methods, data, analyses | Vol. 16(2), 2022, pp. 133-140

Editorial: Vignette Analysis: Methodology and Recent Developments

Lena M. Verneuer-Emre[1], *Stefanie Eifler*[2] & *Hermann Dülmer*[3]

[1] RWTH Aachen University

[2] Catholic University of Eichstätt-Ingolstadt

[3] University of Cologne and University of Passau

The term 'vignette analysis' draws on various disciplinary traditions to refer to various techniques for measuring normative judgements, subjective beliefs, and behavioural intentions on the basis of respondents' answers to (a number of) brief descriptions of hypothetical situations, persons, or objects. The use of vignettes in survey research has been suggested within the framework of the 'indirect measurement movement' in empirical social research (Campbell, 1950) with the intention of bringing social context information into measurement. Numerous methodological studies have been undertaken with the aim of scrutinizing the assumed advantages of using vignettes.

The idea for this special issue was born in 2019, when we had the pleasure of hosting a session at the Conference of the *European Survey Research Association* (ESRA) in Zagreb, that brought together researchers with a special interest in research on vignette analyses. It was here that we again noticed the diversity of findings and the different ways of using vignette analyses, ranging from genuine methodological contributions through to applications of vignettes in the context of substantive research. A similar picture now emerges in this special issue: The contributions present methodological research on vignette analyses and innovative applications of this method, mostly located within the framework of experimental designs like factorial survey experiments, but also in the context of more general applications of vignettes such as anchoring vignettes or conjoint analyses.

The diversity of research findings on vignette analyses is our starting point in this editorial. The overall structure of this special issue of *mda* is as follows: The first chapter starts with a detailed literature review of factorial survey experiments to provide an overview of developments and trends in recent decades. In the same context, the second chapter provides an illustrative example for an application of factorial survey experiments. Subsequently, this special issue discusses two crucial

aspects relating to the application of vignettes – presentation and design resolution: Chapters three to five are dedicated to presentation format of vignettes in the context of factorial survey experiments, conjoint analyses, and anchoring vignettes. Chapter six focusses on design resolutions and the computer-based determination of the resolution IV design in *SAS On Demand for Academics*. The final chapter takes up the rarely used estimation technique of seemingly unrelated models in the context of factorial survey experiments. The papers collected and structured in this way in this special issue are framed within current research topics and findings more precisely below.

We start our framing of the collected papers with what is a truly outstanding contribution to the 'indirect measurement movement' already referred to, i.e., the factorial survey approach, which was introduced by Rossi (1979) as proposed by Paul F. Lazarsfeld (c.f., among others, Wallander, 2009). By transferring the basic principles of the *factorial* design (*multivariate* experimental design) into a sample *survey* (cf. Rossi & Anderson, 1982; Dülmer, 2007), the factorial survey combines both the high internal validity of causal inferences from experimental designs with the principally high external validity of causal inferences from survey research (Sniderman & Grob, 1996; Mutz, 2011; Auspurg & Hinz, 2015), regarding the generalizability of results to the broader population (cf. Sniderman & Grob, 1996; Auspurg & Hinz, 2015). Factorial surveys employ an experimental design that permits general conclusions to be drawn about causal mechanisms even without a random sample of respondents (cf. Auspurg & Hinz, 2015).[1]

The factorial survey approach has been applied widely throughout the social sciences in recent decades. Studies have been undertaken on topics such as choosing the appropriate experimental design in factorial surveys (Atzmüller & Steiner, 2010; Dülmer, 2007; 2016), the effects of order, variation, wording, and presentation mode (Auspurg & Jäckle, 2017; Eifler & Petzold, 2014; Sauer et al., 2020; Shamon et al., 2022), choosing the most appropriate answer scale (Auspurg & Hinz, 2015; Sauer et al., 2020), learning and fatigue effects (Auspurg & Jäckle, 2017; Shamon et. al., 2022), and the susceptibility of vignettes to social desirability response bias (Eifler, 2007; 2010; Eifler & Petzold, 2019; Groß & Börensen, 2009; Petzold & Eifler, 2020; Petzold & Wolbring, 2019). So far, the results of these studies are multifaceted and partly inconclusive, thereby giving rise to further questions.

1 While there are several approaches to the analysis of causal relationships with different research designs, many social scientists consider particularly the group of *experimental designs* as the *silver bullet* to the analysis of causal relationships (Shadish et al., 2002). The reason for this is that, in an experiment, social scientists "manipulate the presumed cause and observe the outcome afterward" (Shadish et al., 2002: 6) instead of considering social phenomena as they naturally occur in order to study causal relationships.

Edgar Treischl and *Tobias Wolbring* draw on the work of Lisa Wallander (2009) to open the special issue with a detailed literature review of factorial survey experiments published between 1982 and 2018. Besides looking at the development of research focussing on factorial survey experiments, the authors also focus on methodological advances as well as open questions in this research field. Their review shows that more and more research has been undertaken in this field over a period of several years, both with regard to attitude research and to issues relating to behavioural research topics. At the same time, the authors identify unresolved methodological challenges concerning the validity of vignettes and related realism issues.

As Treischl and Wolbring as well as the growing research on factorial surveys show, there are no substantive limits to the use of factorial designs. It is – as it is for all empirical analysis – mainly a question of the specific objective of the research and the appropriate implementation that may lead to the application of factorial designs. Generally speaking, the common denominator of factorial surveys is that they all aim to identify the relevant factors for judgements or behavioral intentions while studying social phenomena.

Clemens Maria Schmidt draws on Lucien Karpik's 'Economics of Singularities' to analyze the choice of movies using the Factorial Survey Approach. Due to the subjectivity of such a choice, the uncertainty of judgements is in Schmidt's view best anticipated by applying a factorial survey experiment in a student sample. As well as arriving at the interesting finding that diverse social devices are used to choose a film, Schmidt discusses the advantages of the factorial survey method and in particular how it supports analysis of the causal influence of those devices in situations where a choice has to be made.

Next, we consider the decisions researchers have to make when planning the application of vignettes in a survey: Besides the transformation of theoretical assumptions into situational descriptions, dimensions and levels to be depicted, challenges also arise with approximation to realism and the adequacy of the presented situation when applying vignettes in surveys. One crucial decision when setting up a vignette design concerns *the way vignettes are presented* to respondents. Vignettes were initially and, in most cases, continue to be presented as detailed written situational descriptions or in the form of short statements (e.g., Armacost et al., 1991; Triandis et al., 1998; Wallander, 2009). For some time now, studies have also used photos or videos (e.g., Golden III et al., 2001; Eifler, 2007; Noel et al., 2008; Krysan et al., 2009) as the presentation mode. First attempts have even been made to use virtual reality to present scenarios to respondents and in this way to focus on realism issues using immersive techniques (e.g., van Gelder et al., 2019). Whereas most applications make use of either written or visual vignettes, little research has so far been undertaken on the systematic comparison of different formats and their

effects. The findings that do exist are rarely clear-cut (Rashotte, 2003; Eifler, 2007; van Gelder et al., 2019). We are all the more pleased therefore to have three contributions that focus on presentation format with reference to very recent research contributions – in the context of factorial survey experiments, conjoint analyses, and anchoring vignettes:

Drawing on the theoretical perspectives of broken windows theory and the topic of fear of crime, *Stefanie Eifler* and *Knut Petzold* apply a split ballot experiment to compare different presentation formats of vignettes (written and photo) in a factorial survey. The authors investigate whether the context presented in a photo vignette leads to higher context approximation and thus to more valid answers than when using (classic) written vignettes. Overall, it is shown that the presentation format makes no difference to the assumed level of fear of crime of the vignette-dimensions. The presentation format was only observed to have an effect for setting characteristics (e.g., darkness) in the photo vignette.

Experiments that are closely related to vignette analysis are conjoint analysis (Luce & Turkey, 1964) and choice experiments (McFadden, 1974; cf. also Auspurg & Hinz, 2015). While the term "vignette analysis" prevails in social sciences, the term "conjoint analysis" traditionally dominates in marketing research where researchers are usually interested in the preference order for certain products. However, the basic structure of the experimental design for conjoint analysis and vignette analysis is the same, except that traditional conjoint analysis does not use confounded designs and all factors have to influence the judgement behaviour independently of each other (additive model without interaction terms, cf., Louviere, 1994). *Sophie Cassel*, *Josefine Magnusson* and *Sebastian Lundmark* focus on the presentation format in the context of such conjoint designs. The authors replicate the work of Shamon, Dülmer, and Giza (2019) and extend it to a paired conjoint experiment. Following a direct replication and analysis of the results of the extension, the authors confirm the conclusion that the table format is to be preferred to the text format in conjoint experimental designs.

Mengyao Hu, *Sunghee Lee*, *Hongwei Xu*, *Roberto Melipillán*, *Jacqui Smith*, and *Arie Kapteyn* contribute to the application of anchoring vignettes in health surveys with a special focus on the presentation format of these vignettes. In general, the challenge of inconsistent survey responses may arise due to diverse understandings of the subject in question – a problem that cannot be accounted for after data collection. The application of anchoring vignettes as an additional measurement tool in the process of data collection is one way of accounting for this difficulty: With the help of anchoring vignettes, the proportion of incomparability can be extracted in the process of analyzing the gathered data (cf. King et al., 2004; King & Wand, 2007; Hopkins & King, 2010; van Soest et al., 2011).

Hu et al. propose the use of image anchoring vignettes to overcome problems of complexity and time. By using data from a cross-cultural experiment and

comparing text and image vignettes, the authors conclude that image vignettes can improve respondents' differentiation of intensity levels, response consistency as well as the survey time in general.

Finally, D-efficiency in combination with design resolution and set size is also discussed. The higher the *D-efficiency* of a quota design, the lower the correlations between different vignette dimensions and the more balanced are the levels of each vignette dimension (Kuhfeld, 1997, cf. also Dülmer, 2016). The same applies to interaction terms, provided that they were included when a D-efficient design was generated. A design's *resolution* provides information about the aliasing (confounding) structure within a vignette set and/or about the confounding structure across the different quota sets selected by the researcher: the higher a design's resolution, the more main effects and higher order interaction effects, that are perfectly uncorrelated with other (higher order) interaction effects, can be estimated (McLean & Anderson, 1984; Ryan, 2007; Kuhfeld, 2010; cf. also Dülmer, 2016). Hence, higher design resolutions ensure a better protection from possible biases in the estimated effects than lower design resolutions. The disadvantage of a higher design resolution, however, is usually seen in the higher set sizes that are required for such designs.

Julia Kleinewiese contributes to the crucial topic of design resolution by focusing on quota designs, more precisely on D-efficient designs, and looks closely at the two-way interactions in resolution IV designs as well as the (minimum) number of vignettes (set size) for reaching a D-efficiency above 90 and as closely as possible to 100 (uncorrelated, balanced designs). Driven by the aim of an application-oriented paper, the author compares the aliasing structure of resolution IV designs as defined in the literature with the structure created by *SAS On Demand for Academics*. As well as discussing and reflecting on her finding of a discrepancy between the two, Kleinewiese also draws conclusions for the application of D-efficient designs and suggests, if possible, using resolution V designs as a standard design resolution in the social sciences.

S*trategies of data analyses* are of special interest for researchers who apply factorial surveys. By presenting several situational descriptions with varying dimensions to respondents, the data requires special treatment due to its hierarchical structure. Multilevel modelling is therefore the recommended choice for analyzing data with several ratings per respondent produced by factorial designs (Snijders & Boskers, 2012; Dülmer, 2016). A special case arises for factorial designs that are designed to measure not only different ratings per respondents but that also present several rating options for each vignette and thus produce multiple ratings per vignette.

Alexander Schmidt-Catran draws on this type of data structure and proposes an approach to statistically account for multiple ratings per vignettes with a Seem-

ingly Unrelated Regression framework. This approach – located within Structural Equation Modelling techniques – enables coefficients to be compared across ratings as well as the factor structure underlying such ratings to be analyzed. The author aims to make his proposal accessible to researchers by providing two application examples and the syntax in an online appendix.

References

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

Armacost, R. L., Hosseini, J. C., Morris, S. A., & Rehbein, K. A. (1991). An Empirical Comparison of Direct Questioning, Scenario, and Randomized Response Methods for Obtaining Sensitive Business Information. *Decision Sciences, 22*(5), 1073-1090.

Atzmüller, C., & Steiner, P. M. (2010). Experimental Vignette Studies in Survey Research. *Methodology, 6*(3), 128-138.

Auspurg, K., & Hinz, T. (2015). *Factorial Survey Experiments*. Sage University Paper Series on Quantitative Applications in Social Sciences, 07-175. Los Angeles: Sage.

Auspurg, K., & Jäckle, A. (2017). First Equals Most Important? Order Effects in Vignette-Based Measurement. *Sociological Methods & Research, 46*(3), 490-539.

Campbell, D. T. (1950). The Indirect Assessment of Social Attitudes. *Psychological Bulletin, 47*, 15-38.

Dülmer, H. (2007). Experimental Plans in Factorial Surveys: Random or Quota Design? *Sociological Methods & Research, 35*(3), 382-409.

Dülmer, H. (2016). The Factorial Survey: Design Selection and its Impact on Reliability and Internal Validity. *Sociological Methods & Research, 45*(2), 304-347.

Eifler, S. (2007). Evaluating the Validity of Self-Reported Deviant Behavior Using Vignette Analyses. *Quality & Quantity, 41*, 303–318.

Eifler, S. (2010). Validity of a Factorial Survey Approach to the Analysis of Criminal Behavior. *Methodology, 6*(3), 139-146.

Eifler, S., & Petzold, K. (2014). Der Einfluss der Ausführlichkeit von Vignetten auf die Erfassung prosozialer Einstellungen: Ergebnisse zweier Split-Ballot Experimente. *Soziale Welt, 65*(2), 247-269.

Eifler, S., & Petzold, K. (2019). Validity Aspects of Vignette Experiments: Expected "What-If"-Differences between Reports of Behavioral Intentions and Actual Behaviour. In P. Lavrakas, M. Traugott, C. Kennedy, A. Holbrook, E. de Leeuw, & B. West (Eds.), *Experimental Methods in Survey Research: Techniques that Combine Random Sampling with Random Assignment* (pp 393-416). New York: Wiley.

Golden III, J. H., Johnson, C. A., & Lopez, R. A. (2001). Sexual Harassment in the Workplace: Exploring the Effects of Attractiveness on Perception of Harassment. *Sex Roles, 45*(11/12), 767-784.

Groß, J., & Börensen, C. (2009). Wie valide sind Verhaltensmessungen mittels Vignetten? Ein methodischer Vergleich von faktoriellem Survey und Verhaltensbeobachtung. In P. Kriwy & C. Gross (Eds.), *Klein aber fein! Quantitative Sozialforschung mit kleinen Fallzahlen* (pp 149-178). Wiesbaden: VS Verlag für Sozialwissenschaften.

Hopkins, D. J., & King, G. (2010). Improving Anchoring Vignettes: Designing Surveys to Correct Interpersonal Incomparability. *Public Opinion Quarterly, 74*(2), 201-222.

King, G., Murray, C. J. L., Salomon, J. A., & Tandon, A. (2004). Enchancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research. *American Political Science Review, 98*, 191–207.

King, G., & Wand, J. (2007). Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes. *Political Analysis*, *15*, 46–66.

Krysan, M., Couper, M. P., Farley, R., & Forman, T. A. (2009). Does Race Matter in Neighborhood Preferences? Results from a Video Experiment. *American Journal of Sociology, 115*(2), 527–559.

Kuhfeld, W. F. (1997). Efficient Experimental Designs Using Computerized Searches. In Sawtooth Software (Ed.), *Sawtooth Software*. Research Paper Series (pp 1-14). Retrieved July 13, 2022 (https://homepage.stat.uiowa.edu/~gwoodwor/AdvancedDesign/KuhfeldTobiasGarratt.pdf).

Kuhfeld, W. F. (2010). Experimental Design: Efficiency, Coding, and Choice Designs. In W. F. Kuhfeld (Ed.), *Marketing Research Methods in SAS. Experimental Design, Choice, Conjoint, and Graphical Techniques* (pp 53-241). Retrieved July 13, 2022 (http://support.sas.com/techsup/technote/mr2010c.pdf).

Louviere, J. J. (1994). Conjoint Analysis. In R. P. Bagozzi (Ed.), *Advanced Methods of Marketing Research* (pp 223-259). Cambridge, MA: Blackwell Publishers.

Luce, R. D., & Tukey J. W. (1964). Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement. *Journal of Mathematical Psychology, 1*, 1-27.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp 105-142). New York, NY: Academic Press.

McLean, R. A. & Anderson, V. A. (1984). *Applied Factorial and Fractional Designs*. New York, NY: Marcel Dekker.

Mutz, D. (2011). *Population-Based Survey Experiments*. Princeton: Princeton University Press.

Noel, N. E., Maisto, S. A., Johnson, J. D., Goings, C. D., & Hagman, B. T. (2008). Development and Validation of Videotaped Scenarios: A Method for Targeting Specific Participant Groups. *Journal of Interpersonal Violence, 23*(4), 419–436.

Petzold, K., & Eifler, S. (2020). Die Messung der Durchsetzung informeller Normen im Vignetten- und Feldexperiment. In I. Krumpal, & R. Berger (Eds.), *Devianz und Subkulturen* (pp 167-204). Kriminalität und Gesellschaft. Springer VS: Wiesbaden.

Petzold, K., & Wolbring, T. (2019). What Can We Learn from Factorial Surveys About Human Behavior? A Validation Study Comparing Field and Survey Experiments on Discrimination. *Methodology, 15*, 19–30.

Rashotte, L. S. (2003). Written Versus Visual Stimuli in the Study of Impression Formation. *Social Science Research, 32*(2), 278-293.

Rossi, P. H. (1979). Vignette Analysis: Uncovering the Normative Structure of Complex Judgments. In: R. K. Merton, J. S. Coleman, & P. P. Rossi (Eds.), *Qualitative and Quantitative Social Research. Papers in Honor of Paul F. Lazarsfeld* (pp 176-186). New York: The Free Press.

Rossi, P. H., & Anderson, A. B. (1982). The Factorial Survey Approach: an Introduction. In P. H. Rossi, & S. L. Nock (Eds.). *Measuring Social Judgments. The Factorial Survey Approach* (pp 15-67). Beverly Hills, CA: Sage.

Ryan, Th. P. (2007). *Modern Experimental Design*. Hoboken, NJ: John Wiley & Sons.

Sauer, C., Auspurg K., & Hinz T. (2020). Designing Multi-Factorial Survey Experiments: Effects of Presentation Style (Text or Table), Answering Scales, and Vignette Order. *Methods, Data, Analyses, 14*(2), 195-214.

Shadish, W. R., Cook, T. D., & Campbell, D. D. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Belmont, CA: Wadsworth.

Shamon, H., Dülmer, H., & Giza, A. (2022). The Factorial Survey: The Impact of the Presentation Format of Vignettes on Answer Behavior and Processing Time. *Sociological Methods & Research, 51*(1), 396–438.

Sniderman, P. M., & Grob, D. B. (1996). Innovations in Experimental Design in Attitude Surveys. *Annual Review of Sociology, 22*, 377-399.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling*. Second Edition. Los Angeles: Sage.

Triandis, H. C., Chen, X. P., & Chan, D. K.-S. (1998). Scenarios for the Measurement of Collectivism and Individualism. *Journal of Cross-Cultural Psychology, 29*(2), 275–289.

van Gelder, J.-L., de Vries, R. M., Demetriou, A., van Sintemartensdijk, I., & Donker, T. (2019). The Virtual Reality Scenario Methods: Moving from Imagination to Immersion in Criminal Decision-Making Research. *Journal of Research in Crime and Delinquency, 56*(3), 451-480.

Van Soest, A., Delaney, L., Harmon, C., Kapteyn, A., & Smith, J. P. (2011). Validating the Use of Anchoring Vignettes for the Correction of Response Scale Differences in Subjective Questions. *Journal of the Royal Statistical Society. Series A: Statistics in Society, 174,* 575-595.

Wallander, L. (2009). 25 Years of Factorial Surveys in Sociology: A Review. *Social Science Research, 38*, 505-520.