Social Desirability's Influence on Audience Research

Discerning and Reducing It¹

SIGRID BEKMEIER-FEUERHAHN^A*, PAULA MARIA BÖGEL[®], JÖRG SIKKENGA[⊂], ANDREAS HEINEN[®]

^{AC.D}Institute for Management and Organization, ^BInstitute for Environmental Communication, Leuphana University Lüneburg

The term 'social desirability bias' (SD bias) is used to describe a systematic tendency to adapt one's behavior or communication to conform to perceived social norms. While SD bias, which can affect the validity of empirical studies, has received considerable attention in psychological research, it has yet to be examined in the context of the fine and performing arts. This study is among the first to provide empirical evidence for high trait desirability (TD) in the cultural sector and a resultant SD bias. It shows that due to SD bias, audiences in the cultural sector are more likely to give misleading positive feedback. It is therefore incumbent on researchers in the cultural sector to integrate an awareness of and possibilities for mitigating SD bias at all stages of their studies. This paper offers several specific suggestions on how this can be done.

Keywords

visitor studies, method development, audience development, theater, marketing

1. Introduction

In recent decades, cultural organizations have become more market oriented and increasingly focused on audience development (BAUMGARTH 2009; CAMARERO/GARRIDO 2012; FILLIS 2011; LINDELOF 2014). Success in the global marketplace requires effective strategic decisions, and those decisions, in turn, need to be made based on reliable information on audiences, which can be collected with the help of surveys. Although they are often used, audience studies have been shown to be ineffective in terms of strategic outcome.

* Corresponding Author: bekmeier@leuphana.de

¹ The authors wish to thank Adelheid Mers, Associate Professor, Arts Administration and Policy, School of the Art Institute of Chicago, for her helpful comments and great support on the English translation. The authors also acknowledge the contributions of three reviewers to improve the paper. As shown in table 1, many studies conducted over the last decade have one aspect in common: satisfaction with visits to cultural organizations and events tends to be very high.

Organisation Study		Author	Results		
German Muse- um, Munich	Climate Exhibition	LEWALTER/ GEYER (2003)	85 % good or excellent		
20 theaters in Berlin	Audience Study	TAUCHNITZ (2004)	88 % satisfied		
Natural History Museum, London	Antarctica Exhibition	JAMES (2007)	They all enjoyed it, and most of them gave it a score of 8 or 9 out of 10.		
20 Theaters in Canada	Audience Enga- gement Survey Report	CREATIVE TRUST FOR ARTS & CUL- TURE (2010)	96 % indicted that they enjoy it occasionally or regard it as a vital activity for them		
German Muse- um, Munich	Nanotechnology Exhibition	SPECHT/ LEWALTER (2011)	84,4 % good or excellent		
12 museums in Cologne	Museum Night	KLIMENT (2011)	86 % good or excellent		
Victoria and Albert Museum, London	Case Study Evaluation of Future Plan	PETRIE (2011)	Beautiful (71 %), historic (60 %), informative (49 %), and stimu- lating (49 %) out of a list of 20 words.		
State Theater, Karlsruhe	Audience Study	SIEBENHAAR (2012)	92,8% good or excellent		

Table 1: List of museum, orchestra, theater, and library audience surveys from the United States, Great Britain, and Germany. Each survey asked for overall satisfaction. The selection of studies included here is not representative.

As rewarding as these results may be, they do not motivate decision makers to develop new strategies to meet audiences' actual needs. If all visitors are content, why should anything be adjusted? Thus, these results may lead to inertia. They may also be problematic for another reason: the audiences of most cultural organizations tend to be heterogeneous (FÖHL/GLOGNER-PILZ 2016; FÖHL/NÜBEL 2016), but their evaluations are surprisingly very similar. As audiences dwindle, particularly in the sector of high culture, whether the results of audience studies are reliable and truly meaningful can be questioned.

One explanation for the discrepancy between positive evaluations and the declining number of visitors observed could be that the questionnaires used for evaluations are biased. Both audience and nonaudience studies that rely on direct questions are subject to an array of errors. This paper argues that because of social desirability bias (SD bias), individuals may provide feedback that is more positive than their actual opinion. It is assumed that the cultural sector is particularly vulnerable to SD bias because of an anticipated high trait desirability (TD) associated with the fine and performing arts, which are often associated with 'high culture.'

High culture has long played an important role in marking social status and prestige (VEBLEN 1899; LINTON 1945; BOURDIEU 1979). Individuals who seek out what they regard as the 'better circles' within their class or want to enter a higher class try to impress group members through social prestige. Modern, fragmented society is only partially reflected in Veblen's, Linton's, and partially Bourdieu's linearly conceptualized structure. Lifestyle research, which has been highly influenced by these theories (HAKIM 2000; PATSIAOURAS/FITCHETT 2012; TRIGG 2016), offers more nuanced models that may explain how people differ in terms of their perceived degree of TD concerning certain issues. Whereas Bourdieu's model is, for example, characterized by a hierarchical structure, lifestyle research has challenged the assumption of clear hierarchies by introducing two new dimensions of structuring clusters: discontinuity and expansion.

The term 'discontinuity' is used to refer to the fact that group members can temporarily change their place in a social system by engaging in a particular behavior. Thus, new elites are characterized by 'omnivorousness.' This term was introduced in the academic discourse on culture by PETERSON/SIMKUS (1992), who sought to explain hybrid cultural consumption behavior. They argue that people can switch between reference systems without violating their class borders. As a result, the same person can appreciate high and popular culture. Similarly, Prior (2005) aimed to improve Bourdieu's model of society by integrating the complexity that is inherent in a more nuanced understanding of culture. In consumer research, this behavior is discussed as "hybrid purchasing behavior" (LEPPÄNEN/GRÖNROOS 2009). This kind of behavior has made it increasingly difficult to clearly link individuals to a specific class based on their consumption behavior. While omnivorousness is, as indicated above, regarded as a trait of new elites (ATKINSON 2011), overall cultural capital still plays a major role among reputational criteria (HOLT 1997). Although hybrid consumer behavior is increasingly common, considering distinction behavior is still possible and necessary.

The second dimension, expansion, is a result of social research in the 1980s. Schulze (1992) remarked that distinguishing people based on their demographic and social backgrounds has become increasingly problematic. His research describes social classes as milieus within a two-dimensional space. The construction of these milieus is still based on measures that are similar to those used by Bourdieu, but the former are not arranged in a hierarchy. The social orientation point for behavior that yields prestige is no longer clearly determined but associated with lifestyles consciously or unconsciously adopted by individuals. Thus, uses and especially explicit nonuses of culture may be exploited as markers of distinction at a given level.

Drawing on these theoretical insights, this empirical study aims to clarify the role and importance of SD bias in audience studies. It examines whether research on the cultural sector is particularly prone to SD bias, and if so, to what extent. This study also discusses the results and methodological approaches that could be used to address the problem of social desirability and to increase the quality of responses for future audience studies on cultural organizations.

2. Social Desirability Bias and High Culture as Trait Desirability

The concept of 'social desirability' implies that respondents adapt their responses to the social norm (KRUMPAL 2013; TOURANGEAU/YAN 2007). With regard to perceived social norms, Paulhus (2002: 50) defined 'SD bias' as "the tendency to give overly positive self-descriptions." Steenkamp/Jong/Baumgartner (2010: 200) expanded this definition by explaining why most people tend to behave in this manner:

Socially desirable responses are answers that make the respondent look good, based on cultural norms about the desirability of certain values, traits, attitudes, interests, opinions, and behaviors.

Drawing on these definitions, one can define 'SD bias' as the tendency to embellish to adjust to social norms (JONG/PIETERS/FOX 2010; MICK 1996). Individuals scan their environments for indicators of a proper, socially legitimized opinion that can be presented as needed. The reason for this is that respondents wish to either create a positive self-image or meet the expectations of an interviewer or a specific group. At the same time, they attempt to avoid social sanctions and negative consequences resulting from a truthful answer. Consequently, the actually held attitude or the real behavior will not be revealed. The aim of the respondent is to adapt to a socially desirable behavior and attitude and to limit socially undesirable behaviors (TOURANGEAU/RIPS/RASINSKI 2000).

When seeking to understand the theoretical basis of social desirability, SD bias research has frequently focused on symbolic interactionism (SI) and impression management theory (IM) (GOSEN 2014). Symbolic interactionism, as understood by BLUMER (1969), is defined as the process of interaction in the formation of meanings for individuals. According to Blumer, interaction between individuals is based on autonomous action, which, in turn, is based on the subjective meaning actors attribute to social objects and/or symbols. Thus, individual actors regulate their behavior based on the meaning they attribute to objects and symbols in their interaction with others. Noteworthy in this regard is that Phillips (1971) defined data collection interviews as special forms of social action, as interactions during which respondents seek to optimize the impression they are giving. They aim to maximize both the approval they achieve and their personal satisfaction. This is done by gauging the perceived social appropriateness or desirability of their viewpoint and adapting their responses accordingly. If, by answering candidly, respondents' opinions or behaviors may be perceived by their peers as inappropriate, they will often adapt their response to avoid creating this impression.

The role of social coordination in impression theory is even more significant than in other theories. This theory holds that individuals affect other individuals, society, and their wider environment. This influence is the result of processes of interaction. During these processes, individuals consciously or unconsciously manage the impression they give (MUM-MENDEY/BOLTEN 1993). Impression management is, as Schlenker (1980: 6) emphasized, "the attempt to control images that are projected in real or imagined social interactions."

Against this background, two common approaches concerning different dimensions of 'social desirability' provide important insights (DEMAIO 1984; PAULHUS 2002; TOURANGEAU/YAN 2007). One of these approaches regards social desirability as a 'stable personality characteristic' where the focus is on 'need for social approval' and 'impression management' (CROWNE/MARLOWE 1960; KRUMPAL 2013; TOURANGEAU/YAN 2007; WINKLER et al. 2006). The second approach, which is also taken in this study, describes a context-dependent response strategy that reflects trait desirability (TD) in a particular situation. Seen in this light, the characteristic, the structure, and the comprehensibility of the items and especially the topic that is to be measured (i.e., item content) have an effect on socially desirable responses. In other words, "[this] approach to Social Desirability response bias, perceived desirability of the issue, considers behaviors or traits to be more or less socially desirable and thus discusses Social Desirability in relation to particular items" (RANDALL/FERNANDES 1991: 807).

Indeed, recent studies have identified different context-dependent factors that intensify SD bias. Numerous studies have shown that SD bias plays a role in participants' reactions to the research setting (PODSA-KOFF et al. 2003). Other studies have identified the cultural background of participants as an important factor affecting SD bias (HOPWOOD et al. 2009; MIDDLETON/JONES 2000). In addition, TOURANGEAU/YAN (2007) concluded that SD bias can be perceived especially in controversial issues such as addiction, political radicalism, income, and voting behavior. This kind of desirability vis-à-vis a certain issue is often referred to as 'trait desirability.' BELLI/TRAUGOTT/BECKMANN (2001), for example, found that 20 % of nonvoters claimed to have voted. TOURANGEAU/YAN (2007) concluded that issues that involve social norms and values are particularly prone to SD bias. The trait desirability examined in this paper is high culture.

Veblen's (1899), Linton's (1945), Bourdieu's (1979), and post-Bourdieu's theories of cultural behavior mentioned above have argued that high culture plays an important role in assigning social status and prestige. These approaches suggest that representations of art and, by extension, of cultural knowledge have been common means of achieving social recognition in Western culture. Empirical surveys have shown that human beings value participation in high culture events as an important aspect of socialization (GREATER PHILADELPHIA CULTUR-AL ALLIANCE 2009). In general, high culture is commonly associated with exclusivity and, therefore, an opportunity to increase one's symbolic capital. The Eurobarometer of 2006, for instance, showed how individuals in many different countries attach high social value to the cultural sector (EUROPEAN COMMISSION 2006; EUROSTAT 2011). Regardless of their respective definitions of culture, approximately 77 % of the participants stated that cultural events and organizations are of personal importance to them (65 % in Germany) compared with only 22 % who stated that culture is not important to them.

When analyzing SD bias in this context, the meaning of culture as an active instrument of distinction leads to the following two hypotheses:

H 1: The use of high cultural offerings is perceived to be desirable by society (high culture has a trait desirability).

H 2: Trait desirability has a positive impact on social desirability bias.

As Paulhus et al. (2003) showed, SD can be assessed as a systemic influence. This insight has important implications for measures used to generate both evaluative and behavioral data. Transferred to the cultural sector, this means that SD bias could be observed in an overreporting of responses that cite both satisfaction and frequency of visits to cultural institutions. Reuband (2007) demonstrates that in population surveys that ask about the frequency of opera visits, respondents tend to overestimate the number of visits. He suspects that this finding can be explained by social desirability effects. Writing on their project 'eMotion mapping museum experience,' Tröndle/Kirchberg/Tschacher (2014) noted that male and female participants tended to share their knowledge of art in a different manner. The authors also explained these findings in terms of SD bias. Based on these assumptions, the following hypotheses are derived:

H3: Social desirability affects reported evaluation of cultural institutions.

H4: Social desirability affects reported behavior concerning cultural institutions.



Figure 1 illustrates the model hypothesized and tested in this study.

Figure 1: The proposed model of SD bias in the cultural sector.

3. Methods

3.1 Design

To analyze the problem of social desirability in audience research, conducting research in the context of authentic settings is important. This study focused on the theater sector. More specifically, it examined one theater that presents three different genres: opera, plays, and dance. The participants were asked to visit the theater to see a play. They were assigned to one of four plays: Cat on a Hot Tin Roof, by Tennessee Williams (n=29); Sunset Boulevard, by Billy Wilder (n=31); Lucia di Lammermoor, by Gaetano Donizetti (n=72); and Murder She Wrote, by Agatha Christie (n=21). The participants completed an online survey within four days of their visit. During the first stage of this process, two types of implicit measurements were used: First, an overclaiming test, namely, the overclaiming questionnaire (OCQ), by Paulhus et al. (2003), which functioned as an unobtrusive measure of self-enhancement, was used to gather information on social desirability bias. Then, in an association test (single-category implicit association test), the participants were asked to match evaluative attributes to the plays they had watched. In the second and final stage of the survey, explicit measurements followed. The participants were asked to answer open and closed questions used to measure the social desirability of specific topics, and they reported their evaluation of the play and cultural behavior. The proposed hypotheses were operationalized using an actual visit to a theater to see a play. The following table provides an overview of these operationalizations.

	general formulation	general formulation
H1	The use of high cultural offerings is	The use of high cultural offerings is percei-
	perceived as desirable by society (high	ved as desirable by society (high culture as
	culture as Trait Desirability).	Trait Desirability).
H2	Trait Desirability has a positive im-	The perceived desirability of visits to the
	pact on Social Desirability bias.	theater has a positive impact on overre-
		porting knowledge of theater.
H3	Social Desirability affects the reported	Overreporting limits the reported perfor-
	evaluation of cultural institutions.	mance satisfaction.
H4	Social Desirability affects reported	Overreporting affects the reported behavi-
	behavior concerning cultural	or concerning the theater.
	institutions.	

Table 2: Theater-specified operationalizations.

3.2 Participants

A total of 153 students at a German university participated in this study. Of the participants, 70 % were female. The age of the participants ranged from 18 to 32 years, and the average age was 22.4 years, SD = 2.94. The students were enrolled in different academic programs: The majority of students (55 %) had a background in business administration or related fields (e.g., corporate and business law, business psychology). The other participants were enrolled in cultural studies (16 %), environmental sciences (9 %), and education (11 %). And 7 % of the participants were students in programs that were only mentioned once or twice (e.g., political sciences, digital media). Only 3 % of the participants did not provide any information on their academic background.

The participants were recruited using snowball sampling of students who were enrolled in two seminars on the topic of SD bias in audience research that were offered at the university. Every student from these seminars recruited three to five participants. Participation was voluntary. As the theater visit was part of the study, the participants in the study were granted free entrance to the theater. As students have, however, free entrance to most plays of the theater anyway, because of cooperation between the university and the theater, it can be assumed that this procedure did not result in self-selection processes. All the participants completed the questionnaire and were included in the analysis.

3.3 Trait Desirability

To evaluate the degree to which topics are 'socially loaded,' this study applied the trait desirability approach (PHILLIPS/CLANCY 1972; STOCKÉ/HUNKLER 2007). This approach assumed that trait desirability directly assesses how the respondents perceive the trait under consideration to be evaluated in society. It is based on the assumption that the participants perceive the desirability of certain traits² in society differently. For example, drinking too much alcohol is frowned upon by society. The perceived TD offers an incentive for the participants to adjust their behavior to match social expectations. For example, participants who believe that heavy drinking represents reprehensible behavior will have a strong incentive to underreport the frequency and quantity of their alcohol consumption. The participants in this study were asked to

² Trait is not understood here as a constant attribute of an individual but as a certain issue.

rate the social desirability of four characteristics: frequency of theater visits, frequency of visits to cultural institutions, extent of their knowledge of physics, and extent to which they use the social media website Facebook. For every characteristic, the participants had to rate the social desirability of (a) a strong occurrence³ (ST), (b) a medium occurrence⁴ (ME), and (c) a weak occurrence⁵ (WE) on a nine-point bipolar rating scale with a neutral midpoint. The scale ranges from -4 (perceived negatively by society) to +4 (perceived positively by society). To test hypothesis 2 (correlation between trait desirability and social desirability bias), this study used simple difference scores (SDS) (STOCKÉ/HUNKLER 2006, 2007) to include information on the TD of a weak and a strong occurrence of the characteristics.6 For SDS, the ratings of low and high intensity were transformed to present a range of 0 to 8. In a next step, SDS was calculated as the difference between a high intensity and a low intensity of characteristics: SDS = TDSTRONG - TDWEAK. The resulting scale ranges from -8 (strong incentive to provide a weak characteristic value) to +8 (strong incentive to provide a strong characteristic value).

3.4 Social Desirability

As discussed above, the term 'social desirability' can be defined as the tendency to give overly positive self-descriptions. To measure SD bias, this study applied the overclaiming questionnaire (OCQ), which has been described as "the tendency to claim knowledge about non-existent items" (PAULHUS et al. 2003: 891). The questionnaire contains both actual terms and those specifically created for the purpose of this study. Signal detection theory and, accordingly, signal detection analysis (MACMILLAN/CREELMAN 1991; PAULHUS et al. 2003) are the basis for the analysis of the participants' responses to the OCQ. In our case, an actual term is interpreted as a signal and a fictional one as the absence of a signal. Participants, then, have to distinguish between the two. They do so by indicating their familiarity with each term. The questionnaire contained 14 terms from the sector of theater; of these, seven did exist, and

³ Cronbach's alpha for culture and theater (two items) = .77 (threshold ≥ .4, see ZINN-BAUER/EBERL (2004: 21)).

⁴ Cronbach's alpha for culture and theater (two items) = .75 (threshold ≥ .4, see ZINN-BAUER/EBERL (2004: 21)).

⁵ Cronbach's alpha for culture and theater (two items) = .79 (threshold ≥ .4, see ZINN-BAUER/EBERL (2004: 21)).

⁶ Cronbach's alpha for culture and theater (two items) = .83 (threshold ≥ .4, see ZINN-BAUER/EBERL (2004: 21)).

seven were created for this study. All items are tested in a pretest. Items that were found to be very difficult or very easy were eliminated because they do not reveal any meaningful differences between individuals. The final items included in the questionnaire are listed in Figure 2. The seven terms that were invented for this study (fake terms) are given in italics.

Below you will find a number of terms from the field of theater. Please indicate how familiar you are with the term.						
Kabale und Liebe	I am familiar with it \Box	□ I never heard of it				
Revue	I am familiar with it \Box	□ I never heard of it				
Requisite	I am familiar with it \Box	□ I never heard of it				
Dramaturg [dramatic adviser]	I am familiar with it \Box	□ I never heard of it				
Draußen vor der Tür [Outside the Door]	I am familiar with it 🏼	$\hfill\square$ I never heard of it				
Little Shop of Horror	I am familiar with it 🏼	$\hfill\square$ I never heard of it				
Gewandmeister [costume designer]	I am familiar with it 🏼	$\hfill\square$ I never heard of it				
Langer Abend [A Long Evening]	I am familiar with it \Box	□ I never heard of it				
Transzendentale	I am familiar with it \Box	□ I never heard of it				
Postkompensatorisches Drama	I am familiar with it 🏼	$\hfill\square$ I never heard of it				
Skriptionsregie	I am familiar with it 🏼	$\hfill\square$ I never heard of it				
klimaxiale Umdeutung [reinterpretation]	I am familiar with it 🏼	$\hfill\square$ I never heard of it				
Paraphrasenregie	I am familiar with it 🏼	$\hfill\square$ I never heard of it				
Bühnenunterstieg	I am familiar with it 🛛	□ I never heard of it				

Figure 2: OCQ-theater questionnaire.

The participants were asked to rate their familiarity in a bipolar response system by choosing one of the following responses: 'I am familiar with it' or 'I never heard of it.' Their responses were assigned to one of the four following categories: hits, misses, false alarms, and correct rejections. Based on these dichotomized responses, the hit rate (H) and the false alarm rate (FA) were calculated. H is the proportion of terms recognized vis-à-vis all terms that were included:

H = Hits / (Hits + Misses)

FA is the proportion of falsely recognized terms invented for this study vis-à-vis all fake terms:

FA = Falses / (Falses + Correct Rejections)

To measure SD bias, the response bias c was then calculated as follows: c = (z (H) + z (FA))/2,

with z being the inverse of the cumulative distribution function of the standard normal distribution.

The higher is an individual participant's c, the higher is his or her "stylistic tendency to say 'Yes, I recognize that item' versus 'No, I don't recognize that item'" (PAULHUS et al. 2003: 891).

3.5 Cultural Evaluation: Measurement of Satisfaction

In line with the literature, we used the participants' satisfaction with the play as measurement for cultural evaluation. Two different measurement approaches were used to describe satisfaction: an explicit measurement in the form of direct questions and an implicit measurement in the form of the SC-IAT.

Explicit measures, such as questionnaires including direct questions, may be affected by response bias, for example, SD bias (NEVID/MC-CLELLAND 2010: 990). In contrast, implicit measures are less likely to be influenced by SD bias (KARPINSKI/STEINMAN 2006: 26) and are therefore more reliable in measuring attitudes, especially in the case of socially sensitive topics. For example, implicit measurements of racial attitudes have been shown to be significant predictors of race-related behavior (NEVID/MCCLELLAND 2010: 990). Therefore, we checked the overreporting bias in the culture sector by comparing implicitly measured evaluation with the corresponding explicitly measured self-reported evaluation in the survey.

The most commonly used measure of implicit social cognition is the implicit association test (IAT), which is defined as an "association-based measure of social cognition" (KARPINSKI/STEINMAN 2006: 16). The IAT, a response-time-based method, which attempts to bypass the cognitive filters of respondents, measures the strength of associative links. It consists of two computer-based discrimination tasks and measures the strength of the association between a target concept and an attribute dimension (FAZIO/OLSON 2003; HAINES/SUMNER 2006). The single-category implicit association test (SC-IAT) is a modification of the IAT. In contrast to the classic IAT, the SC-IAT uses only a single target object to measure the strength of association. The aim of this study is, as mentioned above, to measure the strength of association with the single target object, that is, satisfaction with a play. Different studies have provided evidence for the reliability and validity of the SC-IAT (BLUE-MKE/FRIESE 2008; FRIESE et al. 2016; KARPINSKI/STEINMAN 2006; STEINMAN/KARPINSKI 2008). In addition, a meta-analysis by GREENWALD et al. (2009) demonstrated that when examining racial behaviors, the IAT showed significantly higher predictive validity than direct self-report questions.

3.5.1 Implicit measurement of satisfaction: SC-IAT⁷ The SC-IAT consists of a computer-based two-phase allocation task. Both phases include 24 practice trials, followed by 72 counting trials. Using two keys on the keyboard, the participants categorize stimuli from the target category (here, photograph of a scene of the play⁸) and stimuli that are concerned with two opposite characteristic labels (here, good vs. bad) (Figure 3).



Figure 3: Operation of the SC-IAT.

In the first phase, the stimuli⁹ (see Footnote 8 and photographs of scenes of the visited play, e.g., Figure 3) are shown at the center of the screen, and the positive evaluation (good) is located on the left side, whereas the negative one (bad) is shown on the right side. The participants either press the S (positive) or L (negative) key. First, the participants are asked

- 7 A detailed description can be found in the study by KARPINSKI/STEINMAN (2006:18).
- 8 NOSEK/GREENWALD/BANAJI (2007: 270) note that "stimulus items can be presented as words, pictures, sounds, or in a combination of modalities."
- 9 Target words SC-IAT (good): beautiful, celebrating, cheerful, excellent, excitement, fabulous, friendly, glad, glee, happy, laughing, likable, loving, marvelous, pleasure, smiling, splendid, superb, paradise, triumph, and wonderful. Target words SC-IAT (bad): angry, brutal, destroy, dirty, disaster, disgusting, dislike, evil, gross, horrible, humiliate, nasty, noxious, painful, revolting, sickening, terrible, tragic, ugly, unpleasant, and yucky.

to assign the stimuli to the category that is categorized as positive. Then, the combination is reversed: the stimuli of the target category are to be moved into the area that is categorized as negative, and the participants have to press the L key. As Nevid/McClelland (2010: 993) remarked, the SC-IAT is based on the assumption that categorization of target objects will be easier and thus faster when they are paired with an evaluative category consistent with the participants' underlying attitudes than when they are paired with an incongruent category.

The duration of execution of 144 counting trials were measured in milliseconds. The procedure developed by Steinman/Karpinski (2008) was used to calculate the index: Responses that were faster than 350 ms were eliminated, and so were nonresponses. Error responses were replaced with the block mean, plus an error penalty of 400 ms. Additionally, the participants with an error rate of over 20 % were eliminated. The average response times during Phase 1 (play + good) were subtracted from the average response times of Phase 2 (play + bad). The result was divided by the standard deviation of all correct response times in phases 1 and 2.

Negative SC-IAT values indicate that the play is evaluated as bad, whereas positive ones suggest that the play is regarded as good. This implicit measurement method gives an indication of the true evaluation of the play without the cognitive filter of social desirability.

3.5.2 Explicit measurement of satisfaction

The explicit measures of satisfaction were presented after the implicit ones so that the former did not influence the latter (KARPINSKI/STEIN-MAN 2006:18). An inverse direction of this relationship could not, at the time, be observed. Three different explicit measurements in the form of a questionnaire were used:

- Satisfaction via semantic differential: three emotional term couples from the SC-IAT (excellent (= 1) vs. horrible (= 7), marvelous (= 1) vs. revolting (= 7), likeable (= 1) vs. unpleasant (= 7), and stimulating (= 1) vs. boring (=7). These four items were recoded, and their values were added and divided by four. This process resulted in a scale called semantic differential satisfaction.¹⁰
- 2. Satisfaction with the quality of performance, atmosphere, and artistic performance (ranging from 1= very dissatisfied to 7 = very sat-

¹⁰ Cronbach's alpha (four items) = .91 (threshold ≥ .7, see ZINNBAUER/EBERL (2004: 21)).

isfied). The values for these three items were added and divided by three to obtain one item called performance satisfaction.¹¹

 The overall satisfaction was measured via one item (How satisfied are you with the visit?) with a scale ranging from 1= very dissatisfied to 7 = very satisfied).

3.5.3 Cultural behavior

The influence of SD bias on reported cultural behavior was assessed by asking the participants how often they had visited a theater in the previous 12 months. Both an open question and a seven-point Likert scale were used for this purpose.

4. Major Results

To test Hypothesis 1 (general formulation see [Table 2]), the use of high cultural offerings is perceived to be desirable by society (high culture has a trait desirability), the TDs of the four surveyed characteristics (frequency of theater visits, frequency of visits to cultural institutions, extent of their knowledge of physics, and extent to which they use the social media website *Facebook*) were compared. The social desirability of a strong (ST), medium (ME), and weak (WE) occurrence of the four characteristics is shown in Figure 4.



Figure 4: Perceived TD of the analyzed characteristics at the aggregate level (n=153).

11 Cronbach's alpha (three items) = .81 (threshold \geq .7, see ZINNBAUER/EBERL (2004: 21)).

Frequent visits to cultural institutions were regarded as socially desirable by the participants, M ST = 2.83, standard deviation = 1.1. An occasional visit was also perceived as positive, M ME = 1.9, SD = 1.12. In contrast, the participants believed that not using cultural offerings would be unfavorably perceived by society, M WE = -1.90, SD = 1.26. A similar pattern was observed concerning the knowledge of physics. However, these results indicate that the participants regarded the item 'cultural visits' to be much more socially desirable than the item 'knowledge of physics.' The high values of the items 'culture strong' (M = 2.83) and 'theater strong' (M = 2.409) point to trait desirability. Conversely, the excessive use of *Facebook* is considered not desirable.

To determine the statistical significance of trait desirability, a single-sample t-test was conducted to check whether the mean values of culture (Figure 1) among the participants were different from zero. If trait desirability was not given, the characteristics examined here, especially culture, would have a value of zero. The analysis revealed a statistically significant difference between all means¹² (p < .05), supporting Hypothesis 1; that is, the use of high cultural offerings is perceived to be desirable by society (high culture has trait desirability). Hypothesis 1, visits to the theater are perceived to be desirable by society (theater visits as trait desirability), was also supported by the results.¹³

As stated in Hypothesis 2, high TD in the context of high culture, in this case theater, should result in a tendency to answer related questions in a socially desirable manner. To test this assumption, TD, using the SDS Scale introduced by STOCKÉ/HUNKLER (2006, 2007) was computed, and correlation with the response bias c_theater (individual tendency to overclaim knowledge of theater) was calculated. The results showed that the participants who had high positive SDS scores in

- 12 Mean "culture never" (M = -1.902, SD = 1.41) was lower than the score of zero, a statistically significant difference of -1.9 (95% CI, -2.13 to -1.68), t(152) = -16.648, p = .000. Mean "culture sometimes" (M = 1.882, SD = 1.68) was higher than the score of zero, a statistically significant difference of 1.88 (95% CI, 1.68 to 2.09), t(152) = 18.306, p = .000. Mean "culture often" (M = 2.830, SD = 1.05) was higher than the score of zero, a statistically significant difference of 2.83 (95% CI, 2.66 to 2.99), t(152) = 33.339, p = .000.
- 13 Mean "theater never" (M = -1.295, SD = 1.26) was lower than the score of zero, a statistically significant difference of -1.295 (95% CI, -1.50 to -1.09), t(148) = -12.547, p = .000. Mean "theater sometimes" (M = 1.737, SD = 1.12) was higher than the score of zero, a statistically significant difference of 1.737 (95% CI, 1.56 to 1.92), t(151) = 19.116, p = .000. Mean "theater often" (M = 2.409, SD = 1.10) was higher than the score of zero, a statistically significant difference of 2.409 (95% CI, 2.23 to 2.59), t(148) = 26.816, p = .000.

the sector of theater also showed a tendency to overclaim in this sector, r(146) = .139, p = .093. In this psychological context, a p-value < 0.1 is often treated as 'marginally significant.' Different factors could explain this weak correlation. First, the two measurement approaches used in this study focused on different components of social desirable responding. The TD approach examines only how socially desirable a specific trait (e.g., visits to the theater) is perceived. It does not take into account that besides the specific trait desirability, social desirable responding is also a personality characteristic (PAULHUS 2003). Because of the complexity of the construct, comparable studies show similar levels of correlations. For example, similarly low values were also found by Paulhus (2003: 894), who used a regression equation to determine the influence of constructs on the OCQ bias.14 Secondly, methodological limitations have to be taken into account. The approach used in this study to measure TD was a single-item measurement, which is often viewed critically in comparison with multi-item measurements. Against this backdrop, the empirical results of this study still appear to lend support to Hypothesis 2: the perceived desirability of visits to the theater has a positive impact on overreporting knowledge of theater. This high TD seemed to be a motivation for the participants to overreport to avoid presenting themselves in an unfavorable light.

Before Hypothesis 3 was checked, the different satisfaction measurements were compared. This was done, as shown in Figure 5, with the help of box plots. The top of the rectangle shows the upper quartile, the bottom the lower quartile of the distribution. The horizontal line in the rectangle represents the median. The lower box plot whisker extends to the lowest data value that is still within a 1.5 interquartile range of the lower quartile, and the upper one to the highest data value within a 1.5 interquartile range of the upper quartile.

¹⁴ Beta coefficients between OCQ bias and self-deceptive enhancement = .30 (p< .01), OCQ bias and self-deceptive denial = -14 (no significance is given), and OCQ bias and self-monitoring scale = .11 (no significance is given).



Figure 5: Box plots of the different measurements of satisfaction (n=153).

Figure 5 shows the different results as part of the different satisfaction measurements. The values of the SC-IAT were transformed to a sevenpoint scale to allow for a comparison with the other scales. The median of the implicit measurement through the SC-IAT (4.05, mean: 4.06) is lower than those of the other three explicit measurements (semantic differential satisfaction [median = 4.6; mean = 4.50], performance satisfaction [median = 5.67; mean = 5.22], and overall satisfaction [median = 5.0; mean = 4.89]. Also the mean differences¹⁵ suggest that the participants tended to offer positive evaluations in response to explicit measurements and thus gave socially desirable responses.

To explore the SD bias concerning reported cultural evaluation (Hypothesis 3) and cultural behavior (Hypothesis 4), the study tested the influence of the independent variables SD bias (response bias c_theater) and SDS Score on the dependent variables cultural evaluation (overall satisfaction, performance satisfaction, semantic differentiation satisfac-

15 Mean "semantic differential satisfaction" (M = 4.50, SD = 1.21) was higher than the score of 4.06 (mean of the implicit measurement through the SC-IAT), a statistically significant difference of .44 (95 % CI, .24 to .63), t(152) = 4.455, p = .000. Mean "performance satisfaction" (M = 5.22, SD = 1.22) was higher than the score of 4.06 (mean of the implicit measurement through the SC-IAT), a statistically significant difference of 1.16 (95 % CI, .96 to 1.35), t(152) = 11.698, p = .000. Mean "overall satisfaction" (M = 4.89, SD = 1.40) was higher than the score of 4.06 (mean of the implicit measurement through the SC-IAT), a statistically significant difference through the SC-IAT), a statistically significant difference of .83 (95 % CI, .61 to 1.05), t(152) = 7.309, p = .000.

tion, implicit measure of satisfaction) and cultural behavior (number of theater visits [open question and seven-point Likert scale]). The variables age and sex¹⁶ were used as control variables. Six multiple regression analyses were used.¹⁷ The beta coefficients and the significances are given in Table 3. The detailed results are included in the appendix.

	Independent variables, Beta coefficients and significance					
Dependent variables	Response Bias c theater	SDS	age	sex		
Overall satisfaction	.176*	035	146	064		
Performance satisfaction	.190*	012	080	109		
Semantic differentiation satisfaction	.181*	010	122	.023		
Implicit measure of satisfaction	035	080	064	008		
Number of theater visits (7-point-likert-scale)	.278***	.067	086	.236*		
Number of theater visits (open question)	.200*	.031	036	.060		

Table 3: *Beta coefficients of six multiple regression analyses* (more details in the appendix), $*= p \le .05$; $**= p \le .01$; $**= p \le .001$, n= 147.

As shown in Table 3, the response bias c_theater had a significant influence (at least $p \le .05$) on all reported variables. Conversely, there was no significant influence on the variable implicit measure of satisfaction, a nonreported variable.¹⁸

- 16 The variable is interpreted as a dummy variable. The control variables show little influence over the entire analysis.
- 17 Gauss-Markov assumptions were used as the prerequisites for the multiple regression analysis; the tests in this respect were mostly positive-graphical verification of the linearity of the relationships (scatter plot with the dependent variable and the studentized excluded residuals), random sample, verification of autocorrelation by Durbin-Watson statistics (values are all close to two, after which there is no autocorrelation in the residuals), the assumption of homoscedasticity was tested according to GLEJSER (1969), and multicollinearity using the collinearity statistics (tolerance and VIF (variance inflation factor) are consistently close to 1). A normal distribution of the residuals by the Shapiro-Wilk test is to be rejected, which could be altered by the inclusion of further explanatory variables so that the overall low explanatory power of the models could be improved. Therefore, the results should be interpreted with care. LUMLEY et al. (2002) assessed this problem of nonnormal distribution of the residuals as not serious and the use of the methodology as acceptable.
- 18 The remaining independent variables had no significant influence, except for the one for sex on cultural behavior (seven-point Likert scale). TRÖNDLE/KIRCHBERG/ TSCHACHER (2014) already indicated the tendency toward a socially desirable response by women in the cultural field. That a closed question favors this response is to be assumed. The handling of this phenomenon is discussed in more detail in Chapter 6.

Therefore, the results confirmed Hypothesis 3: overreporting limits the reported performance satisfaction. Hypothesis 4 was also confirmed: overreporting affects the reported behavior concerning the theater. These results show that SD can explain positive evaluations in response to, for example, cultural events such as a theatrical performance.

5. Discussion and Limitations

Ouestions concerning SD bias in surveys have received considerable attention in recent years (GOSEN 2014; KRUMPAL 2013; LEE 1993; PAULHUS 2002; PHILLIPS/CLANCY 1972; STOCKÉ/HUNKLER 2007; TOURANGEAU/YAN 2007). When this concept was first introduced, the problem of sensitive topics such as racism, voting behavior, or sexual and religious practices (DICKSON-SWIFT/JAMES/LIAMPUTTONG 2008; LEE 1993; LEE/LEE 2012) was often taken into account. Dimensions such as cultural attitude and behavior were, however, largely ignored. To address this significant gap in the literature, we used the trait desirability approach to measure the 'socially loaded' content of cultural behavior. We also applied the overclaiming questionnaire (OCQ) by examining respondents' tendency to overreport. Finally, we tested the impact of respondents' tendency to overreport in a cultural context on their reported satisfaction and behavior. To measure satisfaction, we used the single-category implicit association test (SC-IAT), an association-based measure of social cognition, as an indirect measurement.

The results of this study show that in the cultural sector, in particular theater, SD bias may play an important role. A substantial body of scholarship supports the assumption that cultural behavior is closely linked to desire for social status and prestige (VEBLEN 1899; LINTON 1945; BOURDIEU 1979; SCHULZE 1992; SHOLT 1997). The findings presented here indicate that the social desirability of culture is significantly related to the tendency to overreport in this context, which triggers positive responses. Of course, paying lip service in evaluations of cultural institutions or cultural behavior during a survey does not require much effort at all. People can easily present themselves as members of a group by adapting their responses to the perceived social norms of such group, for example, in the context of audience research by overreporting the number of theater visits or by giving overly positive evaluations of cultural institutions.

Thus, the extremely positive results of audience surveys in the past need to be questioned because of the inherent SD bias, especially because these kinds of results may lead to wrong conclusions. Since negative opinions, as our research suggests, could be obscured, at least to some extent, because of SD bias, the results of previous studies may not be well suited as a starting point for a critical analysis of the status quo. Instead, the results may be biased and hence not reliable enough to function as a proper basis for strategic development focused on audience development and audience loyalty. Therefore, avoiding, or at least controlling for, SD bias is crucial in future audience research.

Like many other studies, this one also has limitations. Some of these need to be addressed here. The first one concerns the sample; as indicated above, we only included students. Findings of previous studies suggest, however, that attitudes toward theater visits, and thus probably SD bias also, may differ considerably in terms of age, gender, or educational background (cf. FÖHL/LUTZ 2011). A survey of studies on theater audiences by Föhl/Lutz (2011) shows that theater visitors are, on average, approximately 50 years old and, compared with the general public, characterized by higher levels of education. Whereas our sample seems to be at least somewhat representative of theater visitors in general, future studies will need to include participants with different educational backgrounds to test the assumption of Reuband (2007) concerning the influence of education on SD bias, namely, that participants with low levels of education tend to show a stronger SD bias than those with high ones. As recent studies have observed differences in SD bias across cultures, a replication of this study in other countries would also be desirable. Moreover, additional studies that focus on related or different contexts are recommended. This study focused on 'high culture,' and the findings presented here need to be compared with those of studies on other dimensions of culture, such as popular culture.

Another limitation may be the use of an implicit measurement. Limitations apply to the measuring instrument itself and the comparison between implicit and explicit measurement. The SC-IAT is an extension of an IAT, and hence, reservations about the IAT also apply to the SC-IAT. The reliability and reproducibility of results, the inconsistent correlation between implicit und explicit measurements, and the lack of accuracy when predicting behavior are very much subject to debate (EGLOFF/ SCHMUKLE 2002; GREGG/KLYMOWSKY 2013; HOFMANN et al. 2005).

Despite these reservations, according to Gregg/Klymowsky(2013), using an implicit measurement is possible. A study by Steinman/ Karpinski (2008), which examined attitudes toward the clothing brand Gap and sought to predict future visits to retail stores, concluded that a SC-IAT could be used to predict behavior. Faking answers is possible in the IAT and thus presumably also in the SC-IAT (KIM 2003). According to Karpinski/Steinman (2006: 29), attempts to fake in the SC-IAT result in a high error rate. Therefore, participants with a high error rate (more than 20 %) need to be omitted from the sample. In our study, we applied this principle. However, whether this approach is sufficient to rule out faking participants has not been adequately studied yet. Nevid/ McClelland (2010: 991) observed that the difference between implicit and explicit measurement cannot be assigned to a single factor, such as SD bias. For a meta-analysis of the correlation between the implicit association test and explicit self-reporting, measures of significant moderator variables (e.g., topic, order of explicit and implicit measurement, characteristics of self-report measures) see the study by Hofmann et al. (2005). Despite the limitations that restrict the use of the SC-IAT, use of the SC-IAT was emphasized to clarify the difference between implicit and explicit evaluation. The results of this study suggest that this difference can be explained by social desirability.

6. Implications

Although researchers and social science research textbooks do not ignore the problem of the social desirability response set, the efficacy of solutions for overcoming it, and the related empirical evidence, are still unsatisfactory. Also in further research, SD bias is unlikely to be fully eliminated, even if specific methods were used. However, to at least partly address the problem of social desirability and to increase the quality of responses, or rather to obtain more truthful answers from the respondents, employing methods to increase the validity of self-reports and reduce social desirability bias is essential. Herefore most SD bias literature has explored conventional (direct) techniques, such as private setting (interviewer and bystander effect) (AQUILINO 1994; SCHUMAN/CON-VERSE 1971); data collection, such as paper-and-pencil or computer-assisted interviews (HOLBROOK/GREEN/KROSNICK 2003; LEEUW 2001; TOURANGEAU/YAN 2007); question wording, such as neutral, belittling, or defusing formulations (FOWLER JR. 1995; NÄHER/ KRUMPAL 2012; SUDMAN/BRADBURN 1982). SD bias, though, can also be reduced through the use of indirect survey techniques.

More specifically, some recent studies have provided evidence that indirect techniques can be used to collect precise and 'truthful' information from respondents concerning complex socially loaded topics (CHAUDHURI/CHRISTOFIDES 2007; KARLAN/ZINMAN 2012; KRUMPAL 2013; TSUCHIYA/HIRAI/ONO 2007). At the time of writing, two different streams of indirect measurements could be identified: on the one hand, there are individual-oriented surveys, which can be differentiated into verbal and nonverbal indirect measurement. Techniques of indirect verbal questioning seek to capture target information through distracting questions. Some noteworthy strategies include, among others, the following practices:

- Peer-interpretation technique: This technique is based on the assumption that people often project unpleasant or contradictory opinions and behaviors onto others. When discussing difficult issues, people find it easier to attribute socially inappropriate answers to others, society, young people, and so on, rather than admitting they are theirs. Fisher (1993) found that asking individuals to respond to questions indirectly from a peer perspective reduces social desirability bias. Louie/Obermiller (2000) demonstrated this technique in their study on charity donation. The trait desirability scale, as used in this study, adapted from Stocké and Hunkler's original, is also based on this assumption.
- Counterbiasing statements: Questions are prefaced with statements that attempt to justify an answer that may go against social group norms. Raghubir/Menon (1996) used this procedure to request the use of condoms; that is, "recent surveys have indicated that the majority of people struggle to use condoms."
- Bogus pipeline technique (JONES/SIGALL 1971): This technique tries to reduce the number of socially desirable biased responses by having respondents believe that they are connected to an objective procedure (e. g., lie detector) that would show the interviewer the true score of their answer, regardless of whether respondents speak the truth (JONES/SIGALL 1971; TOURANGEAU/YAN 2007). Here, the respondents were consciously deceived into thinking that the researcher had an insight into their inner processes because of fake electrodes, which are often associated with a scientific procedure.

Indirect nonverbal measurement approaches use psychologically oriented survey procedures. These methods measure spontaneous physical reactions (response times, physiological features) and regard them as indicators of effect. A well-known technique in this sector is IAT (GREEN-WALD/MCGHEE/SCHWARTZ 1998). As already mentioned, the test measures the speed of responses when respondents classify stimuli from the target category into positive and negative evaluation categories. In this situation, conscious control of the response process is reduced, thus decreasing cognitive control of the subject in terms of social desirability. A specific version, namely, the SC-IAT, was described above and applied in this empirical study. Overall, indirect nonverbal measurement seems to be a promising method for reducing socially desirable responses (HOUWER 2006). However, using this method in practice-oriented audience research is difficult because it is expensive, takes time, and is rather complex.

That said, group-oriented methods have been shown to be fruitful approaches. Focus group techniques as an interactive, interpretive approach offer particular advantages, as depth detection is one of the primary purposes of this method. Classic examples are the Pollock studies from the 1950s. During the occupation of West Germany after the Second World War, American authorities conducted polls to examine the values and (political) opinions of ordinary Germans. This survey was highly prone to trigger SD bias as few of the participants wished to admit that they had believed or even still believed in the political ideology of the Nazis. They found that the fascist attitudes of the Nazi era had weakened to a large degree. However, researchers from the Frankfurt School criticized the results because they assumed that polls would not be an adequate measurement approach for getting a true and detailed picture of values and attitudes in contemporary Germany. Pollock's later qualitative studies were therefore based on the assumption that group discussion can improve individual communicability. Interaction with others also helps individuals become aware of their latent opinions and articulate their latent and maybe also inconsistent attitudes and opinions toward an issue (LAMNEK 2005, 2008; POLLOCK 1955). If one connects the focus group interview with the concept of symbolic interaction, the underlying principle of these techniques is to increase the diversity of opinions expressed and offer different reference points, which help respondents reduce bias and describe their attitudes, ideas, and behaviors more penetratingly, more vividly, and more truthfully.

In the area of SD bias research, as well as in audience development research, this approach has thus far attracted little attention. STEW-ART/SHAMDASANI (2014) cited pragmatic reasons that limit the use of focus discussions and their reflections in real-world practice. It is worth bearing in mind, however, that in commercial market research, this focus group discussion approach is becoming increasingly important since it provides relatively reliable access to important decision-making information.

This study has provided a critical overview of the theoretical approaches describing social desirability's impact on audience research. Though valuable, theoretical insights have not led to a satisfactory set of real-world practices. To mitigate social desirability's impact on the analysis of the audience, museums, galleries, and other cultural museums should, at the very least, consider implementing the strategies and techniques outlined in the latter pages of this paper.

Authors

Sigrid Bekmeier-Feuerhahn is Professor of Communication and Cultural Management at the Institute for Management & Organization.

Paula Maria Bögel is research assistant at the Institute for Environmental Communication.

Jörg Sikkenga is research assistant at the Institute for Management & Organization.

Andreas Heinen is research assistant at the Institute for Management & Organization.

References

- AQUILINO, William S. (1994): Interview Mode Effects in Surveys of Drug and Alcohol Use: A Field Experiment. – In: *Public Opinion Quarterly* 58, 210.
- ATKINSON, Will (2011): The context and genesis of musical tastes: Omnivorousness debunked, Bourdieu buttressed. – In: *Poetics* 39, 169-186.
- BARGH, John A. (Hg.) (2007): Social Psychology And The Unconscious. The Automaticity of Higher Mental Processes. Frontiers of Social Psychology. New York: Psychology.
- BAUMGARTH, Carsten (2009): Brand Orientation of Museums: Model and Empirical Results. – In: International Journal of Arts Management 11, 30-45.
- BELLI, Robert F./TRAUGOTT, Michael W./BECKMANN, Matthew N. (2001): What leads to voting overreports? Contrasts of overreporters to validated voters and admitted nonvoters in the American National Election Studies. – In: *Journal of Official Statistics* 17, 479-498.
- BLUEMKE, Matthias/FRIESE, Malte (2008): Reliability and validity of the Single-Target IAT (ST-IAT): Assessing automatic affect towards multiple attitude objects. – In: *European Journal of Social Psychology* 38, 977-997.
- BLUMER, Herbert (1969): Symbolic interactionism. Perspective and method. Englewood Cliffs: Prentice-Hall.

- BOURDIEU, Pierre (1979): *La distinction. Critique sociale du jugement*. Paris: Les Éditions de Minuit.
- BRAUN, Henry I./JACKSON, Douglas N./WILEY, David E. (Eds.) (2002): The Role of Constructs in Psychological and Educational Measurement. New York, London: Routledge.
- CAMARERO, Carmen/GARRIDO, Ma J. (2012): Fostering Innovation in Cultural Contexts: Market Orientation, Service Orientation, and Innovations in Museums. – In: Journal of Service Research 15, 39-58.
- CHAUDHURI, Arijit/CHRISTOFIDES, Tasos C. (2007): Item Count Technique in Estimating the Proportion of People with a Sensitive Feature. – In: *Journal of Statistical Planning and Inference* 137, 589-593.
- CREATIVE TRUST FOR ARTS & CULTURE (2010): Audience Engagement Survey Report. <http://www.creativetrust.ca/wp-content/uploads/2011/02/Creative-Trust-Audi ence-Engagement-Survey-Final-Report-December-2010.pdf> [25.07.2016].
- CROWNE, Douglas P./MARLOWE, David (1960): A New Scale of Social Desirability Independent of Psychopathology. – In: Journal of Consulting Psychology 24, 349-354.
- DEMAIO, Theresa J. (1984): Social Desirability and Survey. In: Turner, Charles/Martin, Elizabeth (Eds.), Surveying Subjective Phenomena: Russell Sage Foundation, 257–281.
- DICKSON-SWIFT, Virginia/JAMES, Erica Lyn/LIAMPUTTONG, Pranee (2008): Undertaking Sensitive Research in the Health and Social Sciences. Managing Boundaries, Emotions and Risks. Cambridge et al.: UP.
- EGLOFF, Boris/SCHMUKLE, Stefan C. (2002): Predictive Validity of an Implicit Association Test for Assessing Anxiety. – In: *Journal of Personality and Social Psychology* 83, 1441-1455.
- EUROPEAN COMMISSION (2006): The Europeans, Culture and Cultural Values. Qualitative Study in 27 European Countries. Summary Report. http://ec.europa.eu/culture/ pdf/doc964_en.pdf> [25.07.2016].
- EUROSTAT (2011): *Cultural Statistics*. 2011 Edition. Internet source: http://epp. eurostat.ec. europa.eu/cache/ITY_OFFPUB/KS-32-10-374/EN/KS-32-10-374-EN. PDF (25.07.2016).
- FAZIO, Russell H./OLSON, Michael A. (2003): Implicit Measures in Social Cognition Research: Their Meaning and Use. – In: Annual review of psychology 54, 297–327.
- FILLIS, Ian (2011): The evolution and development of arts marketing research. In: *Arts Marketing: An International Journal* 1, 11–25.
- FISHER, Robert J. (1993): Social Desirability Bias and the Validity of Indirect Questioning. – In: *Journal of Consumer Research* 20, 303.
- FÖHL, Patrick S./LUTZ, Markus (2011): Publikumsforschung in öffentlichen Theatern und Opern: Nutzen, Bestandsaufnahme und Ausblick [Public research in public theaters and operas: use, state-of-the-art and outlook]. – In: Glogner-Pilz, Patrick/Föhl, Patrick S. (Eds.), Das Kulturpublikum [The culture audience]. Wiesbaden: VS, 53-125.
- FÖHL, Patrick S./GLOGNER-PILZ, Patrick (2016): Spartenübergreifende Kulturpublikumsforschung [cross-cutting culture audience research]. – In: Handbuch Kulturpublikum [handbook culture audience]. Wiesbaden: Springer, 175-205.
- FÖHL, Patrick S./NÜBEL, Damaris (2016): Das Publikum öffentlicher Theater [the audience of public theaters]. – In: *Handbuch Kulturpublikum* [handbook culture audience]. Wiesbaden: Springer, 207-253.

- FOWLER JR., Floyd J. (1995): Improving Survey Questions: Design and Evaluation. Thousand Oaks: Sage.
- FREY, Dieter/IRLE, Martin (Eds.) (1993): Theorien der Sozialpsychologie [theories of social psycology]. Berlin: Huber.
- FRIESE, Malte/SMITH, Colin Tucker/KOEVER, Marton/BLUEMKE, Matthias (2016): Implicit Measures of Attitudes and Political Voting Behavior. – In: Social and Personality Psychology Compass 10, 188-201.
- GLEJSER, H. (1969): A New Test for Heteroskedasticity. In: Journal of the American Statistical Association 64, 316.
- GLOGNER-PILZ, Patrick/FÖHL, Patrick S. (Eds.) (2016): Handbuch Kulturpublikum. Forschungsfragen und -befunde [handbook culture audience: research questions and findings]. Wiesbaden: Springer.
- GOSEN, Stefanie (2014): Social Desirability in Survey Research: Can the List Experiment Provide the Truth? http://archiv.ub.uni-marburg.de/diss/z2014/0228/pdf/dsg.pdf [25.07.2016].
- GREATER PHILADELPHIA CULTURAL ALLIANCE (2009): Research into Action. Pathways to New Opportunities. <www.wallacefoundation.org/knowledge-center/audi ence-development-for-the-arts/audience-surveys/Documents/Pathways-to-New-Op portunities.pdf> [25.07.2016].
- GREENWALD, Anthony G./MCGHEE, Debbie E./SCHWARTZ, Jordan L. K. (1998): Measuring Individual Differences in Implicit Cognition: The Implicit Association Test. In: *Journal of Personality and Social Psychology* 74, 1464-1480.
- GREENWALD, Anthony G./POEHLMAN, T. Andrew/UHLMANN, Eric Luis/BANAJI, Mahzarin R. (2009): Understanding and Using the Implicit Association Test: III. Meta-analysis of Predictive Validity. – In: Journal of Personality and Social Psychology 97, 17-41.
- GREGG, Aiden P./KLYMOWSKY, James (2013): The Implicit Association Test in Market Research: Potentials and Pitfalls. – In: Psychology & Marketing 30, 588–601.
- HAINES, Elizabeth L./SUMNER, Kenneth E. (2006): Implicit Measurement of Attitudes, Stereotypes, and Self-Concepts in Organizations: Teaching Old Dogmas New Tricks. – In: Organizational Research Methods 9, 536-553.
- HAKIM, Catherine (2000): Work-lifestyle choices in the 21st century. Preference theory. Oxford et al.: UP.
- HOFMANN, Wilhelm/GAWRONSKI, Bertram/GSCHWENDNER, Tobias/LE, Huy/ SCHMITT, Manfred (2005): A Meta-analysis on the Correlation Between the Implicit Association Test and Explicit Self-report Measures. – In: Personality & Social Psychology bulletin 31, 1369-1385.
- HOLBROOK, Allyson L./GREEN, Melanie C./KROSNICK, Jon A. (2003): Telephone versus
 Face-to-Face Interviewing of National Probability Samples with Long Questionnaires.
 In: Public Opinion Quarterly 67, 79-125.
- HOLT, Douglas B. (1997): Distinction in America?: Recovering Bourdieu's Theory of Tastes from its Critics. In: Poetics 25, 93-120.
- HOPWOOD, Christopher J./FLATO, Claudia G./AMBWANI, Suman/GARLAND, Beth H./ MOREY, Leslie C. (2009): A Comparison of Latino and Anglo Socially Desirable Responding. – In: Journal of Clinical Psychology 65, 769-780.

39

- HOUWER, Jan De (2006): What are Implicit Measures and Why are We Using Them? In: Wiers, Reinout/Stacy, Alan W. (Eds.), *Handbook of implicit cognition and addiction*. Thousand Oaks: Sage, 11-28.
- JAMES, Alison (2007): Antarctica Exhibition Evaluation (2). Under the Sea Ice computer interactive. http://www.nhm.ac.uk/resources-rx/files/antarctica-under-sea-iceevaluation-15933.pdf> [25.07.2016].
- JONES, Edward E./SIGALL, Harold (1971): The Bogus Pipeline: A New Paradigm for Measuring Affect and Attitude. – In: Psychological Bulletin 76, 349-364.
- JONG, Martijn G. de/PIETERS, Rik/FOX, Jean-Paul (2010): Reducing Social Desirability Bias Through Item Randomized Response: An Application to Measure Underreported Desires. – In: *Journal of Marketing Research* 47, 14-27.
- KARLAN, Dean S./ZINMAN, Jonathan (2012): List Randomization for Sensitive Behavior: An application for measuring use of loan proceeds. – In: *Journal of Development Economics* 98, 71-75.
- KARPINSKI, Andrew/STEINMAN, Ross B. (2006): The Single Category Implicit Association Test as a Measure of Implicit Social Cognition. – In: *Journal of Personality and Social Psychology* 91, 16-32.
- KIM, Do-Yeong (2003): Voluntary Controllability of the Implicit Association Test (IAT). In: Social Psychology Quarterly 66, 83.
- KLIMENT, Tibor (2011): Forschungsergebnisse zur Langen Nacht der Museen. [Research Results concerning Museum's Night]. http://www.rfh-koeln.de/aktuelles/meldungen/2011/langenachtergebnisse/index_ger.html> [27.07.2016].
- KRUMPAL, Ivar (2013): Determinants of Social Desirability Bias in Sensitive Surveys: A Literature Review. – In: Quality & Quantity 47, 2025-2047.
- LAMONT, Michèle/FOURNIER, Marcel (Eds.) (1992): *Cultivating Differences. Symbolic Boundaries and the Making of Inequality.* Chicago Ill. et al.: UP.
- LAMNEK, Siegfried (2005): *Gruppendiskussion. Theorie und Praxis* [group discussion. Theory and practice]. Weinheim: Beltz.
- LAMNEK, Siegfried (2008): *Qualitative Sozialforschung* [qualitative social research]. Weinheim, Basel: Beltz.
- LEE, Raymond M. (1993): *Doing Research on Sensitive Topics*. London: Sage Publications, Inc.
- LEE, Yeon-Ok/LEE, Raymond M. (2012): Methodological Research on 'Sensitive' Topics: A Decade Review. – In: Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique 114, 35–49.
- LEEUW, Edith D. de (2001): Reducing missing data in surveys: An overview of methods. In: Quality and Quantity 35, 147–160.
- LEPPÄNEN, Hanna/GRÖNROOS, Christian (2009): *The Hybrid Consumer. Exploring the Drivers of a New Consumer Behaviour Type.* (=Meddelanden från Svenska Handelshögskolan, Band 543). Helsinki: Svenska Handelshögskolan.
- LEWALTER, Doris/GEYER, Claudia (2003): Abschlussbericht der Besucherbefragung in der Sonderausstellung "Klima" des Deutschen Museums München. [Audience research report about the special exhibition "Climate" in the German Museum Munich]. <http:// www.deutsches-museum.de/fileadmin/Content/TRASH/struktur_klima3_end2.pdf> [27.07.2016].

- LINDELOF, Anja Mølle (2014): Audience Development and its Blind Spot: A Quest for Pleasure and Play in the Discussion of Performing Arts Institutions. – In: *International Journal of Cultural Policy* 21, 200-218.
- LINTON, Ralph (1945): *The Cultural Background of Personality*. Oxford: Appleton-Century.
- LOUIE, Therese A./OBERMILLER, Carl (2000): Gender Stereotypes and Social-desirability Effects on Charity Donation. – In: Psychology & Marketing 17, 121-136.
- LUMLEY, Thomas/DIEHR, Paula/EMERSON, Scott/CHEN, Lu (2002): The Importance of the Normality Assumption in Large Public Health Data Sets. – In: *Annual Review of Public Health* 23, 151-169.
- MACMILLAN, Neil A./CREELMAN, C. Douglas (1991): Detection Theory: A User's Guide Cambridge: UP.
- MICK, David Glen (1996): Are Studies of Dark Side Variables Confounded by Socially Desirable Responding?: The Case of Materialism. – In: *Journal of Consumer Research* 23, 106.
- MIDDLETON, Karen L./JONES, Jeri Lynn (2000): Socially Desirable Response Sets: The Impact of Country Culture. In: *Psychology and Marketing* 17, 149-163.
- MUMMENDEY, Hans D./BOLTEN, Heinz-Gerd (1993): Die Impression-Management-Theorie [Impression management theory]. – In: Frey, Dieter/Irle, Martin (Eds.), Theorien der Sozialpsychologie [theories of social psycology]. Berlin: Huber, 57-77.
- NÄHER, Anatol-Fiete/KRUMPAL, Ivar (2012): Asking Sensitive Questions: The Impact of Forgiving Wording and Question Context on Social Desirability Bias. – In: Quality & Quantity 46, 1601-1616.
- NEVID, Jeffrey S./MCCLELLAND, Nate (2010): Measurement of Implicit and Explicit Attitudes Toward Barack Obama. – In: Psychology and Marketing 27, 989-1000.
- NOSEK, Brian A./GREENWALD, Anthony G./BANAJI, Mahzarin R. (2007): The Implicit Association Test at Age 7: A Methodological and Conceptual Review. – In: Bargh, John A. (Ed.): Social Psychology and the Unconscious. The Automaticity of Higher Mental Processes. Frontiers of Social Psychology. New York: Psychology , 265-292.
- PATSIAOURAS, Georgios/FITCHETT, James A. (2012): The Evolution of Conspicuous Consumption. – In: *Journal of Historical Research in Marketing* 4, 154-176.
- PAULHUS, Delroy L. (2002): Socially Desirable Responding: The Evolution of a Construct. – In: Braun, Henry I./Jackson, Douglas N./Wiley, David E. (Eds.): The Role of Constructs in Psychological and Educational Measurement. London: Routledge, 49-69.
- PAULHUS, Delroy L./HARMS, P. D./BRUCE, M. Nadine/LYSY, Daria C. (2003): The Over-claiming Technique: Measuring Self-enhancement Independent of Ability. – In: *Journal of Personality and Social Psychology* 84, 890-904.
- PETERSON, Richard A./SIMKUS, Albert (1992): How Musical Taste Groups Mark Occupational Status Groups. – In: Lamont, Michèle/Fournier, Marcel (Eds.): Cultivating differences. Symbolic Boundaries and the making of inequality. Chicago: UP, 152-168.
- PETRIE, Matthew C. (2011): V&A. Case Study Evaluation of FuturePlan: Medieval and Renaissance Galleries. http://media.vam.ac.uk/media/documents/med_ren_galleries_summative_evaluation_2011.pdf> [25.07.2016].
- PHILLIPS, Derek L. (1971): Knowledge from What?: Theories and methods in social research. Chicago: Rand McNally & Co.

- PHILLIPS, Derek L./CLANCY, Kevin J. (1972): Some Effects of "Social Desirability" in Survey Studies. In: American Journal of Sociology 77, 921-940.
- PODSAKOFF, Philip M./MACKENZIE, Scott B./LEE, Jeong-Yeon/PODSAKOFF, Nathan P. (2003): Common Method Biases in Behavioral Research: a Critical Review of the Literature and Recommended Remedies. – In: *The Journal of Applied Psychology* 88, 879-903.
- POLLOCK, James Kerr (1955): German Democracy at Work: a Selective Study. University of Michigan Press.
- PRIOR, Nick (2005): A question of Perception: Bourdieu, Art and the Postmodern. In: *The British Journal of Sociology* 56, 123-139.
- RAGHUBIR, Priya/MENON, Geeta (1996): Asking Sensitive Questions: The Effects of Type of Referent and Frequency Wording in Counterbiasing Methods. – In: Psychology and Marketing 13, 633-652.
- RANDALL, Donna M./FERNANDES, Maria F. (1991): The Social Desirability Response Bias in Ethics Research. – In: *Journal of Business Ethics* 10, 805–817.
- REUBAND, Karl-Heinz (2007): Partizipation an der Hochkultur und die Überschätzung kultureller Kompetenz [Participation in the high culture and the overestimation of cultural competence]. – In: Österreichische Zeitschrift für Soziologie [Austrian Journal of Sociology] 32, 46–70.
- SCHLENKER, Barry R. (1980): Impression Management: The Self-Concept, Social Identity, and Interpersonal Relations. Belmont: Brooks/Cole Publishing Company Monterey, CA.
- SCHULZE, Gerhard (1992): Die Erlebnisgesellschaft: Kultursoziologie der Gegenwart [The experience society: A contemporary cultural sociology]. Frankfurt/M.: Campus.
- SCHUMAN, Howard/CONVERSE, Jean M. (1971): The Effects of Black and White Interviewers on Black Responses in 1968. In: *Public Opinion Quarterly* 35, 44–68.
- SIEBENHAAR, Klaus (2012): Publikumsstudie am Badischen Staatstheater Karlsruhe. [Audience Study at the State-Theatre Karlsruhe]. http://www.staatstheater.karlsruhe. de/media/docs/Staatstheater_Karlsruhe_Ergebnisdarstellung_Version121106.pdf> [27.7.2016].
- SPECHT, Inga/LEWALTER, Doris (2011): Abschlussbericht der Besucherbefragung zur Ausstellung Nano- und Biotechnologie im Deutschen Museum München. [Audience research report about the Nanotechnology exhibition in the German Museum Munich]. <http://www.deutsches-museum.de/fileadmin/Content/TRASH/Bericht_ZNT_ Besucherstruktur_Februar_2011.pdf> [27.07.2016].
- STEENKAMP, Jan-Benedict E.M/JONG, Martijn G. de/BAUMGARTNER, Hans (2010): Socially Desirable Response Tendencies in Survey Research. – In: Journal of Marketing Research 47, 199-214.
- STEINMAN, Ross B./KARPINSKI, Andrew (2008): The Single Category Implicit Association Test (SC-IAT) as a measure of implicit consumer attitudes. – In: *European Journal* of Social Sciences 7, 32-42.
- STEWART, David W./SHAMDASANI, Prem N. (2014): Focus groups: Theory and practice. Thousand Oaks: Sage.
- STOCKÉ, V./HUNKLER, C. (2006): Measures of Desirability Beliefs and their Validity as Indicators for Socially Desirable Responding. https://ub-madoc.bib.uni-mannheim. de/2603/1/dp06_03.pdf> [25.07.2016].

- STOCKÉ, V./HUNKLER, C. (2007): Measures of Desirability Beliefs and Their Validity as Indicators for Socially Desirable Responding. – In: *Field Methods* 19, 313-336.
- SUDMAN, Seymour/BRADBURN, Norman M. (1982): Asking Questions: a Practical Guide to Questionnaire Design. San Francisco: San Francisco Calif. Jossey-Bass.
- TAUCHNITZ, Jürgen (2004): Publikum im Rampenlicht [Spotlight on Audience]. http://www.marketing-tauchnitz.de/level9_cms/download_user/1928_Ergebnisse.pdf [27.07.2016].
- TOURANGEAU, Roger/RIPS, Lance J./RASINSKI, Kenneth (2000): The Psychology of Survey Response. Cambridge: UP.
- TOURANGEAU, Roger/YAN, Ting (2007): Sensitive Questions in Surveys. In: Psychological Bulletin 133, 859-883.
- TRIGG, Andrew B. (2016): Veblen, Bourdieu, and Conspicuous Consumption. In: Journal of Economic Issues 35, 99-115.
- TRÖNDLE, Martin/KIRCHBERG, Volker/TSCHACHER, Wolfgang (2014): Subtle Differences: Men and Women and Their Art Reception. – In: *The Journal of Aesthetic Education* 48, 65-93.
- TSUCHIYA, Takahiro/HIRAI, Yoko/ONO, Shigeru (2007): A Study of the Properties of the Item Count Technique. In: *Public Opinion Quarterly* 71, 253-272.
- TURNER, Charles/MARTIN, Elizabeth (Eds.) (1984): *Surveying subjective phenomena*. New York: Russell Sage Foundation.
- VEBLEN, Thorstein (1899): The Theory of the Leisure Class. New York: Random House.
- WIERS, Reinout/STACY, Alan W. (Eds.) (2006): Handbook of Implicit Cognition and Addiction. Thousand Oaks: Sage.
- WINKLER, Niels/KROH, Martin/SPIESS, Martin/OTHERS (2006): Entwicklung einer deutschen Kurzskala zur zweidimensionalen Messung von sozialer Erwünschtheit [Development of a German short scale for two dimensional measurement of social desirability]. German Institute for Economic Research. Discussion Papers, 579.
- ZINNBAUER, Markus/EBERL, Markus (2004): Die Überprüfung von Spezifikation und Güte von Strukturgleichungsmodellen: Verfahren und Anwendung. [Verification of the specification and quality of structural equitation models: method and application] Schriften zur Empirischen Forschung und Unternehmensplanung der Ludwig-Maximilians-Universität München, Heft 21/2004 [Publications on Empirical Research and Business Planning at the Ludwig-Maximilians-University Munich, Issue 21/2004].

	Dependent variables								
Indepen- dent va- riables	Overall satisfaction			Performance satisfaction			Semantic differentiation satisfaction		
Step 1	B/ Beta	t	р	B/ Beta	t	Р	B/ Beta	t	р
constant	6.237	6.161	.000	5.614	6.500	.000	5.622	6.440	.000
age	073/ 166	-1.981	.049	038/ 102	-1.213	.227	054/143	-1.695	.092
sex	.206/.068	.815	.417	.300/.117	1.391	.166	.079/.030	.361	.718
	$R^2 = .037$ F(2, 144)=2.748 p=.067			$R^2 = .029$ F(2, 144)=2.146 p=.121			R ² = .023 F(2, 144)=1.703 p=.186		
Step 2	B/ Beta	t	р	B/ Beta	t	р	B/ Beta	t	р
constant	6.143	6.028	.000	5.428	6.322	.000	5.494	6.256	.000
age	064/ 146	-1.747	.083	030/ 080	958	.340	046/122	-1.453	.149
sex	.195/.064	.769	.443	.279/.109	1.294	.198	.058/.023	.268	.789
Respon- se bias c	.561/.176	2.130	.035	.514/.190	2.295	.023	.495/.181	2.183	.031
SDS	026/ 035	424	.672	007/ 012	139	.890	006/010	122	.903
	$R^2 = .067$			$R^2 = .064$			$R^2 = .055$		
	F(4, 142)=2.539			F(4, 142)=2.420			F(4, 142)=2.066		
	p<.05			p=.051			p=.088		
	Durbin-Watson= 2.23			Durbin-Watson=			Durbin-Watson= 2.003		
				2.186					

Appendix (multiple regression analysis with control variables sex and age)

Table 4: Multiple regression analysis for the dependent variables related to cultural evaluation.

	Dependent variables								
Indepen- dent va- riables	Implicit measure of satisfaction			Number of theater visits, 7 point-likert- scale			Number of theater visits, open question		
Step 1	B/ Beta	t	р	B/ Beta	t	Р	B/ Beta	t	р
constant	4.644	5.284	.000	3.695	3.968	.000	2.162	3.037	.003
age	022/ 059	690	.491	049/ 119	-1.463	.146	018/060	711	.478
sex	059/ 023	268	.789	.749/.261	3.222	.002	.168/.080	·947	.345
	$R^2 = .003$ F(2, 144)=.246 p=.782			$R^2 = .095$ F(2, 144)=7.565 p<.01			$R^2 = .012$ F(2, 144)=.878 p=.418		
Step 2	B/ Beta	t	р	B/ Beta	t	р	B/ Beta	t	р
constant	4.812	5.376	.000	3.317	3.659	.000	1.946	2.737	.007
age	024/ 064	742	.460	036/ 086	-1.096	.275	011/036	434	.665
sex	022/ 008	097	.922	.677/.236	3.004	.003	.127/.060	.716	-475
Respon- se bias c	096/ 035	416	.678	.842/.278	3.697	.000	.443/.200	2.412	.017
SDS	051/ 080	939	.349	.047/.067	.858	.392	.031	.719	.473
	$R^2 = .012$			$R^2 = .180$			$R^2 = .058$		
	F(4, 142)=.416			F(4, 142)=7.794			F(4, 142)=2.177		
	p=.797			p<.01			p=.075		
	Durbin-Wat-			Durbin-Watson=			Durbin-W	atson= 1	.759
	son	=1.009		2.030					

Table 5: Multiple regression analysis for the dependent variables related to cultural evaluation and cultural behavior.