

Danish University Colleges

## Qualitative Vignette Experiments: A Mixed Methods Design

Harrits, Gitte Sommer; Møller, Marie Østergaard

*Published in:*  
Journal of Mixed Methods Research

*DOI:*  
<https://doi.org/10.1177/1558689820977607>

*Publication date:*  
2021

*Document Version*  
Peer reviewed version

[Link to publication](#)

*Citation for published version (APA):*  
Harrits, G. S., & Møller, M. Ø. (2021). Qualitative Vignette Experiments: A Mixed Methods Design. *Journal of Mixed Methods Research*, 15(4), 526-545. <https://doi.org/10.1177/1558689820977607>

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

### Download policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

## Qualitative Vignette Experiments: A Mixed Methods Design

Gitte Sommer Harrits (corresponding author)  
VIA University College  
e-mail: [gish@via.dk](mailto:gish@via.dk)  
Tel: +45 87551876

Bio: Gitte Sommer Harrits is Academic Vice Rector of VIA University College. Her research covers social class, categorizations, street-level bureaucracies and professionalism. Methodologically, she primarily works with different types of qualitative data, including semi-structured interviews, focus groups and vignette experiments.

Marie Østergaard Møller  
Department of Politics and Society, Aalborg University  
e-mail: [mol@dps.aau.dk](mailto:mol@dps.aau.dk)

Bio: Marie Østergaard Møller is associate professor in public management at Department of Politics and Society, Aalborg University. Her research covers formations of solidarity, categorization and stereotypes, street-level bureaucracies and professionalism. Methodologically, she primarily works with qualitative interview data. Also, she teaches qualitative methods, including annual courses at the ECPR Winter School in Methods and techniques. The most recent course covers the use of vignettes (and vignette experiments) in qualitative and interpretive research.

### Abstract:

Despite emerging interest in mixing experimental and qualitative methods, a gap exists in integrating qualitative vignettes and experimental design. We present a novel mixed methods design, the qualitative vignette experiment, defined as the use of vignettes designed to introduce experimentally controlled variation in information within qualitative semi-structured interviews. We argue that this design can provide context-rich data on causal mechanisms. This article contributes to the debate on integration in mixed methods research and is useful for scholars studying human practice and the link between macro-structural conditions and micro-interpretive processes. The design builds on an integrated ontology of critical realism and a complementary epistemological foundation of post-positivism and interpretivism. This means that researchers have to adhere to multiple validity criteria.

### Keywords:

Integrated Research Design\*Experiment\*Interpretive Methods\*Vignettes\*Causality\*

## **Introduction**

Discussions of mixed methods combining qualitative and quantitative research have become widespread and mature (e.g. Creswell & Clark, 2018; Greene, 2008; Tashakkori & Teddlie, 2010). However, although emerging, the discussion of mixing experimental methods with qualitative methods is still young (Fetters & Molina-Azorin 2020; O’Cathain, 2016; Plano Clark et al., 2013; Druckman, Green, Kuklinski, & Lupia, 2006).

Some existing studies indeed already combine intervention studies with the collection and analysis of both quantitative and qualitative data (e.g. Maxwell & Sandlow, 1986; Plano Clark et al., 2013; for a recent comprehensive introduction see O’Cathain 2018). In most of these studies, however, the qualitative component remains secondary (Creswell & Clark, 2018). In contrast, we present an integrated mixed methods design, where an experimental logic is embedded within and integrated with qualitative and interpretive interviews. We call this design the qualitative vignette experiment. Our design is thus related to existing mixed methods designs, but it also presents new and innovative elements.

First, by using vignettes designed to introduce experimentally controlled variation in information within qualitative semi-structured interviews, the design allows for a high level of integration in both research problem, vignette construction, case selection, data collection and analytical strategy, while simultaneously building on a complementary understanding of the epistemology of experiments and interpretivism. Second, this integration allows us to link macro-structural and micro-interpretive mechanisms and processes, and it is therefore especially useful when researchers want to study both the presence of causal relations and causal mechanisms, and how these mechanisms unfold in

social practice, in particular when linked to interpretation and meaning-making.

Third, by using vignettes, the design also allows for the application of an experimental logic beyond interventions studies, i.e. in for example social sciences where some characteristics such as e.g. social class or ethnicity is difficult to manipulate in “real life”. The qualitative vignette experiment is thus useful especially for researchers within social sciences studying subject areas where intervention studies are hard to accomplish.

The article makes three main contributions to the mixed methods literature. First, we contribute by giving concrete advice for how to use qualitative vignette experiments in research. Second, we contribute to the discussion of how integration can be achieved at different levels and stages of mixed methods research, while at the same time adhering to a dual epistemological foundation. Finally, we contribute to the literature on vignettes by suggesting an integrated mixed methods setup of experimental vignettes in qualitative interviews.

We begin with a brief discussion of existing literature on mixing experimental and qualitative methods, as well as of the vignette method. Then, we present an existing study using a qualitative vignette experiment, focusing on how public sector professionals make judgements when encountering citizens embedded in different social contexts. Following this example, we carefully describe the method of qualitative vignette experiment, i.e. how to proceed when designing the vignette experiment, and how to collect and analyze data. Finally, we include a discussion of the ontological and epistemological foundation of the qualitative vignette experiment, focusing both on integration and how to assess validity of conclusions and the quality of inferences.

## **Existing methodological practices and discussions**

Designs integrating experimental and intervention studies with qualitative and interpretive methods are emerging but still not widely discussed compared to other mixed methods designs. Creswell and Clark (2018, pp. 168–170) describe the purpose of the so-called “mixed methods experimental or intervention design” as using qualitative data and analyses to “provide personal, contextual, qualitative experiences” (Creswell & Clark, 2018, p. 174) from the people participating in an intervention study.

Also, O’Cathain (2018) in her recent book gives a comprehensive introduction to, and more examples of, the use of qualitative methods in randomized controlled trials, focusing on how to integrate results from qualitative and quantitative data in RCTs. In many discussions, qualitative research is seen as secondary to quantitative data in the randomized controlled trials (e.g. Plano Clark et al., 2013; Sandelowski, 1996). However, O’Cathain (2018) points out numerous ways that qualitative research improve RCTs, such as help understand how the intervention works, understand impact on participants or understanding variation in outcomes to overcome seeing qualitative data as secondary to quantitative data within the mixed methods intervention design.

Still, O’Cathain’s focus seems to be on integrating quantitative and qualitative data and analyses within an overall experimental research. Compared to this, the qualitative vignette experiment seeks integration between qualitative and experimental research, which means that these two logics are seen as interdependent and complementary. The qualitative vignette experiments thus extends previous discussions of mixed methods intervention designs by suggesting a more integrated mix of qualitative and quantitative-experimental logics.

Doing this, the qualitative vignette experiment is more similar to another recent example of a mixed methods experimental design presented by Robinson and Mendelsohn (2012), called the “qualitative experiment”. Here, focus-group and semi-structured interviews were embedded within an experimental design for case-selection (see Maxwell & Sandlow, 1986 for a similar example using ethnographic methods). In other words, a quantitative-experimental logic was introduced during the selection of interviewees who were then exposed to different types of media texts, where after data collection and analyses were conducted using traditional qualitative methods. However, compared to the qualitative experiment, the qualitative vignette experiment does not only apply a quantitative-experimental logic in selection of interviewees and distribution of material to interviewees, but integrates qualitative and quantitative-experimental logics throughout the study. This in opposition to the qualitative experiment, where qualitative methods is considered dominant, the qualitative vignette experiment sees the qualitative and experimental design logics as equal, complementary and integrated.

The construction and use of vignettes, defined as “text, images, or other forms of stimuli [to] which respondents are asked to respond” (e.g. Hughes & Huby, 2004, p. 37), and more specifically, “short stories about hypothetical characters in specified circumstances, to whose situation the interviewee is invited to respond” (Finch, 1987, p. 105) can be a key to reaching this level of integration. We thus refer to the use of vignettes as a method of data collection and data analysis and not as a method of presenting data, or as a an implementation method testing respondents’ reactions to new treatments, as sometimes done in public health research.

The use of vignettes in collecting data is widespread across different fields (e.g.

health, business, social work and political science), and methodological discussions of vignettes in both quantitative and qualitative research are quite extensive (Barter & Renold, 2000; Finch, 1987; Hughes & Huby, 2004; Jenkins et al, 2015; Rossi & Nock, 1982; Steiner et al., 2016). Most scholars agree that vignettes are well suited for the study of beliefs and attitudes (e.g. Hughes and Huby, 2004), especially when issues are sensitive, but that inferences about behavior should be made with caution. Some scholars focus specifically on the use of vignettes in the study of human judgment (e.g. Eskelinen & Caswell, 2006). However, while the methodological advantages of vignettes in the study of judgments are often emphasized, due to the fact that studying responses to realistic scenarios presented in vignettes can be very close to studying “real-life” judgments, the challenges related to the construction and validity of vignettes is not very well covered in the methodological literature. Finally, the use of vignettes especially with quantitative survey experiments has been used to facilitate studies of variables that are not easy to manipulate in “real life”, such as e.g. social class, and the use of vignette is in general praised for the high degree of flexibility and control that it allows the researchers (e.g. Steiner et al., 2016).

As far as we can tell, there are no discussions of mixed vignette designs in the literature. Instead, the use and discussion of the vignette method is isolated within two strands: a quantitative-experimental strand, where vignettes typically are used in an experimental setup (survey experiments or factorial surveys) to test hypotheses on causal relations, and a qualitative-interpretive strand, where vignettes are used to create a contextually sensitive and authentic setting for specific questions. Below, we demonstrate how these two strands of research can be integrated by using vignettes to reach a high

level of integration at all stages of the research process. Also, we demonstrate how the use of vignettes makes it possible to experimentally manipulate characteristics that are normally difficult to manipulate in intervention studies. First, however, we present a study using qualitative vignette experiments.

### **Using the qualitative vignette experiment in a study on professional judgement**

A few examples on the use of qualitative vignette experiments already exist (Ilsvard & Møller, 2015; Møller, 2009, 2016). Here we present a third study using qualitative vignette experiments. This study focused on how professional judgement of families in social risk may be influenced by social class differences (Harrits, 2019; Harrits & Møller, 2014, 2016). We wanted to test the hypothesis that social class differences would result in a tendency for professionals to be more concerned about children from both the upper and lower classes compared to children from the middle classes, i.e. the class background shared by the professionals. Also, we wanted to explore different mechanisms related to making professional judgments and the possible use of social stereotypes, conceptions of normality and private values. Therefore, the study was designed to create rich data on reasoning and reflections about judgments, while at the same time being designed to facilitate systematic comparisons between judgments of social risk in families from different social class backgrounds. We thus developed a qualitative vignette experiment, combining an experimental vignette design and semi-structured interviews.

Case-selection followed a logic of theoretical sampling with focus on sampling professionals from different contexts. First, we chose three comparable professional groups: Teachers, pre-school teachers, and home nurses. These professional groups

implement preventive social and health policy, and they are often required to make judgments about families in social risk. Further, they are quite similar with regard to training, degree of professionalization and the way in which they perform their tasks and interact with children and families. We selected a total of 58 professionals (16 home nurses, 22 school teachers, 20 pre-school teachers).

Second, we chose professionals from eight school districts in four municipalities (i.e. two districts in each municipality). The municipalities were selected based on criteria of size, urbanity, and geographical location, and the school districts were selected based on criteria of social heterogeneity/homogeneity, which was evaluated on the basis of available local socio-economic district data. Within the eight local school districts, institutions and interviewees were then selected in a final step using snowballing and based on criteria of convenience and accessibility, meaning that we chose those institutions and interviewees that were willing to let us talk to them. We made sure to collect a reasonable number of interviews in each professional group to allow for comparative and statistical analysis. However, in this study, we did not make any power calculations (see further below).

The qualitative vignette experiment was embedded within a semi-structured interview divided in three sections, where section one and three used traditional semi-structured questions to collect data on interviewees' work, their professional and social identity and their own understanding of how they make judgments about social risks. In the second section of the interview guide, we introduced a vignette experiment, presenting two vignettes to each interviewee. Each vignette described a situation of a child and family. The first vignette (A), which described a middle-class family, was

introduced to all interviewees, and it functioned as a baseline for the experimental analysis. The second vignette described either an upper-class family (B) or a lower-class family (C), using randomization to decide whether an interviewee should read vignette B or C after having read the first vignette A. The practical organization of the randomization was handled by allocating vignettes in pairs (A+B and A+C) in envelopes and randomly drawing an envelope before each interview, making sure to stratify the randomization within professions and local contexts. Also, the allocation of vignettes in an envelope made sure that the interviewer was not aware of the chosen combination of vignettes before introducing them to the interviewee.

All vignettes were designed to portray a realistic problem which could arguably be seen as both a minor social risk problem not eligible for social services and as a possible social risk problem eligible for social services in some form. This was done to facilitate interviewees' interpretations, which could then be used for subsequent analysis of causal mechanisms. Names as well as information about employment of parents were chosen as cues about social class. As further elaborated below, the vignettes were designed based on existing theoretical and empirical studies, and they were carefully vetted by experts.

Because the interviews were conducted across three professional groups, we had to design a total of nine vignettes, presenting each professional group with problems that fitted their expertise and professional practice. Across the three professional groups, the description of social class background was kept similar, varying only the age of the child and parents to fit the age groups that professionals meet in their lines of work. For example, the pre-school teachers would read stories about kindergarten children, whereas

the teachers would read stories about school children. In table 1, the vignettes presented to pre-school teachers are shown.

[Table 1]

After reading each story, the interviewees were asked to tell what they ‘thought about’ these cases, and how they would act if confronted with such cases. Following this open question, interviewees were probed on aspects of the vignettes and finally, in a closed question, they were asked to make a summary judgment on the degree of social risk of the child and family and their suggestion for how to help the family.

All interviews were transcribed verbatim and organized in NVivo, which we used to perform pattern identification analysis across interviews with interpretative content analysis of each interview. Also, after coding answers to summary judgments on social risk, these were transformed to quantitative data (high/low social risk), and statistical analysis of the differences between summary judgments of vignettes A, B and C were conducted (Mosholm and Fetters, 2017). The statistical analysis tested the hypothesis on whether differences in social class of the families significantly impacted professional judgments of social risk.

Further, qualitative data from the open ended responses to the vignettes were coded, following traditional qualitative guidelines for initial and focused coding and analysis (Charmaz, 2006). The interpretive analysis of professionals’ interpretation and reasoning when confronted with the vignettes helped us demonstrate the content of their judgments, which could account for the differences towards different social classes. Finally, we used the theoretical sampling and coding of the narratives to conduct analyses of contextual

conditions of professional judgments, showing for example how the impact of class difference and social stereotypes were more prevalent in socially homogenous areas (see Harrits, 2019; Harrits & Møller, 2014, 2016 for further presentation of results).

### **Steps and methods in the qualitative vignette experiment**

As can be seen in the example above, the qualitative vignette experiment achieves a high degree of integration at all stages of the research process (Fetters & Freshwater, 2015; Johnson et al., 2019). Thus, the experimental and the interpretive logic support each other throughout the study, including in the crafting of the vignettes. Integration is furthermore supported by an integrated ontology (critical realism) and a complementary epistemology (post-positivism and interpretivism), which results in a complementary set of standards for judging validity and inference quality of conclusions (Onwuegbuzie & Johnson, 2006). We now concentrate on laying out how the different methodological steps (see figure 1) should be followed when using qualitative vignette experiments illustrated in the procedural diagram. In the concluding discussion, we sum up the qualitative vignette experiment design and discuss some limitations.

[Figure 1 about here]

#### *Step 1: Research Problem*

The process of designing a qualitative vignette experiment study starts with crafting an integrated research problem that focuses on both causal relations, interpretation and meaning-making as a causal mechanism. To address causal relations, the research

problem should be based on prior knowledge, such as theory or empirical insights, suggesting hypotheses about specific causal variation. Thus, there is a deductive and theory testing element to the qualitative vignette experiment. However, at the same time, the research question should also focus on interpretive causal mechanisms, and this element could include a more explorative strategy. Most importantly, in order to be suitable for a qualitative vignette experiment, the research problem should frame what we will refer to as a double perspective on causality, i.e. both causal variation and causal mechanisms, and avoid subordinating the experimental logic to the interpretive, or vice versa. In practice, researchers should do a literature review and thus build the research question on existing findings, but at the same time remain open for possible new findings with regard to causal mechanisms. In the study outlined above, the research problem was framed both as a test of a specific hypothesis on class affecting judgements of social risk, and as an exploration of *how* professionals' judgments are affected by social class position of the families they encounter. This was done by reviewing existing research on the use of social stereotypes in professional work, combined with the development of a theoretical model on how social class would impact judgement in a particular professional context.

Ontologically, such research questions are founded in critical realism, which insists on a realist social ontology (Maxwell, 2004). This means that social reality exist 'outside' of our interpretation of it. It also means, though, that human thoughts, reasoning and interpretations are part of social reality, and that, for example, interpretations influence behavior and create 'real' measureable effects. At the same time, critical realism also claims that social reality is so complex that there can be more than one

correct or scientifically legitimate way of understanding and representing it (Lakoff, 1987). The implications of this is that while we can refer to one (integrated) ontological foundation, we need complementary epistemological perspectives to approach this reality (Maxwell & Mittapalli, 2010). The qualitative vignette experiment thus builds on two complementary epistemological stances linked to the study of variation and effects (post-positivism) and micro-interpretive processes (interpretivism) (Bevir, 2006).

This has implications for how we assess validity. Some researchers see validity as linked exclusively to specific methodological traditions, and they find it difficult to apply validity criteria in mixed designs. However, we see validity as a way to evaluate the quality of inferences (Maxwell & Mittapalli, 2010; Onwuegbuzie & Johnson, 2006; Teddlie & Tashakkori, 2003), which means that validity criteria are not linked to specific methodological procedures, but rather to the way procedures relate to the research problem. Below, we address these multiple validity criteria when relevant in each methodological step.

### *Step 2: Vignette construction*

The key element of the qualitative vignette experiment is construction of vignettes, since the benefits of the design integrating a study of causal relations and causal mechanisms is only realized if vignettes are designed and used properly. Vignettes should thus allow for both the introduction of controlled variation in information, and for in-depth interpretation of how this information is perceived and interpreted by interviewees. We start by addressing how to craft the vignette content so that it allows for interpretative analyses, and then we explain how to introduce controlled variation into the vignettes.

First of all, researchers must be careful to design vignettes in authentic ways. If a study fails in presenting a plausible story that very accurately resembles reality, it also fails in capturing valid experiences about the phenomenon under study (Barter & Renold, 2000; Finch, 1987). The first question that should be addressed is therefore: What sources of information should be used when crafting the vignettes? Surely, the answer to this question will depend on the specific research problem. As a general advice, however, researchers should rely on a number of different sources in order to design plausible vignettes.

Theory, earlier research, observations, and socially recognized narratives and representations (stereotypes), policy documents, laws, guidance documents, and media stories all work as possible sources to initiate the design process. Also, a crucial source is pilot observation of the daily practices under study (Barter & Renold, 2000). When crafting the vignettes, researchers should, then, adapt the real-life description of problems from the specific groups studied, such as the pre-school teachers in the example study discussed above. Pre-school teachers deal with child care, and focus especially on development of language, motor skills and social-emotional skills. Therefore, we focused on integrating these problems in the vignettes.

Also, to make vignettes plausible, Barter and Renold emphasize letting vignette content reflect mundane occurrences and avoid eccentric characters. They also emphasize the importance of designing vignettes with enough context to present coherent stories, but without making the stories too complex and exhausting, or too “constructed” and loaded with speculative combinations of factors that will harm the internal consistency of the storyline.

In addition to the importance of drafting plausible story lines, researchers need to concern themselves with constructing vignettes in ways that engage participants both cognitively and emotionally, as if they were engaging with *real* people. This means that vignettes have to be designed to be not only plausible, but also as authentic as possible to make sure that interviewees will actually perceive them as being both relevant and interesting. Therefore the content of the vignettes must not be too strange or exotic, causing interviewees to disregard them as not being worthy of attention, nor the opposite—they should not be so dull or mundane that interviewees lose interest. Further, the form of the vignettes should avoid very technical terms or leaden language, while at the same time maintaining a sufficient precision and familiarity within a specific discourse. Vignettes can become so “technically correct” that interviewees fail to engage in sufficiently comprehensive interpretations. However, there is also the risk of creating vignettes that are so engaging and “lively” that participants lose sight of the main theme of the vignette story. Striking the right balance is thus key to validity of the study.

As already mentioned, the key idea of the qualitative vignette experiment is also the introduction of experimentally manipulated and controlled variation in information. Therefore researchers initially consider each of the building blocks of the information needed in the vignettes, and how to present them. The first building block is information on the variable hypothesized as a cause, and the second building block is information on the situation or practice hypothesized as being affected. In our example, we needed information on class background of the families as well as information on specific problems in the children’s behavior, the interpretation of which we hypothesized would be affected by information on class background.

In general, researchers should be careful to present this information in a way so that it realistically could, but not necessarily would, cause different reactions. In other words, the introduction of variation should be recognizable but not so obvious that it results in social desirability bias or other types of biases in the interviewee's reactions. For example, we varied descriptions of social class background by referring to parental occupations clearly associated with different class positions, as well as to typical Danish "classed" names, i.e. names stereotypically associated with different social class positions in society. Similarly, we described the children's problems in a nuanced way that allowed for different interpretations. Here, literature and empirical studies on children in social risk and on class differences in Denmark were used as inspiration.

Also, the introduction of variation in information needs to be experimentally controlled. Typically, traditional qualitative-interpretive vignette studies introduce only one vignette, whereas quantitative vignette experiments introduce several. The qualitative vignette experiment, introduces two vignettes (baseline and experiment vignette) in the same qualitative interview, and varies information on the independent variable. Information on the dependent variable, that is the situation or problem, is kept as similar as possible. It is not possible, however, to use identical problems or wordings in the two vignettes, since identical wordings of problems would compromise the authenticity and engagement with the storyline as interviewees would see through the experimental purpose.

If needed, further variation can be introduced by varying the experimental vignette, and randomly allocating different versions to the interviewees. As explained above, this is what we did, allocating the A+B vignettes to half the interviewees and the

A+C vignettes to the other half. Finally, if different groups are studied, a set of vignettes should be developed for each group, since each group will need a situation or problem specifically aimed at their practice. Again, this is needed to ensure that vignettes are perceived as authentic, relevant and engaging. In our study, we thus developed a baseline and two versions of the paper vignettes for all three professional groups (nine vignettes in total).

In the methodological literature on vignettes, pilot testing is emphasized as the overall method to strengthen vignettes' plausibility and authenticity. Researchers should therefore have an expert panel, whose members have sufficient knowledge and experience to vet both the presentation and variation in information, as well as the plausibility of the overall story-line. Similarly, researchers should pilot test and vet vignettes using a panel of respondents similar to respondents in the study, thereby making sure that vignettes and the variation created including different wordings of the problem works as intended. This includes testing whether different versions of the vignettes prime different responses, both the differences intended in the experiment, and differences not intended, i.e. where vignettes are crafted to be similar. Finally, it should be tested whether vignettes are sufficiently engaging and produce vivid responses.

In the study described above we used our review of existing research and observations of social interactions in home nursing, kindergartens and schools to design the storylines of the vignettes in plausible terms. When we were done drafting them we had them vetted by field experts (a home nurse, a pre-school teacher and a teacher) and by theoretical experts within the field of prevention and class sociology. Based on their feedback we adapted and changed parts of the stories.

Constructing vignettes in this way integrates an interpretive tradition focusing on context-rich and authentic storylines with an experimental tradition introducing controlled variation in the information contained. When judging the validity of the vignettes, we therefore need to take validity criteria from different traditions into consideration. One set of validity criteria has to do with the presentation of information and the overall storyline. Here, we combine construct validity (Shadish et al., 2002), derived from the experimental and post-positivist tradition, and ecological validity, derived from the interpretive tradition.

We understand construct validity as related to the design of vignettes in ways that carefully and correctly represent the reality they are supposed to represent. This is why we emphasize the construction of vignettes based on existing and detailed empirical knowledge and the use of experts to vet the vignettes. However, as also explained above, vignettes and questions do not only have to be realistic and technically correct, they also have to be authentic in order to cognitively and emotionally engage interviewees and prevent superficial (non-valid) responses to the vignette stories. In other words, qualitative vignette experiments also have to take into consideration ecological validity, which we defined as capturing “the daily life conditions, opinions, values, attitudes, and knowledge base of those we study as expressed in their natural habitat” (Cicourel, 1982, p. 15).

A second set of validity criteria has to do with the introduction of variation and the attempt to make causal inference based on the qualitative vignette experiment. It is a key challenge of causal inference that we can never at the same time observe a situation with a ‘treatment’ present and compare it with an observation of the exact same situation

without the ‘treatment’ present (Holland, 1986; Johnson et al., 2019). Therefore, efforts should be made to isolate effects, preferably by using randomized controlled experiments (Shadish et al., 2002). This way of understanding causality is typically referred to as the variance-oriented model (Maxwell, 2012), and validity of causal inferences is typically referred to as internal validity, focused on the control of variation, possible confounding factors and endogeneity. However, especially within critical realism, causal inference can also be approached by a process-oriented model of causality, where the theoretical validity (Maxwell, 2012) of causal inferences rests mainly on empirically demonstrating how causal mechanisms or processes work.

The qualitative vignette experiment draws on both these understandings of causality and suggest considering both internal and theoretical validity of causal inference. First, internal validity is important when designing variation in information. On the one hand the use of a baseline and an experiment vignette in the same interview strengthens internal validity, since the effects of the varied information on the independent variable can be observed within the same interviewee, resembling a pre-post-test design of laboratory experiments.

However, the use of two vignettes (baseline and experiment) in each interview creates the need for varying the wordings of the problem (the dependent variable), which could potentially introduce confounding variation. Thus, it could be the case that the effects found in our study were caused by differences in the problem descriptions not social class backgrounds. This is why we emphasize that the wordings used to describe problems should be as similar as possible without being identical. However, the qualitative vignette experiment also facilitates the study of causal mechanisms by

creating rich and contextual data on how reasoning and interpretations of situations shape judgments. This facilitates in-depth interpretive analysis of responses to vignettes, which can strengthen theoretical validity and thus the overall causal inferences.

### *Step 3: Case selection and sampling*

Data collection in qualitative vignette experiment is carried out as qualitative interviews, or some other form of qualitative data collection method (see below). Due to constraint of resources, the number of cases or interviewees will thus typically be small or medium (10-200, see also O’Cathain 2018 for a discussion of sample size), as in most qualitative studies. This suggests the use of a theoretical sampling strategy. Please note, however, that if randomization of different versions of the experimental vignette is used, there needs to be more than approximately 30 interviews in total. In some cases, it could even be useful to perform a power calculation, to make sure that sample sizes are large enough to warrant statistical analysis. Theoretical sampling in the tradition of laboratory experiments is not that different (Shadish et al., 2002) from the tradition of qualitative studies (Maxwell, 2012), and both traditions emphasize analytical inference as the main goal, i.e. inference based on a concrete analysis of the differences between the sample and the population or contexts that inferences are made to.

However, when considering analytical inference and deciding on a strategy for theoretical sampling, researchers should be careful to choose only individuals that fall within the studied population. When using vignettes that try to introduce realistic scenarios and storylines to facilitate authentic reasoning, it is important to make sure that interviewees are actually regularly engaged in the type of practices studied. At the same

time, researchers need to select individuals that are situated in different contexts and have different characteristics (e.g. age and social class), in order to facilitate some form of inference, generalization and transference of results. Most often, sampling strategies aiming for either typical cases or cases varying across different contexts and with different characteristics, will be most relevant (Mohr, 1999; Tavory & Timmermans, 2014). In addition to considering analytical inference, theoretical sampling of cases or interviewees in different contexts could also allow for further analysis of contextual conditionings, i.e. of how the causal relations and mechanisms studied with the vignettes unfold under different contextual conditions.

In summary, the sampling process must include criteria that makes saturation and analytical inference beyond the specific sample possible, criteria for selecting relevant individuals that will be able to explain themselves in narrative terms about the phenomenon under study, and criteria to apply a comparative control logic with regard to causal analysis. This illustrates the integration of a post-positivist and an interpretive approach to sampling.

In both experimental and interpretive studies, theoretical sampling and concrete analysis of the differences between the sample and the population is seen as crucial. Following an experimental and post-positivist tradition analytical inference is thus a way of strengthening external validity, although this can never be as high as in studies using random sampling. Following a qualitative and interpretive tradition, external validity is typically described as the ability to transfer results to other contexts. Also, analytical inference can build on e.g. the observation of saturation of patterns across different contexts.

To meet these requirements, in the study described above, we chose interviewees from three comparable professional groups, who we knew were engaged with judgements about families in social risks on a daily basis in similar ways. We selected them from different social contexts to make sure we could qualify whether any observed patterns were generic or bound to particular contexts. Furthermore, the analytical inferences were carefully specified with regard to contextual conditions.

#### *Step 4: Data collection*

In the qualitative vignette experiment, vignettes are the key instrument for data collection. However, vignettes need to be embedded into a data collection method, allowing for interpretive and open questioning and analysis, as well as closed questions that measure responses to the vignettes precisely and allows for statistical analysis of effects. This means that the overall data collection also integrates an interpretive and experimental or post-positivist logic.

In the example above, we chose to embed the vignette experiment in a semi-structured interview, but there are other options such as focus groups, observational studies with follow-up interviews or longer in-depth ethnographies combined with surveying. The advantage of embedding the qualitative vignette experiment in a semi-structured interview is the ability to use different interview questions combining typical interpretive questions with more closed questions. Embedding the qualitative vignette experiment in a semi-structured interview, three methodological issues arise: how to capture responses to vignettes, whether and how to ask additional questions to provide additional data on the phenomenon under study, and how to combine the vignettes and the additional questions.

With regard to questions to the vignettes, it is important to capture both the interviewees' interpretation of and reasoning about vignette information, and their final response to the situation presented in the vignette, i.e. their final judgement. In other words, it is important to capture both the possible causal effects of the vignette, as well as the causal mechanisms, which are here considered to be interpretation and reasoning. The literature on vignettes typically suggests the use of 'think-aloud'-questions in order to let interviewees react to and talk freely about the vignettes (Robinson & Mendelson, 2012). The logic here is similar to that of "ordinary language interviewing" (Schaffer, 2006), where participants' own wordings and associations come forward instead of forcing theoretical concepts or conceptually stretched wording onto the conversation.

However, when it comes to capturing the causal effects and final response, researchers should carefully craft closed questions that force the interviewees to choose a realistic and appropriate response. Further, both open think-aloud questions and closed questions should be followed up by probing for elaboration and clarification. This makes sure that the data is rich and nuanced, while also strengthening the validity of interpretation.

When embedding the qualitative vignette experiment in a semi-structured interview, researchers can utilize the interview to collect further data asking open-ended or narrative questions and letting interviewees themselves describe how they see their own practice can thus provide data that can help corroborate the results of the vignette experiment. As an example, we asked questions about professionals' own experience with clients, children and families from different classes and with judging social risk to

see whether similar reasoning was used here compared to the reasoning used when presented with the vignettes.

One particular concern is avoiding social desirability bias, i.e. the tendency for interviewees to respond in desired ways to appear socially appropriate (Barter & Renold, 2000). Here, the combination of questions to the vignettes and open-ended questions on interviewees own experiences can help. The fictitious, however plausible nature of the vignettes may reduce the potential social desirability bias that can come with client loyalty or lack of client loyalty and give researchers a valid insight into sensitive topics such as interviewees' interpretations of problems and their causes. Such insight is often inaccessible in observations, difficult to access in surveys, and easy to miss in standard interviewing.

Finally, if additional questions beyond the vignettes are included, researchers need to consider how to combine these questions with the vignette questions. If vignettes are introduced in the beginning of the interview, there is a risk of priming narratives about interviewees' own experiences. Thus, additional and more spontaneous narratives to validate analysis of vignettes should be collected before introducing vignettes. In our study, we did this, though additional questions were also asked after the vignettes, this time about interviewees' personal background beyond the social phenomenon under study. These questions are used to check for alternative explanations to the identified patterns of reactions to the vignettes and to perform more contextualized understanding of these.

Also, when qualitative vignette experiment are embedded in a semi-structured interview, the planning and practical organization of interviews needs to follow

traditional methodological advice, on e.g. setting, length, and recording (Kvale & Brinkmann, 2009). Further, vignettes should be brought to the interview site in a suitable form, for example on two sheets of paper with vignettes printed in a font that makes them easily readable to interviewees. Finally, if randomized allocation of different versions of the experiment vignette is chosen, this randomization should be organized before the interviews, for example by organizing vignettes in envelopes and drawing randomly from these before going to an interview. If samples are drawn across different contexts (e.g. municipalities), remember to stratify randomization within these contexts to avoid ending up with an overweight of one combination of vignettes in one municipality. In the study described here, we had a student organize vignettes in envelopes beforehand. The student made sure vignette combinations were evenly presented across the different contexts, and the combination of vignettes were blinded to us until they were introduced in the concrete interview.

In terms of epistemological foundation and validity criteria, issues with regard to data collection resemble the discussion above of construct and ecological validity.

#### *Step 5: Analytical strategy*

Analysis includes organizing and recoding data in an appropriate way, a quantitative analysis of causal effects, an interpretive analysis of causal mechanisms and an integration of the two analyses. We explain each step separately here.

First, data needs to be organized. This means transcribing interviews or organizing audio files, and preferably choosing a software program to support analysis. Also, answers to the closed ended questions, which summarizes responses to the

vignettes, needs to be captured from the qualitative data and quantified. Preferably, this includes a focused coding of the qualitative data, and an interpretation and careful recoding of each response into a quantitative scale (see e.g. Mosholm and Fetters). In the study presented above, we chose to use verbatim transcription, and we organized transcripts and classified interviewees' background information as well as information on the combination of vignettes in NVivo. Also, we coded responses to closed questions, and recoded them into a dichotomous response (high/low social risk).

Using the quantitative data on final vignette responses recoded from the qualitative data, an analysis of causal effects can then be performed, using a traditional chi-square test, Fishers exact test or equivalent to compare groups (see e.g. Angrist and Pischke 2015; Field 2013). In our analysis, we compared responses to the baseline and the experiment vignettes, and to the two versions of the experiment vignettes using Fishers exact test. We found significant differences between judgements of middle-class families (baseline) and both upper- and lower-class families (experiment vignettes), but (as expected) not between upper- and lower class families (Harrits 2019).

The analysis of causal mechanisms is more labor-intensive, and follows general advice for conducting interpretive analysis. We recommend using both open/initial and closed/focused coding techniques, as well as testing the coding and interpretations of the material for inter and intra coding reliability (Charmaz, 2006). In order to fully utilize the qualitative vignette experiment, researchers should conduct both careful within-case analysis of the reasoning of each interviewee, and across-case and pattern-identification analysis of this reasoning, as a minimum across the different vignettes, in order to compare with statistical analysis of causal effects (Miles, Huberman & Saldana 2014).

Also, analysis of the spontaneous narratives on practices can be compared to the interpretation and reasoning about vignettes to further validate results. Finally, comparing different segments of the data can help to explore the impact of contextual conditions.

Next, the integration of the different analyses should be organized by the overall frame of causal effects and causal mechanisms. Thus the key question is, if the expected causal effects can be identified, and further, whether or not the way interviewees reason about vignettes supports (and perhaps elaborates) the overall theory on causal mechanisms. In our study, we could demonstrate that social class background did have an effect on professionals' judgements, in the way that professionals tended to be more worried about children from class backgrounds different from their own. This conclusion was supported by analysis of causal mechanisms, which demonstrated that social stereotypes and conceptions of normality played an important role in professionals' reasoning (Harrits 2019).

The validity of analysis thus also rests upon multiple validity criteria. First, the validity of the quantitative data rests on construct validity of the closed questions asked, as well as on the validity and reliability of the recoding of qualitative to quantitative data. Here, reliability tests could be useful. Second, the interpretive analysis needs to adhere to what Maxwell (2012) calls interpretive validity, which focuses on the adequacy of interpretation. Thus, validity has to be judged by how thoroughly interpretations of the data are grounded and explained. This does not mean that interpretations cannot move beyond the perspective of the interviewees themselves (emic perspective), but only that etic interpretations must, then, be carefully documented and explained (Soss, 2006; Tavory & Timmermans, 2014). Second, validity of the analysis further rests on both

internal and theoretical validity, and on the way analysis of causal effects and causal mechanisms are integrated. In other words, validity must here be judged by how well the analysis of causal mechanisms can be linked to the analysis of causal effects, and thus how solid a convergence can be crafted between internal and theoretical validity. Casual inference can be made only if the analysis of causal effects and the analysis of causal mechanisms converge.

## **Discussion**

In this article, we have presented the qualitative vignette experiment, which is a mixed methods experimental design. We have demonstrated how the qualitative vignette experiment builds on an integrated ontology of critical realism, and on a complementary epistemological foundation of both post-positivism and interpretivism, and that this means that researchers have to adhere to multiple criteria of validity. The qualitative vignette experiment is summed up in table 2 below.

[Table 2]

### *Contribution to the field of mixed methods research methodology*

There is an emerging discussion of mixed methods experimental designs within the community of mixed methods research scholars. The qualitative vignette experiment presented in this article sees the qualitative-interpretive and quantitative-experimental logics as complementary and interdependent, and thus aims for a high level of integration at all stages. Also, the qualitative vignette experiment facilitates moving beyond

intervention studies to phenomena that traditionally are hard to manipulate.

As we have shown, the qualitative vignette experiment achieves a high degree of integration, from crafting the research problem and throughout all subsequent methodological stages. In particular, the key contribution of the design is the double perspective on causality and the integration of analyses of causal effects, using the experimentally controlled variation in the vignettes, as well as the analyses of causal mechanisms, using the context-rich data on interviewees' interpretation of vignettes. In this way, the qualitative vignette experiment is different from existing designs, including both mixed methods interventions designs and qualitative experiments (see table 3 below).

[Table 3 about here]

As we have argued, the qualitative vignette experiments and the integration of experimental and interpretive research designs allows for the testing of causal hypotheses as well as an exploration of causal mechanisms such as demonstrated in the example presented above (Harrits 2019; Harrits & Møller 2014, 2016), where we tested and explored a hypothesis on the impact of social class on professional judgement of families' social problems. Also, the qualitative vignette experiments allowed us to demonstrate, how these differential judgements and decisions were created by biases and stereotypes on social class held by frontline professionals.

Other studies have used qualitative vignette experiments for other research problems, for example showing how general practitioner's private perceptions of health and normal lifestyles impact medical advice given to patients (Ilsvard & Møller 2015),

and how social worker's understanding of worthiness and belonging results in discrimination of clients with diffuse illnesses such as e.g. fibromyalgia (Møller 2009, 2016). Common for these studies is the use of the qualitative vignette experiment as a way to bring forward a meticulous and nuanced analysis of the ways in which social and cultural contexts and categories are causally linked with micro-institutional interpretations and practices.

However, there are also limitations to the application of a qualitative vignette experiment. First of all, the high degree of integration throughout the methodological steps hinges upon the crafting of an integrated research problem, which requires not only solid prior knowledge and pilot-testing, but also an integrated research team, who understands the double perspective of causality. Second, the requirement that researchers must integrate different methodological techniques from the experimental and the interpretive tradition in their crafting of vignettes, sampling of data, collection of data and analysis means that some tension can arise. Especially when crafting vignettes, where attention to detail and realistic representations (construct validity) can come in conflict with attention to the overall storyline, authenticity and engagement (ecological validity). Here, striking the right balance is key. If this is not done convincingly, you risk over-framing your interviewees in ways that will make responses biased and therefore invalid. Third, caution should be taken when making inferences from vignettes responses to behavior. As mentioned above, vignette responses in some instances will be quite close to actual behavior, such as e.g. interpreting a case file. However, even in these situations, researcher should be careful with extrapolating directly from vignette response to actual behavior.

In conclusion, qualitative vignette experiments are thus especially relevant for researchers interested in how macro-structural conditions nest micro-interpretive processes and who work in solid integrated research teams.

## References

- Angrist, J. D. *Mastering Metrics*. Princeton (NJ):Princeton University Press.
- Barter, C., & Renold, E. (2000). "I wanna tell you a story": Exploring the application of vignettes in qualitative research with children and young people. *International Journal of Social Research Methodology*, 3(4), 307–323.
- Bevir, M. (2006). How narratives explain. In P. Schwartz-Shea & D. Yanow (Eds.), *Interpretation and Method. Empirical Research Methods and the Interpretive Turn* (pp. 281–290). Armonk: M.E.Sharpe.
- Charmaz, C. (2006). *Constructing Grounded Theory. A Practical Guide through Qualitative Analysis*. London: Sage.
- Cicourel, A. V. (1982). Interviews, Surveys, and the Problem of Ecological Validity. *The American Sociologist*, 17(1), 11–20.
- Creswell, J. W., & Clark, V. L. P. (2018). *Designing and Conducting Mixed Methods Research*. Londong: Sage.
- Druckman, J. N., Green, D. P., Kuklinski, J. H., & Lupia, A. (2006). The Growth and Development of Experimental Research in Political Science. *American Political Science Review*, 100(4), 627–635.
- Eskelinen, L., & Caswell, D. (2006). Comparison of Social Work Practice in Teams Using a Video Vignette Technique in a Multi-method Design. *Qualitative Social Work*, 5(4), 489–503.
- Fetters, M. D., & Freshwater, D. (2015). The 1 + 1 = 3 Integration Challenge. *Journal of Mixed Methods Research*, 9(2), 115–117.

- Fetters, M.D., & Molina-Azorin, J. F. (2020). Utilizing a Mixed Methods Approach for Conducting Interventional Evaluations. *Journal of Mixed Methods Research*, 14(2), 131–144.
- Field, Andy. (2013) *Discovering Statistics Using IBM SPSS Statistics* (4<sup>th</sup> edition). London: SAGE Publications.
- Finch, J. (1987). The vignette technique in survey research. *Sociology*, 21(1), 105–114.
- Greene, J. C. (2008). Is mixed methods social inquiry a distinctive methodology ? *Journal of Mixed Methods Research*, 2(1), 7–22.
- Harrits, G. S. (2019). Stereotypes in Context: How and When Do Street-Level Bureaucrats Use Class Stereotypes? *Public Administration Review*, 79(1), 93–103.
- Harrits, G. S., & Møller, M. Ø. (2014). Prevention at the Front Line: How home nurses, pedagogues, and teachers transform public worry into decisions on special efforts. *Public Management Review*, 16(4), 447–480.
- Harrits, G. S., & Møller, M. Ø. (2016). *Forebyggelse og bekymring i professionel praksis*. København: Hans Reitzels Forlag.
- Holland, P. W. (1986). Statistics and causal inference: Rejoinder. *Journal of the American Statistical Association*, 81(396), 968–970.
- Hughes, R., & Huby, M. (2004). The construction and interpretation of vignettes in social research. *Social Work & Social Sciences Review*, 11(1), 36–51.
- Ilsvard, S., & Møller, M. Ø. (2015). General Practitioners' Discretion of Preventive Needs. *Professions and Professionalism*, 5(3), 1–16.
- Jenkins, N., Bloor, M. J., Fischer, J., Berney, L., & Neale, J. (2010). Putting it in context: the use of vignettes in qualitative interviewing. *Qualitative Research*, 10(2), 175–

198.

- Jenkins, N., Keyes, S., & Strange, L. (2015). Creating Vignettes of Early Onset Dementia: An Exercise in Public Sociology. *Sociology*, 1–16.
- Johnson, R. B., Russo, F., & Schoonenboom, J. (2019). Causation in Mixed Methods Research: The Meeting of Philosophy, Science, and Practice. *Journal of Mixed Methods Research* 13(2): 143–162.
- Johnson, R. E., Grove, A. L., & Clarke, A. (2019). Pillar Integration Process: A Joint Display Technique to Integrate Data in Mixed Methods Research. *Journal of Mixed Methods Research*. 13(3): 301–320.
- Kvale, S., & Brinkmann, S. (2009). *Interview. Introduktion til et håndværk* (Second). København: Hans Reitzels Forlag.
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things*. Chicago: Chicago University Press.
- Maxwell, J. A. (2004). Using qualitative Methods for causal explanation. *Field Methods*, 16(3), 243–264.
- Maxwell, J. A. (2012). *A Realist Approach for Qualitative Research*. Thousand Oaks: Sage.
- Maxwell, J. A., & Mittapalli, K. (2010). Realism as a Stance for Mixed Methods Research. In A. Tashakkori & C. Teddlie (Eds.), *Mixed Methods in Social & Behavioral Research* (p. 145168). London: Sage.
- Maxwell, J. A., & Sandlow, L. J. (1986). Combining ethnographic and experimental methods in evaluation research: A case study. In D. M. Fetterman & M. A. Pitman (Eds.), *Educational evaluation: Ethnography in theory, practice, and politics*.

Thousand Oaks: Sage.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative Data Analysis. A Methods Sourcebook* (third edition). London: Sage.

Mohr, L. B. (1999). The Qualitative Method of Impact Analysis. *American Journal of Evaluation*, 20(1), 69–84.

Mosholm, E. and Fetters, M.D. (2017). Conceptual models to guide integration during analysis in convergent mixed methods studies. *Methodological Innovations*, 10(2): 1–11.

Møller, M. Ø. (2009). *Solidarity and Categorization: Solidarity Perceptions and Categorization Practices among Danish Social Workers*. Aarhus: Politica.

Møller, M. Ø. (2016). “She isn’t Someone I Associate with Pension”— a Vignette Study of Professional Reasoning. *Professions & Professionalism*, 6(1), 1–20.

O’Cathain, A. (2018). *A practical guide to using qualitative research within randomized controlled trials*. Oxford: Oxford University Press.

Onwuegbuzie, A. J. and Johnson, R. B. (2006). The Validity Issue in Mixed Research, *Research in the Schools*, 13(1): 48-63.

Plano Clark, V. L., Schumacher, K., West, C., Edrington, J., Dunn, L. B., Harzstark, A., ... Miaskowski, C. (2013). Practices for Embedding an Interpretive Qualitative Approach Within a Randomized Clinical Trial. *Journal of Mixed Methods Research*, 7(3), 219–242.

Robinson, S., & Mendelson, A. L. (2012). A Qualitative Experiment: Research on Mediated Meaning Construction Using a Hybrid Approach. *Journal of Mixed Methods Research*, 6(4), 332–347.

- Rossi, P. H., & Nock, S. L. (1982). *Measuring Social Judgments: The Factorial Survey Approach*. London: Sage.
- Saldana, J. (2013). *The Coding Manual for Qualitative Researchers*. (2., Ed.) (second). London: Sage.
- Sandelowski, M. (1996). Using qualitative methods in intervention studies. *Research in Nursing & Health*, 19(4), 359–364.
- Schaffer, F. C. (2006). Ordinary Language Interviewing. In D. Yanow & P. Schwartz-Shea (Eds.), *Interpretation and Method. Empirical Research Methods and the Interpretive Turn*. Armonk: M.E.Sharpe.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin Company.
- Steiner, Atzmüller and Su (2016). Designing Valid and Reliable Vignette Experiments for Survey Research: A Case Study on the Fair Gender Income Gap. *Journal of Journal of Methods and Measurement in the Social Sciences* 7(2): 52-94.
- Soss, J. (2006). Talking Our Way to Meaningful Explanations - A Practice Centered View of Interviewing for Interpretive Methods. In D. Yanow & P. Schwartz-Shea (Eds.), *Interpretation and Method. Empirical Research Methods and the Interpretive Turn* (pp. 127–149). New York: M.E.Sharpe.
- Tashakkori, A., & Teddlie, C. (2010). *SAGE handbook of mixed methods in social & behavioral research*. London: SAGE.
- Tavory, I., & Timmermans, S. (2014). *Abductive Analysis: Theorizing Qualitative Inquiry*. Chicago: University of Chicago Press.

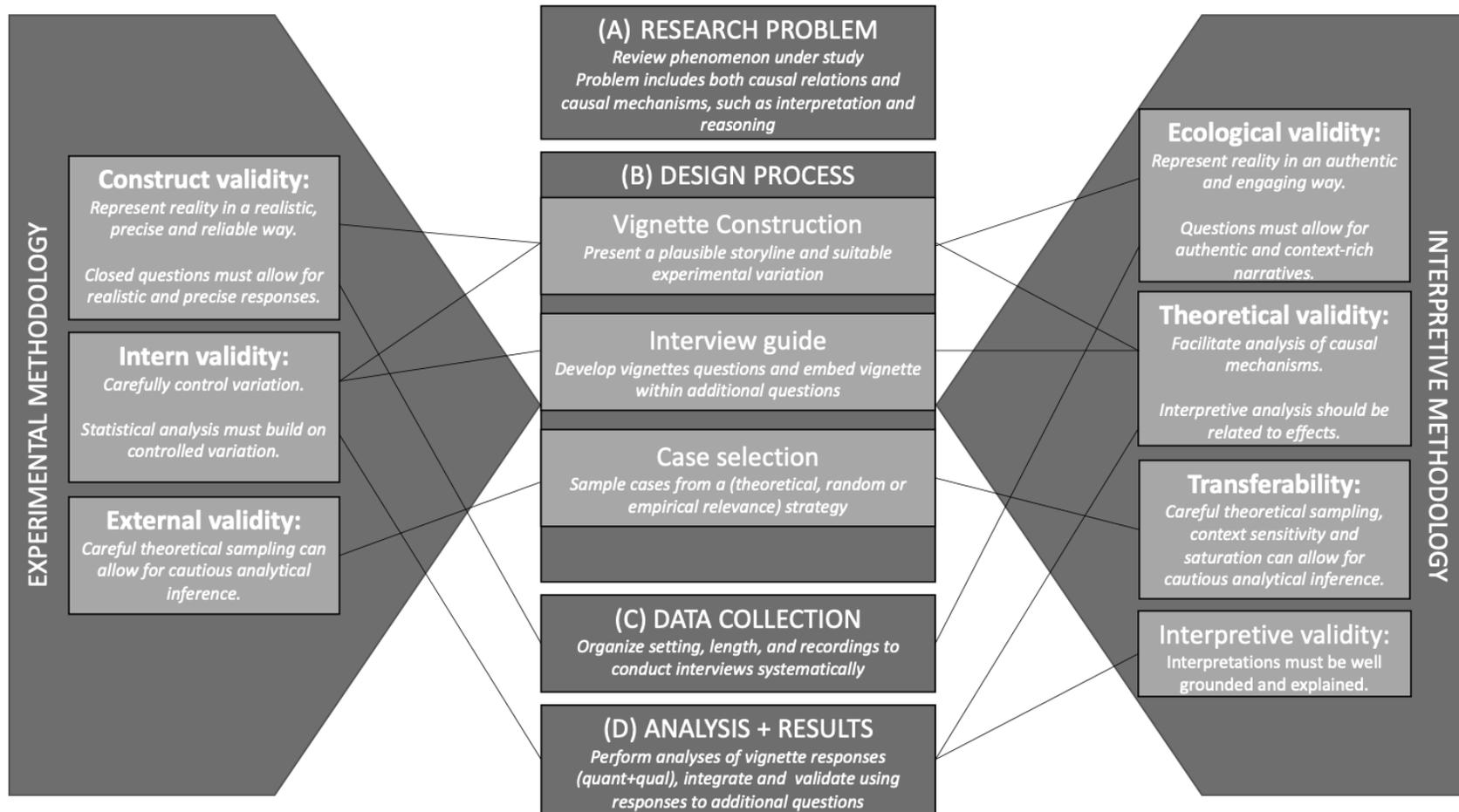
Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in the social and behavioural sciences. In C. Teddlie & A. Tashakkori (Eds.), *Handbook of mixed methods in social and behavioural research* (pp. 3–50). London: Sage.

**Table 1: Vignettes used in the study of professional judgement**

<p>A (Base -line = Middle class)</p>	<p>Imagine a boy in your daycare centre. You are his contact person. His name is <b>Mads</b>, he is four years old, and he started in the daycare only five months ago, when his family moved to town. The mother is 33 and a <b>preschool teacher</b>, the father is 40 and a <b>teacher</b>, and they both <b>work at the local school</b>. Mads is the mother’s only child, while the father has two children from a previous marriage. Mads has had a fine start in the daycare, and it seems as if he has adapted quickly. For example, he says goodbye to his mother without incident in the morning. Within the last month, however, you have noticed that Mads has been in some difficult conflicts with the other children. The conflicts have been of differing character, but typically they center around who gets to decide and on problems sharing toys. In particular, the problem has been that Mads is very aggressive. He shouts very loudly, and in a couple of instances he has hit another boy. You have also experienced that Mads will not eat together with the other children, and that he goes out into the locker room when the lunch is served. The parents are very upset about the situation, and they can tell that Mads has difficulties making friends. The mother tells you that the family moved to the area only five months ago. As a consequence, she has been separated from her mother and older sister, and the family has not yet made many friends in the area. While preparing for a routine meeting with the parents (which is usually placed half a year after the child’s start in the daycare), you have discussed with your colleagues that Mads might have some motor skill problems. For example, he has difficulties catching and throwing a ball. His parents have also told the manager of the daycare that they have tried getting Mads to play football in the afternoon, so that he could get to know some other children. But Mads doesn’t want to play football, and therefore they have stopped.</p>
<p>B/C (Upper/L ower class)</p>	<p>Imagine a boy in your daycare centre. You are his contact person. His name is <b>Vitus / Mike</b>, he is three and a half years old and he started at the daycare half a year ago. The mother is 36 and an <b>MD / nursing assistant</b>, and the father is 39 and an <b>engineer / truck driver</b>. The mother works in a <b>hospital nearby / is temporarily unemployed</b>, and the father is <b>employed in a private company / on sick leave with a back injury</b>. Vitus / Mike has had a rough start at the daycare centre. He has difficulties saying goodbye to his mother in the morning, and she seldom has the time to stay very long. You experience Vitus / Mike as being rather whiny during the day. He is a quiet child, keeping a bit to himself, and it is difficult for him to establish contact with the other children. He also has difficulties concentrating on collective activities, and he often leaves the room when a collective activity begins. During the six months that Vitus/Mike has been at the daycare, he has refused several times to eat the food served in the collective lunch arrangement, which has been democratically decided by the parents in the daycare. The parents have asked if they can bring their own lunch package, <b>because they find the food served in the daycare to be of low quality, and they think that this is why Vitus will not eat it / because they think that it is too expensive now that Mike does not eat the food</b>. While preparing for a routine meeting with his parents (which usually takes place half a year after the child’s start in the daycare), you have discussed with your colleagues that Vitus/Mike may have some difficulties with language development, especially with pronunciation. His parents have told the manager of the daycare that <b>they are worried about Vitus’ language, and that their old daycare centre did not do nearly enough to support the children’s language development / that they can’t see anything wrong with Mike’s language</b>.</p>

Note: Bold text: Different in all three vignettes. The vignettes presented were used in interviews with childcare workers (see Authors 2014 for other vignettes).

**Figure 1: Procedural overview of a qualitative vignette experiment**



**Table 2: Methodological steps, validity issues and integration in the qualitative vignette experiment**

Core elements	Guidelines	Ontological and epistemological foundation and validity criteria	Issues of integration
Research problem	Conduct theoretical and empirical review on the phenomenon under study.	Critical realism: Social reality includes both causal relations and causal mechanisms, such as interpretation and reasoning.	Integrated ontology. Research problem must integrate causal relations and causal mechanisms.
Vignette construction	Present information in realistic and authentic story-lines. Design variation in baseline and experiment vignettes, with variation in independent variable and similarity in situation/problem. Allow for both recognition and interpretation.	Construct validity: Represent reality in a realistic way. Ecological validity: Represent reality in an authentic and engaging way. Double perspective on causality: variation-oriented and process-oriented. Intern validity: Carefully control variation. Theoretical validity: Facilitate analysis of causal mechanisms.	Possible tension between construct and ecological validity. Important to find balance between technically correct and engaging vignettes. Complementarity between internal and theoretical validity.
Case selection and sampling	Use a theoretical sampling strategy. Consider typical and varied contexts and characteristics. Consider suitability of interviewees. Consider possible comparisons between contexts.	External validity: Careful theoretical sampling can allow for cautious analytical inference. Transferability: Careful theoretical sampling, context sensitivity and saturation can allow for cautious analytical inference.	Convergence between external validity and transferability.
Data collection	Embed qualitative vignette experiment in semi-structured interview, focus groups or observation/ethnographic studies. Ask both open-ended and closed questions to vignettes to capture both causal effects and causal mechanisms. Consider asking additional questions for validating analyses.	Construct validity: Closed questions must allow for realistic and precise responses. Ecological validity: Open-ended must allow for authentic and context-rich narratives.	Possible tension between construct and ecological validity. Important to find balance between closed and open questions.
Analytical strategy	Perform statistical analysis of causal effects comparing responses to vignettes. Perform interpretive analysis of interviewees' interpretations and reasoning on vignettes, and other interview data. Organize integration of results by linking causal effects and causal mechanisms.	Interpretive validity: Interpretations must be well grounded and explained. Intern validity: Statistical analysis must build on controlled variation. Theoretical validity: Interpretive analysis of causal mechanisms should be related to effects. Comparisons between segments of data can be used to explore impact of contextual conditions.	Complementary between interpretive, internal and theoretical validity. Analyses of causal effects and causal mechanisms can be used to corroborate each other, and both must concur to achieve causal inference.

**Table 3: Comparison of different mixed methods experimental designs**

	<b>Mixed methods intervention design</b>	<b>Qualitative experiments</b>	<b>Qualitative vignette experiments</b>
Examples	O’Cathain (2018) Creswell and Clark (2018)	Robinson and Mendelsohn (2012)	Harrits, 2019; Harrits & Møller, 2014, 2016; IIsvard & Møller, 2015; Møller, 2009, 2016
Research design	Primarily experimental designs.	Primarily qualitative logic. Randomized division of interviewees in two groups.	Integrating experimental and qualitative-interpretive designs.
Data collection	Quantitative and qualitative data collection, often collected separately.	Qualitative data collection (focus groups).	Qualitative data collection (semi-structured interviews) using vignettes. Use of close-ended (quantitative) question to collect responses to vignettes.
Data analysis	Quantitative and qualitative data analysis, often done separately.	Qualitative data analysis.	Qualitative-interpretive as well as quantitative data analysis, performed complementary.
Level of integration	Depends on the study. Can be highly integrated or less integrated. Integration is mainly done at the level of data results. The dominance and complementarity of qualitative or quantitative component varies.	Low. Qualitative component is dominant.	High. Qualitative-interpretative and experimental components are equal, complementary and integrated in all steps of the research process.