101–116.

Reinemann, C., J. Stanyer, S. Scherr, and G. Legnante. 2012. "Hard and Soft News: A Review of Concepts, Operationalizations and Key Findings." *Journalism* 13 (2): 221–239.

Rossi, P. H. 1979. "Vignette Analysis: Uncovering the Normative Structure of Complex Judgments." In *Qualitative and Quantitative Social Research: Papers in Honor of Paul F. Lazarsfeld*, edited by R. K. Merton, J. S. Coleman, and P. H. Rossi, 176–186. New York: Free Press.

Sanders, K. 2003. *Ethics & Journalism*. Thousand Oaks, CA: Sage.

Sheehan, K. B. 2001. "E-mail Survey Response Rates: A Review." *Journal of Computer-Mediated Communication* 6 (2), JCMC621.

Shoemaker, P. J., and S. D. Reese. 1996. *Mediating the Message*. White Plains, NY: Longman.

Steiner, P., and C. Atzmüller. 2006. "Experimentelle Vignettendesigns in Faktoriellen Surveys [Experimental Vignette Designs in Factorial Surveys]." *Kölner Zeitschrift Für Soziologie Und Sozialpsychologie* 58 (1): 117–146. doi:10.1007/s11575-006-0006-9.

Su, D., and P. Steiner. 2018. "An Evaluation of Experimental Designs for Constructing Vignette Sets in Factorial Surveys." *Sociological Methods & Research* 1, doi:10.1177/0049124117746427.

Thorson, E., R. Wicks, and G. Leshner. 2012. "Experimental Methodology in Journalism and Mass Communication Research." *Journalism & Mass Communication Quarterly* 89 (1): 112–124. doi:10.1177/1077699011430066.

Vos, D. 2016. "How Ordinary MPs Can Make it Into the News: A Factorial Survey Experiment with Political Journalists to Explain the Newsworthiness of MPs." *Mass Communication and Society* 19 (6): 738–757.

Wahl-Jorgensen, K., and T. Hanitzsch. 2009. *The Handbook of Journalism Studies*. London: Routledge.

Wallander, L. 2009. "25 Years of Factorial Surveys in Sociology: A Review." *Social Science Research* 38 (3): 505–520. doi:10.1016/j.ssresearch.2009.03.004.

Weaver, D. H. 2008. "Methods in Journalism Research – Survey." In *Global Journalism Research: Theories, Methods, Findings, Future*, edited by M. Löffelholz, D. H. Weaver, and A. Schwarz, 106–116. Malden: Wiley-Blackwell.

Weaver, D. H., and L. Willnat. 2012. "Journalists in the 21st Century." In *The Global Journalist in the 21st Century*, edited by D. H. Weaver and L. Willnat, 529–551. New York: Routledge.

Weaver, D. H., and L. Willnat. 2016. "Changes in US Journalism: How Do Journalists Think About Social Media?" *Journalism Practice* 10 (7): 844–855.

Wheeler, B. 2014. "AlgDesign: Algorithmic Experimental Design." https://cran.r-project.org/package=AlgDesign.