methods, data, analyses | Vol. 16(2), 2022, pp. 201-234 DOI: 10.12758/mda.2022.01

Fear of the Dark? A Systematic Comparison of Written Vignettes and Photo Vignettes in a Factorial Survey Experiment on Fear of Crime

Stefanie Eifler[1] *& Knut Petzold*[2]

[1] *Catholic University of Eichstätt-Ingolstadt*

[2] *Hochschule Zittau/Görlitz – University of Applied Sciences*

Abstract

Measuring attitudes with vignettes is frequently based on the assumption that the presented context information facilitates a better imagination of topics under study, serving for more valid responses as compared to more usual questionnaire methods. In this study, we focus on the presentation format of vignettes and assume that, in particular, the presentation of photo vignettes facilitates a close context approximation, hereby taking fear of crime from the perspective of broken windows theory as an example of use. A split ballot experiment within the framework of a cross-sectional online survey introduced a variation of the presentation format of a factorial survey experiment and allowed for measuring the difference between using either written vignettes or photo vignettes. While the split ballot experiment used a between-subjects design, each factorial survey experiment used a within-subjects design. The reported level of feelings of unsafety serves as a measure of fear of crime. Results show that, first, all dimensions of the factorial surveys predicted the respective level of fear of crime in both presentation formats, in the direction expected by broken windows theory. Measurement error seems slightly reduced within written vignettes. Second, presentation format-specific differences were observed for dimensions representing physical features of the setting, such as darkness, only, thereby slightly favouring photo vignettes. We finally discuss methodological implications of these results.

Keywords: Factorial Survey Experiments, Presentation format, Written vignettes, Photo vignettes, Broken Windows Theory, Fear of Crime

Our study focuses on factorial survey experiments that are used to measure normative judgements, subjective beliefs or behavioural intentions (cf. Beck & Opp, 2001; Jasso, 2006) through respondent's answers to a number of brief descriptions of hypothetical situations, persons or objects called vignettes (Auspurg & Hinz, 2015; Jasso, 2006; Rossi, 1979; Rossi & Anderson, 1982). Due to their supposed advantages, vignettes have been increasingly applied in surveys (cf. Auspurg & Hinz, 2015; Liebig et al., 2015; Mutz, 2011; Wallander, 2009). First, because of the systematic variation of several features or dimensions, the relative weight of these dimensions with regard to the responses can be determined. Second, effects of a self-selection driven by the respondents' interests can be neutralised through randomisation. Variation and randomisation are also features of a random experiment; thus, third, vignette analyses allow for a causal interpretation of the effects of situational features or vignette dimensions. It is usually stated that, fourth, vignettes comprise more detailed and more concrete information on the phenomena meant and, therefore, facilitate a more standardised imagination of the situation across respondents and less use of general heuristic principles by respondents, hence inducing them to report their *true* opinions (e.g., Shamon et al., 2019).

One main argument for using vignettes is that the presentation of information on the situational context helps to achieve a close proximity to the reality of everyday life. Accordingly, several authors have pointed out that vignettes allow to mirror situations of everyday experience and, thus, to bring individual answers in line with real-life judgement formation or decision-making (cf. Alexander & Becker, 1978; Armacost et al., 1991; Finch, 1987). However, the presentation form of vignettes shapes the results of vignette-based measurements, for example, a detailed or sparse presentation (e.g., Eifler & Petzold, 2014) or the presentation of vignettes in running text or tabular format (Sauer et al., 2020; Shamon et al., 2019). So far, it is still an open question whether factorial survey experiments actually help to improve measurement quality of normative judgments, subjective beliefs or behavioural intentions.

In principle, there are different formats of presenting vignettes within the framework of a survey: the situation can either be described in a written form or presented by visual stimuli, for example, by videos, photos or pictures. While the majority of studies apply written vignettes (Wallander, 2009), some studies use solely video vignettes (Krysan et al., 2009) or solely photo vignettes (Golden III

Acknowledgements

The authors thank two anonymous reviewers for helpful comments to a previous version of this manuscript. Both authors contributed equally to this work.

Direct correspondence to

Stefanie Eifler, Catholic University of Eichstätt-Ingolstadt, Ostenstraße 26, 85072 Eichstätt
E-mail: stefanie.eifler@ku.de

et al., 2001). Another study combines different written and photo information in vignettes (Havekes et al., 2013). Beyond scarce applications, only two studies compared presentation format differences systematically (Eifler, 2007; Rashotte, 2003). Both authors found systematic differences between verbal and visual presentation formats, but they also stated that much more research is required to clearly determine differing results for various types of stimuli.

Against this background, our study is particularly devoted to the presentation format of vignettes. It is a largely open question whether or not written and photo vignettes lead to corresponding or diverging responses and whether or not the effects of situational dimensions in a factorial survey experiment depend upon the presentation format used.

To fill this research gap, we start from psychological approaches which state that different processes of recognition, information processing and remembering verbal and visual information apply. In particular, we apply the *Dual Coding Theory* (DCT) suggested by Paivio (1979) and Sadoski and Paivio (2013) to the systematic analysis of presentation format differences concerning the use of written vignettes or photo vignettes. Taking the example of the broken windows theory, which we employ for the prediction of fear of crime (Keuschnigg & Wolbring, 2015; Keizer et al., 2014; Kelling & Coles, 1996; Wilson & Kelling, 1982), we use vignettes that describe or visualise varying situations of everyday experience. We assign respondents randomly to one of the two presentation formats. By doing so, we demonstrate both presentation format correspondence and presentation format differences.

In the next section, we present a model concerning the role of the presentation format of vignettes and derive testable hypotheses. We analyse the assumptions empirically on the basis of a split ballot experiment among the population of students from a German university. Finally, we discuss our findings critically and consider methodical implications. Overall, our study demonstrates both validity aspects of factorial survey experiments using different presentation formats of vignettes and theoretically predictable differences between these presentation formats.

A Systematic Comparison of Written Vignettes and Photo Vignettes

Just as in the case of answering survey items, we can use the general model of the response process in surveys by Tourangeau (1984) and Tourangeau et al. (2000), in order to delineate the process of responding to vignettes. According to this model, a respondent who is asked a question first has to interpret the question's content (interpretation), subsequently has to retrieve information from the memory

(retrieval), form an opinion (judgement) and then bring the answer into line with the predefined response format (response selection). Transferred to the measurement with vignettes, a subject first has to interpret the situation and the question presented and has to retrieve information from his/her memory referring to it, before he/she can form an opinion and provide an answer.

With regard to factorial survey experiments, several authors have emphasised the idea that using vignettes facilitates a standardised presentation of information about the situations under study (Auspurg & Hinz, 2015; Mutz, 2011; Jasso, 2006; Rossi, 1979; Rossi & Anderson, 1982). In the eyes of Shamon et al. (2019), this leads to a more unified retrieval – called "information intake" (p. 4) by the authors – of relevant information from the memory across subjects. Accordingly, the retrieval stage of the response process is assumed to be characterised by a higher level of interindividual comparability in factorial survey experiments as opposed to survey items.

While the majority of studies apply written vignettes (Wallander, 2009), several researchers have suggested to use visual stimuli within the framework of factorial survey experiments because visual stimuli like video clips, photos or pictures allow for a more natural representation of the situations under study, indicating a clear preference for video vignettes (Caro et al., 2012a, 2012b; Dinora et al., 2020; Golden III et al., 2001; Goyal et al., 2017; Havekes et al., 2013; Hughes & Huby, 2004; Krysan et al., 2009; O'Donnell et al., 2007; Rashotte, 2003).

To our knowledge, only two studies compared observed responses to both written and visual stimuli (Eifler, 2007; Rashotte, 2003). Rashotte (2003) examines what information people receive and use in forming effective responses when observing written versus visual stimuli on social events. In her study, readers of written descriptions of events and viewers of videotapes use different pieces of information in forming impressions based on stimuli type (Rashotte, 2003). While visual cues of nonverbal behaviours appear clearer in videotapes and viewers need less information than readers to get an impression of it, viewers use the same information as readers to evaluate object-persons themselves. The results are consistent with the idea that visual stimuli provide more information and allow for a richer picture of social events. The assumption that visual presentations provide more accurate representations of situations and, thus, evoke more valid responses is also tested by Eifler (2007). Behavioural observations and vignette analyses with visual and verbal material were carried out with regard to three forms of deviant behaviour in everyday life, showing that frequencies of (intended) deviant behaviour were related to the presentation formats. Written vignettes lead to an overestimation of the frequencies of crossing a red traffic light and to an underestimation of the frequencies of cycling through a red traffic light. While deviant behaviour to ignore a 'lost letter' is overestimated by all respondents, the degree of overestima-

tion is smaller in the face of a visual vignette. In both studies, it becomes clear that the role of the presentation format needs more clarification.

Presentation Format and Information Processing

So far, little is known about potential differences between written vignettes and vignettes presenting pictures or photos. In particular, there are – to our knowledge – no systematic theory-guided approaches that would help to explain *why* visual stimuli should be superior to the usual verbal presentations of vignettes. Therefore, we will introduce theoretical ideas from cognitive psychology in order to explain format differences in factorial survey experiments.

We, thereby, start from psychological approaches which state that different processes of recognition, information processing and remembering verbal and visual information apply. In particular, we refer to the DCT suggested by Paivio (1979) and Sadoski and Paivio (2013), which posits the idea that verbal and visual information is coded differently in the human brain.

This approach starts from the idea that there are two coding systems in human memory: one responsible for language or verbal information and the other responsible for pictures or non-verbal information: "In DCT, the linguistic coding system is referred to simply as the verbal code or system, and the nonverbal coding system is often referred to as the imagery code or system because its main functions include the analysis of external scenes and the generation of internal mental images" (Sadoski & Paivio, 2013, p. 29). Both systems overlap and can operate simultaneously in principle. Processing verbal and/or visual information generates "internal mental images" (Sadoski & Paivio, 2013, p. 29) which represent information about situations. It is assumed that mental images of situations match experiences with the same situations (Kosslyn & Pomerantz, 1977; Kosslyn, 1981). Concerning the prediction of systematic presentation format differences, this central assumption would require a specification of particular features of a hypothetical situation with regard to using either written or photo vignettes.

What is crucial with regard to these mental images is that written and visual information about situations is processed by both systems but in a different way: Written information is processed *sequentially* (i.e., by the verbal coding system first and by the non-verbal coding system subsequently), and visual information is processed *simultaneously* by both coding systems at a time (Paivio, 1979; Sadoski & Paivio, 2013). Because of the sequential processing of verbal information, written vignettes can elicit diverging encoding processes by readers, thus leading to diverging visualisations in memory between subjects. In studies on learning and memory, the thesis that verbal information is visualised by readers was supported (Kosslyn, 1981). Because of the simultaneous processing of visual information, photo vignettes facilitate a standardised perception of the concrete situation without any

loop way, thus leading to corresponding mental images of the presented situations between subjects. Correspondingly, Hanna and Loftus (1993) pointed at qualitative differences between verbal and visual information processing. In addition, Harper (2002) emphasised that the parts of the brain that process visual information are evolutionarily older than the parts that process verbal information; thus, images might evoke deeper elements of human consciousness that do words.

While there is much research activity concerning functions of visual memory, like remembering or recalling natural scenes or – more generally – everyday experience (Brockmole, 2009; Findlay & Gilchrist, 2003; Luck & Hollingworth, 2008), there are not more than a handful of studies that are devoted to a systematic comparison of the cognitive processes involved in remembering and recalling both verbal and visual information. Overall, neurophysiological studies have shown that visual stimuli are remembered and recalled more easily than verbal stimuli (Bower, 1970; Shepard, 1967). Correspondingly, a systematic comparison between visual and verbal information revealed that photos are remembered better than words, which was explained in the following way: "(…) pictures contain distinctive cues which make them more discriminable than their labels and this discriminability enhances memory for pictures compared to their labels" (Jenkins et al., 1967, p. 306). McCloud (1994) summarised these differences and stated that verbal information is *perceived,* while visual information is *received*.

From the theoretical considerations presented so far, we conclude that differences in processing verbal and visual information exist. Verbal information requires more extensive information processing by a reader and more background knowledge, whereas visual information presents the information directly. Following this train of thought, photos can be considered to mirror real life (Manghani, 2013, Rose, 2012). According to Barthes (1977, p. 17), a photo is "(…) not the reality but at least it is its perfect *analogon* and it is exactly this analogical perfection which, to common sense, defines the photograph (…): *it is a message without a code*". Accordingly, photos are a concrete point of reference for all who are confronted with them (Collier Jr., 1957; Collier & Collier, 1986). In a similar way, other authors highlight the advantages of presenting photos: "Showing many things at once is a tremendous strength that reflects the all-at-once nature of lived experiences – a reality that is often impossible to communicate through linear textual narratives" (Marion & Crowder, 2013, p. 31).

Therefore, with regard to factorial survey experiments, photo vignettes not only allow for a more realistic presentation of the situations under study but also for evoking the feeling of experiencing the particular situation. While written vignettes facilitate a sequential presentation of information in the form of short stories, photo vignettes present the information simultaneously in the form of pictures, thereby activating visual and verbal mental representations and leading to emotional arousal at the same time.

Explanatory Model and Hypothesis

It follows from the above explicated theoretical ideas, in particular from DCT (Paivio, 1979; Sadoski & Paivio, 2013), that both presentation formats, written vignettes and photo vignettes, lead to mental images that include a visualisation of the presented situation. Therefore, we would expect mostly corresponding results between both presentation formats in a factorial survey experiment with regard to the direction of effects of situational dimensions. In principle, both written vignettes and photo vignettes should facilitate a representation of the same higher order constructs. Nevertheless, as for the simultaneous information processing of visual information, we would expect advantages of using photo vignettes with regard to the strength of effects of situational dimensions.

In order to test these assumptions, we took the broken windows theory (Kelling & Coles, 1996; Keizer et al., 2014; Keuschnigg & Wolbring, 2015; Lewis & Salem, 1986; Skogan, 1990; Wilson & Kelling, 1982) as an example of use. Amongst other topics, this approach has been applied to the analysis of fear of crime. The theory – also referred to as the *Disorder Model* of fear of crime – specifies features which are assumed to be perceived as cues of normative compliance in urban neighbourhoods. Because these features can be both described and pictured, the approach seems particularly suited for a systematic comparison of written vignettes and photo vignettes within the framework of a factorial survey experiment. In addition, visual methods have been used in the analysis of fear of crime because of their feasibility for presenting the context of crime-related cognitions and emotions (Vanderveen, 2018). We tie in with this tradition in principle and extend it to the systematic comparison of presentation format differences in factorial survey experiments.

Within the framework of the disorder model, one refers to the features of urban neighbourhoods that are called "signs of incivility" (Hunter, 1978). These are signs of non-compliance with behavioural norms like littering, graffiti on facades, destruction and decay of buildings or unsupervised youth (Hunter, 1978). In particular, physical signs of disorder, like plaster crumbling of the wall, are distinguished from social signs of disorder, like teenagers hanging around and drinking alcohol (Hunter, 1978; Skogan, 1978; Taylor, 1999). It is assumed that signs of incivility serve as cues for the likelihood of norm enforcement in specific situations. They indicate a failure of informal control processes and call forth perceived victimisation risks, which, in turn, are reflected in higher levels of fear of crime.

For our systematic comparison of written and photo vignettes, we took signs of physical and social disorder, and introduced them as dimensions into factorial survey experiments. In both formats, the situations presented in the vignettes were systematically varied regarding the same dimensions: observability of place, physical decay and littering, unsupervised youth, adult passers-by, video surveillance,

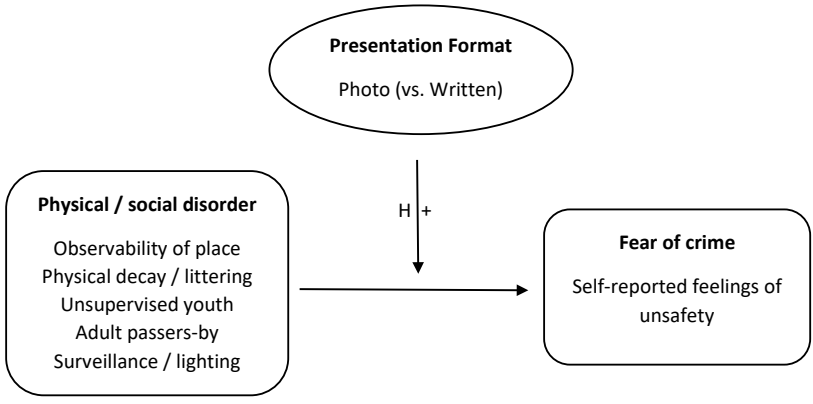

Figure 1 Underlying theoretical model

and lighting. If a setting exhibits physical and social features that indicate a high level of disorder, the level of fear of crime should be increased. We assume that, first, the direction of these influences will be comparable across the presentation formats of the factorial surveys. However, driven by different information processing, we assume stronger effects for the presentation in photo vignettes. Figure 1 shows the vignette dimensions and the hypothesis considered in the model. We suppose that the influences of physical and social features of the setting will be stronger in factorial surveys based on visual stimuli.

Hypothesis:

The effects of physical and social features of the setting that indicate a high level of disorder are stronger in a factorial survey employing photo vignettes compared to a factorial survey employing written vignettes.

Methods

Our empirical examination of presentation format differences in factorial surveys took place within the framework of a split ballot experiment (Benson, 1941) that was part of a survey on attitudes towards safety in public places. We conducted this survey at the Martin-Luther-University Halle-Wittenberg, Germany, in 2014 (Schwarzbach & Eifler, 2020).

Procedure, Data

We set up the study as a web survey using Lime Survey. The survey period was between 9 December 2013 and 17 January 2014. We invited respondents to participate in the survey by an email, providing them with a link to the online questionnaire. The survey followed the "Tailored Design Method" (TDM) (Dillman et al., 2009).

Sample

The survey included the full population of all enrolled students of the Martin-Luther-University Halle-Wittenberg, Germany. We administered the invitation to take part in the survey with the help of the registrar's office, by sending an email to the full population of around N = 20,000 students. An overall number of n = 1,149 students completed the survey. Their mean age was 23.8 years, and 65.8% of them were female. Unfortunately, comparisons with the full population were not possible because no information about age and gender was provided for all enrolled students.

Operationalisation, Measurement

In the following, we describe the experimental design that we used to analyse the effects of using varying presentation formats in factorial surveys. We also describe the design of these factorial surveys. A complete project documentation including the questionnaires is available online for both transparency and replication purposes (Schwarzbach & Eifler, 2020).

Independent Variables

In our study, the subjects responded to a factorial survey using either written vignettes or photo vignettes (i.e., we used a between-subjects design for the split ballot experiment). We randomly assigned each subject to one of the two presentation formats.

To analyse the difference between the presentation formats, we used two factorial surveys based on the same $2^4 \cdot 3$-within-subjects design. The factorial surveys referred to signed of social and physical disorder in urban neighbourhoods that had been used in previous studies (Piquero, 1999; Taylor, 1999). Table 1 illustrates the experimental design.

We pictured the *observability of the place* as either a wide square or a narrow pedestrian underpass. Facades covered with graffiti, empty beer bottles and other garbage around (high physical disorder) versus a clean and tidy environment (low physical disorder) represented *physical decay and littering.* We indicated

Table 1 Experimental Design

	Dimensions	Levels 1	2	3
1	Observability of place	Wide square	Pedestrian underpass	
2	Physical decay, littering	No	Yes	
3	Unsupervised youth	Couple goes for a walk	Teenagers hanging around	
4	Adult passers-by	Passers	No passers	
5	Surveillance / lighting	Bright situation through lighting	Gloomy situation, but video surveillance	Gloomy situation without surveillance

Note: Cartesian product of dimensions and levels 2x2x2x2x3 = 48 unique situations

unsupervised youth by teenagers hanging around (high social disorder) versus a young couple going for a walk (low social disorder). The presence (high social control) or absence (low social control) of *adult passers-by* referred to the respective dimension. *Surveillance and lighting* were part of one dimension including three levels: the presence of CCTV in a gloomy setting (video surveillance), a bright setting through the presence of street lighting (lighting) and the absence of video surveillance and street lighting in a gloomy setting (gloomy setting). We decided to use a dimension comprising three levels because a full combination of lighting and CCTV seemed inappropriate, as CCTV requires sufficient lighting. A group of experts (n = 15), composed of graduate students from the social sciences with a special training in the factorial survey approach, rated the correspondence between written vignettes and photo vignettes to facilitate a test of the presentation format.

From a full combination of the above explained dimensions and their levels, a universe of 48 vignettes was obtained. Given this large number of vignettes, we decided to present vignette sets to our subjects. Thereby, we assured to facilitate an estimation of all main effects of the vignette dimensions. Following the recommendations given by previous methodological studies on factorial surveys (for an overview, see Auspurg & Hinz, 2015), we used six sets of eight vignettes each. We presented one instruction to the respondents for both written vignettes and photo vignettes: "In the following, we ask you to judge a number of situations. We are interested in your feelings of safety or unsafety in these situations. Please put yourself in these situations:" Figure 2 shows an example of two written vignettes and their respective photo vignette counterparts.

Example 1: Photo vignette

Example 1: Written vignette

You are on a wide square. The place is only dimly lit but you will see a sign saying "This area is under video surveillance". The area looks neat and tidy. You realise two teenagers who hang around and drink alcohol. There are some additional adults nearby.

Example 2: Photo vignette

Example 2: Written vignette

You are on a wide square. The place is brightly lit. The area looks neat and well kept. You see a young couple going for a walk. There are some additional adults nearby.

Figure 2 Examples of photo vignettes and written vignettes

Table 2 Quality of randomisation and variation, sample: estimation model

	Total		Written Vignettes		Photo Vignettes	
	N	Percent / M	N	Percent / M	N	Percent / M
Vignette treatments						
Observability of place						
Wide square	4046	49.99	1921	50.34	2143	49.69
Pedestrian underpass	4065	50.01	1895	49.66	2170	50.31
			$\chi^2 = 0.3460$, p = 0.556			
Physical decay, littering						
No	4091	50.38	1952	51.15	2139	49.59
Yes	4038	49.62	1864	48.85	2174	50.41
			$\chi^2 = 1.9679$, p = 0.161			
Unsupervised youth						
Couple goes for a walk	4096	50.39	1924	50.42	2172	50.36
Teenagers hanging around	4033	49.61	1892	49.58	2141	49.64
			$\chi^2 = 0.0029$, p = 0.957			
Adult passers-by						
No	4095	50.38	1919	50.29	2176	50.45
Yes	4034	49.62	1897	49.71	2137	49.55
			$\chi^2 = 0.0217$, p = 0.883			
Surveillance / lighting						
Bright situation, lighting	2727	33.55	1260	33.02	1467	34.01
Gloomy situation, video surv.	2719	33.45	1275	33.41	1444	33.48
Gloomy situation, no surv.	2638	33.01	1281	33.57	1402	33.51
			$\chi^2 = 1.2927$, p = 0.524			
Questionnaire Characteristics						
Vignette set						
1	1320	16.24	587	15.38	733	17.00
2	1636	20.13	732	19.18	904	20.96
3	1300	15.99	623	16.33	677	15.70
4	1356	16.68	651	17.06	705	16.35
5	1169	14.38	546	14.31	623	14.44
6	1348	16.58	677	17.74	671	15.56
			$\chi^2 = 1.6331$, p = 0.897			
Presentation format						
Written vignettes	3816	46.94				
Photo vignettes	4313	53.06				
Total						
$N_{vignettes}$	8129	100.0	3816	100.0	4313	100.0
$N_{probands}$	1019	100.0	479	100.0	540	100.0

Note: Test statistics for age, gender, deck at probands level.

Table 3 Parallelisation of experimental groups, sample: estimation model

	Total		Written Vignettes		Photo Vignettes	
	N	Percent / M	N	Percent / M	N	Percent / M
Respondents' Characteristics						
Gender						
Female	665	65.78	309	64.92	356	66.54
Male	346	34.22	167	35.08	179	33.46
			$\chi^2 = 0.2959$, $p = 0.586$			
Age		23.82		23.90		23.75
			$t = 0.5715$, $p = 0.568$			
Partner						
Yes	571	57.56	286	59.71	285	52.88
No	421	42.44	180	37.58	241	44.71
			$\chi^2 = 5.2303$, $p = 0.022$			
$N_{respondents}$	1019	100.0	479	100.0	540	100.0

Note: Test statistics for age, gender, deck at respondents' level.

Each respondent answered eight vignettes, which resulted in a full estimation sample of n = 8,129 judged vignettes. As for the presentation formats under study, the full estimation sample included n = 3,816 for the written vignettes and n = 4,313 for the photo vignettes. To assess the design's accessibility to systematic group comparisons, we evaluated the randomisation of subjects across the experimental conditions of the split ballot experiment for analysing presentation format differences of the factorial surveys. To do so, we considered whether a parallelisation with regard to the split ballot experiment emerged on the basis of the full estimation model (Table 2).

As depicted from Table 2, there are no substantial differences with regard to the distribution of the vignette dimensions across the two levels of the split ballot experiments (i.e., the presentation of the factorial surveys either using written vignettes or photo vignettes). Subsequently, we examined the vignette dimensions and the respondents' characteristics – age and gender – for a uniform distribution across both vignette presentation modes and show the results in Table 3, which reveals that randomisation of subjects to the presentation formats led to mostly parallel groups with regard to respondents' gender, age and partnership status.

Dependent Variable

The key dependent variable referred to the level of fear of crime when facing the situations described by means of the vignettes. To measure this, we used the stan-

Table 4 Distributions of the dependent variable "Fear of Crime" as reported feelings of unsafety; sample: estimation model

	Both		Written Vignettes		Photo Vignettes	
Feelings of unsafety	N	Percent	N	Percent	N	Percent
Very safe	1880	23.13	876	22.96	1004	23.28
Safe	3619	44.52	1715	44.94	1904	44.15
Unsafe	2028	24.95	971	25.45	1057	24.51
Very unsafe	602	7.41	254	6.66	348	8.07
Total	8129	100.00	3816	100.00	4313	100.00
M	1.166		1.158		1.174	
SD	0.866		0.852		0.878	

T-test: t = -0.813, p = 0.208
U-test: z = -0.471, p = 0.638

dard indicator for fear of crime (i.e., the level of feelings of unsafety). Immediately after presenting each vignette, we asked the following question: "How safe would you feel in this situation?". In response to the question, the subjects used a rating scale (0: very safe; 1: safe; 2: unsafe; 3: very unsafe) for shaping their answer. Table 4 shows the resulting distributions for the full estimation sample.

It follows from Table 4 that, independent of the respective presentation format, most subjects reported a lower level of feelings of safety in public places. A comparison between formats revealed no relevant differences between presentation formats. This reflects that both formats stimulated similar responses on the aggregate level (i.e., across all vignettes).

Method of Analysis

All subsequent analyses refer only to those subjects who considered the factorial surveys as realistic. In conjunction with an evaluation of the online-questionnaire in both formats, we asked the subjects to indicate whether they could imagine themselves in the situation that is presented in the vignettes using a dichotomous response format (0: no; 1: yes). It can be taken from Table A-1 in the appendix that by far, the majority of all subjects evaluated the situations presented in both the written vignettes and the photo vignettes as realistic (n = 1,019), while only a minority did not (n = 188). If the latter group is included in the estimation, only slight differences between all subjects and those who recognised the vignettes as realistic emerged (see Table A-1 in the appendix). For reasons of accurateness, we

decided to use the sample of respondents who evaluated vignettes as being realistic only.

We measured our outcome variable, the level of self-reported feelings of unsafety as an indicator of the level of fear of crime, on a rating scale with four stages, which we interpreted as quasi-metric so that regression models for continuous dependent variables can be applied.

As each respondent assessed a number of vignettes describing varying situations of physical and social disorder, the data structure is hierarchical (Hox et al., 1991; Jasso, 2006). To consider the multi-level structure, we used random intercept fixed slope models, which account for the variation in the outcome variable between respondents (e.g., Snijders & Bosker, 2012). Due to the rather small number of observations at the first level, which is a consequence of the restricted size of the vignette sets, we estimated only the intercept with a random component.

Our primary interest lies in a direct comparison of the factorial surveys across both presentation formats. Accordingly, we estimated a joint model and included multiplicative cross-level-interaction terms between the presentation format at level 2 and all treatment dimensions at level 1. This allows for the estimation of the format's main effect and effects of the vignette dimensions and their levels conditional to the presentation of written or photo vignettes. Accordingly, our estimation strategy can be noted as follows:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + \beta_2 f_j + \beta_3 f_j X_{ij} + \nu_j + \varepsilon_{ij} \quad ; i = 1, ..., n; j = 1, ..., m \qquad \text{(Eq. 1)}$$

Y_{ij}: Reported feelings of unsafety of a respondent j towards a vignette i

X_{ij}: Vector of disorder characteristics varied in vignettes

f_j: Format of presentation to each respondent (photo or written)

ν_j: Error term at respondent level

ε_{ij}: Error term at vignette level

In addition to the analytic strategy presented, the results have undergone a number of checks to prove for their robustness. In a first step, we checked for successful randomisation of the vignettes across respondents, by comparing the dimensions' main effects in models with and without control variables at the respondents' level (see Table A-2 in the appendix). We considered respondents' gender, age and relationship status as sociodemographic covariates, completed by their stated feelings of fatigue when answering the vignettes. All coefficients regarding effects of the vignette dimensions are quite similar between both models, reflecting that randomisation resulted in balanced covariates at the respondents' level. We further compared the random effects model with a fixed effects model using the Hausman test, which revealed only slight and non-substantial differences in coefficients. The

checks indicate that our results are remarkably robust what further confirms that randomisation worked well at both stages of our experimental design. On the basis of these finding, we estimated the presented model without covariates in order to ensure for less missing data due to non-response. As can further be taken from Table 2, the six vignette sets were not assigned to the respondents with exactly the same number during data collection. To account for systematic differences in judgements between vignette sets, we have fixed the effects of the sets in all regression models.

Furthermore, the outcome measurement at a four-point response scale may violate the requirements for linear modelling. Therefore, we have replicated the main effects model in both an ordinal logit model (Table A-3 in the appendix) and a binary logit model (Table A-4 in the appendix). For the latter, we dichotomised the outcome variable using a median split. Although the absolute values of the coefficients cannot be compared directly due to different modelling and scaling, they nevertheless show the same directions and relative strengths within the models. Therefore, the results indicate the robustness of our results. In addition, a comparison between a simple linear regression that neglects the nested data structure and the multilevel model also corroborates our interpretation (Table A-5 in the appendix).[1] We report p-values and confidence intervals to facilitate interpretation.[2]

1 Both, data and codes concerning the analyses strategies will be provided by the authors for replication purposes upon reasonable request.

2 Applying conventional methods of statistical inference is justified even though we did not draw a random sample of respondents for two main reasons: At first, random assignment in experiments reflects data generation through known probability procedures which facilitates formally capturing uncertainties (cf. Berk et al., 1995). Randomisation of subjects to treatments allows for attributing differences between treatments to randomisation error. This justifies testing null hypotheses for treatment effects although statistical inferences apply only to the respondent sample actually used (Edgington, 1966). At second, we invited the total population of students of a German university and consider this population as a realisation from some super population, i.e., a target population which is wider than the actual population under study (Alexander, 2015). On the background of these considerations, we consider our sample population as an equivalent of a random sample which may be analysed on the basis of frequentist methods (Berk et al., 1995).

Results

In this section, we present the results of a random intercept multilevel regression model with interaction terms between the presentation format and all vignette dimensions. The effects of all vignette dimensions conditional on the presentation format are plotted in Figure 3.[3]

As already described above, the mean values of the response scales do not differ across the presentation formats. This finding is reflected by the very small and insignificant coefficient of the main effect of the presentation format in the regression model again. This suggests that both written and photo vignettes generate similar response patterns, at least on the aggregate level including all vignettes. This may be interpreted as a sign of a basic level construct validity in both presentation formats.

It follows from Figure 3 that the effects of the vignette dimensions support the broken windows theory for both presentation formats. The self-reported feeling of unsafety is the stronger the more signs of physical or social disorder are present in a setting as compared to the respective reference categories. The respondents feel more unsafe if a pedestrian underpass is shown instead of a wide square and if there is physical decay and littering indicated by graffiti and garbage lying around instead of a clean and tidy setting. They also report more concern about teenagers hanging around than by a couple walking. Adult passers-by reduce their feelings of unsafety. Compared to a bright scenery resulting from lighting, the respondents feel less safe both in a gloomy situation and when there is video surveillance. The strongest effect is revealed for teenagers hanging around, whereas not being able to overlook a place shows the least effect.

In our hypothesis, we stated that the effects of features of the setting that indicate a high level of physical or social disorder upon the level of fear of crime are stronger in a factorial survey employing photo vignettes compared to a factorial survey employing written vignettes. The conditional effects reveal that there are clear differences in the effects of the vignette dimensions between the two presentation formats, particularly with regard to the dimension of surveillance and lighting. Compared to bright lighting, both video surveillance and a gloomy scenario increase the level of self-reported feelings of unsafety in the photo vignettes significantly stronger than in the written vignettes. Put differently, this means that the level ‘bright situation through lighting’, compared to the levels ‘gloomy situation without surveillance’ and ‘gloomy situation, but video surveillance’, reduces feelings of unsafety much stronger in the photo vignettes than in the written vignettes. Accordingly, the statistically significant coefficients of the interaction terms reveal

3 For reasons of an easier interpretation, the most important results are shown as graphics. The information on the complete regression model can be found in Table A-6 in the appendix.

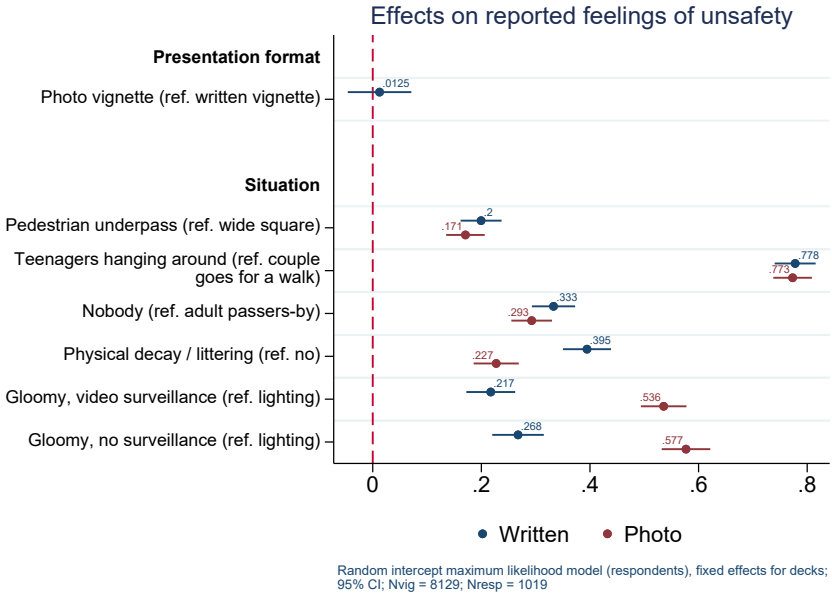

Figure 3 Results of interaction effects model on reported feelings of unsafety (effects conditional to presentation format)

remarkable differences between the two presentation formats both for video surveillance (b = 0.318, p = 0.000) and for a gloomy situation without surveillance (b = 0.309, p = 0.000) by about a third unit at the rating scale. These findings do fully support our hypothesis.

In addition, a significant interaction effect between the level of physical decay/littering and the presentation format is revealed (b = -0.167, p = 0.000). However, the direction of this effect is contrary to our assumption, since physical decay and littering, compared to a clean and tidy environment, increases the self-reported feelings of unsafety more in the written vignettes than in the photo vignettes. This finding, therefore, contradicts our hypothesis.

The remaining signs of disorder of teenagers hanging around instead of a couple walking, the absence of adult passers-by instead of their presence and a narrow pedestrian underpass instead of a wide square show very similar regression weights that do not differ significantly. This indicates that these stimuli evoked comparable response behaviour in both presentation formats.

In summary, we found evidence that the respondents reacted more strongly to the dimension of surveillance/lighting and less to the level of physical decay/littering in the photo vignettes. It can be stated that differences between the presentation formats primarily occurred with regard to signs of physical disorder but hardly

with regard to signs of social disorder. Regarding the latter, written vignettes and photo vignettes seemed to work similarly in both presentation formats.

In addition to differences between treatment effects, we examined possible differences in measurement error by analysing the residuals in separate multilevel models for each presentation format (see Table A-7 in the appendix). Both the log likelihood and information criterion (AIC) show that the model for written vignettes fits data slightly better than the model for photo vignettes. This may suggest that photo vignettes result in data of poorer quality, while written vignettes gain better data. However, a more differentiated comparison of the variance of self-reported feelings of unsafety within and between subjects shows that the poorer model fit is solely due to the variance between respondents, while the variance within respondents is more or less equal for both presentation formats. Accordingly, R^2 coefficients for the variation within subjects hardly differ, while R^2 for the variation between subjects indicates a better fit for the model concerning written vignettes. While the responses differ somewhat more between the respondents in the photo format, the consistency of the responses is roughly the same for both presentation formats. That is, measurement invariance within a respondent is the same regardless of the presentation format, indicating that one of the presentation formats does not force them to give worse answers.

Furthermore, we already assessed the subjective costs of the administration of each format elsewhere (Eifler et al., 2021). These former analyses revealed that dropout rates do not differ between presentation formats, while processing time and self-reported fatigue are reduced when administering a questionnaire including photo vignettes. Concerning dropout rates, results do not indicate that factorial surveys based on photo vignettes would be superior with regard to respondent's willingness to participate. Hence, while the quality of individual responses is largely equal for both presentation formats, evaluating photo vignettes goes along with reduced administration costs in terms of processing time and subjective cognitive demand.

Discussion

With regard to factorial survey experiments, it is often argued that the presentation of information on the situational context when using vignettes allows to mirror the reality of everyday life because it leads to a more standardised imagination of the situation across respondents and less use of general heuristic principles by respondents, which will result in more reliable and valid responses (cf. Alexander & Becker, 1978; Armacost et al., 1991; Finch, 1987; Shamon et al., 2019). Previous studies emphasised the relevance of the presentation format of vignettes for response behaviour (e.g., Eifler & Petzold, 2014; Sauer et al., 2020; Shamon et al.,

2019). So far, there are hardly any empirical studies that have examined the implications of whether vignettes are presented in a written form or by means of visual stimuli, for example, by videos, photos or pictures. Therefore, we were concerned about possible differences between using written vignettes or photo vignettes in factorial survey experiments.

To pursue this question, we used a split ballot experiment employing two factorial survey experiments including either written vignettes or photo vignettes among the population of students from a German university. For our example of use, we referred to the broken windows theory, according to which signs of disorder lead to different levels of fear of crime in a scenario. We used vignettes that describe or display situations with varying signs of physical and social disorder serving as cues for the crime level in a situation. We randomly assigned respondents to one of the two presentation formats and asked them to report their perceived level of safety towards each vignette.

Following Shamon et al. (2019), we assumed that the response process for vignettes is characterised by a more unified retrieval of information from the memory. In line with this, we argued that, on the one hand, the standardised presentation of vignettes will evoke comparable interpretation frames in respondents but, on the other hand, differences between presenting either written vignettes or photo vignettes will occur. Referring to the DCT (Paivio, 1979; Sadoski & Paivio, 2013), we assumed that recall of situational information may depend on the presentation format used in factorial survey experiments. While the verbal information provided by written vignettes is processed in a sequential order, meaning that the verbal information has to be decoded first before visual mental representations are activated, information provided by photo vignettes is processed simultaneously, meaning that both verbal and visual representations are activated concurrently. In our example of use, the respondents were asked to report their feelings of unsafety with regard to everyday situations which provide varying cues of social disorder. We expected that the effects of physical and social signs of disorder should be stronger when presenting photo vignettes.

Our first result is that both vignette formats evoke almost identical distributions of self-reported feelings of unsafety and similar directions of the effects of all vignette dimensions. In accordance with the broken windows theory, cues for signs that indicate a high level of physical and/or social disorder mostly similarly increased the level of fear of crime expressed by the stated feelings of unsafety across both presentation formats. Especially the signs of social disorder, such as teenagers hanging around, show strong effects on the feelings of unsafety. In contrast, the location settings presented in the scenario show the least influence. This result indicates that both presentation formats seem similarly suited to evoke the retrieval of relevant cognitive and affective information from memory and to activate adequate mental representations for the interpretation of a situation.

Our second result is that the effects, conditional on the presentation format, differ significantly for the cues of physical disorder but not for the cues of social disorder. Respondents reacted more strongly to the lighting in a scenario and less to physical decay and littering in the photo vignettes. A reason for this result may be seen in the principal congruence between presenting visual information that particularly triggers the sense of vision as opposed to presenting visual information on physical or social aspects of a situation (see also Vanderveen, 2018). Whether this result may also indicate a higher level of validity of photo vignettes, however, cannot be answered on the basis of our study.

Additional residual analyses revealed that models of the written vignettes fit better when comparing between subjects, while there is a lack of differences when comparing answers of subjects within the presentation format. The mixed results do not indicate clear advantages of one presentation format over the other. Instead, one may gain some and also lose some by choosing either format.

Yet, the results do not fully correspond to our hypothesis, according to which all effects of the signs of disorder should be stronger in the photo vignettes. A possible explanation for the largely similar effect sizes across both presentation formats in the case of signs of social disorder is that they may have a strong activating effect also in the written vignettes. In contrast to signs of physical disorder, signs of social disorder may work as cues for potential social interaction. Humans are social beings and, therefore, may focus strongly on interacting with other people as part of their conditio humana. Through life-long learning of evaluating the social environment, people could develop a strong and inter-individually coherent representation of social situations. That is, when having conversations, or reading books or newspapers, people are trained to imagine other people and groups of people and their activities. This routine could lead to the consequence that social cues in both written vignettes and photo vignettes may evoke comparable cognitive processes so that the same effects are observed. This interpretation would be in line with the assumption that the response process for vignettes is characterised by a more unified retrieval of information from the memory (Shamon et al., 2019).

Moreover, in contradiction with our assumptions, a stronger effect of physical decay and littering on the reported feelings of unsafety was detected with the written vignettes. Yet, the assumption of a more unified retrieval of information from the memory allows for a coherent re-interpretation, as the stronger effect may reflect a methodological artefact. Compared to physical disorder in photo vignettes and to other dimensions in photo vignettes, more information is provided in written vignettes in this dimension. It covers two sentences including comprehensive information about a number of details (graffiti, beer bottles, garbage, wall plaster, asphalt holes). This may have overly activated the retrieval of related mental representations and evoked particular emotional arousal in the written vignettes, and

must be considered as a weakness of our design. Such problems should be avoided in future studies.

The results also contain implications for specific areas of the application of vignettes. First, though both modes seem to obtain meaningful effect estimates, applications with relevant visual information might eventually be better operationalised with photo vignettes. The reason may be that information which shall trigger the sense of vision may not be processed adequately when using written vignettes describing real-world scenarios. This is particularly the case when everyday situations are presented in which aspects of the physical environment are considered. Second, the results of this study also indicate that information on the social environment might be adequately processed also with written vignettes. Furthermore, there are many areas of application of vignettes in which the effects of the physical and, therefore, visible characteristics of a situation or of the physical characteristics of people are not theoretically significant. If no 'visual' dimensions are considered, photo vignettes will most likely not offer any advantage over written vignettes. Third, details may possibly be studied with the written format more comprehensively by actively drawing attention to specific dimensions through dense descriptions. This may appear as an advantage for special questions but may also be a disadvantage if certain factors are focused too strongly and over-activate certain attitudes or norms. In such a case, the importance of this dimension may be overestimated in written vignettes (see number of levels effect). However, photo vignettes may also be used to put a special emphasis on certain dimensions.

The main limitations of our study lie in the external validity of the results. The study was carried out with a homogeneous student sample from a single German university. Although this is sufficient for a method study of this kind, a replication with a more heterogeneous sample is desirable for the future, in order to ensure the robustness of the effects via subgroup analyses or cultural comparisons. In addition, the interpretation is limited by the varied vignette dimensions. It would be conceivable to use further or different operationalisations for social and physical cues in replications. It would also be promising to vary only social or only physical characteristics of the environment in order to increase the external validity. We also measured the stated feelings of unsafety as outcome variable, though the broken windows theory deals with the term fear of crime, which is not identical. Yet, we decided to measure the level of fear of crime using the standard indicator of this construct. The emotional dimension is brought to the fore with this outcome measurement. It would yet be interesting to measure not only attitudes towards the situation but also behavioural intentions, such as leaving the scenario or seeking protection.

As a conclusion, using visual stimuli might have advantages over written stimuli in research situations where visual information is theoretically relevant and particularly triggers the sense of vision. The reason for this is that visual stim-

uli allow to easily display spatial information concerning real-life situations that would be difficult to describe. Nevertheless, photo vignettes are reduced to a two-dimensional representation of aspects of social reality. This means that although they might work better in providing approximations to everyday life than written vignettes, they are not suitable for presenting additional sensory information like noise or smell (Mitchell, 1986). Furthermore, neither written vignettes nor photo vignettes are suitable for presenting information concerning bodily movement or other senses like, for example, the senses of touch, smell or hearing. While using video vignettes might solve some of these problems, as video clips may present both bodily movement and sound, other aspects of real-life situations might possibly only be simulated by means of methods involving virtual realities (Van Gelder et al., 2019; Van Sintemaartensdijk et al., 2020). Future studies concerning the relevance of the presentation format of vignettes should take these considerations into account.

References

Alexander, C. S., & Becker, H. J. (1978). The Use of Vignettes in Survey Research. *Public Opinion Quarterly, 42*(1), 93-104. https://doi.org/10.1086/268432

Alexander, N. (2015). What's More General than a Whole Population? *Emerging Themes in Epidemiology, 12*, 11-11. https://doi.org/10.1186/s12982-015-0029-4

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. https://doi.org/10.1111/j.1540-5915.1991.tb01907.x

Auspurg, K., & Hinz, T. (2015). *Factorial Survey Experiments*. Thousand Oaks: Sage.

Barthes, R. (1977). *Image, Music, Text*. London: Fontana Press.

Beck, M., & Opp, K.-D. (2001). Der faktorielle Survey und die Messung von Normen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 53*(2), 283–306. https://doi.org/10.1007/s11577-001-0040-3

Benson, L. E. (1941). Studies in Secret-Ballot Technique. *Public Opinion Quarterly, 5*, 79-82. https://doi.org/10.1086/265464

Berk, R. A., Western, B., & Weiss, R. E. (1995). Statistical Inference for Apparent Populations. *Sociological Methodology, 25*, 421-458. https://doi.org/10.2307/271073

Bower, G. H. (1970). Imagery as a Relational Organizer in Associative Learning. *Journal of Verbal Learning and Verbal Behavior, 9*, 529-533. https://doi.org/10.1016/s0022-5371(70)80096-2

Brockmole, J. R. (2009). *The Visual World in Memory*. Hove, Sussex: Psychology Press.

Caro, F. G., Ho, T. H., McFadden, D., Gottlieb, A. S., Yee, C., Chan, T., & Winter, J. (2012a). Using the Internet to Administer More Realistic Vignette Experiments. *Social Science Computer Review, 30*(2), 184-201. https://doi.org/10.1177/0894439310391376

Caro, F. G., Yee, C., Levien, S., Gottlieb, A. S., Winter, J., McFadden, D. L., & Ho, T. H. (2012b). Choosing Among Residential Options: Results of a Vignette Experiment. *Research on Aging, 34*(1), 3-33. https://doi.org/10.1177/0164027511404032

Collier Jr, J. (1957). Photography in Anthropology: A Report on Two Experiments. *American Anthropologist, 59*(5), 843-859. https://doi.org/10.1525/aa.1957.59.5.02a00100

Collier, J., & Collier, M. (1986). *Visual Anthropology: Photography as a Research Method*. Albuquerque: University of New Mexico Press.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, Mail, and Mixed-Mode Surveys. The Tailored Design Method*. Wiley: New Jersey.

Dinora, P., Schoeneman, A., Dellinger-Wray, M., Cramer, E. P., Brandt, J., & D'Aguilar, A. (2020). Using Video Vignettes in Research and Program Evaluation for People with Intellectual and Developmental Disabilities: A Case Study of the Leadership for Empowerment and Abuse Prevention (LEAP) Project. *Evaluation and Program Planning, 79*. https://doi.org/10.1016/j.evalprogplan.2019.101774

Edgington, E. (1966). Statistical Inference and Nonrandom Samples. *Psychological Bulletin, 66*(6), 485-487. https://doi.org/10.1037/h0023916

Eifler, S. (2007). Evaluating the Validity of Self-Reported Deviant Behavior Using Vignette Analyses. *Quality & Quantity*, *41*, 303-318. https://doi.org/10.1007/s11135-007-9093-3

Eifler, S., & Petzold, K. (2014). Der Einfluss der Ausführlichkeit von Vignetten auf die Erfassung prosozialer Einstellungen. Ergebnisse zweier Split-Ballot Experimente. *Soziale Welt. Zeitschrift für sozialwissenschaftliche Forschung und Praxis, 2014*(2), 247-270. https://doi.org/10.5771/0038-6073-2014-2-247

Eifler, S., Petzold, K., & Verbeek-Teres, M. (2021). Presentation Format Differences in Factorial Surveys. *Eichstätter Beiträge zur Soziologie, 19*. Retrieved December 13, 2021, from https://edoc.ku.de/id/eprint/29215/

Finch, J. (1987). Research Note. The Vignette Technique in Survey Research. *Sociology, 21*(1), 105-114. https://doi.org/10.1177/0038038587021001008

Findlay, J. M., & Gilchrist, I. D. (2003). *Active Vision: The Psychology of Looking and Seeing*. Oxford: Oxford University Press.

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.

Goyal, N., Wice, M., Kinsbourne, M., & Castano, E. (2017). A Picture Is Worth a Thousand Words: The Influence of Visuospatial and Verbal Cognitive Styles on Empathy and Willingness to Help. *Social Psychology, 48*(6) 372–379. https://doi.org/10.1027/1864-9335/a000318

Hanna, A., & Loftus, G. (1993). A Model for Conceptual Processing of Naturalistic Scenes. *Canadian Journal of Experimental Psychology, 47*(3), 548-569. https://doi.org/10.1037/h0078851

Harper, D. (2002). Talking about Pictures: A Case for Photo Elicitation. *Visual Studies, 17*(1), 13–26. https://doi.org/10.1080/14725860220137345

Havekes, E., Coenders, M., & Van der Lippe, T. (2013). Positive or Negative Ethnic Encounters in Urban Neighbourhoods? A Photo Experiment on the Net Impact of Ethnicity and Neighbourhood Context on Attitudes Towards Minority and Majority Residents. *Social Science Research, 42*(4), 1077–1091. https://doi.org/10.1016/j.ssresearch.2013.02.002

Hox, J. J., Kreft, I. G. G., & Hermkens, P. L. J. (1991). The Analysis of Factorial Surveys. *Sociological Methods & Research, 19*(4), 493-510. https://doi.org/10.1177/0049124191019004003

Hughes, R., & Huby, M. (2004). The Construction and Interpretation of Vignettes in Social Research. *Social Work and Social Sciences Review, 11*(1), 36-51. https://doi.org/10.1921/17466105.11.1.36

Hunter, A. (1978). *Symbols of Incivility*. Paper presented at the Annual Meeting of the American Society of Criminology, November 1978, Dallas.

Jasso, G. (2006). Factorial Survey Methods for Studying Belief and Judgments. *Sociological Methods & Research, 34*(3), 334-423. https://doi.org/10.1177/0049124105283121

Jenkins, J. R., Neale, D. C., & Deno, S. L. (1967). Differential Memory for Picture and Word Stimuli. *Journal of Educational Psychology, 58*(5), 303-307. https://doi.org/10.1037/h0025025

Keizer, K., Lindenberg, S., & Steg, L. (2014). Doing Field Studies: What is it All About? *Group Processes & Intergroup Relations, 17*(3), 404-410. https://doi.org/10.1177/1368430213510750

Kelling, G. L., & Coles, C. M. (1996). *Fixing Broken Windows: Restoring Order and Reducing Crime in Our Communities*. New York: Touchstone.

Keuschnigg, M., & Wolbring, T. (2015). Disorder, Social Capital, and Norm Violation. Three Field Experiments on the Broken Windows Thesis. *Rationality and Society, 27*(1), 96–126. https://doi.org/10.1177/1043463114561749

Kosslyn, S. M. (1981). The Medium and the Message in Mental Imagery: A Theory. *Psychological Review, 88*(1), 46-66. https://doi.org/10.1037/0033-295x.88.1.46

Kosslyn, S. M., & Pomerantz, J. R. (1977). Imagery, Propositions, and the Form of Internal Representations. *Cognitive Psychology, 9*, 52-76. https://doi.org/10.1016/0010-0285(77)90004-4

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. https://doi.org/10.1086/599248

Lewis, D. A., & Salem, G. (1986). *Fear of Crime. Incivility and the Production of a Social Problem*. New Brunswick, NJ: Transaction Books.

Liebig S., Sauer, C., & Friedhoff, S. (2015). Empirische Gerechtigkeitsforschung mit dem faktoriellen Survey. In M. Keuschnigg, & T. Wolbring (Eds.), *Experimente in den Sozialwissenschaften* (pp 321-339). Baden-Baden: Nomos.

Luck, S. J., & Hollingworth, A. (2008). *Visual Memory*. Oxford: Oxford University Press.

Manghani, S. (2013). *Image Studies. Theory and Practice*. London, New York: Routledge.

Marion, J. S., & Crowder, J. W. (2013). *Visual Research. A Concise Introduction to Thinking Visually*. London: Bloomsbury.

McCloud, S. (1994). *Understanding Comics. The Invisible Art*. New York: Harper Perennial.

Mitchell, W. J. T. (1986). *Iconology. Image, Text, Ideology*. Chicago, London: The University of Chicago Press.

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

O'Donnell, A. B., Lutfey, K. E., Marceau, L. D., & McKinlay, J. B. (2007). Using Focus Groups to Improve the Validity of Cross-National Survey Research: A Study of Physician Decision Making. *Qualitative Health Research, 17*(7), 971–981. https://doi.org/10.1177/1049732307305257

Paivio, A. (1979). *Imagery and Verbal Processes*. New Jersey: Lawrence Erlbaum Associates.

Piquero, A. (1999). The Validity of Incivility Measures in Public Housing. *Justice Quarterly, 16*(4), 793–818. https://doi.org/10.1080/07418829900094371

Rashotte, L. S. (2003). Written Versus Visual Stimuli in the Study of Impression Formation. *Social Science Research, 32*(2), 278-293. https://doi.org/10.1016/s0049-089x(02)00050-9

Rose, G. (2012). *Visual Methodologies. An Introduction to Researching with Visual Materials*. London: Sage Publications.

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

Sadoski, M., & Paivio, A. (2013). *Imagery and Text: A Dual Coding Theory of Reading and Writing*. London: Routledge.

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. https://doi.org/10.12758/mda.2020.06

Schwarzbach, H., & Eifler, S. (2020). Einflüsse der Präsentationsform eines faktoriellen Surveys zur Erfassung von Sicherheit im öffentlichen Raum. *Eichstätter Beiträge zur Soziologie, 18*. Retrieved December 13, 2021, from https://edoc.ku.de/id/eprint/25734/

Shamon, H., Dülmer, H., & Giza, A. (2019). The Factorial Survey: The Impact of the Presentation Format of Vignettes on Answer Behavior and Processing Time. *Sociological Methods & Research* (Onlinefirst). https://doi.org/10.1177/0049124119852382

Shepard, R. N. (1967). Recognition Memory for Words, Sentences, and Pictures. *Journal of Verbal Learning and Verbal Behavior, 6,* 156-163. https://doi.org/10.1016/s0022-5371(67)80067-7

Skogan, W. G. (1978). *Victimization Surveys and Criminal Justice Planning.* Washington, DC: National Institute of Law Enforcement and Criminal Justice.

Skogan, W. G. (1990). *Disorder and Decline. Crime and the Spiral Decay in American Neighborhoods.* New York, NY: Free Press.

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

Taylor, R. B. (1999). The Incivilities Thesis: Theory, Measurement, and Policy. In R. H. Langworthy (Eds.), *Measuring what Matters: Proceedings from the Policing Research Institute Meetings* (pp 65–88). Washington, DC: U.S. Department of Justice.

Tourangeau, R. (1984). Cognitive Science and Survey Methods. In T. B. Jabine, M. L. Straf, J. M. Tanur, & R. Tourangeau (Eds.), *Cognitive Aspects of Survey Design: Building a Bridge Between Disciplines* (pp 73–100). Washington, DC: National Academy Press.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press.

Van Gelder, J.-L., De Vries, R. E., Demetriou, A., Van Sintemaartensdijk, I., & Donker, T. (2019). The Virtual Reality Scenario Method: Moving from Imagination to Immersion in Criminal Decision-Making Research. *Journal of Research in Crime and Delinquency*, *56*(3), 451–480. https://doi.org/10.1177/0022427818819696

Van Sintemaartensdijk, I., Van Gelder, J.-L., Van Prooijen, J.-W., Nee, C., Otte, M., & Van Lange, P. (2020). Mere Presence of Informal Guardians Deters Burglars: A Virtual Reality Study. *Journal of Experimental Criminology.* https://doi.org/10.1007/s11292-020-09430-1

Vanderveen, G. (2018). Visual Methods in Research on Fear of Crime: A Critical Assessment. In M. Lee, & G. Mythen (Eds.), *The Routledge International Handbook on Fear of Crime* (pp 170-189). London: Routledge. https://doi.org/10.4324/9781315651781-13

Wallander, L. (2009). 25 Years of Factorial Surveys in Sociology: A Review. *Social Science Research, 38*, 505-520. https://doi.org/10.1016%2Fj.ssresearch.2009.03.004

Wilson, J. Q., & Kelling, G. L. (1982). Broken Windows: The Police and Neighborhood Safety. *Atlantic Monthly, 249*(3), 29-38.

Appendix

Table A-1 Robustness regarding evaluation of realistic vignette descriptions

Reported feelings of unsafety	Evaluation: realistic and non-realistic		Evaluation: only realistic	
Pedestrian underpass (ref. wide square)	0.173***	(14.06)	0.185***	(13.85)
Teenagers hanging around (ref. couple goes for a walk)	0.755***	(61.44)	0.775***	(58.12)
Physical decay / littering (ref. no)	0.295***	(20.50)	0.306***	(19.60)
No passers-by (ref. adult passers-by)	0.295***	(22.81)	0.314***	(22.31)
Gloomy, video surveillance (ref. lighting)	0.382***	(26.18)	0.387***	(24.53)
Gloomy, no surveillance (ref. lighting)	0.427***	(27.68)	0.432***	(25.78)
Photo vignette (ref. written vignette)	0.0155	(0.56)	0.00672	(0.22)
Vignette set (ref. set 1)				
Set 2	-0.0629	(-1.33)	-0.0672	(-1.29)
Set 3	-0.190***	(-3.89)	-0.191***	(-3.55)
Set 4	-0.214***	(-4.40)	-0.222***	(-4.14)
Set 5	-0.200***	(-4.06)	-0.214***	(-4.00)
Set 6	-0.0896	(-1.81)	-0.0593	(-1.08)
Constant	0.259***	(6.21)	0.228***	(4.95)
σ_u	0.432		0.436	
σ_e	0.571		0.569	
Log likelihood	-9,293.0		-7,838.9	
LR-χ^2	4,453.45***		3,948.75***	
AIC	18,615.95		15,707.84	
$N_{Vignettes}$	9,620		8,129	
$N_{Respondents}$	1,207		1,019	

Linear random intercept maximum likelihood estimations.
t statistics in parentheses; * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Table A-2 Robustness with and without covariates

Reported feelings of unsafety	Without covariates		With covariates	
Pedestrian underpass (ref. wide square)	0.184***	(13.69)	0.184***	(13.55)
Teenagers hanging around (ref. couple goes for a walk)	0.776***	(57.82)	0.780***	(57.33)
Physical decay / littering (ref. no)	0.310***	(20.47)	0.311***	(19.52)
No passers-by (ref. adult passers-by)	0.305***	(21.81)	0.319***	(22.26)
Gloomy, video surveillance (ref. lighting)	0.375***	(23.47)	0.382***	(23.77)
Gloomy, no surveillance (ref. lighting)	0.432***	(25.70)	0.434***	(25.38)
Photo vignette (ref. written vignette)	0.0165	(0.53)	0.00768	(0.25)
Vignette set (ref. set 1)				
Set 2			-0.0902	(-1.77)
Set 3			-0.212***	(-3.99)
Set 4			-0.240***	(-4.56)
Set 5			-0.227***	(-4.33)
Set 6			-0.0533	(-0.99)
Age			-0.00206	(-0.57)
Male (ref. female)			-0.289***	(-9.09)
Spouse (ref. no partnership)			-0.0121	(-0.39)
Evaluation: questionnaire fatiguing			0.00143	(0.04)
Constant	0.107***	(3.73)	0.400***	(3.63)
σ_u	0.446		0.416	
σ_e	0.570		0.570	
Log likelihood	-7,605.4		-7,547.3	
LR-χ^2	3,806.7***		3,922.76***	
AIC	15,230.7		15,132.7	
$N_{Vignettes}$	7,855		7,855	
$N_{Respondents}$	985		985	

Linear random intercept maximum likelihood estimations.
t statistics in parentheses; * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Table A-3 Robustness by linear / ordered logit multi-level modelling

Reported feelings of unsafety	Linear multi-level model		Ord. log. multi-level model	
Pedestrian underpass (ref. wide square)	0.185***	(13.85)	0.657***	(13.07)
Teenagers hanging around (ref. couple goes for a walk)	0.775***	(58.12)	2.877***	(46.48)
Physical decay / littering (ref. no)	0.306***	(19.60)	1.154***	(19.29)
No passers-by (ref. adult passers-by)	0.314***	(22.31)	1.171***	(21.66)
Gloomy, video surveillance (ref. lighting)	0.387***	(24.53)	1.454***	(23.52)
Gloomy, no surveillance (ref. lighting)	0.432***	(25.78)	1.618***	(24.76)
Photo vignette (ref. written vignette)	0.00672	(0.22)	0.0266	(0.23)
Vignette set (ref. set 1)				
Set 2	-0.0672	(-1.29)	-0.301	(-1.51)
Set 3	-0.191***	(-3.55)	-0.650**	(-3.17)
Set 4	-0.222***	(-4.14)	-0.827***	(-4.03)
Set 5	-0.214***	(-4.00)	-0.799***	(-3.90)
Set 6	-0.0593	(-1.08)	-0.221	(-1.05)
Constant	0.228***	(4.95)		
Cut 1			1.249***	(7.07)
Cut 2			4.928***	(26.20)
Cut 3			7.897***	(38.48)
σ_u			2.812***	(16.24)
σ_u	0.436		2.812	
σ_e	0.569			
Log likelihood	-7,838.9		-7,495.9	
LR/Wald-χ^2	3,948.75***		2,766.59***	
AIC	15,707.84		15,023.84	
$N_{Vignettes}$	8,129		8,129	
$N_{Respondents}$	1,019		1,019	

Linear random intercept maximum likelihood estimation & ordered logit random intercept maximum likelihood estimation.
t statistics in parentheses; * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Table A-4 Robustness by linear / logit multi-level modelling

Reported feelings of unsafety	Linear multi-level model		Log. multi-level model	
Pedestrian underpass (ref. wide square)	0.185***	(13.85)	0.0906***	(11.24)
Teenagers hanging around (ref. couple goes for a walk)	0.775***	(58.12)	0.356***	(50.06)
Physical decay / littering (ref. no)	0.306***	(19.60)	0.169***	(18.36)
No passers-by (ref. adult passers-by)	0.314***	(22.31)	0.154***	(19.29)
Gloomy, video surveillance (ref. lighting)	0.387***	(24.53)	0.154***	(16.53)
Gloomy, no surveillance (ref. lighting)	0.432***	(25.78)	0.183***	(18.38)
Photo vignette (ref. written vignette)	0.00672	(0.22)	0.00745	(0.53)
Vignette set (ref. set 1)				
Set 2	-0.0672	(-1.29)	0.0101	(0.41)
Set 3	-0.191***	(-3.55)	-0.0394	(-1.62)
Set 4	-0.222***	(-4.14)	-0.0467	(-1.89)
Set 5	-0.214***	(-4.00)	-0.0605*	(-2.54)
Set 6	-0.0593	(-1.08)	0.0276	(1.02)
Constant	0.228***	(4.95)		
σ_u	0.436		1.648	
σ_e	0.569			
Log likelihood	-7,838.9		-3,513.5	
LR/Wald-χ^2	3,948.75***		1,392.5***	
AIC	15,707.84		7,055.06	
$N_{Vignettes}$	8,129		8,129	
$N_{Respondents}$	1,019		1,019	

Linear random intercept maximum likelihood estimation & logistic random intercept maximum likelihood estimation.
t statistics in parentheses; * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Note: Logistic multi-level model with median split of dependent variable; Average Marginal Effects reported (AME).

Table A-5 Robustness by multi-level structure

Reported feelings of unsafety	Linear multi-level model		Linear OLS model	
Pedestrian underpass (ref. wide square)	0.185***	(13.85)	0.185***	(13.02)
Teenagers hanging around (ref. couple goes for a walk)	0.775***	(58.12)	0.776***	(43.22)
Physical decay / littering (ref. no)	0.306***	(19.60)	0.306***	(17.48)
No passers-by (ref. adult passers-by)	0.314***	(22.31)	0.315***	(20.83)
Gloomy, video surveillance (ref. lighting)	0.387***	(24.53)	0.387***	(20.78)
Gloomy, no surveillance (ref. lighting)	0.432***	(25.78)	0.432***	(22.08)
Photo vignette (ref. written vignette)	0.00672	(0.22)	0.0115	(0.38)
Vignette set (ref. set 1)				
Set 2	-0.0672	(-1.29)	-0.0675	(-1.29)
Set 3	-0.191***	(-3.55)	-0.191***	(-3.43)
Set 4	-0.222***	(-4.14)	-0.222***	(-4.13)
Set 5	-0.214***	(-4.00)	-0.215***	(-3.80)
Set 6	-0.0593	(-1.08)	-0.0585	(-1.07)
Constant	0.228***	(4.95)	0.224***	(4.73)
σ_u	0.436			
σ_e	0.569			
Log likelihood	-7,838.9		-8,823.8	
LR-χ^2 / F	3,948.75***		318.10***	
AIC	15,707.84		17,673.65	
$N_{Vignettes}$	8,129		8,129	
$N_{Respondents}$	1,019			

Linear random intercept maximum likelihood estimation & linear ordinary least squares estimations, clustered SE.
t statistics in parentheses; * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Table A-6 Estimation of interactions between presentation format and vignette dimensions

Reported feelings of unsafety		
Photo vignette (ref. written vignette)	0.0958	(1.13)
Pedestrian underpass (ref. wide square)	0.200***	(10.35)
Teenagers hanging around (ref. couple goes for a walk)	0.778***	(40.41)
Physical decay / littering (ref. no)	0.395***	(17.48)
No passers-by (ref. adult passers-by)	0.333***	(16.38)
Gloomy, video surveillance (ref. lighting)	0.217***	(9.47)
Gloomy, no surveillance (ref. lighting)	0.268***	(11.05)
Vignette set (ref. set 1)		
Set 2	0.00311	(0.04)
Set 3	-0.00650	(-0.08)
Set 4	-0.174*	(-2.23)
Set 5	-0.0264	(-0.34)
Set 6	0.0332	(0.42)
Photo vignette * Underpass	-0.0288	(-1.09)
Photo vignette * Teenagers hanging around	-0.00468	(-0.18)
Photo vignette * Physical decay / littering	-0.167***	(-5.40)
Photo vignette * No passers-by	-0.0402	(-1.44)
Photo vignette * Gloomy, video surveillance	0.318***	(10.16)
Photo vignette * Gloomy, no surveillance	0.309***	(9.30)
Photo vignette * set 2	-0.131	(-1.27)
Photo vignette * set 3	-0.346**	(-3.24)
Photo vignette * set 4	-0.0834	(-0.79)
Photo vignette * set 5	-0.349**	(-3.28)
Photo vignette * set 6	-0.165	(-1.53)
Constant	0.181**	(2.87)
σ_u	0.431	
σ_e	0.564	
Log likelihood	-7,755.1876	
LR-χ^2	4,116.22***	
$N_{Vignettes}$	8,129	
$N_{Respondents}$	1,019	

Linear random intercept maximum likelihood estimations.
t statistics in parentheses; * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Note: Conditional effects for presentation format of written vignettes.

Table A-7 Estimation of dimensions' effects by presentation format

Reported feelings of unsafety	Written vignettes		Photo vignettes	
Pedestrian underpass (ref. wide square)	0.200***	(10.68)	0.171***	(9.18)
Teenagers hanging around (ref. couple goes for a walk)	0.778***	(41.68)	0.773***	(41.50)
Physical decay / littering (ref. no)	0.395***	(18.04)	0.227***	(10.45)
No passers-by (ref. adult passers-by)	0.333***	(16.90)	0.293***	(14.93)
Gloomy, video surveillance (ref. lighting)	0.217***	(9.77)	0.536***	(24.41)
Gloomy, no surveillance (ref. lighting)	0.268***	(11.40)	0.577***	(24.64)
Photo vignette (ref. written vignette)	0.00672	(0.22)	0.0115	(0.38)
Vignette set (ref. set 1)				
Set 2	0.00315	(0.04)	-0.128	(-1.84)
Set 3	-0.00647	(-0.08)	-0.353***	(-4.84)
Set 4	-0.174*	(-2.24)	-0.257***	(-3.54)
Set 5	-0.0263	(-0.34)	-0.375***	(-5.19)
Set 6	0.0332	(0.43)	-0.132	(-1.76)
Constant	0.181**	(2.90)	0.277***	(4.79)
σ_u	0.431***		0.430***	
σ_e	0.546***		0.578***	
Log likelihood	-3,534.8		-4,214.6	
LR-χ^2	1,894.0***		2,212.3***	
AIC	7,097.508		8,457.291	
R^2_{within} +	0.414		0.435	
$R^2_{between}$ +	0.206		0.102	
$R^2_{overall}$ +	0.332		0.326	
$N_{Vignettes}$	3,816		4,313	
$N_{Respondents}$	479		540	

Linear random intercept maximum likelihood estimations.
+ from linear random intercept generally least squares estimations.
t statistics in parentheses; * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$