Measuring Subjective Movie Evaluation Criteria

Conceptual Foundation, Construction, and Validation of the SMEC Scales

Verfasser:

Frank M. Schneider

Vom Promotionsauschuss des Fachbereichs Psychologie der Universität Koblenz-Landau zur Verleihung des akademischen Grades Doktor der Philosophie (Dr. phil.) genehmigte Dissertation

Datum der wissenschaftlichen Aussprache: 4.5.2012

Vorsitzender des Promotionsausschusses:

Prof. Dr. Manfred Schmitt

Universität Koblenz-Landau

Berichterstatter:

1. Prof. Dr. Michaela Maier

2. Dr. Uli Gleich

Universität Koblenz-Landau

To Helen and Paulina—you complete me.

"... evaluative responses play a significant role—if not *the* most significant role—for understanding social behavior."

—Bertram Gawronski, *Attitudes can be measured! But what is an attitude?* (2007, p. 579)

"... evaluating films is something that we all do all of the time. Nor do I mean by this merely that we automatically form preferences for some of the films we see over others and rank some of them as better than the rest. As humans, we tend to do this with respect to most of our experiences. But with regard to film viewing, this is not something that simply happens to us automatically. It is something that we avidly pursue. Evaluating films is part of our everyday film culture."

—Noël E. Carroll, *Engaging the Moving Image* (2003, p. 148)

"I have a foolproof device for judging whether a picture is good or bad. If my fanny squirms, it's bad. If my fanny doesn't squirm, it's good. It's as simple as that."

—Harry Cohn (*1891, †1958), President of Columbia Pictures

Contents

List of Tables

List of Figures

Abstract

Audiences' movie evaluations have often been explored as effects of experiencing movies. However, little attention has been paid to the evaluative process itself and its determinants before, during, and after movie exposure. Moreover, until recently, research on the subjective assessment of specific film features (e.g., story, photography) has played a less important role. Adding to this research, this dissertation introduces the idea of subjective movie evaluation criteria (SMEC) and describes the scale construction for their measurement and its validation process. Drawing on social cognition theories, SMEC can be defined as standards that viewers use for assessing the features of films and conceptualized as mental representations of—or attitudes towards—specific movie features guiding cognitive and affective information processing of movies and corresponding evaluative responses. Studies were conducted in five phases to develop and validate scales for measuring and examining the structure of SMEC. In Phase I, open-ended data were categorized and content validated via a modified structure formation technique and items were developed. Subsequently in Phase II, participants completed an online questionnaire including revised and pilot-tested items. Exploratory factor analyses were iteratively applied to explore the latent structure and to select items. The resulting 8-factor model was cross-validated with different samples in Phase III applying confirmatory factor analyses which yielded good fit indices, thereby supporting structural validity. In Phase IV, latent state–trait analyses were carried out to examine the reliability, occasion specificity, common consistency, and method specificity of the eight dimensions. All factors—Story Verisimilitude, Story Innovation, Cinematography, Special Effects, Recommendation, Innocuousness, Light-

heartedness, and Cognitive Stimulation—are reliable and are largely determined by stable individual differences, albeit some of them also show substantial systematic, but unstable effects due to the situation or interaction. These results provide evidence for the substantive validity of the SMEC scales. Finally, in Phase V the nomological network of SMEC was explored (external validity by examining correlations with related constructs like film genre preferences and personality traits). Taken together, whereas the SMEC concept—compatible with contemporary social cognition theories—provides a framework to theorize and address research questions about the role of movie evaluation criteria and evaluative processes, the SMEC scales are the proper tool for investigating the role of these criteria and the processes they are involved in.

Zusammenfassung

Die vorliegende Dissertation beschäftigt sich mit der Messung subjektiver Filmbewertungskriterien (SMEC). SMEC können definiert werden als Standards, die Zuschauer heranziehen, um Eigenschaften von Filmen (z. B. Story, Kamera) zu bewerten. Basierend auf Annahmen aus der Social Cognition Forschung können SMEC als mentale Repräsentationen von – oder Einstellungen gegenüber – spezifischen Filmeigenschaften verstanden werden und spielen eine wichtige Rolle bei der kognitiven und affektiven Informationsverarbeitung und Bewertung von Filmen. In fünf Phasen wurden Skalen zur Erfassung von SMEC entwickelt und validiert. In Phase I wurden mithilfe von offenen Fragen Beschreibungen für Filmbewertungskriterien gesammelt, mittels modifizierter Struktur-Lege-Technik inhaltsvalidiert und schließlich Itemformulierungen abgeleitet. In Phase II wurden die Items reduziert und faktorenanalytisch auf ihre latente Struktur untersucht. Das resultierende 8-Faktoren-Modell wurde anschließend in Phase III mithilfe weiterer Stichproben und konfirmatorischer Verfahren kreuzvalidiert. Latent State–Trait Analysen in Phase IV zeigten, dass es sich um reliable und relativ stabile Dimensionen handelt. Schließlich wurde in Phase V das nomologische Netzwerk der SMEC und verwandter Konstrukte (z. B. Filmbewertungen) untersucht. Insgesamt kann festgestellt werden, dass das SMEC-Konzept einen geeigneten Rahmen bietet, um Forschungsfragen zur Rolle von Filmbewertungskriterien und Bewertungsprozessen bei der Selektion, Rezeption und Wirkung von Filmen zu formulieren und dass die SMEC-Skalen ein geeignetes reliables und valides Instrument zur Untersuchung solcher Fragen darstellen.

Prologue

In the context of the Payne Fund Studies (1929-1932), Louis Leon Thurstone developed *A Scale for Measuring Attitude Toward the Movies* (1930)—one of the first attempts to include a measure accounting for individual differences in attitudes in media effects research on movies (cf. Wartella & Reeves, 1985). Until recently, attitude research and psychological research on movies have both been growing research fields. However, whereas many scholars still agree with Allport's (1935) statement that "the concept of attitude is probably the most distinctive and indispensable concept in . . . social psychology" (p. 798), to date, research on films and movies has remained on the fringes of media psychology. For example, only 1% of media psychological publications between 1989 and 1999 dealt with research on movies or movie theater attendances (Trepte, 1999). Considering the fact that the box office reached 31.8 billion US-dollars worldwide (Motion Picture Association of America [MPAA], 2011), the lack of research is quite remarkable.

The present research embedded in my PhD thesis adds to the accretion of psychological research on movies. In my work, I pursue two objectives: First of all, my primary goal is to construct and validate a measure for subjective movie evaluation criteria. Second, by conceptualizing subjective movie evaluation criteria as attitudes towards specific film features, I aim at bridging the gap between the vast amount of attitude and social cognition research and studies on movie selection, reception, and effects—at least I hope to provide some starting points.

Acknowledgements (or Front Credits)

I am grateful to many individuals whose support made this dissertation possible. First, I would like to thank my advisors for their continual support throughout my scientific education: Michaela Maier, especially for teaching me to efficiently work in interdisciplinary and applied research projects and for funding parts of this dissertation, and Uli Gleich, especially for inspiring me to dedicate important parts of my scientific life to media psychological audience research.

My research would not have been possible without the help of Frieder Schmid who was involved in most of the studies as a student research assistant.

I extend my thanks to my colleagues—Patrick Bacherle, Joachim Marschall, Lukas Otto, Marion Rahnke, Andrea Retzbach, Karin Stengel, and Ines Vogel—especially in their function as members of the AG emKOM as well as to Tobias Rothmund who all gave valuable insights at various stages of this project, and also to Eleonore Hertweck for proofreading the manuscript so conscientiously.

I would also like to thank my parents for supporting me throughout my life and my daughter Paulina for forcing me to escape with her into another world—I wish it could have been more often.

Finally, I thank Helen Hertzsch for not only being a research companion and a critical reviewer who constantly supported and encouraged me at all stages of my dissertation project, but also for being a partner who successfully managed our everyday family life and helped me to keep my life in the proper perspective and balance—you are the right one for experiencing the good and bad movies at the cinema of life.

1.   Introduction

People love to watch movies. In 2010, attendance at German movie theaters reached 126.6 million, and the German box office reached 1.3 billion US-dollars (Spitzenorganisation der Filmindustrie e. V., 2011). The German home video market[1] added another 2.3 billion US-dollars (Bundesverband Audiovisuelle Medien e. V., 2011). In 2009, the average German TV-viewer watched movies or TV-movies 27 minutes a day (Gerhards & Klingler, 2011). No doubt, film viewing is part of our daily life.

But do you remember the last movie you have seen? Was it good or bad and why? Can you give the reasons why you chose exactly this movie? What did you feel while watching the movie? What kind of thoughts did the movie provoke? Was it innovative? After watching it for the first time, did it make you happy? Or was it even so meaningful that it changed something in your life? Can you recommend it? These questions exemplify some central issues in the research field of movie evaluations, but have been rarely examined yet.[2]

Movie evaluations are often explored as a *dependent* variable from a media effects perspective. For instance, experiencing suspense while watching a movie leads to more positive evaluations afterwards (e.g., Zillmann, Hay, & Bryant, 1975). Vorderer (1994),

---

[1] This includes sales and rentals of DVD, Blu-ray/HD-DVD, and VHS as well as video-on-demand/pay-per-view.

[2] At first glance, the recent rise of entertainment research (Bryant & Vorderer, 2006; Oliver & Nabi, 2004; Soto-Sanfiel & Vorderer, 2011; Zillmann & Vorderer, 2000) suggests that this might be the right place to find a plethora of studies dealing with movie evaluations or evaluation criteria. Unfortunately, however, the current state of research is piecemeal and entertainment scholars have just begun to scrutinize the complex, multifaceted, and multidimensional psychological antecedents and consequences of consuming entertainment media (Vorderer, 2011). To come to the point: To date, there is only little research on evaluative processes and evaluation criteria with regard to movies, although these issues are evidently vital to understand entertainment experiences. I will not present the relevant concepts and findings from entertainment research in a separate chapter in this paper, but will subsume them under the respective sections.

however, demonstrated that it can also work the other way around. He manipulated the evaluation of a movie's protagonists by giving biased prior information to the participants: Experiencing suspense depended on how much the protagonist was liked. This emphasizes the perspective on cognitive and affective processes all through the reception of a movie and the role of movie-related evaluations as a *mediating* variable. For instance, during the exposure to a movie, evaluative processes (cognitive appraisals) play an important role in eliciting emotions (cf. Lazarus, 1982; Scherer, 1998). Likewise, movie evaluations—in a more or less automatic manner—might mediate this appraisal process, too. This seems to be consistent with the results of Vorderer (1994) on experiencing suspense as mentioned above. Similarly, the feeling of empathy can only be experienced when recipients positively evaluate the protagonist of a story (Zillmann & Cantor, 1977). This also corresponds with assumptions of disposition-based theories, which focus on the viewers' evaluation of characters and consider these evaluations as prerequisites for enjoyment (e.g., moral evaluations in the context of crime dramas, Raney & Bryant, 2002; for a general overview of disposition-based theories, see Raney, 2006). Furthermore, subjective evaluations of a violent movie's content mediate the effects on the viewers (e.g., aggression; cf. Ekman et al., 1972; Früh, 2001; Grimm, 1999). Finally, we can even think of movie evaluations as *independent* variables. As word-of-mouth recommendations, film critiques, or awards (e.g., De Silva, 1998; for further studies see Chapter 2.1), they inform potential viewers about characteristics of the movie and—because they are evaluation-laden or valenced—provide her or him with cues for movie choice.

By this brief argumentation it should be obvious that research on movie evaluation would foster our understanding of cognitive and affective processes before, during, and after viewing a movie. Before we can analyze the role of movie evaluations

and their consequences, it might be important to pay attention to the formation and antecedents of evaluations or evaluative judgments related to movies. This puts us back to the start where we briefly looked at movie evaluations as a dependent variable. How are movie evaluations measured? What kinds of variables determine movie evaluations? Do we have a set of criteria we can apply when we evaluate a movie? Especially the last question concerning a set of criteria is obviously relevant for professional applications: Audience guides, film theorists, film critics, festival juries, movie associations—all of these apply more or less elaborated criteria in a more or less explicit way. Correspondingly and in line with findings from research on movie preferences (e.g., Rentfrow, Goldberg, & Zilca, 2011; Valkenburg & Cantor, 2000), the conclusion is warranted that even lay audiences develop their own evaluation criteria. How are these criteria related to movie evaluations? Do people differ in their criteria? Do we have to take the criteria into account as important moderators when explaining movie evaluations? To answer these questions, subjective movie evaluation criteria must be conceptualized, operationalized, and measured. Therefore, the two major goals of this paper are to theorize about the conceptualization of subjective movie evaluation criteria and to construct and validate appropriate scales for their measurement.

The remainder of this paper is organized as follows. In Chapter 2, I review empirical findings from movie selection, reception, and effects studies that deal with movie evaluation or related issues. I start with a general review to illustrate the broad scope of movie evaluation research and, therefore, to raise the awareness for the importance of the role of evaluation criteria. Afterwards, I focus on the theoretical framework of the social cognition approach to mental representations in order to provide the cognitive basis for understanding how movie evaluation criteria are organized in the human mind and how these mental structures are related to operating

evaluative processes. Subsequently, both subchapters are integrated into a definition and model of subjective movie evaluation criteria in Chapter 3, where I also discuss the merits and demerits of previous research that has been done on movie evaluation criteria specifically. Further, I outline the present research in Chapter 4, give an overview of the studies I have conducted, and provide a rationale for the organization of these studies in five phases that are presented in Chapters 5 to 10. Finally, I briefly summarize the results, generally discuss the findings and limitations, and conclude with suggestions for future direction.

2. The Relevance of Movie Evaluations and Movie Evaluation Criteria: Theoretical Foundations and Empirical Findings

Before presenting a framework for conceptualizing subjective movie evaluation in Chapter 3, I will lay its foundation in this chapter. Due to the sparse research that has been done on movie evaluation specifically, I will broaden the scope to the research on selection, reception, and effects of movies. In doing so, we will gain some insights into constructs that might be related to movie evaluation criteria. Studies from the research fields of film criticism and television quality will then enhance our understanding of the differences between evaluation criteria, for instance, selection criteria, modes of reception, or enjoyment of specific content. I will end this chapter with an overview of mental representations, thereby attempting to reconnect findings and examples from the first subchapter to theories and models from the research field of social cognition. Thus, in a nutshell, the major aim of this chapter is to provide all the information required to understand the framework of subjective movie evaluation criteria I will outline in Chapter 3.

2.1 Evaluation at the Movies: Insights from Research on Selection, Reception, and Effects of Movies

As mentioned above, evaluation plays an important role before, during, and after watching a film. First, in the case of movie choice, film marketing and uses and gratifications (U&G) approaches are the most common research traditions which have investigated determinants of movie selection. I will briefly review those studies that include evaluative aspects as independent or predictor variables. Second, there are only

a few studies that deal with film evaluation during the reception process. However, there are also some theoretical approaches underscoring evaluation processes during movie watching, in which evaluation is treated as a dependent as well as a mediator or moderator variable. Third, evaluation is often studied from a media effects perspective. Thus, I will discuss the relevant findings and focus on important determinants of movie evaluations as a dependent variable. In addition, relevant findings from entertainment research are incorporated in the several sections.

Selection. Why do people go to the movies? How do they choose from the large number of films available? What are the criteria they apply when deciding to watch a particular movie? Answers to these questions have often been explored from two different research traditions—the macro- and micro-level, or economic and psychological approaches, respectively (cf. Litman & Ahn, 1998). From a macro-level perspective, the most important question is: What determines the success of films? Mostly, film economists are interested in modeling the financial performance of motion pictures (for a classic example see Litman, 1983). More recently, creativity researchers discovered the macro-level approach to investigate the relationship between aesthetics, creativity, and financial success of movies (e.g., Simonton, 2005a, 2005b, 2009a). From a micro-level perspective, the focus is on the question: What variables are important in consumers' decision-making? Communication scholars as well as marketing scientists examine the reasons for moviegoing (instead of alternative leisure activities) or attending a particular film, concentrate on individual viewer characteristics (e.g., needs, preferences), and often conduct their studies in a U&G tradition (viz., motive questionnaires and factor analyzing; e.g., Austin, 1986; Palmgreen, Cook, Harvill, & Helm,

1988). In the following section, I briefly review the macro- and the micro-level approaches and discuss their contribution to film evaluation research.

The macro-level approach. This approach is best described by input–output analysis on the movie level. On the input side, (a) movie characteristics (e.g., genre, stars, director power), (b) studio actions (e.g., production budget, number of screens, advertising, timing policy), and (c) non-studio factors (e.g., movie awards, reviews, consumer ratings) can be distinguished (cf. Hennig-Thurau, Houston, & Walsh, 2007). On the output side, gross box office earnings or attendances are of main interest. To determine their predictive power, input variables are often included in multiple regression models to explain variance in the output variables. The growing body of literature has led to a vast array of findings (for recent reviews see Clement, 2004; Hadida, 2009; Hennig-Thurau, Walsh, & Wruck, 2001). Although in many cases research results are inconsistent (e.g., across different countries, cf. Elberse & Eliashberg, 2003; Hennig-Thurau & Wruck, 2000; Neelamegham & Chintagunta, 1999), it is worthwhile to look at some of the results because at least three kinds of influential variables are evaluation-laden, for instance:

- awards and nominations—because they include jury evaluations (e.g., Desai & Basuroy, 2005; Dodds & Holbrook, 1988; Gemser, Leenders, & Wijnberg, 2008; Nelson, Donihue, Waldman, & Wheaton, 2001; Prag & Casavant, 1994; Simonton, 2009a; Smith & Smith, 1986);

- critical acclaim and professional evaluations (i.e., film critiques and reviews)—because they contain evaluative judgments to argue for or against watching a certain movie (e.g., Basuroy, Chatterjee, & Ravid, 2003; Basuroy, Desai, & Talukdar, 2006; Boatwright, Basuroy, & Kamakura, 2007; Chang & Ki, 2005; Desai & Basuroy, 2005; Gemser, van Oostrum, & Leenders, 2007;

Jansen, 2005; Litman, 1983; Plucker, Holden, & Neustadter, 2008; Reinstein &

Snyder, 2005); and

- word of mouth[3]—because it contains user evaluations (e.g., Chintagunta,

   Gopinath, & Venkataraman, 2010; Cooper-Martin, 1992; Eliashberg, Jonker,

   Sawhney, & Wierenga, 2000; Liu, 2006; Mahajan, Muller, & Kerin, 1984; Moul,

   2007; Reinstein & Snyder, 2005).

Additionally, variables like genre (e.g., Desai & Basuroy, 2005; Litman & Kohl,

1989; Reinstein & Snyder, 2005; Sochay, 1994) and star or director power (e.g., Albert,

1998; Bagella & Becchetti, 1999; De Vany & Walls, 1999; Sawhney & Eliashberg, 1996;

Wallace, Seigerman, & Holbrook, 1993) might be important evaluation criteria.

The research by Simonton is similar to the economic approach, but differs in the

focus on dependent variables like winning an Oscar, movie ratings, and so forth. For

instance, Simonton (2002) was interested in how good different cinematic

components—operationalized via Oscar nominations or honors for different

categories—could predict best picture awards and movie guide ratings. Simonton found

that direction, screenplay, and film editing were the most important predictors for the

best picture award. Screenplay and direction also predicted movie guide ratings, but

with less predictive power. In another study, he factor-analyzed the 16 Oscar variables

and obtained four cinematic creativity dimensions: Dramatic, Visual, Technical, and

Musical (Simonton, 2004). Additionally, he found that best picture honors and movie

guide ratings had nothing in common when the variance they share with the four

creative factors is partialed out. The fact that even critics come to different results

regarding movie evaluation underscores the assumption that not all criteria applied by

---

[3]   To date, in most cases the so-called word of mouse—the online word of mouth—is examined.

critics or experts might be relevant to lay audience (see also Chapter 2.3 in the present paper).

Simonton (2007a) also argues that successful cinematic composition is often unrelated to movie success leading to the distinction between film as business, film as art, and film as music (Simonton, 2005b). This is even more obvious for film songs than for film scores (Simonton, 2007b). Another interesting finding is the independence of movie success measures as operationalized via later domestic gross and movie guide ratings (Simonton, 2009a). This is in line with the distinction between film as art—as appreciated by film critics—and film as business/entertainment—as represented by box-office outcomes (Holbrook & Addis, 2008). This distinction is most apparent in the drama genre. Dramas are more likely to receive best movie awards and critical acclaim, but also have smaller budgets, lower earnings and are distributed on fewer screens (Simonton, 2005b). Another relevant finding of Simonton is that bad art is the opposite of good art (Simonton, 2007c). He analyzed 877 feature films and found that Oscars and Razzies (awards for worst movies or movie-related categories) are correlated with several cinematic attributes in a similar way—with similar weights, but inversed signs. Thus, he concluded that overall cinematic quality can be evaluated on a single good–bad continuum.

Taken together, although most of the studies solely include movie-level data (e.g., aggregated consumer ratings), the macro-level approach suggests possible evaluative factors that might also operate on an individual level (e.g., exposure to critical reviews). Apparently, macro-level studies do not usually address individual motives for viewing motion pictures, nor do they take preferences, attitudes, subjective evaluations, and experiences of filmic content into account. However, the importance of these viewer characteristics has been demonstrated in recent studies, which focused on the indirect

effects of viewers' evaluations. For instance, in a study on the determinants of 331 motion picture's box office sales and profitability, Hennig-Thurau, Houston, and Walsh (2007) included consumer ratings from opening-night polls in a path analysis, thereby modeling direct and indirect influences with the usual suspects of movie success factors as mentioned above. Not surprisingly, consumer ratings turned out to be good predictors for short-term and long-term box office revenue as well as for profitability. Additionally, advertising effects on long-term box office success were mediated by consumer ratings suggesting that advertising is an important variable influencing consumers' movie choices and evaluations. Furthermore, the impact of reviews on short-term and long-term box office success is mediated by consumer ratings and awards. Finally, neither star nor director power was able to directly predict movie success on a statistically significant level. Hennig-Thurau et al. (2007) concluded:

> Consumers' quality perception is shown to be a major determinant of movies' economic success. Influencing this quality perception is a demanding task, with casting stars being an inappropriate, if not counterproductive way to improve customers' assessment of movies' quality. For studios, this finding carries an important implication, namely, that it is not sufficient to use "branded ingredients" (i.e., stars, director) for a film to become a long-term commercial success but to combine these ingredients in a way that corresponds to the *moviegoers' preferences* [emphasis added]. (p. 85–86)

The micro-level approach. The micro-level approach as well as the macro-level approach is dominated by a marketing and consumer perspective. Later in the text, I will focus on the empirical results from this kind of studies. First, I would like to introduce another kind of studies that largely draws on the U&G tradition (for overviews, see Rubin, 2009; Ruggiero, 2000) and is best described by factor analyzing motivation items. Typically, participants—very often US-American undergraduates—are asked why they

go to the movies. Subsequently, their answers are subjected to principal components analyses,[4] and the resulting components are interpreted as need structures. Until now, researchers applying this approach have been more interested in the reasons for cinema attendances than in the viewing of movies in general (e.g., Tesser, Millar, & Wu, 1988). For instance, Palmgreen and others (Austin, 1986; Palmgreen et al., 1988; Palmgreen & Lawrence, 1991) aimed to develop a motivational framework for moviegoing. Consequently, they included items specific to movie theater attendances and also found dimensions covering these aspects (e.g., "because it's a good thing to do socially"). Similar studies were carried out in Germany and replicated several dimensions (cf. Baum, 2003; Benesch, 2004).

Obviously, these dimensions (e.g., Social Utility) assess something other than film-specific evaluation or evaluation criteria; therefore, this makes it difficult to transfer results of moviegoer research to a more general research on movie choice or movie evaluation which also comprises movies on other media (e.g., TV, DVD, etc.). A rare exception to the focus on assessing selection criteria regarding motion pictures at the cinema is the study by Büch (2005), who asked his participants about criteria for selecting films in general and found that the most important criteria were genre, theme or story, and actors.

Regarding video selection, Cohen (1987) was interested in the sources of awareness and borrowing behavior of video renters and investigated their decision-making with observations and structured interviews in Israeli video cassette libraries. Usually, the renters browsed through the shelves and looked at the video titles. Important information sources for the borrowing decision were movie genre and

---

[4] Although in most of the publications reported here details on statistical analyses are often imprecisely or incorrectly labeled (e.g., factor analyses instead of principal components analyses), I will retain the wording used in the original papers for the sake of brevity.

description on the video box. Additionally, interpersonal communication with the librarian and external recommendations (e.g., friends, advertising) were frequently mentioned.

Another kind of studies comes from the diffusion of innovation research (cf. Rogers, 2003). Here researchers are more interested in the sources of information and awareness, but less in needs and motivations, even though they often couple them in their studies.

Austin (1981) surveyed 170 US-American undergraduates and found that word of mouth and theme or content of the movie were the two most frequently mentioned sources of awareness when they were asked who or what drew their attention to the last movie they attended. Furthermore, Austin created 28 items similar to the ones included in economic models (e.g., star, director, awards, reviews, etc.). Participants rated how important these variables were for their attendance decision. The three most important variables were genre, plot, and friends' comments.

In a random-digit telephone survey by De Silva (1998), 366 respondents rated creative (e.g., stars, director, etc.) and promotional (e.g., reviews, awards, advertising, etc.) variables. Together with socio-demographic (e.g., married, children, etc.) and media variables (e.g., cable TV, movie channel subscription, etc.) they were included in regression analyses as independent variables to explain movie attendance and video rental frequency. Genre, word of mouth, stars, previews, reviews, and advertising were rated as the most important for movie attendance. However, regarding the predictive power of the variables in a regression model, the most important predictors were socio-demographic variables (i.e., marital status, income, and age) followed by director, reviews, advertising, and awards. Furthermore, De Silva asked the participants what attracted them to the last video or theater. Despite some similarities (e.g., the most

frequently mentioned variable was stars), there were also differences (e.g., promotional aspects were more important for theater attendance than for home video). Additionally, home video viewers also frequently mentioned the reason to see a movie again. This adds to earlier findings that showed differences in media gratifications dependent on the medium with film as being one of the most need-specific (e.g., Elliott & Quattlebaum, 1979; Katz, Gurevitch, & Haas, 1973; Lichtenstein & Rosenfeld, 1983).

In a study by Faber and O'Guinn (1984), students rated eight sources of movie information on usefulness, importance, frequency of consultation, credibility, and impact on decision-making. For instance, previews and partner's comments were the most important sources of movie information followed by experts' and friends' comments; critical review was the least important source. The same order was found for the rating of the impact of these sources on decision making, respectively. Turning to a rare German study, Rössler (1997b) reported that advertising, word of mouth, previews, film reviews in the daily newspapers, and radio and television were the most important information sources mentioned by moviegoers.

One of the few studies that included movie characteristics items investigated the *Role of Motives and Attributes in Consumer Motion Picture Choice.* Möller and Karppinen (1983) were interested in (a) the criteria moviegoers use when they choose between movies, (b) the predictive power of these criteria regarding the preferences for genres and specific movies, and (c) the variance in these criteria depending on different audiences and genres. They formulated assumptions about the development of film- and genre-specific motives and attribute preferences and their influence on the movie choice decision-making process. Before I report on the most relevant finding, I will describe the basic assumptions (Möller & Karppinen, 1983, pp. 240-241):

(1) Movie choice and attendance are influenced by contextual variables (e.g., culture, general movie supply) and person variables (e.g., values, personality, motives, and life-style).

(2) Cinema-related motives comprise needs and reasons for moviegoing (e.g., change, relaxation, aesthetic experience, suspense, interest in human drama).

(3) External information sources (e.g., availability of movies) and internal sources (i.e., movie-related beliefs and attitudes, both resulting from earlier movie experiences or external information) can be distinguished. In Möller and Karppinen's terminology, "beliefs contain information about movie attributes (such as actors, directors, critiques, characteristics of movie types or genres, popularity etc.) and the attitudes represent fairly stable affective tendencies (preferences) towards both the attributes and movie types" (p. 241). Furthermore, they assume a mutual interplay of movie-related motives and attitudes. For instance, accumulated experience may change the beliefs and attitudes as well as the motives. Likewise, motives can be "used as criteria when the consumer dichotomizes his/her evoked set of movies into acceptable/non-acceptable sets" (p. 241). The complexity of such a set, in turn, depends on the diversity of motives.

(4) Finally, motivation alters the importance viewers attach to different movie characteristics. Thus, both movie characteristics and motives function as criteria for choosing a specific movie.

Their results of canonical correlation analysis indicate statistically significant associations between movie motives and beliefs about important movie attributes. For instance, viewers' tendency to stress characteristics like film directors and critiques was correlated with motives such as interest in human and social issues, art, culture, and

aesthetics, as well as extending views and opinions. This structure was also associated with a preference for human and social drama. Moreover, viewers' tendency to stress characteristics like popularity and media publicity was related to motives such as interest in history and nature, excitement, relaxation, and escape, and maintenance of ideals and values. The association of this structure with genre preferences yielded ambiguous results for adventure-and-thriller and for entertainment movies. This might be due to the broad categories compared to the drama genre. Finally, the authors propose a two-phase decision-making process model for movie choice. First, the viewers judge movies by comparing them according to the available internal and external information: the acceptable/non-acceptable sets mentioned above (rejection-phase). Movie-type motives are assumed to play a dominant role during this phase. In the second phase (preference phase), two or more movie alternatives that fit into the acceptable set can be preference-ordered by taking both movie motives and attributes as criteria into account. The authors assume that when alternative movies share a genre, the movie characteristics will tip the scales.

In sum, integrating characteristics of movies into an explanation of movie choice based on motives and attribute criteria offered valuable insights into the usefulness of such an endeavor. However, on the basis of their correlational design, we can of course only speculate about the causal relationship between motives, attitudes, beliefs, preferences, and attributes—their participants filled out the questionnaires after viewing a movie they selected on their own. Additionally, Möller and Karppinen provided only sparse information about their measures; therefore, we have little knowledge about the attributes—at first glance, they seem to be highly selective. Furthermore, though appealing, they offered a rather "intuitive" approach to the theoretical embedding of all the constructs they mentioned. Finally, they focused mainly

on the movie choice and attendance process in the context of movie theaters. Hence we have to question if this transfers to other media as well as to other phases relevant for movie evaluation.

The usefulness of taking movie characteristics into account when explaining movie choice was underscored in a more elaborate study. In combining an experimental design and a field study, Neelamegham and Jain (1999) found that movie information (i.e., advertising, critic reviews, and word of mouth), viewer expectations, and interest in movies influenced movie choice. Moreover, the viewers' evaluation of peripheral movie characteristics (i.e., sets, costumes, music, special effects), the felt pleasure (i.e., pleased/annoyed, satisfied/unsatisfied, contented/melancholic), and the misfit between gratifications sought and obtained were important predictors for post-receptive overall evaluation. Furthermore, the felt pleasure also raised the probability of recommending the movie to a friend. Using this recommendation as a word-of-mouth indicator and adding it to the choice behavior, Neelamegham and Jain were able to yield more precise market share predictions.

Influence of movie evaluations. Many studies have shown that movie evaluations in the selection phase (e.g., word of mouth, online user reviews, and film critiques) can influence movie choice and post-viewing movie evaluations. For instance, Wyatt and Badger (1984) experimentally showed that film review direction influenced film interest and also post-viewing evaluation. Positive reviews led to higher interest in watching a movie than negative reviews. Furthermore, positive reviews led to more positive evaluations after viewing the corresponding movie than neutral, negative, or no reviews. Additionally, manipulating the amount of information (high vs. low) in a review, the same authors (Wyatt & Badger, 1990) experimentally showed that high amount of information in neutral film descriptions raised interest in film viewing nearly

as much as positive film reviews. However, the direction of the evaluation (positive/mixed/negative) had more impact on the film viewing interest than the amount of information. Further studies also investigated the influence of characteristics of the critics (Chang, 1975; d'Astous & Touil, 1999), the relationship between movie ratings of different critics (Boor, 1990; Simonton, 2009b) or critics and viewers (Austin, 1983; Boor, 1992; Holbrook, 2005; Holbrook & Addis, 2007; Plucker, Kaufman, Temple, & Qian, 2009; Wanderer, 1970), and the inclusion of stars as brands in film plot descriptions (Hennig-Thurau & Dallwitz-Wegner, 2004; Levin, Levin, & Heath, 1997). Delving deeper into the composition of film reviews, Holicki and Krcho (1992) experimentally manipulated eight reviews of a fictitious film to show which elements had the most impact on the overall evaluation of the movie and on the interest in film viewing. They found that the description of involved persons (i.e., actors and director) highly contributed to the overall evaluation of the film. Furthermore, in a path analysis, neither descriptions of persons nor plot influenced the interest in viewing the movie. Furthermore, the only statistically significant predictor connected to film viewing interest was the evaluation of the plot. Büch (2005) experimentally replicated these findings with a real film stimulus in a sample of 15- to 18-year-old pupils. Moreover, he showed that the global evaluations as well as the evaluations of plot, actors, and director were statistically significantly more positive in the positive review condition than in the no-review and in the negative review condition. However, we have to consider the fact that film viewers are usually not exposed to reviews just before they watch a movie; therefore, it is quite possible that the reviews' impact might be due to a priming effect and, thus, only be short-termed.

In a more natural setting, Burzynski and Bayer (1977) used an ingenious design to manipulate prior information. While moviegoers in a multiplex cinema theater waited

until the showing of the previous film ended, their groups were infiltrated by confederates of the experimenters. Information about the movie was manipulated in three ways: positive comments of the confederates, negative comments, or no comments. The confederates talked loudly about the quality of the motion picture that was about to begin. After the participants watched the movie, a theater employee administered questionnaires to randomly selected patrons. The questionnaire contained a 10-point rating scale about the quality of the motion picture they had just seen. Participants in the negative prior information setting rated the movie less positively than participants in the positive prior information setting—Burzynski and Bayer even reported about people exchanging their tickets immediately after their exposure to negative information.

Gutman (1982) found that a previous promotional film—including positive expert interviews about the topic of the movie—and the participants' subjective importance of violent scenes were statistically significant predictors for the overall rating of a violent movie. This means that people with a higher preference for violence rated the movie better than people with a lower preference. Similarly, people who first saw a promotional film rated the movie better than people who were not exposed to the promotional film. Furthermore, participants rated the importance of 25 movie characteristics when they decided to see a dramatic movie. Gutman yielded a six-factor solution: (a) Wholesome Entertainment (e.g., a film for the whole family, relaxing), (b) Violence (e.g., violent, shocked me), (c) Emotional/Caring (e.g., lets you become emotionally involved, has characters you care for), (d) Realism (e.g., realistic, believable), (e) Exciting Action (e.g., action-filled, suspenseful), (f) Makes You Think (e.g., makes you think, informative).

Taken together, experimental research has shown that the selection of a movie and its global evaluation afterwards are highly susceptible to prior evaluative information. This is also consistent with the findings of a German multi-methodological field study that demonstrated that movie recommendations in TV guides are related to the television viewers' movie choices (Hasebrink & Bube, 1998, 1999).

Entertainment research on movie selection. Until recently, the U&G approach and mood management dominated entertainment research (cf. Oliver, 2009). Whereas U&G based research assumes that viewers actively and deliberately choose movies to fulfill their needs, according to mood management theory (Zillmann, 1988, 2000) viewers select movies to regulate their current mood or affective states to achieve an individually optimal level of arousal. However, they are not necessarily aware of such a process. Central to the idea of mood management is the assumption of the hedonistic need to eliminate or reduce negative mood states and enhance or maintain positive mood states (Zillmann, 1988, p. 328). In later research, this assumption was extended by the idea that mood regulation depends on its functional appropriateness (mood adjustment; cf. Knobloch, 2003). Nevertheless, there are still some difficulties to explain, for instance, the (paradoxical) choice of cognitively or emotionally challenging movies that lead to more effortful processing instead of hedonic and light entertainment (e.g., Oliver, 2009; Strizhakova & Krcmar, 2007). One recent approach to explain—for example, the watching of sad dramas—draws on additional motivations. For instance, the concept of eudaimonic motivations (Oliver & Raney, 2011) is based on two distinct forms of happiness—hedonic and eudaimonic (cf. Waterman, 1993). Whereas hedonic happiness can simply be described as pleasure (e.g., "Movies that make me laugh are among my favorites"), eudaimonic happiness is described by feelings of personal expressiveness (Waterman, 1993, p. 679) or meaningfulness (e.g., "I like movies that

challenge my way of seeing the world", Oliver & Raney, 2011). Oliver and Raney (2011) showed that movie viewers not only seek pleasure but also "truth" or "meaning" and also individually differ in their motivations. Although promising, the current state of entertainment research does not (yet) address the relationship between hedonic or eudaimonic motivations, on the one hand, and the preferences for or evaluations of specific movie attributes, on the other hand. I will return to this in the General Discussion. For now, I will end this section with describing the theory of subjective quality assessment, which addresses the effect of evaluating specific media qualities on media selection.

   Theory of subjective quality assessment (TSQA). The TSQA was introduced by Wolling (2004, 2009) and differentiates between desiring and evaluating specific features of a media object. The theory was developed to overcome weaknesses in earlier theories (e.g., the role of message form and content in U&G; for a summary, see Swanson, 1987). For instance, U&G or mood management approaches merely address anticipated gratifications or effects of use that are not related to specific movie features. From a U&G perspective "it is possible to take the various gratifications provided by the media as their features: their capacity to entertain the user, to inform the user, to provide the user with the stuff of conversation, to help the user escape from reality, and so on" (Wolling, 2009, p. 85). From a mood management perspective, Wolling continues, "features might be the product's potential to change a current mood, to raise or lower the user's level of stimulation, or to reinforce a mood" (p. 85). He convincingly concluded that these features are more effects of use and less perceived characteristics of the media product (in the case of films, e.g., story, cinematography, etc.). In contrast to the mentioned approaches, the TSQA is rather based on expectancy x value assumptions and GSGO models (gratifications sought, gratifications obtained; e.g., Palmgreen & Rayburn, 1982,

1985), but the TSQA also differs from these in two major ways: (a) Whereas expectations can be seen as the probability that a media object *will* possess a specific feature, the TSQA postulates desired qualities that can be described as "the recipient's request that a program or other media product *should* possess a certain quality" (p. 88, emphasis added). With regard to the meaning of the word *quality*, it is important to add that Wolling (2009) does not use it in a normative, but in a descriptive way (i.e., synonymously with *feature* or *characteristics*). (b) Distinct from these desired qualities is the second component: the subjective evaluation of the encountered features. In GSGO, evaluations are weighted against expectations; in the TSQA, evaluations are generated through the interplay of the desired or undesired features and the perceived features. Wolling (2009, p. 89–92) provides examples for calculating the assessment of desired and undesired features and explains the differences to expectations x value calculations. Furthermore, he briefly introduces the concept of *intrinsic tensions* (e.g., when a happy and a realistic ending are desired at the same time). Most interestingly, Wolling discusses some "early thoughts" (p. 96). For instance, he describes the relationship between a media product and three quality criteria (i.e., reality, originality, and the effects of the features). However, he states that "there is still much work to be done" (p. 97). In sum, the TSQA seems to be a useful approach that explicitly incorporates parameters that can be related to evaluative processes and criteria. It extends U&G, mood management, and GSGO by relating evaluations not only to the effects of use, but also to specific product features. This distinction is important especially for delineating movie evaluation criteria (see Chapter 3). However, the TSQA focuses rather theoretically on how desired and perceived features affect the viewers' decisions. Until today, only one study on TV series has applied this approach (Wolling, 2004). In the course of time, empirical studies will be able to show if the TSQA can be

applied to media products in general or if additional assumptions are needed to account for potential idiosyncrasies of films. Especially the findings of previous research on movie evaluation criteria that I report in Chapter 3.2 indicate that there may be more than the three criteria for films formulated by Wolling (2009).

Conclusion. What can we learn from evaluation-related research on movie selection? From a macro-level perspective, evaluation-laden, peripheral film variables such as awards, word-of-mouth recommendations, and film critiques play important roles in choosing a movie. However, audience evaluations stay "one of the most understudied variables in movie success research" (Chang & Ki, 2005, p. 252). Additionally, from a micro-level perspective, film-inherent variables including formal features and content of a movie as well as variables accounting for anticipated movie effects or gratifications sought turn out to be crucial in movie viewers' decision-making. Nonetheless, we have to keep in mind that asking "Why do you go to the movies?" or "Why do you choose a specific movie?" might not substitute for "What kind of criteria are important for evaluating a movie?", which is the most relevant question of the present research topic. Moreover, from a methodological point of view, the fact that most of the conducted analyses employed exploratory tools (e.g., principal components analysis, exploratory factor analysis) and applied ad hoc measures is psychometrically unsatisfactory. Finally and from a theoretical point of view, the TSQA as well as the research by Möller and Karppinen (1983) suggest differences between motives, criteria, subjective assessments, and effects of use.

Reception. While watching a movie, at least perceptual and cognitive processes are operating. As several theorists point out, it is plausible that these cognitive processes are—more or less automatically—evaluative. For instance, cognitive theories of emotion

focus on appraisal processes as prerequisites for eliciting emotions (cf. Scherer, Schorr, & Johnstone, 2001; for a recent overview, see Bartsch, Mangold, Viehoff, & Vorderer, 2006). These appraisals do in fact underlie more basic evaluation principles (stimulus evaluation checks concerning an individual's well-being). However, to date it is unknown how appraisals and evaluations of specific movie features are related. Nevertheless, it cannot be denied that while watching a movie appraisal processes, and therefore emotions on the one hand and evaluations regarding specific feature on the other hand, might be influenced by prior attitudes (cf. Chapter 2.2) or predispositions (cf. Chapter 3.1). Therefore, it seems likely that—similar to appraisals—evaluations might mediate the elicitation of emotions. It might also be possible that evaluations trigger different psychological and physiological states during reception or different modes of reception.

   Entertainment research on movie reception. As outlined in the movie selection section, entertainment choice research focused mainly on hedonic ideas. Distinct from this idea is the concept of engagement or involvement (for an overview, see Wirth, 2006). For instance, transportation—perhaps the most widely discussed concept related to narrative entertainment reception (Oliver, 2009)—is "a convergent process, where all mental systems and capacities become focused on events occurring in the narrative" (Green & Brock, 2000, p. 701). First attempts to broaden this concept unraveled underlying dimensions (Busselle & Bilandzic, 2009; de Graaf, Hoeken, Sanders, & Beentjes, 2009) and demonstrated the usefulness of this approach for understanding narrative persuasion. Whereas transportation, narrative engagement, or related concepts like absorption, presence, or flow are mainly concerned with "building mental models of the narrative world" (Oliver, 2009, p. 167), other approaches also include

aspects of thinking about cinematic production. One example is the modes of reception approach that I will briefly summarize in the next paragraph.

Modes of reception. Suckfüll (2004) introduced a multidimensional construct to conceptualize involvement in fictional films—the modes of reception—and to explain individual differences in information processing of the same movies. Suckfüll (2004; Suckfüll & Scharkow, 2009) reviewed early approaches to different reception modes (e.g., Liebes & Katz, 1986; Vorderer, 1992). She concluded that "four potential modes of the reception of fictional films can be derived" (Suckfüll & Scharkow, 2009, p. 368). In a data-driven process, Suckfüll (2004) conducted three studies to develop an appropriate, reliable, and valid measurement instrument—the Modes of Reception Inventory (MoRI; cf. Suckfüll & Scharkow, 2009)—and finally confirmed four factors: (a) Identity Work (e.g., "I try to identify with a film character."), (b) In-Emotion (e.g., "I am completely caught up in the story."), (c) Imagination (e.g., "I make up my own version of the plot."), and (d) Production (e.g., "I often imagine how the film was made."). The MoRI not only distinguishes between dominant modes of reception. In addition, it is assumed to explain genre preferences and movie choice as well as it seems to be related to entertainment experiences.

One of the great advantages of the MoRI—besides its convincing psychometric properties—is that it is conceptualized as independent from specific movie content. However, it cannot serve as an alternative or foundation for assessing movie evaluations because of at least three reasons. First, it aims to describe the involvement of viewers. This descriptive property might be related to or even result from evaluation processes, but does not answer the question about the use or relevance of evaluative criteria. Second, the four factors mainly comprise items referring to subjective effects of the film. Such a dominance of personal experience seems inappropriate for movie evaluation

criteria, or at least this first needs to be proven empirically. Of course, criteria might also include anticipated impressions. Nevertheless, formal features and content as specific properties of a movie have to be considered, too. For instance, evaluating formal film characteristics during a movie might be more important for an aesthetical judgment than evaluating other filmic elements (e.g., topic lines; see Suckfüll, 1997, p. 154). Third, the MoRI was developed to capture receptive processes. Thus, it focuses on what is happening during reception, albeit this might generalize to pre- or post-receptive phases. The items that proved to be as reliable and valid for measuring the reception modes (e.g., "I compare events in a film with my own experience" or "I make up my own version of the plot") might be too narrow to transfer to movie evaluation criteria. As we have seen in the section on movie selection, there is peripheral information (e.g., recommendations, awards, and advertising) that seems to play an important part in the decision-making process, thereby contributing to evaluative processes.

Conclusion. How do recipients process filmic information? Several constructs that address this question have been developed (e.g., transportation, involvement). Whereas all of these constructs deal with single, though related, phenomena, the modes of reception construct provides an integrative perspective. Although multidimensional in nature, the well-developed MoRI is more concerned with involvement than with content and formal features. Although, for instance, there is a factor called Production, it comprises only three items and does not include film features that were found to be crucial in selecting movies. When we think of evaluation as a mediating variable, it seems likely that evaluating cinematic information during watching a movie triggers different modes of reception. From a more speculative viewpoint, evaluations can be integrated into processes of emotion-eliciting and experiencing movies (e.g., entertainment, appreciation, etc.).

Effects. Movie evaluations have often been investigated from an effects perspective. Through this lens, movie evaluations are dependent variables that result from the interplay of characteristics of the person, the film, and the social situation. Examples for individual variables are motives, traits, states, and biological and socio-demographic characteristics. For instance, a person's greater preference for suspenseful films led to a better evaluation of a suspenseful movie (Vorderer, Knobloch, & Schramm, 2001). Movie evaluations were also found to be influenced by positive mood (Schmitz, Knobloch, & Vorderer, 1999) and gender (Oliver, Weaver, & Sargent, 2000). Examples for film-related variables are genre, content, design, characters, and so forth. For instance, sympathy for a protagonist increased the overall evaluation of a movie (Vorderer et al., 2001). In addition, the social situation in which the movie experience takes place (e.g., watching a movie alone or together with friends, at the cinema theater or at home) might have an impact on evaluation as well (e.g., Aelker & Bente, 2011). Although these examples describe the determinants of evaluations after watching a movie—the interplay of person, film, and situation—they are just as valid for evaluations before and during watching a movie.

One problem that arises in the context of considering movie evaluations as effects is the terminological confusion. Nabi and Krcmar (2004) provided a thorough discussion on enjoyment and related terms (i.e., liking, attraction, appreciation, preference, entertainment). One conclusion that can be drawn from this discussion is that evaluations should be seen more as constituents of enjoyment rather than as a synonym for enjoyment and the like. Especially with regard to movies, an overall movie evaluation might encompass many specific evaluations reflecting different film-inherent movie attributes (e.g., story, characters, and photography). Additionally, it might include film-peripheral information (e.g., awards, recommendations). Finally, effects of use can

be evaluated (e.g., affective effects, such as suspense, fun, disgust, and joy, or cognitive effects, such as knowledge, coherence, and mental stimulation). Similarly, Schmitz, Knobloch, and Vorderer (1999) provided a further differentiation of three kinds of evaluation: (a) overall evaluation, (b) quality evaluation, and (c) involvement evaluation. They found that these kinds of evaluation of a crime movie were mainly determined by the fit between preference for suspense and perceived suspense, the higher positive affectivity, and the sympathy for the protagonist. Additionally, quality evaluation was further determined by emotional stability, and involvement evaluation was further determined by empathy, sex, openness to experience, and negative affectivity. Especially the increase of both positive and negative affectivity is striking: Although the other predictors are in accordance with the expectations of mood management theory, this result challenges the dominant assumption of hedonic valence at least in the case of involvement evaluation (for explanations see e.g., Cacioppo & Berntson, 1994; Larsen, McGraw, & Cacioppo, 2001). Similarly in an experiment based on excitation-transfer theory, Vorderer and Bube (1996) manipulated empathic stress of recipients and the ending of a movie. They showed that inducing empathic stress led to an increase in negative emotional states and emotional distress. However, a different result was found for the evaluation of the movie: Experiencing empathic stress led to a better evaluation of the movie regardless of the ending. Although they cautiously interpreted their findings with regard to their ad-hoc movie evaluation scale (ten items, no psychometric properties reported), they concluded that their results questioned the appropriateness of applying excitation-transfer to movie evaluation research and that positive evaluation of a movie might be independent from the felt pleasantness of one's emotional state.

Although the inclusion of single-item evaluation measures in post-viewing questionnaires usually leads to poor psychometric quality (cf. Wegener & Fabrigar,

2004), this procedure is common in media effects research; however, even in the case of single-item measures, insightful analyses of movie evaluations are rarely found. The two studies mentioned above (Schmitz et al., 1999; Vorderer & Bube, 1996) are exceptions, and both indicate that more research is needed, especially research using multiple items and well-founded instruments. A recent effort to delineate entertainment gratifications can be seen as a first step in this direction: Oliver and Bartsch (2010) developed a multi-scale approach for assessing more differentiated audience evaluations. They found four experiential dimensions: Fun, Thought-Provoking, Suspense, and Lasting Impression. Whereas Fun and Suspense can be related to established approaches (e.g., mood management, sensation seeking), the authors also introduced a new concept, namely appreciation, which is defined as "an experiential state that is characterized by the perception of deeper meaning, the feeling of being moved, and the motivation to elaborate on thoughts and feelings inspired by the experience" (Oliver & Bartsch, 2010, p. 76).

Conclusion. Taken together, all approaches that have focused on movie evaluations as effects clearly conceptualized them as cognitive or affective audience responses to movies in general rather than as related to specific movie characteristics. However, according to Möller and Karppinen (1983) or Wolling (2009), this is only one part of the story. Still we know little about how general gratifications interact with specific movie content. One solution might be to investigate whether viewers have general movie evaluation criteria that they apply when evaluating a movie and that can be directly connected to evaluations of specific film features. I will return to this idea in Chapter 2.2 and outline it in Chapter 3 in more detail. Another conclusion that can be drawn from all three research areas, but particularly from the research on movie evaluations as a dependent variable, is the lack of an appropriate instrument for

measuring evaluations of specific movie features. Although the scales by Oliver and Bartsch (2010) are well-developed, validated, and thus promising extensions of single-item overall movie evaluation measures, they have at least three serious shortcomings in reference to my research aim. First, answering the items presupposes that the respondent has seen a certain movie (e.g., "I was moved by this movie."). Although this could be easily solved by rewording the items (e.g., "It is important for me that a movie moves me.") to measure—in the terminology of Wolling (2009)—a desired quality, it would be necessary to repeat the validation process for the reworded items. Second, the items focus only on experiences (i.e., effects of use) and contain neither evaluations of film-inherent features (e.g., story, photography) nor evaluations of film-peripheral features (e.g., awards, recommendations). Third, the scales were developed to explore and distinguish among several facets or dimensions of entertainment experiences (e.g., enjoyment, meaningfulness). These experiences, however, might result from evaluative processes. Again, the evaluations of specific movie characteristics might be the constituents of overall evaluations or related concepts such as enjoyment.

To sum up, on the one hand, we learned from this review that movie evaluations play a crucial role in movie selection, reception, and effects. On the other hand, with only a few exceptions to date, most studies investigated movie evaluations as a by-product. This is reflected in the use of single-item measures or ad hoc scales.

We must carefully differentiate between desired features (cf. Wolling, 2009) or criteria for evaluation that might result from motives, values, or personality traits (cf. Möller & Karppinen, 1983) and perceived features that might result from the interaction of a viewer with a specific movie leading not only to evaluations of multiple experiences (cf. Oliver & Bartsch, 2010) but also to evaluations of film-inherent and film-peripheral features. Whereas the former can be conceptualized as mental representations or

attitudes towards specific movie features, the latter can be seen as evaluative responses resulting from activated representations. I will focus on this idea in the following section.

2.2 Mental Representation of Movie Evaluation Criteria and the Processing of Cinematic Information

Representations are encodings of information in memory (Smith & Queller, 2001, p. 111) or "cognitive structures that reflect acquired knowledge and experience, and that provide the material on which cognitive processes operate" (Carlston, 2010, p. 39). *Mental representation* is an umbrella term for cognitive constructions like memories, concepts, schemata, scripts, mental models, situation models, prototypes, exemplars, associative networks, connectionist networks, attitudes, impressions, stereotypes, and so forth (cf. Bodenhausen, Macrae, & Hugenberg, 2003; Carlston, 2010; Smith & Queller, 2001). Despite several common principles (e.g., representations differ in their accessibility; Carlston, 2010), each item in the list above implies its own theoretical, and sometimes even philosophical, background assumptions.[5] For instance, broadly speaking, schematic representations stress the summation and storage of past experiences in independently organized knowledge structures, whereas associative representations emphasize how knowledge is acquired and used (Smith & Queller, 2001).

In the next paragraphs, I selectively give a brief overview of three models of mental representations: associative network models because they are fundamental to understand the basic principles of memory and information processing (e.g., network of

---

[5] For instance, schema theory has its roots in Kantian philosophy and Gestalt theory, whereas the ideas of associative networks can be traced back to Aristotle and British empiricism (e.g., Hobbes, Hume, and Locke).

concepts, spreading activation), schema theory because of its prominence in media research and its usefulness especially for describing the acquisition of cinematic knowledge or development of criteria, and attitudes because they are directly connected to evaluations, thereby providing the most appropriate way to think of movie evaluation criteria as mental representations.

Associative Networks (AN). The AN approach aims to generally model human memory (cf. Anderson & Bower, 1973).[6] Representations in AN approach—mostly referred to as *concepts*—consist of nodes connected by links that vary in strength. For instance, a concept of James Bond can consist of several nodes such as male, physical attractiveness, armed, technical gimmicks, beautiful women, fast cars, secret service, and so forth. Of course, it can also include the written word "James Bond" as a node. These nodes themselves can be concepts (e.g., the concept of physical attractiveness can consist of further related nodes like specific height and weight, athletic body, etc.). The connections between nodes can vary in strength (e.g., the node of secret service might be stronger associated with the "James Bond" node than the physical attractiveness node). The central operating process in AN is called spreading activation (cf. Collins & Loftus, 1975). Once a node like "James Bond" is activated (e.g., by the perception of the number 007), its activation can spread more easily to nodes that are closer or that have stronger connections (e.g., secret service) than to nodes that are farther away or that have weaker connections (e.g., Jane Austen). Furthermore, the activation of connected nodes might depend on the number of connections (e.g., the node of fast cars or car chases might also be connected to other action movies; cf. Anderson, 1974). If we think

---

[6] Although current research on memory and cognition processes tends to emphasize distributed connectionist models, associative network models are more prevalent and more closely connected to the relevant attitude concepts dealt with in my thesis. See Smith (2009) for a recent review about the appeal of distributed connectionist models for social psychology.

of more organized knowledge structures on a higher, broader level (e.g., comedies), a concept can also be regarded as a main category subsuming subcategories on several sublevels. This is often referred to as a schema.

Schemata. Schema theory has received wide attention in media research (for an overview and critique, see Matthes, 2004). Especially film studies and media psychological research on cognitive processing and comprehension of films, narration, and film genres have dealt with the schema concept (e.g., Bordwell, 1985; Gehrau, 2001; Ohler, 1994; Rusch, 1987; Schwan, 2001; Wuss, 1999[7]). A schema "refers to an active organisation of past reactions, or of past experiences" (Bartlett, 1932, p. 201). This knowledge organization shapes how new stimuli are perceived and interpreted (top-down processing) after it is triggered by incoming stimuli (bottom-up activation). Regarding movies, story schemata and genre schemata are relevant knowledge structures for film comprehension (cf. Bordwell, 1989; Mandler, 1984; Ohler, 1994). For instance, Schwan (1995) demonstrated that activated schematic knowledge structures (i.e., genre schemata) facilitated recalling and appropriately categorizing relevant scenes after viewing a movie. The same schemata might influence the evaluations of the same content in different ways. Evidence comes from a study by Potter, Pashupati, Pekurny, Hoffman, and Davis (2002). They found that viewers judging a violent TV show episode relied more on their personal schema of violence than on the perceived content. Viewers seemed to share a story schema of violence (e.g., with regard to explicitness and graphicness). However, viewers made different judgments, although they watched the same show and applied the same schema. This led Potter et al. to the conclusion that viewers can watch the same content (e.g., a violent TV show), rate it with the same

---

[7]   Wuss, however, also used the term "genre stereotype" (Wuss, 1992, 1999, 2009).

criteria (explicitness), but still end up with different judgments (low violence because of perceived low explicitness or high violence because of perceived high explicitness). The authors offer a plausible explanation: The carefulness of the viewers while watching the TV show might have influenced their subjective interpretations of what they had seen. Another explanation might be that the viewers differed in their schema structure (e.g., even though they applied the same explicitness criterion, some might have attached more importance to this criterion than others).

How and when are story or genre schemata acquired? Schemata develop during life through basic principles of human learning and understanding the world: assimilation and accommodation (i.e., integrating new information into an existing schema and modifying an existing schema to fit new information, respectively, Piaget, 1926). Hence story schemata are thought to already evolve in early childhood (Kintsch, 1977). As Raney (2004) stated, through repeated exposure we learn

> how similar stories are constructed, how typical actions relate to one another, how scenes and settings are constructed, and how themes are repeated, among other things. Over time, a viewer develops various schema structures that are activated when a subsequent media text is encountered. These structures then guide expectations about and interpretation of the ongoing narrative and the characters involved. (p. 353-354)

Schemata (narrative or story schemata, genre schemata) are conceptualized as pure knowledge structures; therefore, they usually contain neither evaluative cues nor connections to affective structures. Drawing on the analyses of Pratkanis (1989; Pratkanis & Greenwald, 1989) about the cognitive structure of attitudes, it can be argued that schemata of stories, genres, or film features might not be sufficient to evaluate movies or movie features—they might not even be necessary because recent conceptualizations of attitudes or associative networks include cognitive knowledge

structures and can mimic schemata (Smith & Queller, 2001, p. 127). However, it is hardly possible to argue against the fact that film-relevant cues (e.g., film critiques, audio-visual film stimuli) indeed do activate knowledge structures, which guide and influence information-processing of movies. According to Smith and Queller (2001) this dilemma might be solved by considering a schema "more as a description of a *function* that can be performed by a learned knowledge representation . . . than a description of an actual entity inside our heads" (p. 127, emphasis in original). Certainly, such a function is one of the reasons for holding an attitude (e.g., Katz, 1960; Pratkanis & Greenwald, 1989; see also next section).

Attitudes. Research on attitudes—one of the most indispensable concepts in social psychology (Allport, 1935, p. 798)—provides a vast amount of empirical findings and theoretical approaches. Surprisingly, if anything, theory building in media psychology and communication science until now has mainly focused on traditional theories and models such as the tripartite model of attitude (e.g., Rosenberg & Hovland, 1960), expectancy–value theories (e.g., Fishbein & Ajzen, 1975), or classic dual-process theories of persuasion (e.g., Chaiken, 1980; Petty & Cacioppo, 1986). Although some explicit attempts have been made to theoretically conceptualize media phenomena as attitudes—for instance, media enjoyment (Krcmar & Renfro, 2005; Nabi & Krcmar, 2004), media gratifications (Palmgreen & Rayburn, 1982, 1985), preferences for TV program types (Doll & Hasebrink, 1990), or sad film paradox (Mills, 1993)—they rely heavily on traditional views on attitude. However, a recent search on the PsycINFO-database with Boolean operator "attitud*" yielding over 132,000 hits between 2001 and

2010[8] (1991–2000: approx. 75,000) clearly indicates that the current research in attitudes is rapidly evolving—directly accompanied by theoretical and methodological advances. Taking these new approaches into account might foster audience and movie research as well.

Although a plethora of attitude definitions exist (for a recent overview and categorization, see Bohner & Dickel, 2011), the umbrella definition of an attitude as "*a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor*" (Eagly & Chaiken, 1993, pp. 1, emphasis in original) has received widest acceptance (e.g., Johnson, Maio, & Smith-McLallen, 2005, p. 618). Early attempts to link attitude and social cognition research (i.e., conceptualize an attitude as a mental representation) date back to the time between 1965 and 1985 (cf. McGuire, 1986). By the 1980s, this integrated perspective was widely accepted. For instance, from a structural perspective, the cognitive representation of an attitude was viewed as consisting of an object-category, an evaluative summary of the corresponding object, and a supporting knowledge structure (Pratkanis, 1989; Pratkanis & Greenwald, 1989) or simply as object–evaluation associations (e.g., Fazio, Chen, McDonel, & Sherman, 1982; for a recent overview see Fazio, 2007). Fazio and colleagues' provided evidence that accessible attitudes (i.e., strong object–evaluation associations) in particular guide information processing and behavior (for overviews see Fazio, 1995, 2000). Therefore, chronically accessible attitudes toward specific film features are supposed to play an important role in movie choice, reception, and effects. Such attitudes could become activated during these three phases when relevant, salient information is encountered. For instance, a film poster displaying a monster from a creature feature horror movie

---

[8] For the same period of time, a search with Boolean operators "(attitud*) AND (movie OR film)" led to 777 hits (1991–2000: 294), however, most of them referred to film effects on attitudes, which are only related to story content (retrieved from PsycINFO on 29th September 2011).

might trigger a related attitude when choosing among movies (e.g., *A Nightmare on Elm Street*), whereas an unexpected appearance of such a creature (e.g., in *From Dusk Till Dawn*) might trigger the attitude during watching the movie, and thinking about the final scene suggesting that all horror was only a dream of the protagonist (e.g., *Dead of Night*) might trigger a totally different attitude but only after viewing the movie. Attitude accessibility is just one indicator for attitude strength (for a review, see Petty & Krosnick, 1995). Strong attitudes are assumed to be stable over time, resistant to persuasion, influential in information processing, and predictive of behavior (cf. Krosnick & Petty, 1995). Another key indicator for attitude strength is attitude importance (e.g., Boninger, Krosnick, & Berent, 1995; Boninger, Krosnick, Berent, & Fabrigar, 1995; Eaton & Visser, 2008; Festinger, 1957; Holbrook, Berent, Krosnick, Visser, & Boninger, 2005; Karpinski, Steinman, & Hilton, 2005; Krosnick, 1988a, 1988b; Rosenberg, 1956; Sherif & Hovland, 1961; Starzyk, Fabrigar, Soryal, & Fanning, 2009; Zuwerink & Devine, 1996). Attitude importance can be described as a meta-attitude (Bassili, 1996) or a belief that links an attitude to an attribute (i.e., varying personal significance). It is assumed to originate from self-interest, social identification, and values and is typically measured by self-reports (cf. Boninger, Krosnick, Berent, & Fabrigar, 1995; Eaton & Visser, 2008). One practical feature of the attitude importance construct lies in its general applicability. For instance, a strong attitude toward horror movies should be easily reportable, even without a present horror film stimulus. In sum, attitude strength plays a crucial role in understanding attitudes (Bassili, 2008; Petty & Krosnick, 1995; Visser, Bizer, & Krosnick, 2006).

Despite the evident appeal of the conceptualization of an attitude and its constituents (i.e., evaluative summary, object category, and knowledge structure) as mental representations, traditional perspectives have focused on cognitive, affective,

and behavioral components of an attitude (e.g., Breckler, 1984; Rosenberg & Hovland, 1960). The contemporary view, however, distinguishes an attitude from affect, cognition, and behavior, but regards them, on the one hand, as antecedents or informational bases from which evaluative summaries are derived (e.g., Cacioppo, Petty, & Geen, 1989; Crites, Fabrigar, & Petty, 1994; Fabrigar, MacDonald, & Wegener, 2005; Zanna & Rempel, 1988) and on the other hand, as consequences or expressive responses (e.g., Eagly & Chaiken, 1993; Zanna & Rempel, 1988). The investigation of the antecedents and consequences of attitudes in this way is called *neotripartite* analysis by some authors (e.g., Eagly & Chaiken, 2007). For instance, research provides evidence that there are stable individual differences in attitude structure. Some attitudes have been found to be based more on affective information and expressed by affective responses, whereas others were based more on cognitive information and expressed by cognitive responses (Crites et al., 1994; Giner-Sorolla, 2001; Haddock & Zanna, 1998; Huskinson & Haddock, 2004, 2006). Consistently, mental representations include connected cognitive, affective, and behavioral knowledge or concepts (Bassili & Brown, 2005; Conrey & Smith, 2007). For instance, in the case of horror movies the formation of an attitude toward terrifying creatures can be based on past behavioral (e.g., biting nails), affective (e.g., fright), or cognitive (e.g., mistaking tree trunks for monsters when walking through a forest) responses to a horror movie.

But why do we hold attitudes? Most importantly, attitudes fulfill a knowledge or object appraisal function (cf. Katz, 1960; Smith, Bruner, & White, 1956). This refers to what Fazio called "the general utility of simply holding an attitude, regardless of its valence" (2000, p. 3). In this sense, attitudes towards movie features help to organize cinematic knowledge and guide or influence information processing (e.g., during reception), which also resembles the functions of a schema (see above), as well as

approach- and avoidance-behavior (e.g., movie selection). Attitude-object functions are often confused with attitude functions. Kruglanski and Stroebe (2005) argued that "functions served by the attitude object refer to goals that the attitude object may help to attain" (p. 341). To clarify this: Forming or maintaining an attitude towards, for instance, features of horror movies helps to structure knowledge about these features and respond faster when they appear (object appraisal function). In contrast, the reason for forming a *negative* attitude toward these features (e.g., to avoid sleepless nights) refers to the attitude object (e.g., horror movies). Several attitude-object functions have been identified across attitude objects (e.g., social adjustment, ego-defense, value expression, etc.; Maio & Olson, 2000).

This brief summary of three metaphors of mental representations has provided some insights about how cinematic information can be stored and retrieved. Because attitudes include evaluative information by definition, the attitude metaphor seems to be the most appropriate model for conceptualizing movie evaluation criteria. However, simply defining an attitude is not sufficient to account for empirical phenomena. For instance, the attitude concept must be embedded in an attitude theory in order to have explanatory power and deal with at least two current major problems: (a) the implicit–explicit distinction and (b) context sensitivity (cf. Bassili & Brown, 2005). Although several solutions have been proposed (for a recent overview, see Gawronski, 2007), dual-process theories seem to be the most appropriate ways to handle the two problems mentioned above. Thus, the next section deals with an example of a dual-process theory, the Associative-Propositional-Evaluation (APE) model.

Dual-process theories. In the last two decades of the 20th century, theorizing in social and cognitive psychology was strongly influenced by domain-specific dual-process models (Payne & Gawronski, 2010; for an overview, see Chaiken & Trope,

1999). Since then, several attempts have been made to integrate these domain-specific approaches into generalized dual-process models (e.g., Gawronski & Bodenhausen, 2006a; Rydell & McConnell, 2006; Smith & DeCoster, 2000; Strack & Deutsch, 2004), which distinguish between associative and propositional processes. Besides the large body of evidence from experimental social and cognitive psychology provided by the above mentioned authors, this distinction is also supported by findings from social neuroscience and physiology (e.g., Cunningham, Johnson, Gatenby, Gore, & Banaji, 2003; Cunningham, Packer, Kesek, & van Bavel, 2008; Cunningham, Raye, & Johnson, 2004; Cunningham & Zelazo, 2007; Lieberman, 2007).

APE model. One of the recently most influential models that conceptualize attitudes as mental representations and describe two underlying processes for evaluative responses is Gawronski and Bodenhausen's APE model (2006a, 2006b, 2007, 2009, 2011; Gawronski, Strack, & Bodenhausen, 2008). Gawronski and Bodenhausen focus on the processes—not on the representations (cf. Smith & DeCoster, 2000; Strack & Deutsch, 2004). They argue that the mere activation of an associative pattern is associatively processed, thereby leading to a first kind of process-specific evaluative response—affective reactions (or implicit evaluations). These affective reactions are neither true nor false. In the case of watching a creature feature horror movie, for instance, affective reactions become obvious in physiological reactions (e.g., facial expression, skin conductance). Of course, these reactions and accompanied feelings are totally unfounded because an adult recipient knows that these creatures are fictitious and only on the screen. However, these two aspects of reasoning and validating activated patterns refer to the propositional process in which the recipient thinks about what's on his or her mind. The propositional processing of an activated pattern also leads to a second kind of process-specific response—evaluative judgments (or explicit

evaluations). As might be clear from this example, both kinds of process-specific responses are not only conceptually different—they also need to be measured with different methods. Whereas the spontaneous responses (i.e., affective reactions) that are triggered by associative processes might best be assessed by applying indirect measures (e.g., physiological indicators, response times), the more deliberative responses (i.e., evaluative judgments) that result from propositional reasoning are usually assessed by applying direct measures (e.g., interviews, questionnaires). Although conceptually different, both processes do not work independently from each other (Gawronski & Bodenhausen, 2006a, Strack & Deutsch, 2004). In the default mode, propositional processes are fully mediated by associative processes. Thus, affective reactions and evaluative judgments should be in line with each other.[9] Finally, it should be noted that Gawronski and Bodenhausen (2011, p. 104) clearly state that they assume a single underlying mental representation. Hence, they argue against dual-attitude models (e.g., Wilson, Lindsey, & Schooler, 2000).

Conclusion. Mental representations—and related terms—are useful metaphors that can be applied to movie evaluation research in general. Specifically, regarding movie evaluation criteria, the attitude concept seems to be appropriate for conceptualization. Looking at movie evaluation criteria from a social cognition perspective has advantages due to the vast amount of empirical research and the variety of theoretical progress in this field. Additionally, it allows for borrowing the advanced measurement methods used in implicit social cognition as well as social neuroscience for the conduction of media psychological research on movies. The APE model offers a

---

[9] In this case, the reasons for non-corresponding outcomes—the so-called implicit–explicit inconsistency—do not lie in diverging processes. However, they might be due to moderator effects (e.g., Gschwendner, Hofmann, and Schmitt, 2006; Hofmann, Gschwendner, Nosek, and Schmitt, 2005). See Gawronski and Bodenhausen (2006a, 2011) for detailed examples on interactions between, direct and indirect influences on, and combined effects of associative and propositional processes.

useful approach for examining evaluative responding with regard to current discussions about the context sensitivity and implicit–explicit distinction. I will return to these topics later in the General Discussion.

3.   Toward a Model of Subjective Movie Evaluation Criteria (SMEC)

The guiding questions for this chapter are: How can the insights from the preceding chapters be integrated into a model of subjective movie evaluation criteria (SMEC)? What exactly are SMEC? How can SMEC be conceptualized? Is there any previous research that can contribute to this conceptualization? And if so, what can we learn from the findings with regard to the measurement of SMEC?

3.1   Conceptualizing Subjective Movie Evaluation Criteria

The following conceptualization of SMEC shares some of the ideas outlined by Möller and Karppinen (1983; see Chapter 2.1), but tries to enrich and extend them. I define SMEC as standards that viewers use for assessing the features of films. This definition consists of six components. Let us look at them in more detail (components are in bold and italics):

- ▪ ***Films*** might be replaced by similar terms like motion pictures, movies, TV movies, cinematic pieces, and so forth—albeit it is the most general category. This list is neither exhaustive nor disjunctive and leaves some space for interpretation. For instance, it comprises films you can view at the cinema, on the TV, home entertainment center, and mobile device. It does not explicitly exclude documentaries (e.g., *Bowling for Columbine*), art films (e.g., *The Cremaster Cycle*), short films (e.g., *Balance*), or music videos (e.g., Madonna's video clip *Bad Girl*, directed by the renowned David Fincher and featuring Oscar-winner Christopher Walken). Although the lines between categories are somewhat blurry, all examples contain narrative and filmic elements,

which can be considered essential to the definition. Nevertheless, the English term movie and the German term *Spielfilm* are at the core of the definition.

▪ ***Features*** are characteristics of the movie (or *qualities*; Wolling, 2009) and can be described as film-inherent features (e.g., story, characters), film-peripheral features (e.g., production costs, awards), or effects (e.g., a feature that causes suspense, a feature that requires cognitive effort to cope with it).

▪ ***Viewers*** are human beings that have seen at least one movie in their life. I assume they continued watching movies, albeit even individuals who stopped watching movies still have access to the standards they have built-up.

▪ ***Standards*** are mental representations. The most appropriate metaphor of those provided in Chapter 2.4 is the attitude-metaphor. From a structural perspective, standards include at least knowledge about the attitude object, namely the feature, and an evaluative summary, namely a favor or disfavor for the feature. Because we cannot directly observe attitudes as they are latent constructs, we have to observe their manifestations (e.g., expressed opinions, filled-out attitude scales, response latencies). This process—from the activation of an associative pattern to the response measurement—might be best described by means of the APE model (see Chapter 2.4).

▪ The ***use*** of standards suggests a deliberative, utilitarian way of (propositional) processing. Although one could think of spontaneous, automatic ways of (associative) processing as well, processing mental representations in this way does not qualify them as standards. Standards vary in their strength. For example, there might be strong standards (e.g.,

aversion to bombastic film scores)[10] that easily come to mind as well as weak standards (e.g., realistic scenes) that only emerge when a stimulus triggers them (e.g., item or film feature). The strength of standards could be operationalized via meta-judgmental or operative indices (e.g., self-report or response latencies; cf. Bassili, 1996; Wegener, Downing, Krosnick, & Petty, 1995). Especially, measuring importance as an indicator of the strength of a standard might be directly related to the standard's impact on evaluative processes.

- **_Assessing_** means "to make a judgment about a person or situation after thinking carefully about it" (Longman Dictionary of Contemporary English, 2009). Regarding the present topic, it is an evaluative judgment about the film characteristic. Most importantly, this assessment or evaluation (or appraisal) can already occur when one is faced with the selection of a movie (e.g., due to advertised or anticipated features). It can also occur during watching a movie (e.g., when taking part in a continuous response measurement study or because of salient triggers such as the swelling sound of a bombastic string orchestra). In the latter case, either stopping to watch the movie or switching the mode of reception might be possible. Finally, post-receptive judgments might occur under different circumstances (e.g., when talking with others, answering an item, asking oneself). They might be determined by comparing expectations and gratifications obtained including cognitive dissonance processing. In addition, however, follow-up communication might invite many more competing processes (e.g., social judgment processes, social comparisons, impression management). Strongly held attitudes towards

---

[10] I am indebted to my partner Helen for this straightforward criterion.

specific film features should prevail in post-viewing conversations, whereas weak standards are susceptible to change (e.g., because of persuasive communication). To set things straight: Standards—albeit strong—do not have to be strong influencers of an overall post-viewing evaluation. For instance, they have no effect when the feature does not exist (e.g., no film score), the standard is not triggered (e.g., jazz score), or other standards are more activated (e.g., the twist ending).

In addition to the previous definitions, such a conceptualization should also address developmental questions (e.g., where do SMEC come from, how do they change, how do they develop over the life-span). These questions are beyond the scope of this paper, but it seems to be a worthwhile endeavor to deepen our understanding about evaluating movies across the life-span, thereby gaining insights into movie socialization. I will return to this issue in the General Discussion. Another set of questions concerns personality and individual differences (e.g., are SMEC stable and cross-situationally consistent traits?). However, to keep it as parsimonious as possible, it appears advisable to leave the plethora of personality variables outside of the SMEC concept. Instead, theorizing about and empirically investigating the relationship between SMEC and further characteristics of the person should be pursued. This will be an important prerequisite for exploring the nomological network of SMEC in Chapter 9.

Personality, Individual Differences, and Movie Preferences. Crystallized, stable SMEC are supposed to interact with movies before, during, and after exposure like any other trait-like construct (cf. Krcmar, 2009). Additionally, I assume that individuals differ in the importance they attribute to their SMEC (e.g., for some viewers it might be important when evaluating a movie that a movie contains thought-provoking content; for others it might be unimportant; cf. Bartsch & Oliver, 2011, p. 15; Klimmt, 2011,

p. 36). Subsequently, these differences may (a) influence movie choices (Möller & Karppinen, 1983; Wolling, 2009), (b) affect the (evaluative) information processing of movies in a similar way as assumed in the modes of reception approach (Suckfüll, 2004), and (c) moderate effects of movie characteristics (e.g., on enjoyment; cf. Eliashberg & Sawhney, 1994).

From a U&G perspective, "personality factors influence needs, which in turn influence viewing practices" (Krcmar, 2009, p. 239); therefore, SMEC might interact with or—even more likely—be determined by further trait-like constructs (e.g., openness to experience, need for cognition). However, the relationships between personality traits and SMEC have not been investigated yet, and surprisingly little research on the relationship between personality traits and movie content preferences has been done so far. Nevertheless, some studies have dealt with the impact of personality traits on viewing motives and behavior and also examined the correlation with content preferences. Although much of this research was carried out in the context of television, in some cases these results might also apply to movies (e.g., violent stimuli or preference for horror). Thus, in the following paragraphs, I will focus only on research that includes data on the relationship between movie content or movie genre preferences and traits, needs, and further stable concepts.[11] To come to the point: Sensation seeking (Zuckerman, 1979, 1994) turned out to be the most studied personality construct related to movies and films; in contrast to research in personality psychology, the Big Five (e.g., Digman, 1989; Goldberg, 1993) have received less attention in media research to date (cf. Böcking & Fahr, 2009).

Sensation Seeking. "Sensation seeking is a trait defined by the need for varied, novel, and complex sensations and experiences and the willingness to take physical and

---

[11] For reviews see Krcmar (2009), Oliver and Krakowiak (2009), or Schmitt (2004).

social risks for the sake of such experience" (Zuckerman, 1979, p. 10). The sensation seeking trait is biochemically based and manifests itself in individually different levels of optimal arousal. To establish these individually different optimal levels, high sensation seekers search for stimulating sensations to avoid boredom, whereas low sensation seekers rather enjoy calm environments instead of excitement. Research has often connected sensation seeking and media choice or content. For instance, high sensation seekers attended more horror and sex movies (Zuckerman & Litle, 1986), liked horror, suspenseful, violent, action, or sexual content more (e.g., Aluja-Fabregat, 2000; Aluja-Fabregat & Torrubia-Beltri, 1998; Burst, 1999; Edwards, 1991; Gleich, Kreisel, Thiele, Vierling, & Walther, 1998; Hirschman, 1987; Schierman & Rowland, 1985), watched horror movies for gore (Johnston, 1995), and preferred high-arousal films (Banerjee, Greene, Krcmar, Bagdasarov, & Ruginyte, 2008). Low sensation seekers, for instance, preferred musical movies and romantic fiction (Schierman & Rowland, 1985), watched more light films (Hall, 2005b), drama (Potts, Dedmon, & Halford, 1996), or comedies (Schierman & Rowland, 1985), and avoided violent and sexually explicit content (Zuckerman & Litle, 1986). Other researchers found more ambiguous patterns, although they controlled for gender or age differences (e.g., only correlation of subdimensions of sensation seeking with preference for violent or action-oriented content or low or statistically non-significant correlation coefficients; cf. Aluja-Fabregat & Torrubia-Beltri, 1998; Bagdasarov et al., 2010; Hall, 2005b; Krcmar & Greene, 1999; Slater, 2003; Tamborini & Stiff, 1987; Tamborini, Stiff, & Zillmann, 1987). Finally, in a study by Eliashberg and Shawney (1994), individual differences in sensation seeking had high predictive power for individual differences in movie enjoyment.

Big Five Personality Factors. The five dimensions—Neuroticism, Extraversion, Openness to Experience, Conscientiousness, and Agreeableness—describe personality

traits on an abstract and broad level (e.g., Digman, 1989; Goldberg, 1993). In general, Batinic (2005) found that the Big Five are more strongly correlated with film preferences than with actual film choice.

Neurotic (or emotionally instable) individuals are more likely to experience such negative affective states as anxiety, anger, or depressed mood. They are vulnerable to stress, self-conscious, and impulsive. Neuroticism was found to be negatively correlated with preferences for action-adventures and comedy (Weaver, 1991) and suspenseful fiction (Burst, 1999). Anxiety was also related to horror and violent action (Grimm, 1993, 1997, 1999). However, some studies found no effects on genre preferences (e.g., Hall, 2005a; Weaver, Brosius, & Mundorf, 1993), and others indicated a complex relationship with different facets (i.e., angry hostility, vulnerability, and impulsiveness; cf. Krcmar & Kean, 2005).

Extraverts tend to be friendly, gregarious, assertive, active, excitement-seeking, and cheerful. In a few studies, Extraversion was found to be positively associated with preferences for comedies (Weaver, 1991), sexual-comedies (Weaver et al., 1993), violent films among boys (Aluja-Fabregat, 2000), horror film attendance of females (Zuckerman & Litle, 1986), and a higher level of enjoyment of violent films (Aluja-Fabregat & Torrubia-Beltri, 1998; Krcmar & Kean, 2005). However, most of them showed that relationships were no longer statistically significant after controlling for socio-demographics (e.g., Hall, 2005a).

People who are open to experiences appreciate aesthetics, reflect on their emotional states, are interested in adventurous activities, are intellectually curious and liberal, and have a vivid imagination. Besides their preferences for information-based programs (Burst, 1999), they also seem to like violent content if it was of aesthetical

value (Krcmar & Kean, 2005). Most recently, Silvia and Berg (2011) demonstrated that Openness to Experience is related to expert knowledge about movies.

Agreeable people think that most other people are trustworthy. They are friendly, empathetic, generous, and helpful. Agreeableness is negatively correlated with Psychoticism (a trait from Eysenck's model of personality; cf. Eysenck & Eysenck, 1976). People who scored high on Agreeableness or low on Psychoticism preferred light-hearted and entertaining programs, such as romances and comedies (Burst, 1999; Hall, 2005a; Weaver, 1991), whereas low Agreeableness or high Psychoticism was related to viewing or liking action-oriented, suspenseful, violent, dramatic, or horror movies (Aluja-Fabregat & Torrubia-Beltri, 1998; Burst, 1999; Krcmar & Kean, 2005; Weaver, 1991; Weaver et al., 1993).

Conscientious people can be described as self-efficacious, orderly, dutiful, achievement-striving, self-disciplined, and cautious. Conscientiousness is negatively related to the preference of suspenseful fictional programs and slightly positively to entertaining programs (Burst, 1999).

Need for cognition (NFC). The need for cognition is described as the "tendency to engage in and enjoy thinking" (Cacioppo & Petty, 1982, p. 130). Hawkins et al. (2001) found a negative correlation between NFC and paying attention to the TV program when watching the drama genre. Oliver (2008) assumed that participants with a higher NFC might have anticipated a higher enjoyment of sad films, but found that NFC did not predict the anticipated enjoyment of movies from the sad film, suspense, or comedy genre. Knobloch-Westerwick and Keplinger (2008) experimentally demonstrated that people with high need for cognition enjoyed complex narratives. Krakowiak (2008) showed that need for cognition predicted the enjoyment of ambiguous characters.

Suckfüll und Scharkow (2009) found statistically significantly positive correlations between NFC and the modes of reception Identity Work, Production, and Imagination.

   Need for affect (NFA). The need for affect is the "general motivation of people to approach or avoid situations and activities that are emotion inducing for themselves and others" (Maio & Esses, 2001, p. 585). With regard to movies, Maio and Esses (2001) found that the Approach dimension of NFA was statistically significantly and slightly positively correlated with choosing emotional over unemotional films and happy over neutral films. However, they found no statistically significant correlation between the Approach dimension and choosing sad over neutral films. Furthermore, the Avoidance dimension showed no statistically significant correlations with film choice at all. Vogel (2007) found a positive correlation between attitude toward sad films and need for emotion (Raman, Chattopadhyay, & Hoyer, 1995). Appel (2008) found a small correlation between Approach and the frequency of watching TV romances. Additionally, he demonstrated that an Approach x Gender interaction statistically significantly predicted the females' preferences for watching a movie that triggers negative emotions. However, neither main effects nor interactions for the Avoidance dimension were statistically significant. Bartsch, Appel, and Storch (2010) showed that the Approach dimension was highly related to experiencing intense emotions while watching horror or drama movies as well as to more negative feelings, higher levels of discrete negative emotions, and more ambivalent emotions. The Avoidance dimension was statistically significantly related to lower meta-emotional enjoyment. Finally, Maio and Esses (2001)—as well as Appel (2008)—found a moderate correlation between NFA and NFC. However, only the Approach dimension in the Maio and Esses' study was statistically significant and moderately correlated with NFC, whereas in Appel's study

this was only true for the Avoidance dimension. Taken together, these findings are rather inconsistent.

Film genre preferences. The preference for one or more movie genres is crucial for understanding choice, viewing, and effects of movies. Why? There are at least two reasons. First, a film's plot, story, or genre[12] was found to be the most important reason for going to the movie theater or attending a specific film (e.g., Austin, 1981; Baum, 2003; Benesch, 2004; Gazley, Clark, & Sinha, 2011). To subsume a movie under a specific genre category might help the potential viewers, for instance, to choose among a vast selection of movies according to their preferences, thereby reducing uncertainty (e.g., select comedy and avoid horror; cf. Gehrau, 2003) and to facilitate the comprehension of the plot during or after watching a new movie (e.g., interpreting a misfortune in a comedy as funny, in a drama as tragic, or interpreting open endings; see also Chapter 2.2 for functions of a genre schema). Second, the importance of genres is also commercially reflected in the film industry (e.g., Hollywood is a mainly genre-based production studio system; cf. Wuss, 2009, p. 248). Some findings from research on movie genre preferences and broader personality traits or needs were presented above and showed that individuals differ in their movie preferences. In addition, researchers are interested in how distinct film genres are related to each other and whether broader categories can be formed by combining more narrow ones (e.g., Austin & Gordon, 1987, p. 12). For instance, Hall (2005b) asked her participants to rate how often they viewed 11 film genres and found three dimensions—Action-oriented, Light, and Serious. Similarly, using questionnaires, experiments, and content analysis of TV program guides, Gehrau (2001, 2009) found three basic genres—comedy, drama, thriller. Most recently, in an

---

[12] Austin and Brown (1987, p. 12) argued that plot, story, and genre are used synonymously, especially among lay audiences: "People use genre labels as a handy, convenient, and easy method for categorizing movie stories and to discriminate among story types. Such labeling may serve as either a substitute for or further elaboration and clarification of film stories".

analysis with over 3,000 participants, Rentfrow, Goldberg, and Zilca (2011) found five broad entertainment-preference dimensions: Communal (film genres: e.g., romance, family), Aesthetic (e.g., foreign, classics), Dark (e.g., horror, cult), Thrilling (e.g., action, science-fiction), and Cerebral (e.g., documentary). These preferences can be explained by socio-demographics as well as personality traits. Their findings also clearly demonstrated that people seek media stimuli that reflect and reinforce facets of their personalities (Rentfrow et al., 2011, p. 251).

Although film genre preferences are usually included in research on movie selection, reception, and effects, they have been criticized as being too "generic", unable to "distinguish between different movies within the same genre", or to have only low predictive power, for instance, with regard to enjoyment (Eliashberg & Sawhney, 1994, p. 1168). However, even though viewers like films of a preferred genre better than films of less preferred genres, a more important predictor might be, of course, the actual quality of the movie. Thus, general movie evaluation criteria, which are related to specific film features rather than to general genres, might moderate how specific movie characteristics influence movie evaluations.

To conclude, previous research on the relationship between traits and film-related constructs focused on special topics (e.g., sensation seeking and preference for violent content, need for cognition and enjoying complex plots, or need for strong emotional experience and interest in sad films). However, research is piecemeal, findings are rather inconsistent, and the use of broad genre categories as attitude objects might not foster our understanding of how personality traits could be related to specific film features. For instance, *Wag the Dog* and *Dumb and Dumber* share the same genre, albeit obvious differences in story, sophistication, cinematography, or performance of involved actors. Taking individual differences in preferences for or

attitudes toward such film features into account would be a more fruitful way to delineate the connection between personality traits and movie research. Thus, exploring SMEC as a link between traits, needs, motives, values, and so forth, on the one hand, and specific film features that are yet to be discovered, on the other hand, appears to be a promising endeavor.

3.2   Previous Research on Movie Evaluation Criteria

> "A good film is when the price of the dinner, the theatre admission and the babysitter were worth it."
>
> —ascribed to Alfred Hitchcock (*1899, †1980), Film director and producer

In this section, I will introduce some existing approaches to systematize or measure movie evaluation criteria. Because there is no approach to the SMEC of lay audiences, I will first shed some light on criteria found in the context of communicator research (i.e., film critics). Then, I will report on selected work from TV quality research. Although television research has accumulated a large body of literature, it is limited to the medium TV and thus cannot be merged with film audience research: Watching television comprises many more formats and genres irrelevant for movie evaluation (e.g., news, talk shows, and ads). Furthermore, it is supposed that TV is associated with different evaluations and viewing motives (Finn, 1997). Despite these constraints, we can gain insights into research methods applied to quality assessments that might be useful for the purpose of the present research. Moreover, the TV quality studies I will focus on include narrative and fictional stimuli. Thus, if carefully interpreted with regard to movies, these findings allow us to gain a first impression of the research done so far.

Movie Evaluation Criteria of Film Experts. From a historical perspective, Handel (1950)—a research director at MGM from 1942 until 1951—can be seen as a pioneer in movie audience research (Cox, 1983). For instance, he applied the Lazarsfeld–Stanton Program Analyzer—a precursor of recent and technically more advanced real-time response measurement dialers—to movie research (i.e., to assess the evaluation and comprehension of movie scenes). Handel pointed out that story, cast, and title are the most important attributes of a movie. However, as these attributes were fixed by the producers rather than predicted from audience research, Handel's remark can be seen as a historical starting point of research on movie evaluation criteria, but provides no further information concerning the criteria of a lay audience.

A pragmatic approach was developed by Stegert (1993) who aimed at providing a textbook for film critics. He distinguished four categories of evaluation criteria: (a) formal-aesthetical criteria and norms (e.g., filmic, plausibility, well-done, avant-garde), (b) effect-based criteria and norms (e.g., enjoyment, distraction, emotion, imaginary journey, suspense), (c) ideological and socio-critical criteria and norms (i.e., criticism, realism, education), (d) ethical criteria and norms (i.e., humanitarianism, respecting the audience). All of these criteria were more or less arbitrarily chosen and put together—some related to film theories, some not. In other words, they are neither deduced from theory nor empirically tested.[13] However, as we will see in the empirical part of this thesis (Phase I, see below), most of them were also generated by recipients, albeit on a more concrete level. This raises the question whether consumers or recipients simply reproduce criteria proposed by other authorities (film theorists, film critics, etc.) or if

---

[13] In media science, similar efforts have been made to categorize criteria for interpretative film analysis (cf. Mikos, 2003). Qualitative content analysis is an empirical way to derive criteria. For instance, in an argument analysis of evaluative texts (e.g., film critiques, press releases, etc.) about controversial fictional TV content, several criteria similar to the ones mentioned above were found (cf. Bolik, 1999; Bolik & Schanze, 2001).

they say what they themselves think. I will come back to this question in the paragraph after the next.

Another category system of movie evaluation criteria evolved from Wyatt and Badger's (1988) US nationwide survey on newspaper film critics. The critics' importance ratings on 22 movie characteristics were factor-analyzed yielding a four-factor solution: (a) Sex/Violence (e.g., vulgarity, sex, violence, nudity, and MPAA rating), (b) Production Elements (e.g., editing, design, and cinematography), (c) Performance Elements (e.g., lead actor, director, plot, and screenplay), and (d) Production Difficulty (e.g., production difficulty, cost, and genre). The most important characteristics were lead actor, plot, and screenplay—all part of Performance Elements. Additionally, critics rated the importance of 13 stylistic and content review elements. Factor-analysis yielded three factors: (a) Personal Impression/Judgment (e.g., subjective responses, evaluative adjectives), (b) Objective Reporting (e.g., facts of the film), (c) Audience Reaction (e.g., discussing audience reaction, viewers who might enjoy). The most important review element was the overall evaluation. Finally, critics rated the importance of 25 functions of a movie. A five-factor solution emerged: (a) Aesthetic Experience (e.g., artistic experience, aesthetic experience, and self-education), (b) Entertainment (e.g., diversion, escape, and relaxation), (c) Arousal (e.g., danger, emotional arousal, and sexual stimulation), (d) Subject of Conversation (e.g., basis of conversation, what's talked about), and (e) Ethical Value (e.g., model of behavior, reinforce values). The most important film functions were entertainment, experience expansion, escape, and diversion. Although the approach of Wyatt and Badger is fully exploratory, data-driven, and somewhat problematic in its interpretation of statistical results (e.g., high cross-loadings of several items are not interpreted), it offers some interesting insights into the professional perspective on film criticism and movie evaluation. However, it is questionable whether these results can be

applied to laypeople, novices, or to the ordinary movie audience in the US as well as in other countries.

A study from the FiT[14] project (cf. Rössler, 1997a, 1997c) examined this question more closely. In FiT, Rössler analyzed film reviews in German daily newspaper and—among other studies—surveyed German newspaper film critics and compared the results to a study he conducted with moviegoers (for details see Rössler, 1997b). Rössler asked film critics and moviegoers to rate the importance of the (normative) movie evaluation criteria proposed by Stegert (1993, see above). Following the methodological approach by Wyatt and Badger (1988, see above), he obtained a four-dimensional description of film critics' evaluation criteria: (a) Technical Qualities (i.e., sound, camera, music, and editing), (b) Plot (i.e., director, actors, dramaturgy, and plausibility), (c) Educational Engagement (i.e., political ambition, social relevance), and (d) Entertainment (i.e., entertainment value, stars, and special effects). The most important evaluation criteria were director, actors, entertainment, dramaturgy, and plausibility. Hence, Rössler concluded that the Plot dimension might be crucial to critics' evaluation of a movie. When moviegoers were asked, 85% rated items about entertainment (e.g., humor and suspense) as most important for evaluating movies. However, only items from the dimensions Entertainment, Educational Engagement, and one item to assess the importance of aesthetic photography were included in this part of the survey. Therefore, it does not make sense to compare the results of the moviegoers with those of the film critics. Finally, there are further methodological issues that question Rössler's interpretation of the audience survey results (i.e., sampling, participants' movie preferences, or relying only on descriptive statistics). In reference to the decision-making process of moviegoers, they rated the story's theme, film score, stars, and special

---

[14] FiT is the abbreviation for *Film in der Tageszeitung* which translates as "Film in daily newspapers".

effects as most important. Interestingly, this is consistent with the rank order Rössler (1997b) found when asking film critics how the audience chooses movies. However, in a US-American study, there was only a low correlation between the ratings of film critics and consumers. Furthermore, consumers and critics differed in their standards (Holbrook, 1999).

Although the results above regarding the importance and dimensions of criteria may strongly remind us of the findings outlined in the chapter on moviegoing motives, they can only be interpreted as criteria or movie attributes for *selecting* movies—not necessarily for *evaluating* movies. To my best knowledge, Linton and Petrovich (1988) conducted the only study that, amongst other questions, asked students to rate the importance of 15 movie attributes when evaluating a movie. Further, the participants were asked to rate several statements about moviegoing experiences. The authors conducted factor analysis on the movie attributes and found two underlying dimensions: Foreground (e.g., storyline, characters, acting, scenery, music, etc.) and Background (e.g., editing, photography, casting, director, etc.). However, and among other theoretical and methodological shortcomings which are discussed by the authors in their discussion section, they neither reported the loadings of their factor analysis nor provided information about how the number of factors was determined. Thus, we have to be cautious about their conclusion that "the study also tends to confirm the relative lack of importance assigned to the 'background' technical attributes" (Linton & Petrovich, 1988, p. 37).

In the following section, I will summarize some insights into evaluation criteria that come from the research field of TV quality.

Movie Evaluation Criteria in TV Quality Research. Only a few studies that have dealt with TV evaluation criteria focused on fictional formats (i.e., TV series). For instance, in an exploratory study by Himmelweit, Swift, and Jaeger (1980), participants rated TV series on 17 stylistic attributes (e.g., realistic, lighthearted, funny, informative, complicated, and violent). Himmelweit et al. demonstrated that "viewers . . . are sensitive to stylistic attributes of individual programs and that these affect their liking" (p. 81). Additionally, the researchers assessed the participants' stylistic preferences and obtained five dimensions from a factor analysis: (a) Approach to/Avoidance of Potentially Upsetting Stimuli (e.g., "A sad ending just makes me miserable"), (b) Aloofness or Involvement in Programs (e.g., "I prefer programs that appeal to my heart"), (c) Preferences for Low/High Cognitive Effort (e.g., "I prefer programs that you have got to make a real effort to understand"), (d) Preference for Real vs. Fantasy Content/Presentation (e.g., "I prefer programs set in everyday surroundings"), and (e) Liking for Thriller/Action Content (e.g., "What makes a plot interesting is the action"). They conclude that the viewers'

> willingness and ability to rate programs on a wide range of attributes showed that although the industry tends to think in terms of genres, the public looks at each program in its own right and not merely as a member of a family of programs. (p. 93)

Finally, the researchers also proposed applying criteria derived from the viewers' own suggestions in future studies.

Greenberg and Busselle (1996) developed an instrument to assess TV program quality with a sample of over 1,300 US undergraduates. After viewing one of four situation comedies or one of four action adventure programs, the participants rated the TV shows on 44 adjective pairs regarding general attributes of the program (e.g.,

interesting–dull, *ordinary–exceptional*, or *true to life–not true to life*) and on four items regarding the quality, liking, and repeated viewing of the program. Greenberg and Busselle factor-analyzed the attributes and obtained five factors for the comedies—(a) Appreciation (e.g., enjoyable, entertaining), (b) Real vs. Funny (e.g., realistic, serious–light), (c) Fairness (e.g., fair, gentle), (d) Modern (e.g., new, modern), (e) Originality (e.g., unusual, original)—and four factors for action–adventures—Appreciation, Realism, Humor, Originality. The scores of the single factors were regressed on an average of the four quality items. Interestingly, they found that the Appreciation factor explained 85% of the variance in overall quality in comedy and 54% in action adventure programs. They concluded that this scale provides an alternative measure for assessing overall quality ("They [the items] do not explain quality", p.194). The remaining scales—when Appreciation was excluded from the regression equation—explained 60% of the overall quality variance in comedies and 30% in action–adventures. The factors Realism, Humor, and Originality emerged in comedy as well as in action adventure shows; Fairness and Modernity emerged in comedies only.

Similar studies were conducted with TV shows in general (Gehrau, 2008) or other fictional TV formats like soap operas (Schenk & Rössler, 1987), crime series (Schenk, Büchner, & Rössler, 1986), and TV series of different genres (Wolling, 2004).

The major drawback of these studies lies in the simultaneous measurement of predictor and criterion variables—both are explicitly related to the program watched. Hence, it is impossible to draw any conclusions about the causal relationship between general evaluation criteria and the specific evaluation of a particular film stimulus. What is even worse, we cannot infer anything about an independent mental representation of criteria or their consistency from these results. Furthermore, Gunter (1997) questioned the content validity of adjective lists. Instead, he suggested to "go directly to viewers

themselves even before the attribute scales are produced as a source of insights concerning what those scales might be" (p. 14).

In his own research on soap operas, Gunter (1997) developed genre-specific scales in two steps consisting of different methods. In the first step, participants were divided into four small groups and watched an episode of one of four popular soap operas while simultaneously rating their enjoyment via electronic handhelds. Afterwards, they were provided with the mean graphs and talked about their subjective explanation of the graphs' peaks and troughs in a group discussion. From the key statements of the focus group, Gunter qualitatively derived 26 statements covering eight dimensions: (a) Verisimilitude (e.g., "The characters were true to life"), (b) Established Characters (e.g., "All the characters clearly belonged to the local area"), (c) Tension/Drama (e.g., "There were a number of problems going on that surprised and intrigued me"), (d) Entertainment/Involvement (e.g., "I couldn't wait to find out what happened next"), (e) Coherence/Cohesion (e.g., "Each of the storylines in this episode had a clear beginning, middle, and end"), (f) Technical Professionalism (e.g., "The camera work balanced long and close-up shots"), (g) Contrast and Balance (e.g., "It included both light and serious storylines"), (h) Plot/Setting (e.g., "The setting was recognizable and reasonably familiar"). In the second step, he administered the scales to 3,000 participants of a British TV research panel who watched soap operas and found that most of the dimensions were statistically significant predictors for watching one of the four soap operas. In sum, the beta-weights depended on the criterion and on the soap opera. Tension/Drama, Entertainment/Involvement, and Technical Professionalism were most consistently related to viewing, whereas some other attribute scales (e.g., Characters) barely achieved significance as predictor in any of the regression models. Unfortunately, Gunter fails to report important statistics (i.e., the exact number of

participants and the explained variance in the regression models). Furthermore, he provided only little psychometric information about the scales (only internal consistencies ranging from .65 to .80, but no further evidence for factorial validity). Hence, the results must be interpreted cautiously.

A few studies have focused on the criteria children use to evaluate TV programs. For instance, Snoek and Bouwman (1995) showed that aesthetic quality of the images, interestingness, excitement, and humor were important predictors for the overall evaluation of cartoons. Unfortunately, they administered the criteria items after their 8- to 13-year-old participants watched cartoon clippings. Hence, we have the same problem as mentioned above—mingling general criteria with specific evaluations. Furthermore, although the authors derived the criteria items from previously conducted qualitative studies, they failed to examine the dimensionality and structural validity of the criteria. This, however, has been done in probably the most comprehensive approach to illuminate the quality of children's television, namely from the perspectives of mothers, children, program makers, and television critics in the Netherlands (Nikken, 1999). Nikken and colleagues (e.g., Nikken & van der Voort, 1997; Nikken, van der Voort, & van Bochove, 1996) found seven types of quality standards when asking mothers how a good children's TV program should be (e.g., comprehensibility, aesthetic quality, involvement). These types plus two additional types expected to be relevant were then used to assess children's quality standards for TV programs (Nikken & van der Voort, 1997). The authors constructed nine scales from the quality dimensions: (a) Credibility (e.g., "A children's program should give a balanced image of reality."), (b) Comprehensibility (e.g., "A children's program should be understandable for all children."), (c) Entertainment (e.g., ". . . should be funny."), (d) Aesthetic Quality (e.g., ". . . should contain beautiful images."), (e) Involvement (e.g., ". . . should capture a child's

attention."), (f) Presence of Role Models (e.g., " . . . should feature persons a child wants to be like."), (g) Innocuousness (e.g., ". . . should not make a child sad."), (h) Restfulness (e.g., ". . . should set a child at ease."), (i) Thought Provocation (e.g., ". . . should make a child curious."). The results resembled those of the mothers-study, especially when taking only fictional TV program genres into account. As Valkenburg and Janssen (1999) argued, the study is limited in that Nikken and Van der Voort (1997) largely used items from the list from the mothers-study:

> It is possible, however, that the quality standards mothers use to evaluate programs are different than the standards applied by children. It could be, for instance, that mothers define the quality of children's television in terms of its impact on children's cognitive and emotional development (with the result that they might be more focused on the innocuousness of a children's program), whereas children might consider innocuousness as an irrelevant quality standard that they would never mention spontaneously. (Valkenburg & Janssen, 1999, p. 7)

Therefore, Valkenburg and Janssen (1999) used a focus group approach to collect children's criteria in a first step and conducted a paper-and-pencil questionnaire study in a second step. Their principal components analysis resulted in an eight-dimensional solution: (a) Interestingness (e.g., "show somebody I would really like to be"[15]), (b) Romance (e.g., "be about love"), (c) Realism (e.g., "show things that are real"), (d) Violence (e.g., "be about brave and strong heroes"), (e) Humor (e.g., "be full of jokes"), (f) Innocuousness (e.g., "not show any violence that children can imitate"), (g) Comprehensibility (e.g., "be easy to understand"), (h) Action (e.g., "contain a lot of action"). The sample consisted of Dutch and US children and results were quite similar for both countries.

---

[15] This item clearly shows the correspondence to the Role Model factor in the Nikken and Van der Voort study. Further items are similar to the Thought Provoking ("make me think") factor.

Conclusion. Taken together, findings from film criticism and TV quality research show the following limitations:

- Film critics as experts might have different criteria than laypeople.

- TV programs contain specific features. Their application to movies hardly seems justifiable.

- Even children seem to already have developed sophisticated, multidimensional criteria for assessing TV programs, although this might just reflect social norms (e.g., of their mothers).

- Adjective lists seem to be inappropriate as a measurement instrument because they might be too abstract to capture specific film features.

- Simultaneous assessment of criteria importance and evaluation of stimuli already watched might lead to erroneous conclusions.

- Most of the instruments applied were ad hoc scales.

- The use of exploratory data analysis (e.g., principal components analysis) for dimensionalizing the construct space (i.e., subscale development) was not followed by further confirmation or validation procedures.

- In sum, we have no sound instrument for the measurement of subjective movie evaluation criteria.

Although there are serious shortcomings in previous research, there are some ideas for improvement. So, what can we learn from the findings and research methods? What recommendations do scholars provide (not all are reported in detail in this section)?

- Qualitative pilot studies within the target population are recommended.

- The individual steps in developing a measurement instrument should be conducted in separate studies (e.g., item development, scale construction).

- Of course, the topic should be movies, not TV shows.

- Confirmatory and construct validation studies need to be conducted.

- Detailed reporting on the psychometric properties of the scales is warranted.

- In sum, we need a sound instrument for the measurement of subjective movie evaluation criteria, and such an instrument must be developed in accordance with psychological standards of measuring latent constructs (e.g., attitudes).

3.3   Conclusion

Let me briefly summarize the main ideas of the present and preceding chapters. Both knowledge and evaluative information are part of a mental representation of or an attitude toward a specific film feature. Such an attitude guides our behavior and information processing before, during, and after watching a movie. Hence, an attitude might influence the choice of a movie, the mode of reception, and the effects of a movie. Regarding the last aspect, an overall evaluation of a movie might reflect aggregated evaluations of film features, thereby offering a detailed analysis of movie enjoyment or entertainment experiences. It is important to note that such an influence might be especially true for strong attitudes (i.e., when they are more important or accessible than others). Attitude and attitude structure are connected to further trait-like constructs such as personality traits, needs, motives, or values. Because people individually differ in their attitudes and attitude structures, we can assume that this is also true for attitudes toward specific film features. Although previous research has accumulated evidence for the existence of subjective evaluation criteria, we still do not know their number, scope, or stability—and we do not have any proven measure. In Chapter 4, I will outline the present research and the most important first step for continuing research on SMEC: the construction of a psychometrically sound measure.

4. The Present Research

The major aim of this dissertation is the construction of a standardized instrument for reliably and validly measuring SMEC. This is crucial for conducting any kind of research in the field of movie evaluation, but especially when SMEC are assumed to take the role of a moderating or mediating variable (cf. Hoyle & Robinson, 2004, p. 220).

Construct validity is "*the central concern* in psychological measurement" (John & Soto, 2007, pp. 475; emphasis in original). I will shortly introduce construct validity as an umbrella term which includes several aspects of psychometric quality. Subsequently, I will give a brief overview of the research questions and the studies conducted. Although several traditional and integrated approaches to construct validity exist (e.g., Campbell & Fiske, 1959; Cronbach & Meehl, 1955; Loevinger, 1957), in their summary, John and Benet-Martinéz (2000, pp. 351–357) point out the advantages of Messick's (e.g., 1981, 1989, 1995) integrated approach that distinguishes six types of construct validity (see Table 1), albeit they do overlap somewhat. This approach is an integrated one because it includes all aspects of validity and also incorporates different approaches to reliability as facets of generalizability (Cronbach, Rajaratnam, & Gleser, 1963; Shavelson, Webb, & Rowley, 1989). From the construct validity view, "validity is considered a property of the interpretation of a measure, rather than a property of the measure itself" (John & Benet-Martínez, 2000, p. 352).

Table 1
Types of Construct Validity

Type	Description
Generalizability	Evidence that score properties and interpretations generalize across populations, groups, settings, and tasks (e.g., reliability and replication)
Content validity	Evidence of content relevance, representativeness, and technical quality of items (e.g., expert judgments and review)
Structural validity	Evidence that the internal structure of the measure reflects the internal structure of the construct domain (e.g., exploratory and confirmatory factor analysis)
External validity	Evidence that the measure relates to other measures and to nontest criteria in theoretically expected ways (e.g., criterion correlation, multi–trait multimethod matrix)
Substantive validity	Evidence that measurement scores meaningfully relate to theoretically postulated domain processes (e.g., mediation analysis)
Consequential validity	Rationale and evidence for evaluating the intended and unintended consequences of score interpretation and use, including test bias and fairness

Note. Adapted from John & Benet-Martínez (2000, p. 352), John & Soto (2007, p. 476), and Messick (1995)

In addition, an economical and short instrument for measuring SMEC should be preferred to avoid tiring the respondents (Burisch, 1984). Research steps in scale construction and validation are straightforward and have often been described in handbooks (e.g., John & Benet-Martínez, 2000; Wegener & Fabrigar, 2004), thus, I will illustrate the following phases of instrument development by briefly mentioning only the central research goals: In Phase I, following the recommendations of Gunter (1997), I collected descriptions of SMEC via open-ended questions and categorized them as the first step in item development. Because we have little knowledge about SMEC and the existing scales are of poor quality (see Chapter 3.2), this inductive, data-driven process seems to be the most appropriate strategy to gain a deeper understanding of the content domain. This phase refers to content validity. In Phase II, I administered an online-questionnaire including the developed items for measuring SMEC to explore the latent structure and to exclude poor items. The discovered latent structure was validated in Phase III. Studies in Phases II and III were conducted to examine structural validity. By applying latent state–trait theory, Phase IV focused on aspects of generalizability (i.e., the reliability of measures and the generalization across two measurement points). Moreover, investigating the consistency of the measure as well as the situational or

interactional influences and method specificity contributed to substantive validity. In Phase V, my goal was to develop and explore the nomological network of SMEC (external validity); that is, to investigate convergent and discriminant relationships to external constructs or in other words, "what the measured construct is" and "what the measured construct is NOT" (cf. Wegener & Fabrigar, 2004, p. 160). For an overview of the five phases of scale construction, see Table 2.[16]

Table 2
Phases of Scale Construction in the Present Research

Phase	Aim	Method	Study	*N* (complete)
I	Content validity, base for item construction	Open-ended online + paper&pencil questionnaires	Pilot	258
	Categorization, categories for item construction	Structure formation technique + focus group		12
	Reliability of category coding	Content analysis		2 coders
II	Item comprehensibility	Paper&pencil pilot test	Study 1	14
	Item comprehensibility, technical check	Online pilot test using cognitive survey technique		8
	Exploring the latent structure, item reduction (EFA)	Online questionnaire		659 (500)
III	Item comprehensibility, technical check	Online pilot test	Study 2a	11
	Generalizability and Structural validity (E/CFA and CFA)	Online questionnaire	Study 2a	849 (587)
		Paper&pencil questionnaire	Study 3	152 (147)
IV	Substantive Validity and Generalizability: Reliability, consistency, occasion specificity (Latent state–trait analyses)	Online questionnaire, repeated measurement	Study 2a+2b	282 (273)
V	External validity (CFA)	Online questionnaire	Study 2a	see above

---

[16] Consequential aspects of validity concern "issues of bias, fairness, and distributive justice" (Messick, 1995, p. 745). They are more relevant to applied research (e.g., educational testing or personnel selection) than to basic research as is here the case (cf. John & Soto, 2007). Thus, this type of construct validity is not examined further in the present research.

5.   Phase I: Collecting and Categorizing Descriptions for Movie Evaluation Criteria and Item Development

Phase 1 consisted of two steps. In the first step, I conducted a questionnaire study with open-ended questions to collect descriptions of movie evaluation criteria. Subsequently, I analyzed and standardized the answers to these open-ended questions. In Step 2, the findings were further structured and content-validated with a modified structure formation technique. This pilot work served as a starting point for building a construct map in form of a coding scheme, exploring content validity, operationalizing movie evaluation criteria, and deriving the items' wording (cf. Borg & Shye, 1995; Wilson, 2005).

5.1   Step 1: Data Collection and Reduction

If a sound theoretical basis is lacking and it is not possible to draw on the literature for the appropriate amount and wording of item stems and response categories, the use of open-ended questions is recommended (Converse & Presser, 1986). As described and discussed in Chapter 3.2, besides methodological shortcomings, the criteria found by previous research have often been established on a rather abstract level (e.g., adjective lists, semantic differentials) or formulated in terms more familiar to film theorists or film critics (cf. Mikos, 2003; Stegert, 1993). Thus, they might be difficult to understand for laypeople. Additionally, empirical support for the assumption that these criteria are really appropriate to describe subjective perspectives on movie evaluations is lacking. To take subjective perspectives into account, previous research rather suggests deriving criteria or corresponding everyday language terms by (qualitatively) exploring the target group's responses to open-ended questions (cf.

Gunter, 1997; Nikken & van der Voort, 1997; Oliver & Bartsch, 2010; Suckfüll, 2004; Valkenburg & Janssen, 1999). Therefore, I used open-ended items to collect everyday language terms of movie evaluation criteria.

Method. To collect such terms, 156 undergraduates from the University of Koblenz-Landau, Campus Landau, filled out paper-and-pencil questionnaires with an open-ended item ("In general, what do you think what kind of criteria can be applied to evaluate a movie?") allowing for a maximum of 20 entries.[17] Furthermore and to obtain a more heterogeneous sample, 102 mainly non-student participants from a convenience sample completed the same open-ended item in an online-questionnaire.

Results. On average, the 258 participants (age: M = 26, SD = 9.7, range 18–73; 73% female[18]) gave 5.5 answers (SD = 2.50; range = 1–15). The total amount of 1,419 answers comprised 586 different terms for movie evaluation criteria. To get a quick overview and easily standardize and structure this large number of answers, two phases of reduction were carried out (cf. Eisermann, 2004, p. 98). First, if the terms differed only in spelling or grammar (e.g., "Authenzität [*sic*]" instead of "Authentizität", "zB [*sic*]" instead of "z. B.", etc.), orthographical or grammatical errors were corrected or—if the terms were correct—then both versions were standardized. Second, the remaining 502 terms were checked for semantic equivalence. For example, "suspenseful" and "suspense" were subsumed under the term "suspense". This finally resulted in 274 terms for movie evaluation criteria resulted (see Table A1 of Appendix A).

---

[17] "In general" was used to include not only the personal view, but also the criteria participants think other people might apply.

[18] In 2009, 70% of all students at the University Koblenz-Landau (Campus Landau) were female.

5.2   Step 2: Content Validation and Categorization

In the second step my goal was to find out whether the terms describe the construct space for movie evaluation criteria exhaustively. Moreover, because 274 terms are still far too many in order to formulate appropriate items for each individual term, I was also interested in further reducing the number of terms, primarily by subsuming them under broad but distinct categories. Therefore, I employed a modified structure formation technique (MSFT).

Method. The MSFT can be best described as a dialogue–consensus approach to subjective theories (e.g., Groeben, 1990; Groeben & Scheele, 2000) and includes elements of structure formation techniques (e.g., Scheele, 1992; Scheele & Groeben, 1988) and focus groups (cf. Merton, 1987).

Participants and procedure. I used the MSFT in Prof. Maier's colloquium for graduation candidates, which was held at the University of Koblenz-Landau, Campus Landau, in July 2009 and took approximately two and a half hour. Members were one professor, two post-docs, three doctoral candidates, and six diploma candidates—all from the research fields of psychology, education, or social sciences. Most of the 12 participants were unfamiliar with the research on movie evaluation. Therefore, they were labeled 'novices'. Three participants had already conducted media psychological research on movies. Therefore, they were labeled 'experts'. Three groups were formed from the 12 participants —each consisting of four persons. Each group was equipped with a set of 274 cards; each card contained one movie evaluation criterion term from Step 1 (see Table A1 of Appendix A). Every member in each group was provided with one-fourth of the set (i.e., 68 or 69 cards). Each group member was asked to lay out the cards by herself or himself while placing semantically similar terms close to each other. After this procedure, the four group members formed two pairs. Each pair was asked to

assimilate and adapt their individual structural maps to one single map for the respective pair (see Figure A1, A of Appendix A). This process was repeated with the two pairs producing one final structural map per group (see Figure A1, B of Appendix A). In every step, group members were allowed to write new criteria on blank cards, build duplicates, or label main categories. Finally, the three final structural maps were visualized on a flip chart and presented by each group. During the whole process, I was assisted by a student in photo-documenting the whole process and taking notes of the difficulties arising from developing and merging maps.

Results. To sum up the results of the protocols, the participants wrote down main categories, but did not add new terms. Furthermore, they found it difficult to develop a complete category system out of the 274 cards within this short period of time. Regarding the main categories only, the processes in all three groups resulted in very similar final maps (see Figure A1, C, and Figure A2, B & C in Appendix A). Interestingly, despite labeling the categories in a similar way (e.g., "cognitive" or "story") or subsuming some of the same cards under the same category labels (e.g., "upsetting", "touching", "emotions" under "affective"/"emotional" or "budget", "production location" under "production"), the links between main categories and subcategories as well as the mapping of the cards into the main categories were quite different. Although "eyeball" inspection is quite useful to get a first impression, it is not an appropriate method for examining the results of the MSFT in more detail. To make the main categories and the assigned cards of one group comparable to the other groups, I inductively developed a coding scheme (see Table A2 in Appendix A) and applied it to the results of MSFT (on card-level). Namely, each card was coded for every single group. After doing so, I compared every coded card of one group to the corresponding coded card of the two other groups. In addition, a second person, who did not participate in the MSFT session,

also coded the cards and compared the three groups. Krippendorff's α (cf. Hayes & Krippendorff, 2007) was .82, which indicates satisfactory intercoder reliability. Results of the group comparisons were as follows: First, for 21% of the cards (n = 57) all three groups applied the same categories (in sense of the coding scheme). Second, for 42% of the cards (n = 115) only two groups applied the same categories. Finally, for the remaining 37% cards (n = 102) there was no inter-group correspondence.

Discussion. The first step of Phase I led to 274 terms for movie evaluation criteria. In the second step, these terms were structured and categorized via MSFT. Participants consisting of experts and novices did not add any new terms. Although at first glance some of the main categories, which the group assigned to the terms, appeared to be similar, content analysis showed that the groups differed in their assignment of categories to the terms.

What can we learn from this with regard to content validity and item development? First of all, the terms seem to describe the construct space of movie evaluation criteria exhaustively. Therefore, they seem to be relevant and representative for the construct of movie evaluation criteria. Second, broad categories were found, under which many terms could be subsumed. However, they were not always distinct from each other. The conclusion is that 274 terms are far too many and the approximately 20 categories (e.g., Figure A1 in Appendix A) are too broad and too fuzzy to directly derive items from these categories. As a compromise for item construction, I used my coding scheme that was developed from the main categories and subcategories of all three groups. It comprises 73 distinct categories (see Table A2 in Appendix A). Because the categories varied in their breadth (i.e., in their number of subcategories), I tried to create items that reflect every facet of the construct space. In Phase II, these items were put to the test.

6.   Phase II: Exploring the Latent Structure and Item Selection

The pilot work in Phase I was conducted to get a deeper understanding of the construct space and content of movie evaluation criteria. Furthermore, wording ideas were gathered to facilitate the item construction process. The aim of Phase II was to reduce the number of items and explore the underlying structure. It is important to note that all items were worded in a way that allows the participants to assess the subjective importance of a criterion in general (i.e., not only related to a specific film they have just seen). As discussed in Chapter 3.2, post-viewing questionnaires with film-specific and general criteria items might produce biased results. When asking participants about their criteria in general, it is not possible to ask them for an evaluation because an evaluation is always related to a specific movie (e.g., you cannot ask how good or bad the camera work of a film is in general). However, because we are interested in the importance of criteria (as an indicator for attitude or criteria strength, cf. Chapters 2.2 and 3.1), we can ask the participants to rate the criteria on a corresponding scale. Therefore, every item consisted of three components: (a) "When evaluating movies, how important for you personally is/are...", (b) a phrase including the criterion (e.g., an item concerning the verisimilitude of a film's story was worded "...that the story the film tells is realistic?"), and (c) a rating scale (see Table B1 in Appendix B). All items were derived from the coding scheme (see Phase I) and positively phrased to avoid artifacts only due to the negative wording (Marsh, 1996). Following the recommendations of Krosnick and Presser (2010) for constructing rating scales, the scale I chose had the following characteristics: (a) *unipolar* because a criterion's importance can only range between *not at all important* and *extremely important*; there is no "negative importance" (e.g.,

Krosnick & Fabrigar, 1997); (b) *five points* because research showed that "ratings tend to be more reliable and valid when five points are offered for unipolar dimensions" (Pasek & Krosnick, 2010, p. 36); (c) *all* five points with *construct-specific labels* because using a common Likert-scale (e.g., combining statements with disagree–agree labels) is problematic (cf. Krosnick & Presser, 2010): 0 (*not at all important*), 1 (*slightly important*), 2 (*moderately important*), 3 (*very important*), 4 (*extremely important*); (d) display of *only the labels* but not the numbers (e.g., Krosnick & Presser, 2010).

6.1   Method

Preliminary Stages. More than 100 items were constructed and then pretested in Prof. Maier's colloquium for graduation candidates held at the University of Koblenz-Landau in April 2010. Members (N = 14) were one professor, one post-doc, three doctoral candidates, and nine diploma candidates—all from the research fields of psychology, education, or social sciences. Results led to minor wording revisions and the exclusion of some items. The final version—including 93 items (see Table B1) for measuring SMEC—was created as an online-questionnaire and subjected to another pilot test. This pilot test was conducted as a mix of thinking-aloud-technique (n = 2) and exit interviews (n = 6) by two interviewers. Think-alouds were performed via Skype or face-to-face communication. Exit interviews were conducted by telephone or e-mail. The interviewees were asked to report difficulties in comprehending questions or giving answers. The interviewers' discussion of the results led to minor changes in item wording and the inclusion of an example how to fill out the online questionnaire (see Appendix B) because four of the eight interviewees recommended this.

Participants and Procedure. Cover letters including the link to an online questionnaire were distributed via mailing lists to students and colleagues of the

University of Koblenz-Landau. Participants were offered a chance to win shopping vouchers for Amazon as an incentive to take part in the study. To further increase the number of participants, all recipients were asked to forward the link to other persons they knew. The "Start"-button on the welcome page was clicked by 659 people; 506 of them completed all SMEC items (Sample 1), 500 participants completed the entire questionnaire (age: M = 30.4, SD = 11.9, range = 16–77; 65% female;[19] 93% had at least finished secondary school).

Measures. Besides socio-demographic and other items not reported here, the major part of the online questionnaire consisted of 93 items to measure SMEC and two items to check for methodological issues (i.e., perceived difficulty of responding to questions, specific reference movie(s) for answering the items). The response scale has already been described above.

Data Analytic Procedure. Prior to analysis and using SPSS, all 93 variables were examined for multivariate outliers, univariate normality by assessing skewness and kurtosis (skew < 2, kurtosis < 7; West, Finch, & Curran, 1995), and item difficulties (cut-off criteria for exclusion were < .20 or > .80; cf. Dahl, 1971). By using Mahalanobis distance with p < .001, at least four cases were identified as multivariate outliers. However, individually inspecting their responses did not allow for straightforward interpretations. Therefore, and because hardly any theoretical background is available for guiding the interpretation, no outliers were excluded from Sample 1. The distributions of 13 items were severely skewed, had a highly positive kurtosis, or showed unsatisfactory item difficulties; therefore, these items were excluded. Using Sample 1 (N = 506), the remaining 80 criteria items were analyzed with CEFA software (Comprehensive Exploratory Factor Analysis; Browne, Cudeck, Tateneni, & Mels, 2010).

---

[19] In 2010, 64% of all students at the University Koblenz-Landau were female.

Following the recommendations of Fabrigar, Wegener, MacCallum, and Strahan (1999), exploratory factor analysis (EFA; maximum likelihood [ML] extraction, oblique geomin rotation, no row standardization; cf. Browne, 2001) was run to assess the dimensionality of the movie criteria and identify reasonable corresponding indicators.[20] RMSEA, RMSEA "close" fit (CFit; RMSEA ≤ .050), expected cross-validation index (ECVI), and parallel analysis were investigated to determine the number of factors. Items without salient loadings (less than |.40| or less than |.50| for the upper bound of the 90% CI) or with multiple salient cross-loadings were removed considering 90% CI for the loadings in the pattern matrix.

6.2   Results

For the 80 items included in EFA, a good model fit was obtained for a 10-factor solution (RMSEA = .047, 90% CI of RMSEA [.045, .048], CFit = .999; ECVI = 13.25; parallel analysis also suggested a 10-factor solution). However, 25 items had no salient loading on any of these factors. Thus, these items were excluded, and EFA was run again with the remaining 55 items. This procedure was repeated until adequate fit, salient loadings, and non-salient cross-loadings were achieved. After a total of eight processing cycles, an adequate fit was achieved for an eight-factor solution with 32 items (RMSEA = .052, 90% CI of RMSEA [.046, .057], CFit = .291; ECVI = 2.28; parallel analysis also suggested an eight-factor solution). The geomin-rotated pattern matrix of the eight-factor solution is presented in Table 3.

---

[20] The Kaiser-Meyer-Olkin coefficient was good (.85), measure of sample adequacy coefficients were all above .70, and the Bartlett test of sphericity was significant, thus indicating applicability of EFA (Bühner, 2011). Mardia's coefficient for multivariate skewness Mardia (1970) was not statistically significant, thus indicating no violation of multivariate normal distribution. The coefficient for multivariate kurtosis was statistically significant. However, the pattern matrix of the maximum likelihood extraction correlated over .98 with the pattern matrix of the principal axis factor analysis, which requires no distributive assumptions, thus indicating appropriateness of applying maximum likelihood extraction.

Table 3

Latent Structure of Movie Evaluation Criteria: Descriptive Statistics (M, SD), Results of ML-EFA (Loadings in the Pattern Matrix, CI, and Communalities h^2; N = 506)

No.[a,b]	Wording	M	SD	SV	SI	CI	FX	RE	IN	LH	CS	h^2
sv1 [7]	that the story the movie tells is based on true facts (e.g., based on a historic event or tells the life story of a real person)?	0.86	0.98	**.73** [.66, .81]								.55
sv2 [8]	that the story the film tells is realistic?	1.72	1.21	**.65** [.58, .72]								.43
sv3 [9]	that the movie addresses contemporary issues?	0.88	0.91	**.56** [.49, .64]								.40
si1 [28]	that the movie tells a story in a novel way?	1.98	1.07		**.71** [.61, .82]							.56
si2 [29]	that the story the film tells is unusual?	2.12	1.06		**.69** [.59, .80]							.54
ci1 [38]	the film's camera work and the shots?	2.54	1.04			**.90** [.87, .93]						.79
ci2 [37]	the way how the film is cut or how individual shots are cut together?	2.43	1.08			**.83** [.80, .86]						.71
ci3 [39]	the illumination and lighting in the movie?	2.11	1.10			**.79** [.75, .83]						.64
ci4 [40]	the color scheme in the movie (e.g., the use of black-and-white, red as a symbol or signal color, etc.)?	2.27	1.09			**.58** [.52, .64]						.49
fx1 [45]	that the production of the film was extravagant?	1.12	1.03				**.68** [.62, .74]					.50
fx2 [43]	the digital (post)editing of the film?	1.42	1.13			.20 [.13, .27]	**.67** [.60, .73]					.57
fx3 [42]	the movie illusions (e.g., special effects such as fire, explosions, stunts, combat scenes)?	1.73	1.17				**.66** [.60, .73]					.54
fx4 [46]	the technical design of the film altogether?	1.91	1.08			.23 [.16, .31]	**.64** [.57, .71]					.53
re1 [58]	the award(s) for the film (e.g., Oscar, Golden Globe, Golden Palm, Golden Bear, German Film Award, Grimme Award)?	1.03	0.98					**.79** [.73, .84]				.61
re2 [59]	the reviews of the film in the "media" (e.g., press, radio & TV, movie sites in the Internet)?	1.19	1.04					**.71** [.65, .77]				.47
re3 [57]	the opinions of friends, acquaintances, etc. about the film?	1.72	1.04					**.58** [.51, .64]				.39
re4 [59]	how the film is advertised (e.g., on TV or movie trailers, posters, newspaper, and magazine ads)?	1.02	0.96				.25 [.16, .33]	**.51** [.44, .58]				.38
re5 [93]	the fact that the film is considered a classic movie?	1.14	1.08					**.45** [.38, .52]				.30

(*table continues*)

Table 3 (*continued*)

No.[a]	Wording	M	SD	SV	SI	CI	FX	RE	IN	LH	CS	h²
in1 [79]	that the movie is free of scenes that you find disgusting?	1.81	1.53						**.92** [.89, .95]			.83
in2 [81]	that the movie is free of scenes that make you angry?	1.41	1.35						**.83** [.79, .87]			.73
in3 [77]	that the movie is free of scenes that you find frightening?	1.43	1.36						**.73** [.69, .78]			.61
in4 [16]	that the movie is free of scenes containing violence?	1.10	1.21	.24 [.15, .32]					**.52** [.45, .58]			.39
lh1 [67]	that the film puts you in a cheerful mood?	2.24	1.18							**.83** [.78, .89]		.70
lh2 [62]	that you find the movie humorous?	2.37	1.00							**.70** [.63, .77]		.49
lh3 [72]	that you find the movie entertaining?	3.03	0.84							**.59** [.52, .67]		.43
lh4 [65]	that you find the movie relaxing?	2.21	1.17							**.57** [.50, .65]		.38
lh5 [83]	that the movie takes your mind off everyday things?	2.60	1.13							**.43** [.34, .51]		.26
cs1 [87]	that the movie is thought-provoking?	2.73	1.00								**.82** [.78, .87]	.69
cs2 [86]	that the movie is intellectually challenging?	2.41	1.03								**.82** [.76, .87]	.67
cs3 [88]	that the movie communicates values?	1.98	1.16								**.57** [.50, .64]	.43
cs4 [89]	that watching the film broadens your knowledge?	2.15	1.08	.20 [.11, .29]							**.58** [.50, .65]	.47
cs5 [84]	that the movie motivates you to do something (e.g., politically, socially, search for information)?	1.65	1.10								**.56** [.49, .63]	.41

Note. ML-EFA = exploratory factor analysis with maximum likelihood extraction; factor loadings < .20 are omitted; factor loadings > .40 are in bold. 90% CI are reported in brackets. SV = Story Verisimilitude, SI = Story Innovation, CI = Cinematography, FX = Special Effects, RE = Recommendation, IN = Innocuousness, LH = Light-heartedness, CS = Cognitive Stimulation.

[a]The number in brackets refers to the corresponding German wording in Table B1 in Appendix B. The items were introduced with the phrase "When evaluating movies, how important for you personally is/are…"

The first factor is labeled *Story Verisimilitude* (SV) and comprises three items reflecting correspondence to (contemporary) reality (e.g., "…the story the movie tells is based on true facts"). The second factor is labeled *Story Innovation* (SI) and consists of two items reflecting the originality of the story (e.g., "…the story the film tells is unusual"). The third factor is labeled *Cinematography* (CI) and comprises four items reflecting core cinematic techniques (e.g., "…the film's camera work and the shots"). The fourth factor is labeled *Special Effects* (FX) and consists of four items also reflecting cinematic techniques, but focusing more on technical effects (e.g., "…the digital editing of the film"). The fifth factor is labeled *Recommendations* (RE) and comprises five items reflecting external resources for film evaluation (e.g., "…the award(s) for the film"). The sixth factor is labeled *Innocuousness* (IN) and consists of four items reflecting lack of potentially unpleasant characteristics (e.g., "…the movie is free of scenes containing violence"). The seventh factor is labeled *Light-heartedness* (LH) and comprises five items reflecting amusement and escapism (e.g., "…you find the movie humorous" or "…that the movie takes your mind off everyday things"). Finally, the eighth factor is labeled *Cognitive Stimulation* (CS) and consists of five items reflecting the viewer's cognitive processes such as cogitation or learning (e.g., "…the movie is thought-provoking").

As can be seen from the latent factor correlations in Table 4, some factors are positively moderately correlated (i.e., SV and CS, LH and FX, LH and IN, LH and RE, CI and FX), while the rest of them show only low positive or no correlations (one exception is the low negative correlation between SI and IN; *Mdn* of all latent correlations is .14). Taken together, most of the factors are not independent of each other. However, they share only little variance. Therefore, it seems to be justified to interpret them as distinct dimensions. Because none of the factors is highly correlated with another, it is also

unlikely that there are underlying higher factors. Nevertheless, this is an empirical question that could be addressed in future research.

Table 4
Correlations Between Latent Factors, Confidence Intervals, and Raykov's Rho

	SV	SI	CI	FX	RE	IN	LH	CS
SV	*.71*							
SI	-.04 [-.13, .05]	*.76*						
CI	-.01 [-.09, .07]	**.14** **[.04, .24]**	*.82*					
FX	-.02 [.11,. 07]	**.16** **[.05, .26]**	**.25** **[.17, .32]**	*.75*				
RE	**.19** **[.10, .27]**	**.17** **[.08, .26]**	.06 [-.02, .14]	**.13** **[.04, .21]**	*.72*			
IN	**.15** **[.07, .23]**	**-.10** **[-.18, -.01]**	-.02 [-.09, .06]	.06 [-.02, .14]	.08 [.00, .16]	*.86*		
LH	**.11** **[.02, .19]**	**.17** **[.08, .25]**	.04 [-.04, .12]	**.35** **[.27, .42]**	**.28** **[.19, .36]**	**.34** **[.27, .41]**	*.78*	
CS	**.36** **[.28, .44]**	**.20** **[.11, .28]**	**.17** **[.10, .25]**	-.02 [-.10, .07]	**.15** **[.06, .23]**	-.03 [-.11, .05]	.08 [.00, .16]	*.79*

Note. N = 506. ML factor analysis, geomin rotation; Raykov's rho in the diagonal (Raykov, 2001); latent factor correlations (90% CI in brackets) below the diagonal; statistically significant correlations in bold. For abbreviations see Table 3.

6.3 Discussion

Two items were dropped during item analysis—actor performance and story theme—because of poor psychometric properties (e.g., low variances < .80). Because these items were rated as most important, they seem to be crucial in the evaluation of movies. However, it goes without saying that we would not ask people how important moving images are to them when they evaluate a movie—an item that is analogous to the two items above. Questions of this kind—as well as the items regarding actor performance and story theme—seem to be the sine qua non for perceiving a film as *cinematic* (e.g., in contrast to *photographic*).[21] To put it another way, the results suggest that the general importance of actor performance and story theme might not perform well in differentiating film viewers' evaluation criteria. Of course, actor performance and

---

[21] See Carroll (2003) for a discussion of film as film and the classical film theories of Arnheim (1932) and Kracauer (1960), both related to film evaluation.

story theme might perform excellently when it comes to evaluating a specific film because these aspects are seen as important. Thus, future research that seeks to develop scales for the post-viewing evaluation of movie characteristics should include these aspects, whereas a general instrument—such as the one developed here—would gain little knowledge from including these criteria.

Throughout the iterative process of item reduction, the number of factors remained relatively constant. However, a ninth factor, which included four items especially referring to identification with protagonists, vanished because these items showed low salient loadings or cross-loadings mainly on other factors and, therefore, were finally dropped from analyses. How can the low loadings and disappearance of these items be interpreted and justified? Identification with characters is an important aspect of understanding audience responses (e.g., Cohen, 2001, 2006). However, when looking at the items researchers used to study identification, it is obvious that most of them were connected to specific characters or a specific movie.[22] Consequently, these items can only be answered *after* watching a movie. This is very similar to approaches that measure narrative engagement or transportation (e.g., Busselle & Bilandzic, 2009; Green & Brock, 2000). To think about how important it is to identify oneself with protagonists with unknown characteristics might be possible in general, but certainly not as easy as to state how important it is to avoid disgusting scenes, like car stunts, or puzzle over a whodunit. This seems to be consistent with low loadings of the identification items on RE suggesting that perhaps the importance of prior information about the movie, the actors, or the characters is related to the importance of identification. Again, this question should be addressed by future research.

---

[22] For example, "I think I have a good understanding of character X." (Cohen, 2001), for a story about asylum seekers in the Netherlands: "I felt bad for the asylum seeker." (de Graaf et al., 2009), or "I identified with the characters." after watching a specific movie such as *The Sea Inside* (Igartua, 2010).

Finally, for measuring SI more accurately it seems advisable to include a third item. This was done in the next study described in the section below.

7.   Phase III: Validating the Latent Structure

In Phase I, domain-relevant criteria were gathered, categorized, and prepared for item development, thus referring to content validity. In Phase II, responses to these items were examined by means of item analysis and EFA, which provided insights into the latent structure of SMEC. The aim of Phase III was to cross-validate the latent structure discovered in Sample 1 with another sample and to further support structural validity and aspects of generalizability (e.g., replication).

7.1   Method

Participants and Samples. For the purpose of validating the latent structure, two samples (Sample 2a and 3) were combined. Similar to Sample 1 (Study 1), Sample 2a also resulted from an online-study (Study 2); however, the recruiting strategy differed in three ways. First, 10 undergraduates received course credit for distributing a link to an online questionnaire via their social network profiles or mailing lists. Second, a press release including a link to the study was distributed (e.g., via the news ticker of the Informationsdienst Wissenschaft; idw, http://idw-online.de/en). Third and most importantly, I contacted a variety of special interest online platforms (e.g., critic.de, moviepilot.de, kino.de, kino-zeit.de, tvspielfilm.de) asking for support of my study. Fortunately, some of them posted the link on their news website, in their weekly newsletter, on Facebook, Twitter, or online forum. The "Start"-button on the welcome page was clicked by 849 people; 659 of them completed all SMEC items (Sample 2a), 587 of them completed the entire questionnaire (age: M = 29.7, SD = 10.6, range = 14–72; 55% female; 86% had at least finished secondary school; Mdn of cinema attendances

per month was 1). In an additional study (Study 3), 152 students of the University of Koblenz-Landau (Campus Landau) were recruited to participate in an experimental study (for further details see Bacherle, Schneider, & Retzbach, 2011, Exp. 2). Before they were exposed to different experimental conditions, they filled out questionnaires including the SMEC items; 147 completed all SMEC items (Sample 3, age: $M = 22.7$, $SD = 4.3$, range = 18–47; 64% female; over 85% were undergraduates enrolled in psychology or social sciences courses). Together, Sample 2a and Sample 3 yielded a total of 806 valid cases. The total sample was randomly split into two Samples R1 and R2 (R = random half) each comprising 403 cases.

Data Analytic Procedure. Following the rationale described by Brown (2006), exploratory factor analysis within the confirmatory factor analysis framework (E/CFA; e.g., Brown, White, Forsyth, & Barlow, 2004; Jöreskog & Sörbom, 1979) was conducted for analyzing Sample R1 ($n = 403$). One of the reasons for not directly moving into a more restrictive framework of analyzing the latent structure (e.g., CFA; for a comprehensive discussion see Brown, 2006, pp. 193–202) was the question if really all cross-loadings and error covariances could be fixed to zero—as it is done in CFA. Because the latent structure is far from being established at this early stage of scale development (e.g., the latent structure has not yet been validated in a different sample), E/CFA is seen as an appropriate intermediate step between EFA and CFA. To briefly illustrate the conceptual differences between EFA, E/CFA, and CFA, Figure 1 provides example path diagrams for each approach. In EFA, every item is allowed to freely load on both factors; both factor variances are fixed to 1. In E/CFA, anchor or marker items (black squares in Figure 1, B) are only allowed to load on their corresponding factor (their cross-loadings are fixed to zero; not displayed in Figure 1, B), whereas all other items are allowed to freely load on both factors; both factor variances are fixed to 1. In

CFA, every item is only allowed to load on its corresponding factor; cross-loadings are fixed to zero; to set the metric of the latent factor one item for each factor is fixed to 1. In all three models the factors are correlated, and the measurement error is random.
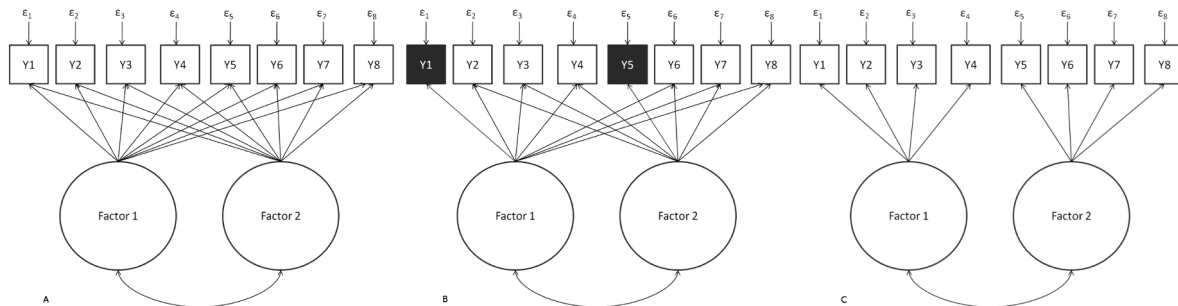
Figure 1. Path diagrams for (A) EFA (oblique rotation), (B) E/CFA, and (C) CFA. See text or Brown (2006) for details.

EFA and E/CFA provide the same degrees of freedom and the same overall fit indices. However and in addition to EFA, E/CFA provides further tools for inspecting model misspecifications (e.g., a variety of fit indices, standardized residuals, or modification indices). To avoid capitalization on chance, the second random sample (R2; *n* = 403) was used to replicate the findings from E/CFA in Sample R1. All analyses were conducted with EQS 6.1 (Bentler & Wu, 2005). Table 5 reports common cutoff values for assessing model fit.

Table 5
Indices and Cutoff Values for Assessing Model Fit in E/CFA and CFA

Model fit[a]	χ^2/df	RMSEA (90% CI)	SRMR	CFI
Good	≤ 2.00	≤ 0.06 (0.00 incl.)	≤ 0.08	0.950–1.000
Acceptable	2.01–3.00	≤ 0.08	< 0.11	0.900–0.949
Poor	> 3.00	> 0.08	≥ 0.11	< 0.900

Note. RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; CFI = comparative fit index.

[a]See Brown and Cudeck (1992), Hu and Bentler (1999), Marsh, Hau, and Web (2004), and Schweizer (2010).

Measures. All questionnaires included the final 32 items from Study 1. Additionally, one item (Item si3) that was supposed to load on Story Innovation was

added as a third indicator for this factor.[23] Besides socio-demographic items, the questionnaires also included measures to investigate external validity. They are not of interest in this part of the validation process, but will be reported in detail in Phase V.

7.2   Results and Discussion

E/CFA and CFA with robust ML estimation were conducted because of violation of multivariate normality. Therefore, model fit was evaluated with robust test statistics. Table 6 provides model fit indices for four models.

Table 6
Fit Indices for Model Comparison of EFA, E/CFA, and CFA Models of SMEC

Model	Analysis	Sample (*N*)	SB-χ^2/*df*	rRMSEA	rRMSEA 90% CI	SRMR	rCFI
1a	EFA	1 (506)	2.35	.052	.046, .057	-	-
1b	E/CFA	R1 (403)	1.39	.031	.024, .038	.020	.977
2	E/CFA	R1 (403)	1.16	.020	.000, .031	.017	.993
2	CFA	R2 (403)	1.92	.048	.042, .053	.061	.926

Note. SB-χ^2/*df* = Satorra–Bentler χ^2/*df;* rRMSEA = robust root mean square error of approximation; SRMR = standardized root mean square residual; rCFI = robust comparative fit index.
Model 1a: ML, 8 factors, 32 items.
Model 1b: Robust ML; 8 factors, 33 items: Item si3 added;
Model 2:   Robust ML; 8 factors, 28 items: Item si3 added; Items ci3, fx4, re3, lh5, and cs2 removed

Model 1 refers to the latent structure and number of items found in Sample 1. For comparison, Model 1a represents the findings of ML-EFA from Phase II. Note that it contained only 32 items. Item si3 was added in Studies 2 and 3 because the factor SI consisted of only two items. Hence, the corresponding Model 1b (E/CFA, Sample R1) consisted of eight factors and 33 items. Although Model 1b clearly replicated the pattern found in Model 1a (and even fitted the data from Sample R1 better than Model 1a fitted the data from Sample 1), modification indices (not reported here) suggested that freely estimating some error covariances would improve the model. Additionally, several

---

[23]   German wording for Item sv3 was "dass ein Film etwas zeigt, was vor ihm noch kein anderer gezeigt hat?" It translates as ". . . that a movie shows something that has never been shown in a film before?"

cross-loadings were salient. However, freely estimating error covariance parameters or excluding items needs theoretical justification.

On the one hand, the reasons for correlated error variances seem to lie in redundant item content. For instance, Item **cs1** ("that the movie is thought-provoking") and Item cs2 ("that the movie is intellectually challenging") are obviously synonymous. In a similar vein, Items ci3 and **ci4** as well as **lh4** and lh5 seemed to be semantically related (for item wording see Table 4). Based on their psychometric properties (e.g., magnitude of factor loadings, salient loadings on other factors), the items in bold were preferred, while the competing ones were dropped. On the other hand, two items (fx4 and re3) were excluded because their loading patterns were ambiguous. For instance, re3 had only low loadings on RE ($\lambda = .29$), but also loaded similarly on LH ($\lambda = .19$). One reason might be that friends provide information concerning the movie's entertainment value, whereas the remaining items of the RE dimension cover aspects of the movie itself. The Item fx4 seems to be a summary item for the FX dimension; it also statistically significantly cross-loaded on CI ($\lambda = .36$) and RE ($\lambda = .12$) and showed the lowest loading on FX ($\lambda = .57$) compared to the other FX-items. Therefore, it was excluded, too. The E/CFA of Model 2 (Sample R1) includes the remaining 28 items and shows improved model fit (e.g., ΔrCFI = .016; rRMSEA includes zero as lower bound CI).

Further support for structural validity comes from the replication of Model 2 via CFA in Sample R2 (see Table 6). Again, nearly all fit indices indicate good model fit, except the rCFI, which shows only acceptable model fit. However, this difference can be explained by a slightly distorted simple structure (i.e., additional, statistically significant cross-loadings, which can only be detected with E/CFA but are unspecified in CFA). Comparing the factor loadings of the E/CFA-Model 2 with the ones in the CFA-Model 2 in Table 7 clearly indicates such a slightly distorted simple structure. In a simulation

study on fit indices, Beauducel and Wittmann (2005) demonstrated that χ^2/df, RMSEA, and SRMR are less affected by low main loadings and slight distortions of the simple structure than incremental fit indices like the CFI. Thus, the CFA results support the findings from E/CFA in an independent sample.

The correlations between the latent SMEC factors (CFA with the total sample $N = 806$) are shown in Table 8. Again (see Phase II), the magnitude of the correlations clearly indicates distinct dimensions, although the positively moderate correlations between SV and CS, CI and SI, CI and FX, FX and LH, and IN and LH also show that these dimensions are not fully independent of each other.[24] Raykov's rho and Cronbach's α range from rather low (SV) to satisfactory (e.g., IN) reliability estimates (see diagonal in Table 8).

---

[24]   Although the median of the latent correlations (.16; computed from Table 8) is slightly higher than the corresponding median in Sample 1 (.14), this difference is statistically not significant ($z = .29$, $p > .05$, two-tailed). Despite the danger of inflated correlations when ignoring substantive cross-loadings (Marsh et al., 2009), this seems not to be the case here. Otherwise, comparing the patterns of latent correlations between Table 8 (CFA) and Table 4 (EFA) shows some differences. However, these differences are difficult to interpret because of the different sample sizes, the different number of items, and the different statistical analyses.

Table 7

Latent Structure of the SMEC Scales: Exploratory Factor Analysis within the Confirmatory Factor Analysis Framework (E/CFA) and Confirmatory Factor Analysis (CFA)

	SV		SI		CI		FX		RE		IN		LH		CS	
Item	E/CFA[a]	CFA[b]	E/CFA	CFA	E/CFA	CFA	E/CFA	CFA	E/CFA	CFA	E/CFA	CFA	E/CFA	CFA	E/CFA	CFA
sv1	**.72**	**.71**														
sv2	**.57**	**.61**					-.21									
sv3	**.65**	**.55**														
si1			**.71**	**.70**												
si2			**.82**	**.66**					-.21				.20			
si3	.15		**.74**	**.70**												
ci1					**.90**	**.88**										
ci2					**.80**	**.89**										
ci4					**.58**	**.62**										
fx1							**.62**	**.72**	.13							
fx2					.23		**.64**	**.77**								
fx3							**.65**	**.76**								
re1									**.71**	**.70**						
re2	.17		.23						**.69**	**.75**			-.17			
re4									**.51**	**.59**			.20			
re5									**.54**	**.52**						
in1											**.84**	**.82**				
in2											**.59**	**.71**				
in3											**.84**	**.83**				
in4	.20										**.68**	**.81**				
lh1													**.82**	**.82**		
lh2													**.74**	**.81**		
lh3	-.23												**.70**	**.52**		
lh4													**.70**	**.70**		
cs1	-.25														**.71**	**.65**
cs3											.22				**.80**	**.75**
cs4															**.78**	**.72**
cs5			-.22												**.88**	**.79**

Note. Factor loadings with $p > .05$ are omitted; factor loadings $> .40$ are in bold. For abbreviations of the SMEC dimensions see text.

[a]$n_{R1} = 403$.

[b]$n_{R2} = 403$.

[89]

Table 8

Correlations Between the Latent Factors and Reliability Estimates (N = 806)

	SV	SI	CI	FX	RE	IN	LH	CS
SV	*.66/.65*							
SI	**.07**	*.74/.74*						
CI	-.06	**.32**	*.86/.83*					
FX	.06	**.20**	**.34**	*.80/.80*				
RE	**.13**	**.15**	.06	**.16**	*.72/.72*			
IN	**.18**	-.03	**-.09**	.04	**.10**	*.86/.86*		
LH	**.17**	.04	**-.10**	**.36**	**.12**	**.38**	*.83/.83*	
CS	**.36**	**.14**	**.12**	.04	**.18**	**.14**	**.11**	*.82/.82*

Note. Diagonal: Raykov's rho/Cronbach's alpha; below diagonal: latent factor correlations; $p < .05$ in bold. For abbreviations of the SMEC dimensions see text. ML-CFA model fit: Satorra-Bentler χ^2/df = 2.82, Robust RMSEA = .048, 90% CI of Robust RMSEA [.044, .051]. SRMR = .059, Robust CFI = .924.

To sum up Phase III, structural validity of the SMEC was supported by replicating the structure with a different sample. One item was added, and five items were removed, thus resulting in eight dimensions comprising 28 items. Some of the eight dimensions are not independent from each other. Although fit indices of the E/CFA indicate that cross-loadings account for the data better than the more restrictive CFA model, the latter also shows an acceptable model fit. For the sake of parsimony, the CFA is preferred, albeit the factor correlations might be somewhat inflated (cf. Marsh et al., 2009). In Phase IV, I investigate the stability of SMEC to answer the question whether SMEC are more stable or more transient constructs.

8. Phase IV: Investigating the Stability of SMEC— Reliability, Common Consistency, Occasion Specificity, and Method Specificity[25]

Traditional trait approaches have been repeatedly criticized in differential and personality psychology; this has led to several alternative conceptualizations: situationism (e.g., Mischel, 1968), use of aggregation to defend the trait concept (e.g., Epstein & O'Brien, 1985), moderator approach (e.g., Bem & Allen, 1974); or interactionism (e.g., Bowers, 1973; Endler & Magnusson, 1976; Sarason, Smith, & Diener, 1975). It is generally understood that measurement always takes place in a situation, and a measured construct does not only contain components of stable and error variance, but also situational or interactional variance. However, if situational effects substantially contribute to a construct, it would be difficult to argue for conceptualizing it as a stable trait. Therefore, it is important to examine whether SMEC should be best conceptualized as rather stable attitudes (high proportion of trait component, low proportion of situational component) or rather temporary constructions (high proportion of situational component, low proportion of trait component). However, this question cannot be solved by means of classical test theory (Lord & Novick, 1968). For instance, a participant who answered LH-items on a first measurement occasion might answer these items differently on a second measurement occasion. It would be premature to interpret the differences as changes in LH—or even as evidence for an unreliable measure—because it is also possible to think of situational circumstances accounting for a temporary shift (e.g., mood states or remembering a

recently seen movie when answering the items). Latent state–trait (LST) theory (e.g., Steyer, 1987; Steyer, Ferring, & Schmitt, 1992; Steyer, Schmitt, & Eid, 1999) is an extension of classical test theory and—in contrast to generalizability theory (cf. Cronbach et al., 1963; Shavelson et al., 1989)—provides a non-experimental framework to answer the question about stability and situational or interactional effects by taking the different sources of variance into account.

The most common LST assumptions reflected in one of the basic models in LST theory is the latent state–trait model with method factors (see Figure 2 for such a model with two occasions and two indicators; cf. Steyer et al., 1992; Steyer et al., 1999). It contains (a) a latent trait ξ that reflects perfectly stable individual differences in the measured construct and is common to all measurement instruments and occasions (thus without index), (b) occasion-specific factors ζ_k that reflect systematic situational influences on measurement occasion k and are common to all instruments, (c) method-
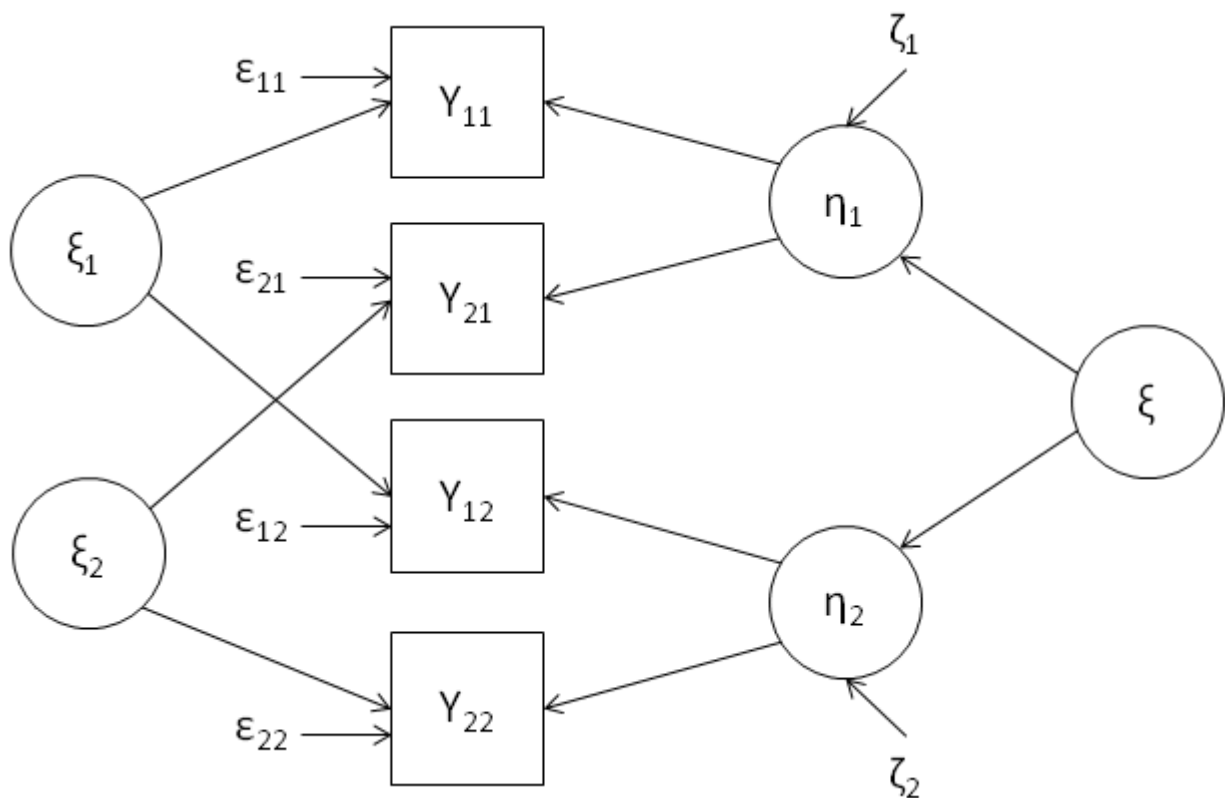
Figure 2. Latent state–trait model with method-specific factors for two occasions and two instruments of measurement (e.g., test halves). All factor loadings and effects are set equal to 1. See text for details.

specific factors ξ_i that reflect systematic variance of the measurement instrument i and are common to all occasions.

Moreover, LST theory defines several coefficients: common consistency (*cCon*), occasion specificity (*Spe*), method specificity (*mSpe*), and reliability (*Rel*). Their definitions are as follows (Equations 1–4; for exact formulas see Steyer et al., 1992):

$$cCon\ (Y_{ik}) = \mathrm{Var}\ (\xi)/\mathrm{Var}\ (Y_{ik})\,, \tag{1}$$

$$Spe\ (Y_{ik}) = \mathrm{Var}\ (\zeta_k)/\mathrm{Var}\ (Y_{ik}), \tag{2}$$

$$mSpe\ (Y_{ik}) = \mathrm{Var}\ (\xi_i)/\mathrm{Var}\ (Y_{ik}), \tag{3}$$

$$Rel\ (Y_{ik}) = [\mathrm{Var}\ (\xi) + \ \mathrm{Var}\ (\zeta_k) + \ \mathrm{Var}\ (\xi_i)]/\mathrm{Var}\ (Y_{ik}) \tag{4}$$

$$= \ \mathrm{Var}\ (\eta_k)/\mathrm{Var}\ (Y_{ik})$$

$$= \ cCon\ (Y_{ik}) + \ Spe\ (Y_{ik}) + \ mSpe\ (Y_{ik}).$$

All coefficients can be directly calculated from observed correlation or covariance matrices (Hagemann & Meyerhoff, 2008; Steyer & Schmitt, 1990). However, analyzing LST models within the CFA framework offers the opportunity to test assumptions about the measurement model and obtain fit indices of competing models.

To assess trait consistency, occasion specificity, method specificity, and reliability of the eight SMEC, LST analyses were conducted using data from two measurement points.

8.1   Method

Participants and Procedure. Exactly four weeks after completing the questionnaire from Study 2a (t1: $k = 1$), participants, who wanted to take part in the lottery ($n = 371$), were invited for the second part of the study (Study 2b; t2: $k = 2$). On average, the time between t1 and t2 was 31 days ($SD = 6.8$). The "Start"-button on the

welcome page was clicked by 282 people (participation rate: 76%); 275 of them completed all SMEC items (Sample 2b; age: M = 29.9, SD = 11.0, range = 14–70; 60% female; 88% had at least finished secondary school); 273 participants completed the questionnaire.

Measures. Study 2b contained all 33 SMEC items from Study 2a. Only the final 28 items from Phase III are used in the following analyses. Further measures were included to investigate construct validity and will be reported in detail in Phase V. Testing an LST model with method factors (cf. Figure 2) requires two manifest indicators for each of the eight latent variables (SMEC). Two test halves for each SMEC were generated by randomly assigning the indicators of each SMEC construct to either one or the other half (see Table C1 in Appendix C).

8.2   Results and Discussion

Table 9 shows fit indices for every LST model applied to each SMEC scale. Several restrictions had to be imposed to obtain good fitting models (for details see note to Table 9). Based on the estimated variance components of these models, the LST coefficients in Table 10 were calculated.

Table 9
Fit Indices for SMEC Latent State–Trait Models

Scale	χ^2	df	χ^2/df	RMSEA [90% CI]	SRMR	CFI
SV	0.95	1	0.95	0 [.00, .16]	.015	1
SI	0.36	1	0.36	0 [.00, .13]	.010	1
CI	6.79	4	1.70	.050 [.00, .11]	.032	.996
FX	0.22	1	0.22	0 [.00, .12]	.007	1
RE	1.26	5	0.25	0 [.00, .02]	.016	1
IN	7.52	5	1.50	.043 [.00, .10]	.032	.997
LH	5.05	4	1.38	.037 [.00, .10]	.023	.998
CS	9.07	5	1.81	.054 [.00, .11]	.050	.994

Note. N = 275. LST models were specified as follows: Error variances were freely estimated for SV, SI, and FX, constrained on each occasion for CI and LH and constrained on all occasions for RE, IN, and CS. Both method factors were freely estimated for SV and SI and constrained for RE and IN. One method factor was fixed to zero due to lack of significance for CS (test half 1) and for CI, FX, and LH (test half 2).

Table 10
Latent State–Trait Coefficients of the SMEC Scales for Both Measurement Occasions

Scale	Reliability		Common Consistency		Occasion Specificity		Method Specificity	
	t1	t2	t1	t2	t1	t2	t1	t2
SV	.81	.82	.57	.53	.12	.18	.12	.11
SI	.83	.87	.61	.65	.15	.14	.07	.08
CI	.88	.94	.59	.69	.24	.20	.05	.05
FX	.86	.89	.66	.66	.15	.18	.05	.05
RE	.84	.84	.54	.58	.20	.15	.10	.11
IN	.93	.94	.78	.70	.11	.20	.04	.04
LH	.88	.91	.67	.77	.18	.11	.03	.03
CS	.88	.88	.69	.67	.14	.16	.05	.05

Note. N = 275. Total scales (aggregation across test-halves within occasions; for formulas see Deinzer et al., 1995, and Steyer & Schmitt, 1990)

First, all SMEC scales demonstrate high reliabilities across both measurement occasions. On average, 88% is due to systematic variance indicating that the scales are reliable. Second, reliabilities are largely determined by stable interindividual differences (i.e., common consistencies were 65% on average). Third, indicators contain small, but statistically significant proportions of test-half-specific variance (i.e., method specificities were 6% on average). Inspecting the test halves in more detail (not reported in Table 10) leads to the conclusion that especially the one-item test halves of the three-item scales (i.e., SV and SI) might be optimized by adding further items. Finally, on average, 16% variance is due to systematic, but unstable effects of the situation or interaction. For instance, occasion specificities for IN and LH vary about 10 percentage points between t1 and t2. Additionally, SV, CI, and RE show high proportions of occasion specificity compared to common consistency. Obviously, this could be interpreted as the impact of movies just seen before t2 on these SMEC. To sum it up, these results clearly demonstrate that the scales reliably measure stable individual differences in SMEC that depend to a small extent on systematic situational or interactional effects present at the measurement occasion.

9.   Phase V: Examining the Nomological Network of SMEC

Phase IV provided support for generalizability (across two measurement occasions) and first insights into substantive validity (i.e., SMEC are largely due to stable individual differences and less to situational influences). The aim of Phase V was to locate the SMEC in the nomological network of related constructs by examining their discriminant, convergent, and criterion correlations as part of support for external validity. Although nothing is yet known about possible relationships to other constructs relevant to SMEC, it stands to reason that the different SMEC dimensions have their own particular relationships with (a) film-specific constructs like aesthetic fluency in film, film expertise, genre preferences, and concrete evaluations of movies (criterion validity), (b) thematically related constructs from specific domains like trendsetting (assumed association with SI and RE) and political interest (assumed association with SV and CS but dissociation with FX, IN, and LH), (c) more general personality traits, such as sensation seeking, the Big Five, need for cognition, and need for affect that all might contribute to specific SMEC (see Chapter 3.1), and (d) traits like socially desirable responding (desire for social approval) and need to evaluate, which both should correlate only low with all SMEC scales to demonstrate high discriminant validity.

9.1   SMEC and Related Constructs

Before describing the measurement of related constructs and nontest criteria, their theoretically assumed relationship with specific SMEC is provided in this section.

Story Verisimilitude. SV shows no or negative associations with aesthetic, artistic, or artificial aspects of a movie as they are reflected in almost completely

unrelated SMEC dimensions, such as SI, CI, and FX (see Tables 4 and 8). Therefore, on the one hand, negative correlations with film expertise and aesthetic fluency in film (Silvia & Berg, 2011) are expected.[26] On the other hand, the importance of realism and current themes as reflected in SV might be positively correlated with interest in current issues as facets of political interest and conscientiousness (e.g., O'Hara, Walter, & Christopher, 2009). Furthermore, SV should be negatively correlated with genre preferences that reflect unrealistic, fantastic, or science fiction stories.

Story Innovation. SI reflects original and novel aspects of a film's story. This should lead to positive associations with traits like sensation seeking, Openness to Experiences, and Extraversion. SI is also assumed to be positively correlated with film expertise and aesthetic fluency in film. In addition, diffusion research (e.g., Rogers, 2003) suggests that experts (e.g., film critics) and opinion leaders influence others in adopting innovations (e.g., viewing a specific recent movie; cf. Venkatraman, 1989). Trendsetters adopt innovations earlier than others, focus on the most important parts of innovations, and recommend innovations to others (Batinic, Haupt, & Wieselhuber, 2006; Batinic, Wolff, & Haupt, 2008). Thus, trendsetting should be positively correlated with SI. Innovative stories are often based on ideas about the future or include some kind of "experimental" adaptation of archetypical content (e.g., a hero who saves the world, the Romeo and Juliet love story). Therefore, SI should be positively correlated with preferences for science fiction and avant-garde genres, on the one hand, and show no correlations with preferences for other genres because they could be innovative or traditional, on the other hand.

---

[26] Film expertise and similar concepts such as film analyticity (Robinson, 1975), aesthetic fluency in film (Silvia & Berg, 2011), cineliteracy (British Film Institute, 1999), or film competence/education (e.g., Arbeitskreis Filmbildung, 2009, VisionKino) have been proposed, but there is little research that includes them. Two constructs—aesthetic fluency in film and film expertise—are described in more detail in the Method section and in the Appendices D and E, respectively.

Cinematography. CI focuses on visual aspects of movies like pictures, cutting, or colors and is moderately correlated with SI. This obvious association with aesthetic, artistic, and visual stimuli should be reflected in positive relationships to the same constructs as mentioned with regard to SI (i.e., Openness to Experiences, Extraversion, sensation seeking, film expertise, aesthetic fluency). Additionally, CI should also be positively correlated with preferences for avant-garde films and other genres in which visualization plays an important role (e.g., science fiction, animation), whereas there should be no correlation with genres that usually do not focus on visual aesthetics (e.g., comedy or romance). Furthermore, because CI and FX are positively correlated, it is assumed that slight differences in the magnitude of correlations might turn out, thereby reflecting the nuances of these dimensions (e.g., CI might be more related to serious films, whereas FX might be more related to light movies as reflected in the correlations of both dimensions with other SMEC such as LH and CS).

Special Effects. FX also reflects visual aspects but focuses more on additional technical effects. Special effects are often assumed to be innovative parts of a movie (e.g., movie trailers contain the most eye-catching special effects). Thus, FX should be positively correlated with trendsetting as well as with sensation seeking and Openness. However, FX might be negatively correlated with constructs that reflect more cognitive effort (e.g., need for cognition, political interest). It is also likely that FX positively correlates with the score of the Fun scale (Oliver & Bartsch, 2010) reflecting light entertainment. Positive relationships with genre preferences for action, animation, adventure, science fiction, comedy, and horror are expected. In contrast, negative relationships with more serious genres like drama, avant-garde, or documentary are assumed.

Recommendation. RE takes the assumed influence of peripheral information about films into account. Therefore, it should be related to trendsetting (especially, information input) as well as to film expertise and aesthetic fluency in film because for all these constructs background information plays an important role. RE is expected to be positively correlated with genres that receive critical acclaim in the media and awards. As Simonton (2005b) demonstrated, this especially applies to the drama genre. In addition, the importance of peripheral information might also be reflected in a positive correlation with the score on the Appreciation scale (Oliver & Bartsch, 2010) as viewers adopt the judgments of experts.

Innocuousness. IN means that there are no aversive scenes (i.e., film elements that are disgusting, frightening, violent, or enraging) and should be negatively associated with Emotional Stability and sensation seeking. Furthermore, negative relationships with preferences for film genres that usually include scenes with explicit content or that might elicit strong negative emotions are assumed (e.g., thriller, horror, erotic, tragedy), whereas positive relationships with preferences for not offensive, harmless genres like romance and (most often) comedy—as also reflected in the moderately positive correlation between IN and LH—are expected. Consequently, IN should also be positively correlated with the Fun scale.

Light-heartedness. LH contains items that clearly reflect the content of the Fun scale (Oliver & Bartsch, 2010). Therefore, a positive relationship between these two scales is assumed. Furthermore, this should be true for the relationship with preferences for light entertainment genres (e.g., comedy, romance), but not for serious genres (e.g., avant-garde, drama). With regard to personality traits, negative correlations with sensation seeking, Openness to Experiences, and need for cognition are expected, whereas positive correlations with Agreeableness and Conscientiousness

(Burst, 1999; Hall, 2005a) are assumed. Moreover, LH should also negatively correlate with political interest and cinematic, aesthetic knowledge.

Cognitive Stimulation. CS consists of items that reflect the content of the Appreciation scale (Oliver & Bartsch, 2010). Although it extends it by adding aspects such as broadening knowledge, communicating values, and taking action, CS should be positively correlated with Appreciation as well as with preferences for serious genres (e.g., drama, avant-garde, and documentary). As it is obvious from the factor's name, CS should also be positively related to constructs that reflect cognitive effort in information processing like need for cognition, trendsetting, and political interest as well as to traits that reflect engaging in stimulating activities, such as sensation seeking or Openness to Experiences.

In addition, I was also interested in the relationship between SMEC and the need for affect. However, the sparse and rather inconsistent findings do not allow the formulation of hypotheses (see Chapter 3.1).[27]

Finally, all SMEC scales should show at most only low correlations with socially desirable responding and the need to evaluate. Socially desirable responding (SDR) is "the tendency to give answers that make the respondent look good" (Paulhus, 1991, p. 17). This individual tendency has received wide attention in social psychological and personality research and was conceptualized in different ways—especially to examine bias in responses to self-report questionnaires (for an overview see Paulhus, 1991). Paulhus (e.g., 1984, 2002; Paulhus & Trapnell, 2008) distinguishes between self-deceptive enhancement (SDE; honest, but positively biased self-reporting) and

---

[27] For example, because the need for affect reflects approaching or avoiding emotions with only respect to strength but not valence, it remains unclear whether participants with a high score on the Approach subscale would rate absent negative emotions as important because they only pursue the experiencing of strong *positive* emotions (resulting in a positive correlation) or rate this as unimportant because they pursue the experiencing of *strong* emotions independent of their valence.

impression management (IM; consciously faking self-presentation to others). Surprisingly, media psychological and communication research have actually explored only the desirability of messages and media effects related to the third-person effect (e.g., Perloff, 1999; Reid & Hogg, 2005) and not paid much attention to social desirability as a response style (cf. Möhring & Schlütz, 2010). With regard to the measurement of SMEC, a high correlation with SDR would allow for alternative explanations, for instance, that participants have fulfilled their need for social approval—thus, the SMEC scales would have only measured SDR but not the importance of SMEC. Although this case seems to be unlikely because (a) the study is highly anonymous, (b) the dimensions seem to be very clearly interpretable, and (c) the topic is rather straightforward than senstive, empirical studies should be conducted to rule out this possibility. Likewise, the need to evaluate (NTE) construct—the "chronic tendency to engage in evaluative responding" (Jarvis & Petty, 1996, p. 185)—should correlate only low with the SMEC scales; otherwise, this would allow for the alternative explanation that participants' scores on SMEC scales depend only on the NTE score (e.g., the participants would score high on SMEC only because they have the tendency to build and hold strong attitudes toward any object; cf. Bizer et al., 2004; Britt, Millard, Sundareswaran, & Moore, 2009).

9.2 Method

Participants and Procedure. The samples used for the present analyses came from Studies 2a (Sample 2a: $N = 587$) and 2b (Sample 2b: $N = 273$), which have already been described in the Method sections of Phases III and IV. To estimate the latent correlations between SMEC and the other constructs, corresponding measures were included in both studies.

Measures. Following the recommendations of several scholars (e.g., Levine, Hullett, Turner, & Lapinski, 2006; Slaney & Maraun, 2008), results of CFA from the present research for the already published scales are also reported. These CFA results can also serve as a basis for judging how appropriate the measures are for inclusion in a simultaneous CFA with SMEC.

Subjective movie evaluation criteria (SMEC). The eight SMEC scales consisted of 28 items, the final version resulting from Phase III (see previous phases for psychometric properties). All items were included in Studies 2a and 2b.

Film genre preferences. Study 2a included 14 items for measuring film genre preferences (e.g., action, comedy, romance, drama, horror; for the complete list, see Table 11). Although some researchers combined several genres into broader categories (e.g., Gehrau, 2001; Hall, 2005b; Rentfrow et al., 2011), the results were rather inconsistent (see Chapter 3.1). Thus, it came as no surprise that an EFA resulted in many one-item factors, and thus, each genre was measured by a single item.[28]

Aesthetic fluency in film (AF). AF is a very recent construct introduced by Silvia and Berg (2011). The Aesthetic Fluency in Film scale is the corresponding self-report measure that aims to assess film expertise as a knowledge base for facilitating aesthetic experiences. It includes 22 terms reflecting important names related to film production, history, and theory (Silvia & Berg, 2011). Cronbach's α was .97; AF was moderately positively correlated with gallery visiting, Openness to Experiences, and film viewing.

For the present study, I used German terms for the names and adapted the response scale. In addition, other items were added to take the cultural differences into account (e.g., "The Paramount decision" or "Errol Morris" seem to be too specific to US-

---

[28] To conduct CFA, each single item measure was included as a perfectly reliable, single indicator for a corresponding latent variable. Although this assumption is questionable, it allows for measuring the SMEC on the latent level and correcting for measurement errors, while assessing the fit of the measurement model at the same time.

American film culture). A total of 22 items were included in Study 2a. To retain the one-dimensional structure proposed by Silvia and Berg (2011), item analysis and CFA were conducted and psychometrically problematic items were excluded under consideration of item difficulties, factor loadings, residual sizes, and modification indices. The 10 remaining items showed a Cronbach's α of .90 and provided a good fitting model (S-B χ^2/df = 2.32, robust CFI = .984, SRMR = .028, robust RMSEA = .048, 90% CI of robust RMSEA [.03, .06]). Furthermore, Pearson correlations with film expertise (r = .60, N = 273), film viewing frequency (r = .30, N = 563), knowledge of recent film awardees (r = .51, N = 273), and preferences for avant-garde films (r = .42, N = 587) were moderately to highly positive, thus providing support for criterion validity. Details on the psychometric properties of the German AF scale are provided in Appendix D.

Film expertise (FE). FE was measured in Study 2b by six newly developed items (see Appendix E for details). A second-order CFA showed good model fit (N = 273; χ^2/df = 1.42, CFI = .994, SRMR = .028, robust RMSEA = .04, 90% CI of robust RMSEA [.00, .09]) and satisfactory reliability (Raykov's rho = .85 for the total scale). FE and AF were highly correlated (see above).

Film-specific evaluation scales. For measuring film viewer's entertainment gratifications from specific movies, Oliver and Bartsch (Oliver & Bartsch, 2010) developed three scales: Appreciation (e.g., "I found this movie to be very meaningful."), Fun (e.g., "It was fun for me to watch this movie."), and Suspense (e.g., "This was a heart-pounding kind of movie."). Each scale consists of three items and Cronbach's alphas across several studies ranged from .75 to .93.

In the present research, a German version (Bartsch & Oliver, 2010) with similar Cronbach alphas as the English version was included in Study 2b. Before filling out the scales, participants were asked to recall the last movie they had seen and respond to the

gratification items with that movie in mind. Although this procedure has the disadvantage that film genres and content might be unequally distributed, it is the most economical approach and provides a very conservative test due to the heterogeneity of the included films.[29] Because SMEC were already obtained in Study 2a, evaluations concerning the last movie and SMEC at t1 are not confounded. A CFA led to an acceptable fit of the three-factorial model (e.g., S-B-χ^2/df = 2.97, CFI = .976, SRMR = .040, RMSEA = .085). Cronbach's alphas were .90 for Appreciation, .89 for Suspense, and .95 for Fun.

Big Five personality traits. The five dimensions—Extraversion, Emotional Stability, Openness to Experience, Conscientiousness, and Agreeableness—are often assessed using personality scales, such as the Revised NEO Personality Inventory (NEO-PI-R; Costa & McCrae, 1992), the Big Five Inventory (BFI; John, Donahue, & Kentle, 1991), or the International Personality Item Pool (IPIP; Goldberg et al., 2006). These instruments comprise very many items. For instance, the NEO-PI-R consists of 240 items. For the sake of practicability, the Ten-Item Personality Inventory (TIPI; Gosling, Rentfrow, & Swann, 2003) was used in Study 2a. It is a very brief measure with only two items per factor. The small number of items is one reason for the low internal consistencies for Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience (Cronbach's α = .68, .40, .50, .73, and .45 respectively). However, the test–retest reliabilities are within an acceptable range (r_{tt}s = .77, .71, .76, .70, .62 respectively) as well as convergent correlations with more-item measures (i.e.,

---

[29] For example, when a participant saw a serious drama, she or he probably would not rate it highly on the Fun scale, whereas a participant who viewed a comedy might rate highly on this scale. A less strict test would only examine appropriate movies (e.g., Fun only in relation to comedies). The only economically practical way would have been genre group analyses. However, firstly, there were too few cases to conduct powerful analyses. Secondly, even when genres are coded, this is no guarantee that the specific content was included in the movie (cf. *Dumb and Dumber* vs. *Wag the Dog* in Chapter 3.1). To explore the relationship between the SMEC scales and the ones by Oliver and Bartsch, experimentally manipulated groups should be preferred (cf. Oliver & Bartsch, 2010).

BFI, NEO-PI-R). Ehrhart and colleagues (2009) found support for the factor structure and the convergent validity of the TIPI. The German version of the TIPI (TIPI-G) was translated and validated by Muck, Hell, and Gosling (2007). Alphas for Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Openness to Experience were .57, .42, .66, .67, and .54 respectively. They also found support for the five-dimensional structure via CFA (i.e., χ^2/df = 1.56, CFI = .95, RMSEA = .06).

In the present research, the model fit of the Big Five measured with TIPI-G provided an acceptable fit (e.g., robust CFI = .91, SRMR = .058, robust RMSEA = .07).

Sensation seeking. This trait is often assessed by scales like the Sensation Seeking Scale Form V (SSS-V; e.g., Zuckerman, 1994) or the Arnett Inventory of Sensation Seeking (AISS; Arnett, 1994). Similar to the assessment of the Big Five personality traits, a short form was used for external validation. The Brief Sensation Seeking Scale (BSSS; Hoyle, Stephenson, Palmgreen, Lorch, & Donohew, 2002) consists of eight items comprising four facets—Experience Seeking (ES), Boredom Susceptibility (BS), Thrill and Adventure Seeking (TA), and Disinhibition (DI).[30] The internal consistencies for the BSSS are acceptable. Hoyle et al. (2002) reported Cronbach's alphas between .74 and .76 for the eight-item scale. Moreover, a one-dimensional CFA provided an acceptable model fit (e.g., CFI = .93, RMSEA = .07).

In the present study, the BSSS was transferred into German via translation–back-translation method. The scale was administered in Study 2a (t1) as well as four weeks later in Study 2b (t2). Cronbach's alpha for t1 was .77 (N = 587) and .78 (N = 273) for t2. Test–retest reliability was very high (r_{tt} = .90). The one-dimensional CFA model fit was

---

[30] The BSSS has another great advantage: It does not contain any items related to movie exposure or film preferences. The fact that the SSS-V contains items referring to the behavior that should be predicted by sensation seeking (e.g., drug abuse, sexual behavior) has been criticized for circularity (e.g., Arnett, 1994; Roth & Herzberg, 2004). Interestingly, to my best knowledge, this has not yet been discussed when the SSS-V was applied in media psychological research, even though it contains three items about movie preferences.

not as good as the findings reported by Hoyle et al. (2002). However, a four-dimensional model taking the four facets into account provided a good model fit (e.g., robust CFI = .95, SRMR = .042). This is consistent with more recent findings (Stephenson, Velez, Chalela, Ramirez, & Hoyle, 2007). Therefore, latent correlations were calculated for the subscales only.

Need for cognition (NFC). Cacioppo, Petty, and Kao (1984) derived an 18-item short version from the initial 34-item NFC scale (Cacioppo & Petty, 1982). They report a Cronbach's α of .90 and postulate a one-factor solution. The German version of the NFC was translated and explored by Bless, Wänke, Bohner, Fellhauer, and Schwarz (1994). Their one-dimensional short version (NFC-S) consists of 16 items with a Cronbach's α of .83. Recently, Betrams and Dickhäuser (2010) were the first who examined the factorial structure and reliability of the German NFC-S in three studies. Their confirmatory factor analysis provided support for the one-dimensional factor structure of the German NFC-S (i.e., SRMR = .059, RMSEA = .045); internal consistencies varied between .81 and .89, and test–retest reliability was .83.

In the present study, Cronbach's α was .87. However, CFA indicated a poor model fit for a one-factor solution: S-B-χ^2/df = 5.14, robust CFI = .82, SRMR = .068, robust RMSEA = .084, robust AIC = 326.41. Taking method artifacts into account, a one-factor model with correlated uniqueness among all negatively worded items (cf. Marsh, 1996) yielded a much better fit: S-B-χ^2/df = 2.46, CFI = .96, SRMR = .034, RMSEA = .050, robust AIC = 27.16. An alternative model with an additional second method factor accounting for the positively worded items was also tested (cf. Bors, Vigneau, & Lalande, 2006), but did not yield better fit indices (e.g., robust AIC = 114.15). Although a simple one-factor 16-item solution was not supported, this result is consistent with other findings considering method effects as an alternative explanation instead of a second substantive

factor (cf. Forsterlee & Ho, 1999; but see Lord & Putrevu, 2006, and Tanaka, Panter, & Winborne, 1988, for other multidimensional factor solutions). Consequently, the one-dimensional structure was retained. Additionally, building four random parcels out of the 16 items yielded an excellent model fit (χ^2/df = 1.06, CFI = 1, SRMR = .007, RMSEA = .010, AIC = -1.88).

Political interest. Political interest is the "degree to which politics arouses a citizen's curiosity" (van Deth, 1990, p. 278). To assess political interest more reliably than with usual single-item measures, a new scale—Short Scale Political Interest (SSPI) including five items—was recently introduced (Otto & Bacherle, 2011). The authors reported a Cronbach's α of .88 and the results of their CFA support one-dimensionality (e.g., χ^2/df = 1.63, CFI = .99, RMSEA = .06).

In the present study, the SSPI was administered in Studies 2a (t1) and 2b (t2). Cronbach's α for t1 was .94 (N = 587) and .79 (N = 273) for t2. Test–retest reliability was very high (r_{tt} = .88).[31] The model fit was excellent (t1; S-B χ^2/df = 1.60, robust CFI = .999, SRMR = .009, robust RMSEA = .03, 90% CI of robust RMSEA [.00, .07], Raykov's rho = .95).

Trendsetting (TDS). The TDS questionnaire consisting of 27 items and representing three facets—Input (I), Throughput (T), and Output (O)—was developed and validated by Batinic et al. (2006). Although the authors postulate three facets—or steps—of TDS, the one-dimensional model fitted their data best (e.g., CFI = .95, RMSEA = .049; Batinic et al., 2006). Batinic et al. (2008) derived a one-dimensional 9-item short scale with three items representing each facet (TDS-K). They reported Cronbach's alphas between .85 and .91 for four studies and a test–retest reliability of .72.

---

[31] The study I conducted to assess test–retest reliability is identical with the one Otto and Bacherle (2011) reported in their paper.

In the present study, the TDS-K was used for economical reasons. Cronbach's α for the TDS-K was .87 (.61 for I, .71 for T, .89 for O). The fit for the one-dimensional model led to ambiguous results: SRMR (.07) showed a good fit, robust CFI (.90) an acceptable fit, and robust RMSEA (.108) as well as S-B χ^2/df (7.85) only poor fit (robust AIC = 157.87). Therefore, a three-factorial measurement model—with each factor representing a facet of TDS—was fitted to the data and improved the model fit (e.g., robust CFI = .94, robust RMSEA = .09, 90% CI of robust RMSEA [.076, .105], SRMR = .05, robust AIC = 90.65). Because the factors were highly correlated (≥.70), a second-order factor was introduced (the fit was identical to the three-factorial first-order model). Although the model fit is far from excellent, taking the three-factorial structure into account seems to be an alternative way to model TDS.

Need for affect (NFA). The NFA scale developed by Maio and Esses (2001) consists of two subscales: Approach and Avoidance, each comprising 13 items. The German version was translated and validated by Appel (2008). Appel, Gnambs, and Maio (in press) derived a 10-item short version for the English as well as for the German NFA scale. They reported the following reliability estimates: For the two 5-item subscales and the full 10-item scale Cronbach's alphas varied between .72 and .82 across four samples. For the full 10-item scale, test–retest reliability was .87 (the reliability estimates obtained via latent state–trait analysis were .83 and .85 for t1 and t2, respectively, indicating only low situational specificity). In their scale development studies, however, the authors found little support for a one-dimensional measurement model. With or without method factor, the two-dimensional models yielded better fit indices. Appel et al. (in press) obtained acceptable fit only by allowing all items to cross-load on both factors (e.g., robust CFI between .93 and .98, SRMR between .03 and .05 in

four samples). Additionally, the factor correlations between Approach and Avoidance were only moderately high (between -.34 and -.46).

The same was found to be true in the present study using E/CFA to allow for cross-loadings (robust CFI = .95, SRMR = .04, robust RMSEA = .07; latent correlation between the two factors φ = -.36). Therefore, latent correlations are only calculated for the subscales.

Socially desirable responding (SDR). SDR was conceptualized sensu Paulhus (e.g., 1984) and measured with an adaption of two German short forms of the Balanced Inventory of Desirable Responding (BIDR-S-Ger; Musch, Brockhaus, & Bröder, 2002; Winkler, Kroh, & Spiess, 2006), which consists of two scales—Self-Deceptive Enhancement (SDE) and Impression Management (IM)—each comprising three items. Winkler et al. (2006) developed the BIDR-S-Ger as an economical instrument for conducting large-scale surveys (e.g., they used the German Socio-Economic Panel for development and validation). They reported internal consistencies (SDE: Cronbach's α = .60, IM: Cronbach's α = .55) and findings supporting the validity of the subscales. Bertrams and Dickhäuser (2009) reported Cronbach's α of .51 for SDE and .53 for IM.

In the present study, the six items were administered with the other measures described in this section at two measurement points: t1 and t2 (four weeks later). Participants rated the items on a five-point response scale (0 = *completely disagree* to 4 = *completely agree*). Cronbach's α was .58 for SDE and .50 for IM at t1 (N = 584) and .67 and .57 at t2 (N = 271), respectively.[32] Test–retest reliability was acceptable (r_{tt} = .74 for SDE and .72 for IM, respectively). The two-dimensional approach provided a good model fit (S-B-χ^2/*df* = 1.95, robust CFI = .98, SRMR = .04, robust RMSEA = .04). The

---

[32] Because the assumption of tau-equivalence is often violated, Cronbach's alpha usually misestimates scale reliability. The alternative for congeneric measures, Raykov's rho, was .60 and .58 for t1 for SDE and IM, respectively (for further discussion, see Raykov, 2001; or for a similar approach, McDonald's omega, see McDonald, 1999).

latent factors—SDE and IM—were only moderately positively correlated (φ = .39); therefore, latent correlations are provided for the subscales only.

Need to evaluate (NTE). The NTE scale consists of 16 items. Jarvis and Petty (1996) reported internal consistencies from .82 to .89 and a 10-week test–retest reliability of .84. Similar Cronbach's alphas were reported by others (e.g., Briñol, Petty, & Wheeler, 2006; Cronley, Mantel, & Kardes, 2010; Tormala & Petty, 2001). However, there are other studies showing mixed results. For instance, in studies used to develop the Spanish version of the NTE scale (Horcajo, Díaz, Briñol, & Gandarillas, 2008), parallel analysis, principal components analysis (PCA), and CFA showed better solutions when some items were excluded and two dimensions were considered—Need to Evaluate and Need for Neutrality. Cronbach's α reached .73 for the unidimensional scale, .73 for the Need to Evaluate subscale, and .70 for the Need for Neutrality subscale; test–retest reliability was .74, .72, and .73, respectively. According to studies using a translated version of the NTE scale in Germany, the picture is the same. Collani (2010)—who translated and validated the German version—reported a Cronbach's α of .69. Using PCA, he found that two components were more appropriate and recommended using two subscales—Need to Evaluate and Need for Neutrality (for both Cronbach's α was .70). Nuszbaum, Voss, Klauer, and Betsch (2010), however, could not replicate the two-dimensional structure and used the 16 translated items from Collani (2010) in a unidimensional way; in their study, Cronbach's α was .77. Brömer (2000) used a different German version of the NTE (translated on his own) and obtained a Cronbach's α of .58.

In this study, Cronbach's α of the German NTE scale from Collani (2010) only reached .71. However, after splitting the scale into two dimensions, α was even worse: .66 for Dimension 1 (Need to Evaluate) and .48 for Dimension 2 (Need for Neutrality).

Furthermore, using a CFA approach to test for method artifacts (i.e., correlated uniqueness among the negatively worded items; cf. Marsh, 1996) was not successful. Therefore, the scale was optimized for two purposes. First, confirmatory factor analysis should yield an acceptable model fit indicating at least tau-equivalence. Thus, Cronbach's α might be an appropriate estimator for reliability. Second, factor loadings should reach at least a minimum of .40. This was established by conducting exploratory and confirmatory factor analysis. In the end, a one-factor solution including only six items yielded an excellent model fit: χ^2/df = 0.44, CFI = 1, SRMR = .012, RMSEA = 0. Furthermore, factor loadings reached from .44 to .77, and Cronbach's α was .73. The scale mean of the six-item scale was highly correlated (r = .90) with the scale mean of the 16-item scale, thus showing equivalence and justifying the use of the shorter version.

9.3   Results and Discussion

Table 11 presents the latent correlations between the SMEC and the constructs described in the previous section. Most of the associations between SMEC dimensions and related constructs were as expected. However, some unexpected findings from Table 11 have to be mentioned and discussed for each SMEC scale.

Table 11
Latent Correlations Between the SMEC Scales and External Criterion Constructs

Study (*N*)	Validation Construct	SV	SI	CI	FX	RE	IN	LH	CS
	Film Genre Preferences								
2a (587)	Romance	.07	-.07	**-.10**	.05	**.16**	**.35**	**.27**	.00
2a (587)	Action	**-.17**	.05	**.11**	**.46**	.01	**-.11**	**.21**	**-.19**
2a (587)	Animation	**-.23**	.08	**.12**	**.26**	.01	-.02	**.11**	-.04
2a (587)	Thriller	-.05	**.11**	**.13**	**.16**	.06	**-.29**	**-.09**	-.04
2a (587)	Crime	-.06	.01	**.11**	**.15**	.09	**-.17**	.00	-.03
2a (587)	Adventure	-.03	.01	.05	**.27**	.04	.02	**.17**	.03
2a (587)	Science-Fiction/Fantasy	**-.33**	**.16**	**.16**	**.22**	-.04	**-.15**	.01	**-.13**
2a (587)	Drama	.05	.06	**.09**	**-.25**	**.14**	**-.11**	**-.32**	**.21**
2a (587)	Tragedy	.04	**.13**	**.11**	**-.19**	**.15**	**-.20**	**-.33**	**.18**
2a (587)	Comedy	.03	.03	.00	**.21**	.01	**.12**	**.46**	-.08
2a (587)	Documentary	.09	-.01	.04	**-.14**	-.09	**-.13**	**-.20**	**.17**
2a (587)	Avant-garde	-.05	**.21**	**.24**	**-.34**	.03	**-.14**	**-.43**	**.17**
2a (587)	Horror	**-.15**	**.12**	**.11**	**.12**	-.08	**-.51**	**-.13**	**-.16**
2a (587)	Erotic	-.02	**.18**	.08	.04	**.11**	**-.16**	-.07	-.03
	Film Competence								
2a (587)	Aesthetic Fluency in Film	**-.16**	**.26**	**.39**	**-.13**	**.14**	**-.29**	**-.49**	.04
2b (273)	Film Expertise	**-.20**	**.35**	**.49**	.09	**.18**	**-.40**	**-.35**	.08
	Film-specific Evaluation scales								
2b (273)	Appreciation	**.26**	**.14**	.02	.01	**.15**	**.21**	-.01	**.32**
2b (273)	Fun	-.02	**.19**	.02	**.16**	.07	.04	**.17**	-.01
2b (273)	Suspense	**.20**	.12	-.02	**.17**	.09	**.14**	**.15**	**.22**
	Big Five								
2a (587)	Extraversion	.06	**.11**	**.18**	.07	.05	-.02	.08	**.19**
2a (587)	Agreeableness	.04	**.13**	.06	**.11**	.01	**.11**	**.16**	**.10**
2a (587)	Conscientiousness	**.18**	-.10	-.04	**.13**	-.02	**.14**	**.15**	.01
2a (587)	Emotional Stability	-.06	.04	.00	-.01	-.12	**-.20**	-.05	-.04
2a (587)	Openness to Experiences	-.01	**.26**	**.23**	-.07	-.08	-.02	-.12	**.30**
	Sensation Seeking								
2a (587)	Experience Seeking	.09	**.27**	**.21**	-.09	-.02	.00	**-.19**	**.25**
2a (587)	Boredom Susceptibility	.02	**.27**	**.20**	.02	-.05	**-.14**	-.07	**.20**
2a (587)	Thrill and Adventure Seeking	-.08	.11	.08	**.18**	-.05	**-.39**	-.09	-.02
2a (587)	Disinhibition	-.08	.10	.08	-.03	-.01	**-.32**	-.08	.09
	Trendsetting								
2a (587)	Total scale	.03	**.19**	**.19**	**.19**	.02	**-.12**	-.01	**.12**
2a (587)	Input	-.08	**.24**	**.23**	.03	.06	**-.18**	**-.15**	**.26**
2a (587)	Throughput	.00	**.19**	**.16**	**.20**	.00	-.08	.03	.11
2a (587)	Output	.07	**.16**	**.18**	**.19**	.02	**-.13**	-.01	.09
2a (587)	Political Interest	**.11**	.01	.05	**-.19**	.04	**-.14**	**-.26**	**.16**
2a (587)	Need for Cognition	-.08	.04	**.16**	**-.18**	**-.14**	**-.14**	**-.18**	**.19**
2a (587)	Need to Evaluate	**.11**	**.12**	**.13**	.05	-.08	-.07	-.10	**.16**
	Need for Affect								
2a (587)	Approach	**.13**	**.14**	.07	-.08	.10	**.21**	.09	**.29**
2a (587)	Avoidance	-.09	.03	.08	-.08	-.07	.00	.00	.04
	Socially Desirable Responding								
2a (587)	Self-Deceptive Enhancement	**.15**	.10	.00	.08	-.05	-.01	.11	.03
2a (587)	Impression Management	**.13**	.02	.01	.09	-.12	**.18**	**.18**	.08

Note. Significant correlations (*p* < .05) in bold. All model fit indices indicated acceptable to good fit (see Table F1 in Appendix F for details).

Story Verisimilitude. The highest positive correlation for SV was found between SV and Appreciation (φ = .26), whereas the highest negative correlation was between SV and preferences for Science Fiction/Fantasy (φ = -.33). Although SV was also negatively correlated with Animation, Action, and Horror as expected, there was a low positive, but not statistically significant correlation with Documentary. In addition, relationships with Conscientiousness and Political Interest were positive, but only low. One reason for this might be that the story in a film is considered being mainly fictional and narrative. Although it seems to be important that such stories are based on historical facts, are realistic, and address contemporary issues, this does not mean that these stories need to be real. Persons who are interested in politics generally prefer real news and documentations of real life.[33] Another explanation lies in the different facets of SV—as already demonstrated by the low reliability estimates in Phase III (see Table 8) and the substantial method specificity coefficients in Phase IV (see Table 10): On the one hand, Item sv1 and sv2 reflect aspects like "based on true facts" and "realistic", which are correlated with Conscientiousness, but not with Political Interest; on the other hand, Item sv3 reflects contemporary issues and is correlated with Political Interest, but not with Conscientiousness.[34] Thus, future research on SV should develop and include additional items to explore the facets (or maybe even subdimensions) and their relation to each other in more detail. Finally, the low but statistically significant correlations with both SDR dimensions as well as with need to evaluate are rather marginal. Nevertheless, they warrant further investigation.

Story Innovation. The highest positive correlation for SI was found between SI and Film Expertise (φ = .35); no negative correlation was statistically significant.

---

[33] This is supported by a Pearson correlation between PI and preference for documentary of .24 (p > .05; N = 585) and the negative latent correlation between PI and FX of -.19.

[34] Pearson r = .13 for C and sv1, .11 for C and sv2, .14 for PI and sv3 (ps > .05; Ns = 585).

Correlations between SI and related constructs were as expected. All relationships with film-specific evaluation scales were positive—a finding that underscores the importance of innovative stories for movie evaluation. Interestingly, SI was moderately positively correlated with only two of the four sensation seeking subscales, ES and BS. With regard to the further results, for instance, the moderately negative correlation of IN with the two other sensation seeking subscales, TAS and DI, it seems safe to conclude that inspecting the subscales of sensation seeking offers valuable and more differentiated insights than just looking at the total score. Thus, the present analysis also supports the discriminant validity of the BSSS8.

Cinematography. The highest positive correlation for CI was found between CI and Film Expertise (φ = .49), whereas the highest negative correlation, between CI and preference for Romance, was only φ = -.10. Further correlations with genre preferences as well as with film competence, the Big Five, and sensation seeking scales clearly draw a distinction between CI and FX as expected. Surprisingly, CI did not correlate with any of the film-specific evaluation scales. One explanation might be that there were no movies with outstanding cinematography in the sample. For instance, CI showed only a low relationship with preference for Drama and no relationship with Comedy; however, over 30% of the respondents' last movies can be categorized as dramas or comedies. Hence, it would be a good idea to relate CI with the evaluation of music videos or short films because aesthetic and artistic concepts might be realized in a more sophisticated way in short films.

Special Effects. The highest positive correlation for FX was found between FX and preference for Action (φ = .46), whereas the highest negative correlation was between FX and preference for Avant-garde (φ = -.34). FX is related to Fun and Suspense but not to Appreciation, which emphasizes the distinction between FX and CI.

Remarkably, FX is low positively correlated with only one of the sensation seeking scales—the TAS. It did not load to a statistically significant level on the other subdimensions nor on Openness to Experiences. However, these findings should be interpreted cautiously because of the narrow scope of the BSSS8 and the TIPI (only two items per each dimension).

Recommendation. The highest positive correlation for RE, between RE and Film Expertise, was only $\varphi = .18$; the highest negative correlation, namely between RE and Need for Cognition, was only $\varphi = -.14$. The former was in line with my assumptions; the latter can be explained by considering recommendations as a kind of spoiler. For those people who prefer analyzing a movie, enjoy being challenged, or like to puzzle over a whodunit, prior knowledge about story, plot, and evaluation might be unimportant or even detrimental to their enjoyment of effortful thinking. Taken together, the results in Table 11 indicate that RE has rather little in common with the included constructs. Although there are positive correlations with preference for Drama, Appreciation, or film competence scales, these are all rather low. Surprisingly, RE did not correlate with trendsetting. On the one hand, this demonstrates that RE and trendsetting are independent constructs, especially the RE and the TDS-K Input scale. On the other hand, the question is where do trendsetters get their information (about movies) from, if not from peripheral sources available *before* the product is released? One explanation might be that the items in the RE scale refer more to current (advertising, reviews) or past (awards, film is a classic) than to new information. Hence, trendsetters might not care about these kinds of recommendations as information sources. Previous research has shown that aspects of recommendation (awards, reviews, ads) are crucial to our understanding of movie choice and success (see Chapter 2.1). Therefore, future work should elaborate on RE and explore the substantive and external validity of this scale.

Moreover, adding more items to investigate the uniqueness of the facets of the scale might be an effective way to cope with the substantial method specificity found in Phase IV.

Innocuousness. The highest positive correlation for IN was found between IN and preference for Romance (φ = .35), whereas the highest negative correlation was between IN and preference for Horror (φ = -.51). As expected, participants who tend to rate IN as important at all have less knowledge of and are less interested in film as art. Additionally, they score lower on Emotional Stability and sensation seeking. Contrary to expectations, IN correlated only low with preference for Comedy and not at all with Fun. Despite drawing a distinction between IN and LH, this result is difficult to explain. Perhaps this is so because comedies can also include scenes in which disgusting things happen, schadenfreude is evoked, or violence is celebrated (e.g., *Pulp Fiction*). This genre category might be too general.[35]

Light-heartedness. The highest positive correlation for LH was found between LH and preference for Comedy (φ = .46), whereas the highest negative correlation was between LH and Aesthetic Fluency in Film (φ = -.49). Nearly all correlations are in line with previous assumptions and reflect the semantic closeness to U&G dimensions like Entertainment or Escapism as well as to the Fun scale. At the same time, the magnitudes are rather low or moderate. Thus, the results point in the right direction but clearly draw a distinction between the constructs. A statistically significant positive correlation between Impression Management as a SDR subscale and LH hints to the fact that light entertainment is socially approved as reflected in film business and box office figures.

---

[35] Pearson correlations (all Ns = 585, all ps < .05) show that even though Comedy is moderately positively correlated with Romance (r = .32), it is also positively associated with Action (.19), Adventure (.19), and Animation (.16), but not negatively correlated with Horror (r = 0, p > .05). This indicates that comedy as a basic genre (cf. Gehrau, 2009) has blurry boundaries and a vague meaning.

Cognitive Stimulation. The highest positive correlation for CS was found between CS and Appreciation (φ = .32), whereas the highest negative correlation was between CS and preference for Action (φ = -.19). CS is associated with constructs that reflect cognitive challenges as well as broader aspects of stimulation (e.g., Openness to Experience, sensation seeking), albeit these correlations are rather moderate. Surprisingly, the correlation between CS and the Approach scale of need for affect was the highest NFA correlation. Thus, CS might also reflect some kind of emotional challenge (e.g., experiencing intense and ambivalent emotions; cf. Bartsch et al., 2010). This is in line with the Appreciation scale, which also includes an affective item ("I was moved by this movie."; Pearson r = .25, N = 273, p < .05) that is correlated with CS in the same way as a rather cognitive item ("The movie was thought provoking."; Pearson r = .28, N = 273, p < .05).

As mentioned above, latent correlations of all SMEC scales with SDR and NTE were assumed to be low, thus indicating discriminant validity. Self-deceptive enhancement showed only one statistically significant latent correlation (i.e., with SV). Impression management was statistically significantly positively correlated with SV, IN, and LH. The magnitudes of the latent correlations are small, and thus, the SMEC and SDR are distinct constructs.

Taken together, nearly all examined relationships between related constructs and SMEC showed only small to moderate magnitudes. Hence, most of the constructs share only little proportions of variance with SMEC. This is especially true for RE. The low latent correlations indicate discriminant validity. Furthermore, they help to draw distinctions between the SMEC themselves (e.g., SV and CS, CI and FX, and IN and LH). However, these findings must be interpreted with caution because of at least two

reasons: (a) the level of correspondence (cf. Ajzen & Fishbein, 1977)[36] and (b) the unresolved question how occasion and method specificity might contribute to the analyses. In their seminal article on attitude–behavior relations, Ajzen and Fishbein (1977) pointed out that high correlations can be obtained only by establishing highly corresponding units of attitudinal and behavioral measurement (for a similar remark regarding the correspondence of sought and obtained gratifications at different levels of media abstraction see Palmgreen, 1984, pp. 34–35). For instance, it is quite evident that measuring the evaluative importance of thought-provoking movies has a very different level of specificity than such a broad personality trait like Openness to Experiences. Therefore, in future studies a higher level of correspondence (e.g., between CI and the Openness-facet Aesthetics) could be achieved, for instance, by measuring personality traits with the 30 facet scales of the NEO-PI-R (Costa & McCrae, 1992) that assesses the Big-Five.

In Phase IV, partitioning true score variance into consistency, occasion specificity, method specificity, and error variance showed that between 9% and 24% of variance can be explained by situational or interactional effects. Thus, it would be interesting to see how the estimation of the correlations between latent traits is altered by conducting longitudinal research and simultaneously submitting SMEC and related constructs into multitrait–multioccasion analyses (cf. Dumenci & Windle, 1998). Of course, additionally investigating method effects by using a multitrait–multimethod-multioccasion design will also yield valuable insights (Courvoisier, Nussbeck, Eid, Geiser, & Cole, 2008).

---

[36] In the context of personnel selection, this is also known as the bandwidth–fidelity tradeoff (cf. Cronbach & Gleser, 1965).

10.   General Discussion

The two major goals of this dissertation were to theorize about the conceptualization of subjective movie evaluation criteria and to construct and validate appropriate scales for their measurement. First, I will briefly summarize the scale development and discuss the facets of construct validity. This also includes thoughts about limitations and future research directions. Second, I will turn to a broader discussion referring to the conceptualization of SMEC and its contribution to psychological science.

10.1   Facets of Construct Validity

Content Validity. In the present research, the content aspect of construct validity refers to the content-relevance and representativeness of criteria in the construct domain of movie evaluation. Because the focus was on subjective movie evaluation criteria, the process of selecting appropriate terms was not based on a deductive, but on an inductive, data-driven approach. A clearly qualitative focus was adopted due to the limited theoretical and empirical knowledge available on SMEC. Open-ended questions, structure formation technique, focus groups, content analysis, think-alouds, interviews—all these methods were applied to deal with the tasks of item wording, pilot-testing items, and establishing some kind of interrater agreement. Although the application of the modified structure formation technique did not lead to the clear-cut criteria categories I had hoped for, the whole process helped to delve deeper into the subjective ideas of how movie evaluations manifest themselves and what kind of criteria seem to play an important role. It is important to note that the

[119]

content validity of the final SMEC scales is somewhat constrained. On the one hand, due to a lack of theoretical knowledge, items were excluded mainly because of statistical properties. Now that the scales exist, it would be possible to explore the constructs they measure in more detail (e.g., the role of identification with characters as discussed in Chapter 6.3). On the other hand, an economical and short instrument can help avoid exhausting the participants (Burisch, 1984). Surely, this is a compromise. With regard to huge item pools such as the IPIP (Goldberg et al., 2006), we should consider the fact that the items might function differently depending on the level of specificity and on the aim of the measurement. Thus, establishing a comprehensive item pool would help to specify the boundaries of the construct domain (with some items located more at the core and some items located at the boundaries) and to further explore the underlying structure.

Structural Validity. The results of EFA, E/CFA, and CFA clearly indicated strong support for the structure of the SMEC discovered. The rather low correlations between the eight SMEC does not suggest a hierarchical structure on a second level relevant to all factors. However, some dimensions (e.g., IN and LH, SV and CS, CI and FX) showed moderate correlations and also fit together theoretically. Are they facets or subdimensions of a higher-ordered factor? Supplementary analyses are warranted, albeit criterion-correlations revealed clear distinctions between the scales.

External Validity. Examining the nomological net of SMEC and related constructs revealed rather low correlations, although some moderate to high correlations were as expected (e.g., CI and Film Expertise, LH and preference for Comedy, or FX and preference for Action). Furthermore, these relationships offered valuable insights into the distinctiveness of the SMEC scales, despite similarities. For instance, CI and FX correlated quite differently with film competence scales.

Future research could investigate the external relationships between SMEC and constructs not considered in this initial attempt. Especially the need for entertainment (Brock & Livingston, 2004), eudaimonic and hedonic motivations (Oliver & Raney, 2011), and the modes of reception seem to be relevant candidates. For instance, with regard to the modes of reception, the SMEC scales were developed in a manner comparable to the MoRI (e.g., qualitative pilot studies, exploratory and confirmatory factor analyses; Suckfüll & Scharkow, 2009). However, whereas the MoRI construction focused on what is happening while watching a movie, the SMEC scale development concentrated on the importance of criteria for evaluating a movie in general. Hence, the goals and questions asked were entirely different, and it comes as no surprise that both instruments ended up with different dimensions. Nevertheless, the SMEC might also be fruitfully applied to explain reception processes, especially given that evaluative processes are assumed to be essential for further reception processes or effects. In this case, a closer look at the relationship between SMEC and modes of reception in the nomological network in the future might foster our understanding about criteria as mental representations and modes of reception as operating mental processes. This also refers to another important aspect of construct validity—substantive validity.

Substantive Validity. Substantive validity refers to the theoretical basis of the construct. Latent state–trait analyses revealed that the SMEC are more stable than transient constructs, thus supporting theoretical assumptions about the nature of SMEC. Another aspect that was only briefly mentioned in Chapters 2.2 and 3.1 is the question about the formation and development of SMEC. The developmental psychological perspective is often overlooked in media psychological research. We can only speculate how SMEC are developed during the life-span. We know that cinema and movies are important issues for 6- to 13-year-old children (Medienpädagogischer

Forschungsverbund Südwest, 2010). Film viewing usually begins at this age and is accompanied by the development of evaluative criteria. Of course, there might be some exceptions where viewing movies starts later in life (cf. Hobbs, Frost, Davis, & Stauffer, 1988), but usually—at least in modern civilizations—it is reasonable to assume that this happens in early childhood. As soon as a child sees her or his first movie, subjective movie evaluation criteria start to evolve through basic principles of learning (e.g., classical and operant conditioning, imitation, etc.). Consequently, the complexity of SMEC depends on the developmental level. For instance,

- a limited cognitive capacity constrains the processing and understanding of complex information (Piaget, 1937; Sturm, 1984), thus children start to prefer fast-paced and unfamiliar content at the age of five or six (cf. Valkenburg & Cantor, 2000);

- acquiring knowledge about film characteristics and production promotes fantasy–reality distinction, thereby probably changing the effects of and preferences for cartoons and special effects (cf. Valkenburg & Cantor, 2000);

- younger children focus more attention on and are more influenced by the visual appearance of information (cf. Acuff & Reiher, 1997; Hoffner & Cantor, 1985), thereby liking something beautiful, but maybe dangerous, more than something ugly, but harmless (cf. Valkenburg & Cantor, 2000);

- socio-emotional abilities—for example, understanding others' emotions and role-taking—improve throughout childhood, thus older children think more about the credibility of actors' performances as well as the identification with specific characters (cf. Valkenburg & Cantor, 2000);

- peer-group pressure is assumed to have a strong impact on preferences, especially among school children (Valkenburg & Cantor, 2000);

- finally, gender differences in preferences emerge by the time children are two or three—for example, girls get more interested in real-life drama and innocuous and easily comprehensible programs, boys in action-oriented films (Valkenburg & Cantor, 2000; Valkenburg & Janssen, 1999)—and may continue over the life-span (cf. Hoffmann & Schwender, 2007).

Evidently, it seems plausible that during maturation preferences become stronger and criteria crystallize. However, the life-stages hypothesis (cf. Visser & Krosnick, 1998) proposes a curvilinear relationship between age and susceptibility to attitude change. On the one hand, with regard to SMEC, this suggests that criteria might be less influenced by situational effects in middle adulthood than in childhood, adolescence, or old age. On the other hand, movie viewing frequency is highest among adolescents, young adults, and early middle-aged adults (Bundesverband Audiovisuelle Medien e. V., 2011; FFA Filmförderungsanstalt, 2011; Gerhards & Klingler, 2011; MPAA, 2011; Prommer, 1999). Does the emergence of criteria precede the increased exposure to (preferred) movies or does the increased viewing behavior result in more stable criteria? Presumably, there is a complex, dynamic, and mutual interplay between choosing movies, viewing them, and consolidating SMEC.

Taken together, research provides some evidence that media preferences crystallize during different developmental levels. During socialization, we do not only develop likes and dislikes for film genres. Rather, it seems plausible that we even develop fairly stable likes and dislikes (evaluations) for specific film features. Moreover, in the course of time we gain knowledge about movies, film production, and cinematic experiences.

In the present research, I assumed that processing film features for the first time leads to the formation of attitudes towards these film features. Does this happen fast or

slowly, does it follow a linear growth curve dependent on the frequency of viewed movies, does it finally culminate in a fixed set of criteria? It is yet too early to answer these questions. Furthermore, film features are not independent of each other; specific features frequently appear together with certain other features because the movies are produced in correspondence with a specific film genre category (e.g., a romantic comedy combines a happy ending, funny scenes, less innovative story, no special effects, etc.). Therefore, the development of attitudes towards specific film features is accompanied by the development of mental representations of genre categories and genre preferences. Moreover, distinguishing between SMEC and genre preferences is a creative endeavor. Applying psychological theories that deal with cognitive as well as socio-emotional development might help to disentangle both constructs and clarify their relationships. Finally, answering questions about how SMEC develop would contribute considerably to the substantive validity of the construct.

Generalizability. To which extent do score properties and interpretations generalize to and across groups, occasions, and tasks (cf. Messick, 1995)? This also includes thinking about measurement errors. The factorial structure of the SMEC was tested and replicated in three different samples. Latent state–trait analyses demonstrated high reliabilities across two measurement occasions, which were mainly determined by stable individual differences (indicated by high common consistency coefficients). Future studies should aim to increase the number of measurement occasions and attend to the aspect of measurement invariance across different groups. A further aspect that is related to generalizability is sampling.

Sampling. At first glance, there seem to be severe limitations; therefore, sampling is discussed in more detail. Regarding the samples utilized in this paper, three points have to be considered. First, due to economic constraints, scale development was

totally based on nonprobability sampling, which might preclude from generalizing the findings to (i.e., drawing statistical inferences from), for instance, a representative population of German movie viewers (but see Blanton & Jaccard, 2008, Farber, 1952, Mook, 1983, for the difference between representing and generalizing). Second, the samples largely consisted of higher-educated participants (e.g., mainly students) leading to a sample bias and presumably to systematic bias (cf. Peterson, 2001). Third, although the phases of item generation, content validation, and structural validation included the administration of paper-and-paper questionnaires, most studies presented in this paper relied heavily on online-samples. This adds to the sample bias, but also urges to think about potential effects of the modes of administration. Summing up these potentially serious drawbacks, one might assume that this reduces the validity of the present findings. However, referring to the ever recurring debate on the usefulness of convenience—and especially student—samples (e.g., Courtright, 1996; McNemar, 1946; Sears, 1986), research clearly indicates that although this kind of sampling limits generalizability, it has proven to be useful, otherwise it would no longer be common practice today (cf. Henry, 2008). So, why is it useful? First,

> any research that tests the relationship between factors allows for the falsification of theories if the expected relationships do not hold. As a result (regardless of whether the data are gathered from a fully representative sample of the population or from subsamples), if the relationship only holds for a portion of the sample, or is contingent upon some other condition occurring, then examining the multivariate relationship provides a test of the theory across these circumstances. (Basil, Brown, & Bocarnea, 2002, p. 505)

This does not limit the present findings, but certainly calls for further studies, which might be conducted under other circumstances (e.g., in a field study at a cinema theater) and include different samples with different socio-demographic or ethnic

backgrounds. Moreover, cross-cultural validations are warranted because movie preferences (e.g., Weaver et al., 1993) as well as genres as predictors for movie success (e.g., Hennig-Thurau & Wruck, 2000) vary across countries (i.e., Germany and USA).

Second, students are easily available, inexpensive, and can be of great usefulness in advancing science because (a) they possess interesting or relevant characteristics and "provide an efficient means to develop theories before testing their generalizability in broader samples" (Dasgupta & Hunsinger, 2008, p. 94). Henry (2008) argues that research using student samples is not wrong "simply because it uses student samples. However, this research . . . may be describing phenomena or mechanisms as they exist only in a unique context" (p. 61). This context, however, is highly relevant in examining SMEC and film reception. In 2010, the highest frequency of movie attendances was found among adolescents and students (i.e., approx. 27% of the German population of moviegoers, but only 12% of the German total population, FFA Filmförderungsanstalt, 2011; similar relevance of corresponding video-renter age groups corroborate these findings, Bundesverband Audiovisuelle Medien e. V., 2011). Furthermore, more than two-thirds of the moviegoers received post-secondary or higher education (FDW Werbung im Kino e.V., 2010). Finally, more women went to the movies (55%, FFA Filmförderungsanstalt, 2011). Surely, we cannot draw conclusions about all film viewers from moviegoer statistics, but the samples in the present research appear to go in the right direction.

Third, nowadays online-based surveys have become a powerful data collection tool for the social sciences. Conducting research with online-questionnaires produces results highly comparable to those of traditional paper-and-pencil assessments (Chuah, Drasgow, & Roberts, 2006; Gosling, Vazire, Srivastava, & John, 2004). Thus, whereas the

sampling discussed above is more crucial to generalizability, the mode of administration does not appear to cause any problems.

Finally, a further extension of generalizability should be achieved by applying multigroup analyses to address questions about measurement invariance in future (e.g., Do US-American and German participants, female and male respondents, or film novices and experts interpret the SMEC scales in a similar manner? cf. Vandenberg & Lance, 2000).

Consequential Validity. Consequential validity was not an issue in the present research (see Footnote 16). However, with regard to applied research, test bias and fairness become important aspects. What kinds of applications are possible? Surely, the film business could apply the scales to tailor movies to their target groups or obtain more relevant information from test screening audiences. Another area of application was already mentioned in Chapter 9—film competence. Serious attempts have been made to establish film education as a fixed part of the curriculum (e.g., Arbeitskreis Filmbildung, 2009; British Film Institute, 1999). Even though the SMEC was not directly developed for an educational context, evaluating film competence programs is within the scope of the SMEC scales. When used in the film business or in an educational context, one must be aware of the consequences of measuring subjective criteria with the SMEC scales. For example, one must make sure that there are no negative effects due to invalid tests.

10.2   Contributions, Limitations, and Future Directions

Media Psychological and Entertainment Research. The present approach broadens the scope of current media psychological research on movies to explore the building blocks of selection, reception, and effects—attitudes and evaluative processes.

In the theoretical part of the paper, I described the relevance of evaluative processes and evaluative criteria. Both allow for integrating basic principles of social cognition with approaches to explain, for instance, enjoyment and appreciation of entertaining and meaningful cinematic pieces. The present paper theorized that attitudes towards specific film features (i.e., SMEC) might function as a missing link between global values, motives, and personality traits, on the one hand, and the specific content of movies, on the other hand. In short, attitudes as mental representations might be more closely related to the recipient's information processing than other constructs (e.g., genre preferences). This, however, is a question of substantive validity (see above) and an important future task for establishing the SMEC scales. A possible first step in the right direction might be to examine the relationship between hedonic motivations, LH, and Fun, on the one hand, and eudaimonic motivations, CS, and Appreciation, on the other hand (Bartsch & Oliver, 2011; Oliver & Bartsch, 2010; Oliver & Hartmann, 2010). However, because Fun and Appreciation are two relatively homogeneous constructs, it is important to develop film-specific evaluation items that can be directly linked to the SMEC items. First attempts to develop and administer such items have been made recently (Bacherle et al., 2011). The difficulty lies in newly developing three kinds of measures at the same time: (a) for assessing how important specific film features are (e.g., the importance of FX, as has been done in the present research), (b) for assessing the extent or amount of the relevant features (e.g., *no movie illusions at all* to *very many movie illusions*), and (c) for assessing the evaluation of the respective extent or amount (e.g., bipolar: *extremely positive* to *extremely negative*; unipolar: e.g., *not at all positive* to *extremely positive* and *not at all negative* to *extremely negative*; evaluative space grid as a single-item measure for positivity and negativity, cf. Larsen, Norris, McGraw, Hawkley, & Cacioppo, 2009). The first kind (a) focuses on stable individual differences, the second

(b) on the movie-specific content. Similarly, Wolling (2009) combined desired/undesired features and perceived features. The major differences lie in the conceptualization of attitudes and desires as well as in the theoretical background. Whereas Wolling focuses on expectation x value theory (i.e., combining [a] and [b]), I prefer to connect the idea of movie evaluation criteria as attitudes to current advances in social cognition, namely conceptualizing attitudes as mental representations and describing and explaining evaluative responses by means of the APE model. Thus, it would be interesting to see whether (a) and (b) might explain (c), or whether there are further factors relevant for a feature-specific evaluation. Additionally, this research would help unravel the processes underlying global movie evaluations.

Attitude Theory and Measurement. I conceptualized SMEC as mental representations of film features or attitudes towards specific movie characteristics. First, thinking about movies as attitude objects might advance measurement methods, for instance, when conventional ways of assessing implicit attitudes have to be adapted to deal with complex, dynamic film stimuli. Or, broadly speaking: "attitude theory and measurement are completely intertwined, and advances in one contribute to (and benefit from) advances in the other" (Ostrom, 1989, p. 11). Certainly, it is far too early to speak of a movie evaluation *theory*. However, this rather broad conceptualization of SMEC offers the opportunity to incorporate them into a plethora of partly competing attitude theories (e.g., APE model, Gawronski & Bodenhausen, 2006a; iterative reprocessing model, Cunningham, Zelazo, Packer, & van Bavel, 2007; connectionist network models, Conrey & Smith, 2007; meta-cognitive model, Petty, Briñol, & DeMarree, 2007; etc.) and to put these models to the test, thereby fostering theoretical progress and suggesting to investigate film features as a new kind of unexplored attitude objects.

Self-Reports. "If we want to know how people feel: what they experience and what they remember, what their emotions and motives are like, and the reasons for acting as they do—why not ask them?" (Allport, 1942, p. 37). Allport's quote is still relevant today and summarizes the aims of perhaps the most frequently applied method in the social sciences—self-reports. However, we also have to be aware of the limitations of self-reports. Although the SMEC scales showed only low correlations with socially desirable responding, there are numerous other factors that might lead to biased self-reporting (e.g., Gschwendner, Hofmann, & Schmitt, 2006; Hofmann, Gschwendner, Nosek, & Schmitt, 2005). For instance, asking a respondent about her or his SMEC might confront her or him with the fact that she or he has never thought about any criteria before. Thus, the items might trigger evaluative responses and the construction of attitudes on-the-spot—some researchers even question whether these phenomena can be called attitudes at all or whether they might only be temporary constructions due to (stable) contextual cues (cf. Schwarz, 2007; Schwarz & Bohner, 2001; Wilson & Hodges, 1992). Indirect measures that neither force the respondents to deliberatively process information nor require full awareness of the processes could be applied or developed to investigate in a multitrait-multimethod analysis how different methods contribute to the external validity of the SMEC scales. On the one hand, this is more deeply connected to social cognition research, provides a real challenge for movie evaluation research, and might even foster our understanding of inconsistencies between direct self-reports and indirect observations (e.g., observational or physiological measures) concerning the measurement of emotions during movie reception. On the other hand, "indeed, people have been known to die in support of their explicit attitudes, but to date there is no evidence that people are willing to passionately defend their unendorsed automatic associations" (Petty & Briñol, 2006,

p. 741). As Petty and Briñol (2006, pp. 741, Footnote 2) have added, implicit attitudes are not unimportant. However, the point here is that I am also interested in the effect of evaluations on movie selection and recommendations. Hence, it is more likely that people who express their attitudes towards a movie (characteristic) will also argue for seeing or recommending a movie. If they "only" hold an implicit attitude, this might not be the case—maybe such a "semi-active" representation might thus be not important? This assumption is corroborated by meta-analytical findings that showed that attitudes were better predictors of future behaviors when attitudes were (a) easily accessible, (b) stable over time, (c) formed by direct experiences with the attitude object, and (d) frequently reported (Glasman & Albarracín, 2006). This is also in line with results that demonstrated that attitude strength and attitude importance moderated implicit–explicit consistency; this means that for people with highly accessible, strong, or important attitudes, the correlations between implicit and explicit measures are higher (e.g., Hofmann, Gschwendner, & Schmitt, 2005; Nosek, 2005). However, the movie choice of an undecided viewer, who probably does not yet hold any attitude toward specific film features, might be more determined by implicit associations (cf. Galdi, Arcuri, & Gawronski, 2008). Thus, future investigations of implicit movie evaluation criteria would definitely add to the self-report data and is a worthwhile research goal.

Personality and Individual Differences. Finally, film features as attitude objects clearly reflect everyday life experiences. As Rentfrow and Gosling (2003) stated, "there is a growing concern that the breadth of topics studied by many research psychologists is too narrow and excludes many important facets of everyday life that are worthy of scientific attention" (p. 1236).[37] They argued that music is one such facet.

---

[37] Or as Neisser (1976) stated more candidly: "A psychology that cannot interpret ordinary experience is ignoring almost the whole range of its natural subject matter" (p. 4).

Film is another. In the present research, for instance, it might not come as a surprise that Openness to Experience is related to SI, CI, and CS. However, prior research on movie preferences did not have the opportunities to go beyond the broad and fuzzy categories of film genre which are often open to subjective speculation and interpretation. To my best knowledge, there is no research that examined relationships between Openness to Experience and movie content; moreover, the TV research here is misleading, because people high in Openness to Experience do not watch TV, but read books and go to the movies (cf. Finn, 1997). Now, this example has shown that the SMEC offer a more fine-grained analysis to investigate how personality traits and attitudes towards movie characteristics are related, thereby contributing to research on personality and individual differences.

10.3   Conclusion

In the present research, subjective movie evaluation criteria were defined as standards that viewers use for assessing the features of films and conceptualized as mental representations of or attitudes towards specific movie features guiding cognitive and affective information processing of movies and corresponding evaluative responses. This broad conceptualization of SMEC offers the opportunity to incorporate them into elaborated models within the social cognition framework (e.g., APE model).

In five phases comprising five studies (Pilot Study, Study 1, Study 2a, Study 2b, Study 3), the SMEC scales were constructed and validated. The eight SMEC scales—Story Verisimilitude, Story Innovation, Cinematography, Special Effects, Recommendation, Innocuousness, Light-heartedness, and Cognitive Stimulation—consist of 28 items (3–4 for each scale) and are ready for application.

Epilogue

Mark Snyder spent over 20 years modifying, refining, and measuring the self-monitoring construct (Gangestad & Snyder, 2000; Snyder, 1987), and Jane Loevinger devoted over 30 years of her life to the development and construct validation of the Washington University Sentence Completion Test for measuring ego development (Loevinger, 1998). Neither would I dare to compare my work to theirs, nor would I assume that the SMEC scale will have such an impact on psychology—but hopefully in the future, a number of researchers—me included—will elaborate on the SMEC construct and scales, improve them, and see them prevail.

Let me end with some perfect words from John and Benet-Martínez (2000) on what they call the *modern-oriented scale construction* that I am committed to:

> Questionnaire construction, like measurement more generally, involves theory-building and thus requires an iterative process. It begins with (a) generating hypotheses; (b) building a model and plausible alternatives; (c) generating items using construct definitions, generalizability facets, and content validation procedures as guides; . . . (d) gathering and analyzing data; (e) confirming and disconfirming the initial models; and (f) generating alternative hypotheses leading to improved models, additional and more content-valid items, more data gathering, and so on. The cycle continues, until a working model has been established that is "good enough" – one that the investigator can live with, at least for a while, given the constraints and limits of real-life research. (p. 363)

References (or End Credits)

Acuff, D. S., & Reiher, R. H. (1997). *What kids buy and why: The psychology of marketing to kids*. New York, NY: Free Press.

Aelker, L., & Bente, G. (2011). Emotional contagion while watching movies together: The impact of physical and mediated emotional cues on shared emotions. In Ö. Odağ, M. Schreier, & Y. Thies-Brandner (Eds.), *Media psychology. Focus theme: Cognitive and emotional involvement during media reception. Proceedings of the 7th conference of the Media Psychological Division of the German Psychological Society* (pp. 14–15). Lengerich: Pabst.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*(5), 888–918.

Albert, S. (1998). Movie stars and the distribution of financially successful films in the motion picture industry. *Journal of Cultural Economics*, *22*(4), 249–270.

Allport, G. W. (1935). Attitudes. In C. A. Murchison (Ed.), *A handbook of social psychology* (pp. 798–844). Worcester, MA: Clark University Press.

Allport, G. W. (1942). *The use of personal documents in psychological science*. New York, NY: Social Science Research Council.

Aluja-Fabregat, A. (2000). Personality and curiosity about TV and films violence in adolescents. *Personality and Individual Differences*, *29*(2), 379–392.

Aluja-Fabregat, A., & Torrubia-Beltri, R. (1998). Viewing of mass media violence, perception of violence, personality and academic achievement. *Personality and Individual Differences*, *25*(5), 973–989.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, *6*(4), 451–474.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory*. Washington, DC: Winston & Sons.

Appel, M. (2008). Manche mögen's heiß: Ergebnisse der deutschsprachigen Version eines Instruments zur Erfassung des Emotionsmotivs (need for emotion/need for affect) [Some like it hot. Results of the German language version of an instrument aimed at the assessment of the need for affect/need for emotion]. *Diagnostica, 54*(1), 2–15.

Appel, M., Gnambs, T., & Maio, G. R. (in press). A short measure of the need for affect. *Journal of Personality Assessment.*

Arbeitskreis Filmbildung (2009). *Filmbildung: Kompetenzorientiertes Konzept für die Schule* [Film education: Competency-oriented curriculum for schools]. Retrieved from http://www.visionkino.de/WebObjects/VisionKino.woa/media/3355

Arnett, J. (1994). Sensation seeking: A new conceptualization and a new scale. *Personality and Individual Differences*, *16*(2), 289–296.

Arnheim, R. (1932). *Film als Kunst* [Film as art]. Berlin: Ernst Rowohlt.

Austin, B. A. (1981). Film attendance: Why college students chose to see their most recent film? *Journal of Popular Film and Television*, *9*(1), 44–49.

Austin, B. A. (1983). Critics' and consumers' evaluations of motion pictures: A longitudinal test of the taste culture and elitist hypotheses. *Journal of Popular Film and Television*, *10*(4), 156–167.

Austin, B. A. (1986). Motivations for movie attendance. *Communication Quarterly*, *34*(2), 115–126.

Austin, B. A., & Gordon, T. F. (1987). Movie genres: Toward a conceptualized model and standardized definitions. In B. A. Austin (Ed.), *Current research in film. Audiences, economics and law* (Vol. 3, pp. 12–33). Norwood, NJ: Ablex.

Bacherle, P., Schneider, F. M., & Retzbach, A. (2011). Distracted by dynamic measures? Real-time response measurement or camera observation do not impair affective and cognitive information processing. In Ö. Odağ, M. Schreier, & Y. Thies-Brandner (Eds.), *Media psychology. Focus theme: Cognitive and emotional involvement during media reception. Proceedings of the 7th conference of the Media Psychological Division of the German Psychological Society* (p. 54). Lengerich: Pabst.

Bagdasarov, Z., Greene, K., Banerjee, S. C., Krcmar, M., Yanovitzky, I., & Ruginyte, D. (2010). I am what I watch: Voyeurism, sensation seeking, and television viewing patterns. *Journal of Broadcasting & Electronic Media*, *54*(2), 299–315.

Bagella, M., & Becchetti, L. (1999). The determinants of motion picture box office performance: Evidence from movies produced in Italy. *Journal of Cultural Economics*, *23*(4), 237–256.

Banerjee, S. C., Greene, K., Krcmar, M., Bagdasarov, Z., & Ruginyte, D. (2008). The role of gender and sensation seeking in film choice: Exploring mood and arousal. *Journal of Media Psychology: Theories, Methods, and Applications*, *20*(3), 97–105.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.

Bartsch, A., Appel, M., & Storch, D. (2010). Predicting emotions and meta-emotions at the movies: The role of the need for affect in audiences' experience of horror and drama. *Communication Research*, *37*(2), 167–190.

Bartsch, A., Mangold, R., Viehoff, R., & Vorderer, P. (2006). Emotional gratifications during media use – An integrative approach. *Communications*, *31*(3), 261–278.

Bartsch, A., & Oliver, M. B. (2010, June). *Media enjoyment and appreciation. The role of perceived cognitive and emotional challenge in entertainment experience*. Paper presented at the Annual Conference of the International Communication Association, Singapore.

Bartsch, A., & Oliver, M. B. (2011). Making sense of entertainment: On the interplay of emotion and cognition in entertainment experience. *Journal of Media Psychology: Theories, Methods, and Applications*, *23*(1), 12–17.

Basil, M. D., Brown, W. J., & Bocarnea, M. C. (2002). Differences in univariate values versus multivariate relationships: Findings from a study of Diana, Princess of Wales. *Human Communication Research*, *28*(4), 501–514.

Bassili, J. N. (1996). Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength. *Journal of Personality and Social Psychology*, *71*(4), 637–653.

Bassili, J. N. (2008). Attitude strength. In W. D. Crano & R. Prislin (Eds.), *Attitudes and attitude change* (pp. 237–260). New York, NY: Psychology Press.

Bassili, J. N., & Brown, R. D. (2005). Implicit and explicit attitudes: Research, challenges, and theory. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 543–574). Mahwah, NJ: Erlbaum.

Basuroy, S., Chatterjee, S., & Ravid, S. A. (2003). How critical are critical reviews? The box office effects of film critics, star power, and budgets. *Journal of Marketing*, *67*(4), 103–117.

Basuroy, S., Desai, K. K., & Talukdar, D. (2006). An empirical investigation of signaling in the motion picture industry. *Journal of Marketing Research*, *43*(2), 287–295.

Batinic, B. (2005, September). *Ein empirischer Beitrag zur Diskrepanz zwischen Kinofilmpräferenz und der Kinofilmwahl* [An empirical contribution to the mismatch between movie preference and movie choice]. Vortrag auf der 4. Tagung der Fachgruppe Medienpsychologie der Deutschen Gesellschaft für Psychologie (DGPs), Erfurt.

Batinic, B., Haupt, C. M., & Wieselhuber, J. (2006). Validierung und Normierung des Fragebogens zur Erfassung von Trendsetting (TDS) [Validity and norms for a questionnaire to assess trendsetting (TDS)]. *Diagnostica*, *52*(2), 60–72.

Batinic, B., Wolff, H.-G., & Haupt, C. M. (2008). Construction and factorial structure of a short version of the Trendsetting Questionnaire (TDS-K): A cross-validation using multigroup confirmatory factor analyses. *European Journal of Psychological Assessment*, *24*(2), 88–94.

Baum, H. (2003). *Entscheidungsparameter bei der Filmauswahl von Kinogängern in Deutschland. Kommunikationsstrukturen und Meinungsführerschaft bei Kinofilmrezipienten. Eine Pilotstudie im Kommunikationsraum Berlin* [Decision parameters of German moviegoers for the selection of theatrical films. Communication structures and opinion leadership among movie recipients] (Doctoral dissertation). Retrieved from urn:nbn:de:kobv:188-2003002585

Beauducel, A., & Wittmann, W. W. (2005). Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. *Structural Equation Modeling*, *12*(1), 41–75.

Bem, D. J., & Allen, A. (1974). On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review*, *81*(6), 506–520.

Benesch, A. (2004). *Kinofilmrezeption. Nutzungsmotive und Entscheidungskriterien bei der Filmauswahl* [Film reception at the cinema. Gratifications sought and decisive criteria for movie choice] (Unpublished master thesis). Universität Koblenz-Landau, Landau.

Bentler, P. M., & Wu, E. J. C. (2005). EQS 6 structural equations program manual [Manual]. Encino, CA: Multivariate Software.

Bertrams, A., & Dickhäuser, O. (2009). High-school students' need for cognition, self-control capacity, and school achievement: Testing a mediation hypothesis. *Learning and Individual Differences*, *19*(1), 135–138.

Bertrams, A., & Dickhäuser, O. (2010). University and school students' motivation for effortful thinking: Factor structure, reliability, and validity of the German Need for Cognition scale. *European Journal of Psychological Assessment*, *26*(4), 263–268.

Bizer, G. Y., Krosnick, J. A., Holbrook, A. L., Wheeler, S. C., Rucker, D. D., & Petty, R. E. (2004). The impact of personality on cognitive, behavioral, and affective political processes: The effects of need to evaluate. *Journal of Personality*, *72*(5), 996–1028.

Blanton, H., & Jaccard, J. (2008). Representing versus generalizing: Two approaches to external validity and their implications for the study of prejudice. *Psychological Inquiry*, *19*(2), 99–105.

Bless, H., Wänke, M., Bohner, G., Fellhauer, R. F., & Schwarz, N. (1994). Need for Cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben [Need for cognition: A scale measuring engagement and happiness in cognitive tasks]. *Zeitschrift für Sozialpsychologie*, *25*(2), 147–154.

Boatwright, P., Basuroy, S., & Kamakura, W. (2007). Reviewing the reviewers: The impact of individual film critics on box office performance. *Quantitative Marketing and Economics*, *5*(4), 401–425.

Böcking, T., & Fahr, A. (2009). Television program avoidance and personality. *Communications*, *34*(3), 323–344.

Bodenhausen, G. V., Macrae, C. N., & Hugenberg, K. (2003). Social cognition. In I. B. Weiner (Series Ed.) & T. Millon & M. J. Lerner (Vol. Eds.), *Handbook of psychology: Vol. 5. Personality and social psychology* (pp. 257–282). Hoboken, NJ: Wiley.

Bohner, G., & Dickel, N. (2011). Attitudes and attitude change. *Annual Review of Psychology*, *62*, 391–417.

Bolik, S. (1999). Qualitätsargumente. Zwischenergebnisse empirischer Medienwertungsforschung [Quality arguments: Interim results of empirical media evaluation research]. In P. Ludes & H. Schanze (Eds.), *Medienwissenschaften und Medienwertung* [Media sciences and media evaluation] (pp. 99–116). Opladen: Westdeutscher Verlag.

Bolik, S., & Schanze, H. (Eds.). (2001). *Medienwertung* [Media evaluation]. München: Wilhelm Fink.

Boninger, D. S., Krosnick, J. A., & Berent, M. K. (1995). Origins of attitude importance: Self-interest, social identification, and value relevance. *Journal of Personality and Social Psychology*, *68*(1), 61–80.

Boninger, D. S., Krosnick, J. A., Berent, M. K., & Fabrigar, L. R. (1995). The causes and consequences of attitude importance. In R. E. Petty & J. A. Krosnick (Eds.). *Attitude strength: Antecedents and consequences* (pp. 159–189). Hillsdale, NJ: Erlbaum.

Boor, M. (1990). Reliability of ratings of movies by professional movie critics. *Psychological Reports*, *67*(1), 243–257.

Boor, M. (1992). Relationships among ratings of motion pictures by viewers and six professional movie critics. *Psychological Reports*, *70*, 1011–1021.

Bordwell, D. (1985). *Narration in the fiction film*. London: Methuen.

Bordwell, D. (1989). A case for cognitivism. *Iris*, *5*(2), 11–40.

Borg, I., & Shye, S. (1995). *Facet theory: Form and content*. Thousand Oaks, CA: Sage.

Bors, D. A., Vigneau, F., & Lalande, F. (2006). Measuring the need for cognition: Item polarity, dimensionality, and the relation with ability. *Personality and Individual Differences*, *40*(4), 809–828.

Bowers, K. S. (1973). Situationism in psychology: An analysis and a critique. *Psychological Review*, *80*(5), 307–336.

Breckler, S. J. (1984). Empirical validation of affect, behavior, and cognition as distinct components of attitude. *Journal of Personality and Social Psychology*, *47*(6), 1191–1205.

Briñol, P., Petty, R. E., & Wheeler, S. C. (2006). Discrepancies between explicit and implicit self-concepts: Consequences for information processing. *Journal of Personality and Social Psychology*, *91*(1), 154–170.

British Film Institute. (1999). *Making movies matter: Report of the Film Education Working Group*. London: British Film Institute.

Britt, T. W., Millard, M. R., Sundareswaran, P. T., & Moore, D. (2009). Personality variables predict strength-related attitude dimensions across objects. *Journal of Personality*, *77*(3), 859–882.

Brock, T. C., & Livingston, S. D. (2004). The need for entertainment scale. In L. J. Shrum (Ed.), *The psychology of entertainment media: Blurring the lines between entertainment and persuasion* (pp. 255–274). Mahwah NJ: Erlbaum.

Brömer, P. (2000). Einstellungsbildung: Erlebte Ambivalenz, subjektive Schwierigkeit und Antwortlatenz bei evaluativ inkonsistenten, konsistenten und neutralen Informationen [Attitude formation: Experienced ambivalence, perceived difficulty, and response latency when information is either evaluatively inconsistent, consistent, or neutral]. *Zeitschrift für Experimentelle Psychologie*, *47*(2), 115–128.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Brown, T. A., White, K. S., Forsyth, J. P., & Barlow, D. H. (2004). The structure of perceived emotional control: Psychometric properties of a revised Anxiety Control Questionnaire. *Behavior Therapy*, *35*(1), 75–99.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, *36*(1), 111–150.

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, *21*(2), 230–258.

Browne, M. W., Cudeck, R., Tateneni, K., & Mels, G. (2010). CEFA: Comprehensive Exploratory Factor Analysis (Version 3.04) [Software]. Retrieved from http://faculty.psy.ohio-state.edu/browne/

Bryant, J., & Vorderer, P. (Eds.). (2006). *Psychology of entertainment*. Mahwah, NJ: Erlbaum.

Büch, B. (2005). *Der Einfluss von manipulierten Filmkritiken auf die Bewertung von Spielfilmen* [The impact of manipulated film critiques on the evaluation of movies] (Unpublished master thesis). Universität Koblenz-Landau, Landau.

Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* [Introduction to test and questionnaire construction] (3., aktual. und erw. Aufl.). München: Pearson Studium.

Bundesverband Audiovisuelle Medien e. V. (2011). *Video Market 2010: BVV-Business-Report*. Retrieved from http://www.bvv-medien.de/jwb_pdfs/JWB2010.pdf

Burisch, M. (1984). Approaches to personality inventory construction: A comparison of merits. *American Psychologist*, *39*(3), 214–227.

Burst, M. (1999). Zuschauerpersönlichkeit als Voraussetzung für Fernsehmotive und Programmpräferenzen [Personality structure and media use: The influence of personality traits on television viewing motivation and preferences]. *Zeitschrift für Medienpsychologie*, *11*(3), 157–181.

Burzynski, M. H., & Bayer, D. J. (1977). The effect of positive and negative prior information on motion picture appreciation. *Journal of Social Psychology*, *101*(2), 215–218.

Busselle, R., & Bilandzic, H. (2009). Measuring narrative engagement. *Media Psychology*, *12*(4), 321–347.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin*, *115*(3), 401–423.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131.

Cacioppo, J. T., Petty, R. E., & Geen, T. R. (1989). Attitude structure and function: From the tripartite to the homeostasis model of attitudes. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 275–309). Hillsdale, NJ: Erlbaum.

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, *48*(3), 306–307.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81–105.

Carlston, D. (2010). Models of implicit and explicit mental representation. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 38–61). New York, NY: Guilford Press.

Carroll, N. E. (2003). *Engaging the moving image*. New Haven, CT: Yale University Press.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*(5), 752–766.

Chaiken, S., & Trope, Y. (Eds.). (1999). *Dual-process theories in social psychology*. New York, NY: Guilford Press.

Chang, B.-H., & Ki, E.-J. (2005). Devising a practical model for predicting theatrical movie success: Focusing on the experience good property. *Journal of Media Economics*, *18*(4), 247–269.

Chang, W. H. (1975). A typology study of movie critics. *Journalism Quarterly*, *52*(4), 721–725.

Chintagunta, P. K., Gopinath, S., & Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science*, *29*(5), 944–957.

Chuah, S. C., Drasgow, F., & Roberts, B. W. (2006). Personality assessment: Does the medium matter? No. *Journal of Research in Personality*, *40*(4), 359–376.

Clement, M. (2004). Erfolgsfaktoren von Spielfilmen im Kino. Eine Übersicht der empirischen betriebswirtschaftlichen Literatur [Critical success factors for movies in the cinema. An overview of the empirical literature on business management]. *Medien und Kommunikationswissenschaft*, *52*(2), 250–271.

Cohen, A. A. (1987). Decision making in VCR rental libraries: Information use and behavior patterns. *American Behavioral Scientist*, *30*(5), 495–508.

Cohen, J. (2001). Defining identification: A theoretical look at the identification of audiences with media characters. *Mass Communication & Society*, *4*(3), 245–264.

Cohen, J. (2006). Audience identification with media characters. In J. Bryant & P. Vorderer (Eds.), *Psychology of entertainment* (pp. 183–197). Mahwah, NJ: Erlbaum.

Collani, G. von (2010). Eine deutsche Skala zum Bedürfnis nach Bewertung (Need to Evaluate). In A. Glöckner-Rist (Ed.), *Zusammenstellung sozialwissenschaftlicher Items und Skalen. ZIS Version 14.0* [Compilation of social science items and scales, version 14.0]. Bonn: GESIS.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428.

Conrey, F. R., & Smith, E. R. (2007). Attitude representation: Attitudes as patterns in a distributed, connectionist representational system. *Social Cognition*, *25*(5), 718–735.

Converse, J. M., & Presser, S. (1986). *Survey questions: Handcrafting the standardized questionnaire*. Thousand Oaks, CA: Sage.

Cooper-Martin, E. (1992). Consumers and movies: Information sources for experiential products. In J. F. Sherry, Jr. & B. Sternthal (Eds.), *Advances in Consumer Research* (Vol. 19, pp. 756–761). Provo, UT: Association for Consumer Research.

Costa, P. T., & McCrae, R. R. (1992). *The Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Ressources.

Courtright, J. A. (1996). Rationally thinking about nonprobability. *Journal of Broadcasting & Electronic Media*, *40*(3), 414–421.

Courvoisier, D. S., Nussbeck, F. W., Eid, M., Geiser, C., & Cole, D. A. (2008). Analyzing the convergent and discriminant validity of states and traits: Development and applications of multimethod latent state-trait models. *Psychological Assessment*, *20*(3), 270–280.

Cox, E. P., III. (1983). Hollywood looks at its audience: A report of film audience research [Book review]. *Journal of Marketing Research*, *20*(4), 453–454.

Crites, S. L., Jr., Fabrigar, L. R., & Petty, R. E. (1994). Measuring the affective and cognitive properties of attitudes: Conceptual and methodological issues. *Personality and Social Psychology Bulletin*, *20*(6), 619–634.

Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions* (2nd ed.). Urbana, IL: University of Illinois Press.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302.

Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology*, *16*(2), 137–163.

Cronley, M. L., Mantel, S. P., & Kardes, F. R. (2010). Effects of accuracy motivation and need to evaluate on mode of attitude formation and attitude–behavior consistency. *Journal of Consumer Psychology*, *20*(3), 274–281.

Cunningham, W. A., Johnson, M. K., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2003). Neural components of social evaluation. *Journal of Personality and Social Psychology*, *85*(4), 639–649.

Cunningham, W. A., Packer, D. J., Kesek, A., & van Bavel, J. J. (2008). Implicit measurement of attitudes: A physiological approach. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes. Insights from the new implicit measures* (pp. 485–512). New York, NY: Psychology Press.

Cunningham, W. A., Raye, C. L., & Johnson, M. K. (2004). Implicit and explicit evaluation: fMRI correlates of valence, emotional intensity, and control in the processing of attitudes. *Journal of Cognitive Neuroscience*, *16*(10), 1717–1729.

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: A social cognitive neuroscience perspective. *Trends in Cognitive Sciences*, *11*(3), 97–104.

Cunningham, W. A., Zelazo, P. D., Packer, D. J., & van Bavel, J. J. (2007). The iterative reprocessing model: A multilevel framework for attitudes and evaluation. *Social Cognition*, *25*(5), 736–760.

Dahl, G. (1971). Zur Berechnung des Schwierigkeitsindex bei quantitativ abgestufter Aufgabenbewertung [Computing a difficulty index for rating scales]. *Diagnostica*, *17*, 139–142.

Dasgupta, N., & Hunsinger, M. (2008). The opposite of a great truth is also true: When do student samples help versus hurt the scientific study of prejudice? *Psychological Inquiry*, *19*(2), 90–98.

D'Astous, A., & Touil, N. (1999). Consumer evaluations of movies on the basis of critics' judgment. *Psychology & Marketing*, *16*(8), 677–694.

De Silva, I. (1998). Consumer selection of motion pictures. In B. R. Litman (Ed.), *The motion picture mega-industry* (pp. 144–171). Needham Heights, MA: Allyn & Bacon.

De Vany, A., & Walls, W. D. (1999). Uncertainty in the movie industry: Does star power reduce the terror of the box office? *Journal of Cultural Economics*, *23*(4), 285–318.

Deinzer, R., Steyer, R., Eid, M., Notz, P., Schwenkmezger, P., Ostendorf, F., & Neubauer, A. (1995). Situational effects in trait assessment: The FPI, NEOFFI, and EPI questionnaires. *European Journal of Personality*, *9*(1), 1–23.

Desai, K. K., & Basuroy, S. (2005). Interactive influence of genre familiarity, star power, and critics' reviews in the cultural goods industry: The case of motion pictures. *Psychology & Marketing*, *22*(3), 203–223.

Digman, J. M. (1989). Five robust trait dimensions: Development, stability, and utility. *Journal of Personality*, *57*(2), 195–214.

Dodds, J. C., & Holbrook, M. B. (1988). What's an Oscar worth? An empirical estimate of the effects of nominations and awards on movie distribution and revenues. In B. A. Austin (Ed.), *Current research in film. Audiences, economics and law* (Vol. 4, pp. 72–88). Norwood, NJ: Ablex.

Doll, J., & Hasebrink, U. (1990). Zur Programmauswahl von Fernsehzuschauern: Die Bedeutung von Einstellungen gegenüber Sendungstypen [Program selection by television viewers: The meaning of attitudes towards types of broadcast]. *Rundfunk und Fernsehen*, *38*(1), 21–36.

Dumenci, L., & Windle, M. (1998). A multitrait–multioccasion generalization of the latent trait–state model: Description and application. *Structural Equation Modeling*, *5*(4), 391–410.

Eagly, A. H., & Chaiken, S. (Eds.). (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.

Eagly, A. H., & Chaiken, S. (2007). The advantages of an inclusive definition of attitude. *Social Cognition*, *25*(5), 582–602.

Eaton, A. A., & Visser, P. S. (2008). Attitude importance: Understanding the causes and consequences of passionately held views. *Social and Personality Psychology Compass*, *2*(4), 1719–1736.

Edwards, E. D. (1991). The ecstasy of horrible expectations: Morbid curiosity, sensation seeking, and the interest in horror movies. In B. A. Austin (Ed.), *Current research in film. Audiences, economics and law* (Vol. 5, pp. 19–38). Norwood, NJ: Ablex.

Ehrhart, M. G., Ehrhart, K. H., Roesch, S. C., Chung-Herrera, B. G., Nadler, K., & Bradshaw, K. (2009). Testing the latent factor structure and construct validity of the Ten-Item Personality Inventory. *Personality and Individual Differences*, *47*(8), 900–905.

Eisermann, J. (2004). *Strukturerkennende Verfahren bei Daten aus offenem Antwortformat. Ein Beitrag zur Psychologie des Wunsches* [Methods of revealing structure in data of open question form. A contribution to the psychology of the wish] (Doctoral dissertation). Retrieved from http://www.diss.fu-berlin.de/2004/97/index.html

Ekman, P., Liebert, R. M., Friesen, W. V., Harrison, R., Zlatchin, C., Malmstrom, E. J., & Baron, R. A. (1972). Facial expressions of emotion while watching televisid violence as predictors of subsequent aggression. In G. A. Comstock, E. A. Rubinstein, & J. P. Murray (Eds.), *Television and social behavior. Reports and papers, Volume V: Television's effects: Further explorations. A technical report to the Surgeon General's Scientific Advisory Committee on Television and Social Behavior* (pp. 22–58). Washington, DC: U.S. Government Printing Office.

Elberse, A., & Eliashberg, J. (2003). Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures. *Marketing Science*, *22*(3), 329–354.

Eliashberg, J., & Sawhney, M. S. (1994). Modeling goes to Hollywood: Predicting individual differences in movie enjoyment. *Management Science*, *40*(9), 1151–1173.

Eliashberg, J., Jonker, J.-J., Sawhney, M. S., & Wierenga, B. (2000). MOVIEMOD: An implementable decision support system for pre-release market evaluation of motion pictures. *Marketing Science*, *19*(3), 226–243.

Elliott, W. R., & Quattlebaum, C. P. (1979). Similarities in patterns of media use: A cluster analysis of media gratifications. *Western Journal of Speech Communication*, *43*(1), 61–72.

Endler, N. S., & Magnusson, D. (1976). Toward an interactional psychology of personality. *Psychological Bulletin*, *83*(5), 956–974.

Epstein, S., & O'Brien, E. J. (1985). The person–situation debate in historical and current perspective. *Psychological Bulletin*, *98*(3), 513–537.

Eysenck, H. J., & Eysenck, S. B. G. (1976). *Psychoticism as a dimension of personality*. London: Hodder and Stoughton.

Faber, R. J., & O'Guinn, T. C. (1984). Effect of media advertising and other sources on movie selection. *Journalism Quarterly*, *61*(2), 371–377.

Fabrigar, L. R., MacDonald, T. K., & Wegener, D. T. (2005). The structure of attitudes. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 79–125). Mahwah, NJ: Erlbaum.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299.

Farber, M. L. (1952). The college student as laboratory animal. *American Psychologist*, *7*(3), 102.

Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Hillsdale, NJ: Erlbaum.

Fazio, R. H. (2000). Accessible attitudes as tools for object appraisal: Their costs and benefits. In G. R. Maio & J. M. Olson (Eds.), *Why we evaluate. Functions of attitudes* (pp. 1–36). Mahwah, NJ: Erlbaum.

Fazio, R. H. (2007). Attitudes as object-evaluation associations of varying strength. *Social Cognition*, *25*(5), 603–637.

Fazio, R. H., Chen, J.-m., McDonel, E. C., & Sherman, S. J. (1982). Attitude accessibility, attitude-behavior consistency, and the strength of the object-evaluation association. *Journal of Experimental Social Psychology*, *18*(4), 339–357.

FDW Werbung im Kino e.V. (2010). *Kinobesucher in der MA 2010: Analyse der Berichterstattung der MA 2010*. Retrieved from http://www.uia-europe.com/fdw-research/pdf/ma2010.pdf

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson, and Company.

FFA Filmförderungsanstalt. (2011). *Der Kinobesucher 2010: Strukturen und Entwicklungen* [German Federal Film Board: The Moviegoers 2010]. Berlin. Retrieved from http://www.ffa.de/downloads/publikationen/kinobesucher_2010.pdf

Finn, S. (1997). Origins of media exposure: Linking personality traits to TV, radio, print, and film use. *Communication Research*, *24*(5), 507–529.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Forsterlee, R., & Ho, R. (1999). An examination of the short form of the Need for Cognition scale applied in an Australian sample. *Educational and Psychological Measurement*, *59*(3), 471–480.

Früh, W. (2001). *Gewaltpotentiale des Fernsehangebots: Programmangebot und zielgruppenspezifische Interpretation* [Violence potentials of television]. Wiesbaden: Westdeutscher Verlag.

Galdi, S., Arcuri, L., & Gawronski, B. (2008). Automatic mental associations predict future choices of undecided decision-makers. *Science*, *321*(5892), 1100–1102.

Gangestad, S. W., & Snyder, M. (2000). Self-monitoring: Appraisal and reappraisal. *Psychological Bulletin*, *126*(4), 530–555.

Gawronski, B. (Ed.). (2007). What is an attitude? [Special Issue]. *Social Cognition*, *25*(5).

Gawronski, B., & Bodenhausen, G. V. (2006a). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*(5), 692–731.

Gawronski, B., & Bodenhausen, G. V. (2006b). Associative and propositional processes in evaluation: Conceptual, empirical, and metatheoretical issues: Reply to Albarracín, Hart, and McCulloch (2006), Kruglanski and Dechesne (2006), and Petty and Briñol (2006). *Psychological Bulletin*, *132*(5), 745–750.

Gawronski, B., & Bodenhausen, G. V. (2007). Unraveling the processes underlying evaluation: Attitudes from the perspective of the APE model. *Social Cognition*, *25*(5), 687–717.

Gawronski, B., & Bodenhausen, G. V. (2009). Operating principles versus operating conditions in the distinction between associative and propositional processes. *Behavioral and Brain Sciences*, *32*(2), 207–208.

Gawronski, B., & Bodenhausen, G. V. (2011). The associative-propositional evaluation model: Theory, evidence, and open questions. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology, Vol. 44* (pp. 59–127). San Diego, CA: Academic Press.

Gawronski, B., Strack, F., & Bodenhausen, G. V. (2008). Attitudes and cognitive consistency: The role of associative and propositional processes. In R. E. Petty, R. H. Fazio, & P. Briñol (Eds.), *Attitudes. Insights from the new implicit measures* (pp. 85–117). New York, NY: Psychology Press.

Gazley, A., Clark, G., & Sinha, A. (2011). Understanding preferences for motion pictures. *Journal of Business Research*, *64*(8), 854–861.

Gehrau, V. (2001). *Fernsehgenres und Fernsehgattungen: Ansätze und Daten zur Rezeption, Klassifikation und Bezeichnung von Fernsehprogrammen* [Television genres and television forms: Approaches and data relating to the reception, classification and designation of television programs]. München: Reinhard Fischer.

Gehrau, V. (2003). (Film-) Genres und die Reduktion von Unsicherheit [(Film) Genres and the reduction of uncertainty]. *Medien und Kommunikationswissenschaft*, *51*(2), 213–231.

Gehrau, V. (2008). *Fernsehbewertung und Fernsehhandlung: Ansätze und Daten zu Erhebung, Modellierung und Folgen von Qualitätsurteilen des Publikums über Fernsehangebote* [Television evaluation and television action: Approaches and data relating to the recording, modeling and consequences of quality assessments of television programs by viewers]. München: Reinhard Fischer.

Gehrau, V. (2009). Basisgenres. Ein Ansatz zur Strukturierung medienvermittelter fiktionaler Öffentlichkeit am Beispiel des Fernsehangebots [Basic genres]. In K. Merten (Ed.), *Konstruktion von Kommunikation in der Mediengesellschaft. Festschrift für Joachim Westerbarkey* [Construction of communicaion in the media society: A festschrift for Joachim Westerbarkey] (pp. 31–43). Wiesbaden: VS Verlag für Sozialwissenschaften.

Gemser, G., Leenders, M. A. A. M., & Wijnberg, N. M. (2008). Why some awards are more effective signals of quality than others: A study of movie awards. *Journal of Management*, *34*(1), 25–54.

Gemser, G., van Oostrum, M., & Leenders, M. A. A. M. (2007). The impact of film reviews on the box office performance of art house versus mainstream motion pictures. *Journal of Cultural Economics*, *31*(1), 43–63.

Gerhards, M., & Klingler, W. (2011). Sparten- und Formattrends im deutschen Fernsehen: Rückblick auf das Programmjahr 2009 [Format trends in German television in 2009]. *Media Perspektiven*, *42*(1), 36–54.

Giner-Sorolla, R. (2001). Affective attitudes are not always faster: The moderating role of extremity. *Personality and Social Psychology Bulletin*, *27*(6), 666–677.

Glasman, L. R., & Albarracín, D. (2006). Forming attitudes that predict future behavior: A meta-analysis of the attitude-behavior relation. *Psychological Bulletin*, *132*(5), 778–822.

Gleich, U., Kreisel, E., Thiele, L., Vierling, M., & Walther, S. (1998). Sensation-Seeking, Fernsehverhalten und Freizeitaktivitäten [Sensation-seeking, television viewing behavior and leisure activities]. In W. Klingler, G. Roters, & O. Zöllner (Eds.), *Fernsehforschung in Deutschland. Themen - Akteure - Methoden* [Television research in Germany: Themes - actors - methods, Vol. 1] (Teilband 1, pp. 661–688). Baden Baden: Nomos.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, *48*(1), 26–34.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*(1), 84–96.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., JR (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist*, *59*(2), 93–104.

Graaf, A. de, Hoeken, H., Sanders, J., & Beentjes, H. (2009). The role of dimensions of narrative engagement in narrative persuasion. *Communications*, *34*(4), 385–405.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, *79*(5), 701–721.

Greenberg, B. S., & Busselle, R. W. (1996). Audience dimensions of quality in situation comedies and action programmes. In S. Ishikawa (Ed.), *Quality assessment of television* (pp. 169–196). Luton, UK: University of Luton Press; John Libbey Media.

Grimm, J. (1993). Der kultivierte Schrecken? Erlebnisse von Horrorfilmen im Rahmen eines Zuschauerexperiments [Cultivated horror? Experiences of horror films during a viewer experiment]. *Publizistik*, *38*(2), 206–217.

Grimm, J. (1997). Physiologische und psychosoziale Aspekte der Fernsehgewalt-Rezeption: TV-Gefühlsmanagement zwischen Angst und Aggressionen [Physiological and psychosocial aspects of television violence reception: TV emotion management between anxiety and aggression]. *Zeitschrift für Medienpsychologie*, *9*(2), 127–166.

Grimm, J. (1999). *Fernsehgewalt: Zuwendungsattraktivität, Erregungsverläufe, sozialer Effekt. Zur Begründung und praktischen Anwendung eines kognitiv-physiologischen Ansatzes der Medienrezeptionsforschung am Beispiel von Gewaltdarstellungen* [Violence on television: Attraction – excitation processes – social effect. Basis and practical application of a cognitive physiological approach to media reception research, using portrayals of violence as an example]. Opladen: Westdeutscher Verlag.

Groeben, N. (1990). Subjective theories and the explanation of human action. In G. R. Semin & K. J. Gergen (Eds.), *Everyday understanding. Social and scientific implications* (pp. 19–44). London: Sage.

Groeben, N., & Scheele, B. (2000). Dialogue-hermeneutic method and the "Research Program Subjective Theories". *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *1*(2), Art. 10, http://nbn-resolving.de/urn:nbn:de:0114-fqs0002105.

Gschwendner, T., Hofmann, W., & Schmitt, M. (2006). Moderatoren der Konsistenz implizit und explizit erfasster Einstellungen und Persönlichkeitsmerkmale [Moderators of the consistency of implicitly and explicitly assessed attitudes and personality traits]. *Psychologische Rundschau*, *57*(1), 13–33.

Gunter, B. (1997). An audience-based approach to assessing programme quality. In P. Winterhoff-Spurk & T. H. A. van der Voort (Eds.), *New horizons in media psychology: Research cooperation and projects in Europe* (pp. 11–34). Wiesbaden: Westdeutscher Verlag.

Gutman, J. (1982). The impact of advertising at the time of consumption. *Journal of Advertising Research*, *22*(4), 35–40.

Haddock, G., & Zanna, M. P. (1998). Assessing the impact of affective and cognitive information in predicting attitudes toward capital punishment. *Law and Human Behavior*, *22*(3), 325–339.

Hadida, A. L. (2009). Motion picture performance: A review and research agenda. *International Journal of Management Reviews*, *11*(3), 297–335.

Hagemann, D., & Meyerhoff, D. (2008). A simplified estimation of latent state-trait parameters. *Structural Equation Modeling*, *15*(4), 627–650.

Hall, A. (2005a). Audience personality and the selection of media and media genres. *Media Psychology*, *7*(4), 377–398.

Hall, A. (2005b). Sensation seeking and the use and selection of media materials. *Psychological Reports*, *97*(1), 236–244.

Handel, L. A. (1950). *Hollywood looks at its audience: A report of film audience research*. Champaign, IL: University of Illinois Press.

Hasebrink, U. & Bube, H. (1998). *Spielfilmbewertungen in Programmzeitschriften und ihr Zusammenhang mit dem Publikumsverhalten* [Movie evaluations in TV program guides and their relationship to audience behavior]. Unpublished report, Hans Bredow Institute, Hamburg.

Hasebrink, U., & Bube, H. (1999, April). *Die Rolle der Spielfilmempfehlungen in Programmzeitschriften für die Filmauswahl der Fernsehzuschauer* [The role of film recommendations in TV program guides for TV audience movie choice]. Vortrag auf der 41. Tagung experimentell arbeitender Psychologen (TeaP), Leipzig.

Hawkins, R. P., Pingree, S., Hitchon, J., Gorham, B. W., Kannaovakun, P., Kahlor, L., . . . Schmidt, T. (2001). Predicting selection and activity in television genre viewing. *Media Psychology*, *3*(3), 237–263.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77–89.

Hennig-Thurau, T., & Dallwitz-Wegner, D. (2004). Zum Einfluss von Filmstars auf den Erfolg von Spielfilmen [On the impact of stars on the success of motion pictures]. *MedienWirtschaft – Zeitschrift für Medienmanagement und Kommunikationsökonomie*, *1*(4), 157–170.

Hennig-Thurau, T., Houston, M. B., & Walsh, G. (2007). Determinants of motion picture box office and profitability: An interrelationship approach. *Review of Managerial Science*, *1*(1), 65–92.

Hennig-Thurau, T., Walsh, G., & Wruck, O. (2001). An investigation into the factors determining the success of service innovations: The case of motion pictures. *Academy of Marketing Science Review*, *1*(6). Retrieved from http://www.amsreview.org/articles/hennig06-2001.pdf

Hennig-Thurau, T., & Wruck, O. (2000). Warum wir ins Kino gehen: Erfolgsfaktoren von Kinofilmen [Why we go to the movies]. *Marketing ZFP*, *22*(3), 241–258.

Henry, P. J. (2008). College sophomores in the laboratory redux: Influences of a narrow data base on social psychology's view of the nature of prejudice. *Psychological Inquiry*, *19*(2), 49–71.

Himmelweit, H. T., Swift, B., & Jaeger, M. E. (1980). The audience as a critic: A conceptual analysis of television entertainment. In P. H. Tannenbaum (Ed.), *The entertainment functions of television* (pp. 67–106). Hillsdale, NJ: Erlbaum.

Hirschman, E. C. (1987). Consumer preferences in literature, motion pictures, and television programs. *Empirical Studies of the Arts*, *5*(1), 31–46.

Hobbs, R., Frost, R., Davis, A., & Stauffer, J. (1988). How first-time viewers comprehend editing conventions. *Journal of Communication*, *38*(4), 50–60.

Hoffmann, D., & Schwender, C. (2007). Biographical functions of cinema and film preferences among older German adults: A representative quantitative survey. *Communications*, *32*(4), 473–491.

Hoffner, C., & Cantor, J. (1985). Developmental differences in responses to a television character's appearance and behavior. *Developmental Psychology*, *21*(6), 1065–1074.

Hofmann, W., Gschwendner, T., Nosek, B. A., & Schmitt, M. (2005). What moderates implicit-explicit consistency? *European Review of Social Psychology*, *16*(10), 335–390.

Hofmann, W., Gschwendner, T., & Schmitt, M. (2005). On implicit-explicit consistency: The moderating role of individual differences in awareness and adjustment. *European Journal of Personality*, *19*(1), 25–49.

Holbrook, A. L., Berent, M. K., Krosnick, J. A., Visser, P. S., & Boninger, D. S. (2005). Attitude importance and the accumulation of attitude-relevant knowledge in memory. *Journal of Personality and Social Psychology*, *88*(5), 749–769.

Holbrook, M. B. (1999). Popular appeal versus expert judgments of motion pictures. *Journal of Consumer Research*, *26*(2), 144–155.

Holbrook, M. B. (2005). The role of ordinary evaluations in the market for popular culture: Do consumers have "good taste"? *Marketing Letters*, *16*(2), 75–86.

Holbrook, M. B., & Addis, M. (2007). Taste versus the market: An extension of research on the consumption of popular culture. *Journal of Consumer Research*, *34*(3), 415–424.

Holbrook, M. B., & Addis, M. (2008). Art versus commerce in the movie industry: A two-path model of motion-picture success. *Journal of Cultural Economics*, *32*(2), 87–107.

Holicki, S., & Krcho, M. (1992). Filmkritik, Filmbeurteilung und der Wunsch, den Film zu sehen. Zum Einfluß von Struktur und Inhalt einer Filmrezension [Film criticism, film analysis and the desire to see a film: The influence of structure and content of a film review]. *Publizistik*, *37*(3), 361–372.

Horcajo, J., Díaz, D., Briñol, P., & Gandarillas, B. (2008). Necesidad de evaluación: Propuesta para su medición en castellano. *Psicothema*, *20*(4), 557–562.

Hoyle, R. H., & Robinson, J. C. (2004). Mediated and moderated effects in social psychological research: Measurement, design, and analysis issues. In C. Sansone, C. C. Morf, & A. T. Panter (Eds.), *The Sage handbook of methods in social psychology* (pp. 213–233). Thousand Oaks, CA: Sage.

Hoyle, R. H., Stephenson, M. T., Palmgreen, P., Lorch, E. P., & Donohew, R. L. (2002). Reliability and validity of a brief measure of sensation seeking. *Personality and Individual Differences*, *32*(3), 401–414.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55.

Huskinson, T. L. H., & Haddock, G. (2004). Individual differences in attitude structure: Variance in the chronic reliance on affective and cognitive information. *Journal of Experimental Social Psychology*, *40*(1), 82–90.

Huskinson, T. L. H., & Haddock, G. (2006). Individual differences in attitude structure and the accessibility of the affective and cognitive components of attitude. *Social Cognition*, *24*(4), 453–468.

Igartua, J.-J. (2010). Identification with characters and narrative persuasion through fictional feature films. *Communications*, *35*(4), 347–373.

Jansen, C. (2005). The performance of German motion pictures, profits and subsidies: Some empirical evidence. *Journal of Cultural Economics*, *29*(3), 191–212.

Jarvis, W. B. G., & Petty, R. E. (1996). The need to evaluate. *Journal of Personality and Social Psychology*, *70*(1), 172–194.

John, O. P., & Benet-Martínez, V. (2000). Measurement: Reliability, construct validation, and scale construction. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 339–369). Cambridge, UK: Cambridge University Press.

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory—Versions 4a and 54*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.

John, O. P., & Soto, C. J. (2007). The importance of being valid: Reliability and the process of construct validation. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 461–494). New York, NY: Guilford Press.

Johnson, B. T., Maio, G. R., & Smith-McLallen, A. (2005). Communication and attitude change: Causes, processes, and effects. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 617–669). Mahwah, NJ: Erlbaum.

Johnston, D. D. (1995). Adolescents' motivations for viewing graphic horror. *Human Communication Research*, *21*(4), 522–552.

Jöreskog, K. G., & Sörbom, D. (1979). *Advances in factor analysis and structural equation models*. Cambridge, MA: Abt Books.

Karpinski, A., Steinman, R. B., & Hilton, J. L. (2005). Attitude importance as a moderator of the relationship between implicit and explicit attitude measures. *Personality and Social Psychology Bulletin*, *31*(7), 949–962.

Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly*, *24*(2), 163–204.

Katz, E., Gurevitch, M., & Haas, H. (1973). On the use of the mass media for important things. *American Sociological Review*, *38*(2), 164–181.

Kintsch, W. (1977). On comprehending stories. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 33–62). Hillsdale, NJ: Erlbaum.

Klimmt, C. (2011). Media psychology and complex modes of entertainment experiences. *Journal of Media Psychology: Theories, Methods, and Applications*, *23*(1), 34–38.

Knobloch, S. (2003). Mood adjustment via mass communication. *Journal of Communication*, *53*(2), 233–250.

Knobloch-Westerwick, S., & Keplinger, C. (2008). Murder for pleasure: Impacts of plot complexity and need for cognition on mystery enjoyment. *Journal of Media Psychology: Theories, Methods, and Applications*, *20*(3), 117–128.

Kracauer, S. (1960). *Theory of film: The redemption of physical reality*. London: Oxford University Press.

Krakowiak, K. M. (2008). *When good characters do bad things. Examining the effect of moral ambiguity on enjoyment* (Doctoral dissertation). Retrieved from http://etda.libraries.psu.edu/theses/approved/WorldWideFiles/ETD-2903/Dissertation_with_Vita_%28Maja%29.pdf

Krcmar, M. (2009). Individual differences in media effects. In R. L. Nabi & M. B. Oliver (Eds.), *The Sage handbook of media processes and effects* (pp. 235–250). Thousand Oaks, CA: Sage.

Krcmar, M., & Greene, K. (1999). Predicting exposure to and uses of television violence. *Journal of Communication*, *49*(3), 24–45.

Krcmar, M., & Kean, L. G. (2005). Uses and gratifications of media violence: Personality correlates of viewing and liking violent genres. *Media Psychology*, *7*(4), 399–420.

Krcmar, M., & Renfro, S. L. (2005, May). *Media enjoyment as affect, cognition, and behavior: Developing a scale to assess media enjoyment*. Paper presented at the 55th Annual Conference of the International Communication Association, New York, NY.

Krosnick, J. A. (1988a). Attitude importance and attitude change. *Journal of Experimental Social Psychology*, *24*(3), 240–255.

Krosnick, J. A. (1988b). The role of attitude importance in social evaluation: A study of policy preferences, presidential candidate evaluations, and voting behavior. *Journal of Personality and Social Psychology*, *55*(2), 196–210.

Krosnick, J. A., & Fabrigar, L. R. (1997). Designing rating scales for effective measurernent in surveys. In L. E. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 141–164). New York, NY: Wiley.

Krosnick, J. A., & Petty, R. E. (1995). Attitude strength: An overview. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 1–24). Hillsdale, NJ: Erlbaum.

Krosnick, J. A., & Presser, S. (2010). Question and questionnaire design. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (2nd ed.). Bingley, UK: Emerald Group Publishing Limited.

Kruglanski, A. W., & Stroebe, W. (2005). The influence of beliefs and goals on attitudes: Issues of structure, function, and dynamics. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 323–368). Mahwah, NJ: Erlbaum.

Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology*, *81*(4), 684–696.

Larsen, J. T., Norris, C. J., McGraw, A. P., Hawkley, L. C., & Cacioppo, J. T. (2009). The evaluative space grid: A single-item measure of positivity and negativity. *Cognition and Emotion*, *23*(3), 453–480.

Lazarus, R. S. (1982). Thoughts on the relations between emotion and cognition. *American Psychologist*, *37*(9), 1019–1024.

Levin, A. M., Levin, I. P., & Heath, C. E. (1997). Movie stars and authors as brand names: Measuring brand equity in experiential products. In M. Brucks & D. J. MacInnis (Eds.), *Advances in consumer research* (Vol. 24, pp. 175–181). Provo, UT: Association for Consumer Research.

Levine, T., Hullett, C. R., Turner, M. M., & Lapinski, M. K. (2006). The desirability of using confirmatory factor analysis on published scales. *Communication Research Reports*, *23*(4), 309–314.

Lichtenstein, A., & Rosenfeld, L. B. (1983). Uses and misuses of gratifications research: An explication of media functions. *Communication Research*, *10*(1), 97–109.

Lieberman, M. D. (2007). Social cognitive neuroscience: A review of core processes. *Annual Review of Psychology*, *58*, 259–289.

Liebes, T., & Katz, E. (1986). Patterns of involvement in television fiction: A comparative analysis. *European Journal of Communication*, *1*(2), 151–171.

Linton, J. M., & Petrovich, J. A. (1988). The application of the consumer information acquisition approach to movie selection: An exploratory study. In B. A. Austin (Ed.), *Current research in film. Audiences, economics and law* (Vol. 4, pp. 24–45). Norwood, NJ: Ablex.

Litman, B. R. (1983). Predicting success of theatrical movies: An empirical study. *Journal of Popular Culture*, *16*(4), 159–175.

Litman, B. R., & Ahn, H. (1998). Predicting financial success of motion pictures: The early '90s experience. In B. R. Litman (Ed.), *The motion picture mega-industry* (pp. 172–197). Needham Heights, MA: Allyn & Bacon.

Litman, B. R., & Kohl, L. S. (1989). Predicting financial success of motion pictures: The '80s experience. *Journal of Media Economics*, *2*(2), 35–50.

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing*, *70*(3), 74–89.

Loevinger, J. (1957). Objective tests as instruments of psychological theory [Monograph]. *Psychological Reports*, *3*, 635–694.

Loevinger, J. (Ed.). (1998). *Technical foundations for measuring ego development: The Washington University Sentence Completion Test*. Mahwah, NJ: Erlbaum.

Longman Dictionary of Contemporary English (5th ed.). (2009). Essex, UK: Pearson Education Limited.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Lord, K. R., & Putrevu, S. (2006). Exploring the dimensionality of the Need for Cognition scale. *Psychology & Marketing*, *23*(1), 11–34.

Mahajan, V., Muller, E., & Kerin, R. A. (1984). Introduction strategy for new products with positive and negative word-of-mouth. *Management Science*, *30*(12), 1389–1404.

Maio, G. R., & Esses, V. M. (2001). The need for affect: Individual differences in the motivation to approach or avoid emotions. *Journal of Personality*, *69*(4), 583–615.

Maio, G. R., & Olson, J. M. (Eds.). (2000). *Why we evaluate: Functions of attitudes*. Mahwah, NJ: Erlbaum.

Mandler, J. M. (1984). *Stories, scripts, and scenes: Aspects of schema theory*. Hillsdale, NJ: Erlbaum.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*(3), 519–530.

Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology*, *70*(4), 810–819.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*(3), 320–341.

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling*, *16*(3), 439–476.

Matthes, J. (2004). Die Schema-Theorie in der Medienwirkungsforschung: Ein unscharfer Blick in die "Black Box"? [The schema theory in media effects research: A blurred view into the "black box"?]. *Medien und Kommunikationswissenschaft*, *52*(4), 545–568.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.

McGuire, W. J. (1986). The vicissitudes of attitudes and similar representational constructs in twentieth century psychology. *European Journal of Social Psychology*, *16*(2), 89–130.

McNemar, Q. (1946). Opinion-attitude methodology. *Psychological Bulletin*, *43*(4), 289–374.

Medienpädagogischer Forschungsverbund Südwest (MPFS). (2010). *KIM-Studie 2010: Kinder + Medien, Computer + Internet. Basisuntersuchung zum Medienumgang 6- bis 13-Jähriger* [2010 KIM study on the media use of children in Germany]. Retrieved from http://www.mpfs.de/fileadmin/KIM-pdf10/KIM2010.pdf

Merton, R. K. (1987). The focussed interview and focus groups: Continuities and discontinuities. *Public Opinion Quarterly*, *51*(4), 550–566.

Messick, S. (1981). Constructs and their vicissitudes in educational and psychological measurement. *Psychological Bulletin*, *89*(3), 575–588.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York, NY: Macmillan.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741–749.

Mikos, L. (2003). *Film- und Fernsehanalyse* [Film and television analysis]. Konstanz: UTB.

Mills, J. (1993). The appeal of tragedy: An attitude interpretation. *Basic and Applied Social Psychology*, *14*(3), 255–271.

Mischel, W. (1968). *Personality and assessment*. New York, NY: Wiley.

Möhring, W., & Schlütz, D. (2010). *Die Befragung in der Medien- und Kommunikationswissenschaft: Eine praxisorientierte Einführung* [Surveys in media and communication sciences. A practical introduction] (2., überarb. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.

Möller, K. E. K., & Karppinen, P. (1983). Role of motives and attributes in consumer motion picture choice. *Journal of Economic Psychology*, *4*(3), 239–262.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, *38*(4), 379–387.

Motion Picture Association of America (2011). *MPAA Theatrical Statistics 2010*. Retrieved from http://www.mpaa.org/policy/industry

Moul, C. C. (2007). Measuring word of mouth's impact on theatrical movie admissions. *Journal of Economics and Management Strategy*, *16*(4), 859–892.

Muck, P. M., Hell, B., & Gosling, S. D. (2007). Construct validation of a short five-factor model instrument: A self-peer study on the German adaptation of the Ten-Item Personality Inventory (TIPI-G). *European Journal of Psychological Assessment*, *23*(3), 166–175.

Musch, J., Brockhaus, R., & Bröder, A. (2002). Ein Inventar zur Erfassung von zwei Faktoren sozialer Erwünschtheit [An inventory for the assessment of two factors of social desirability]. *Diagnostica*, *48*(3), 121–129.

Nabi, R. L., & Krcmar, M. (2004). Conceptualizing media enjoyment as attitude: Implications for mass media effects research. *Communication Theory*, *14*(4), 288–310.

Neelamegham, R., & Chintagunta, P. (1999). A Bayesian model to forecast new product performance in domestic and international markets. *Marketing Science*, *18*(2), 115–136.

Neelamegham, R., & Jain, D. (1999). Consumer choice process for experience goods: An econometric model and analysis. *Journal of Marketing Research*, *36*(3), 373–386.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. New York, NY: Freeman.

Nelson, R. A., Donihue, M. R., Waldman, D. M., & Wheaton, C. (2001). What's an Oscar worth? *Economic Inquiry*, *39*(1), 1–16.

Nikken, P. (1999). *Quality in children's television* (Doctoral dissertation). Retrieved from http://www.opvoedingsondersteuning.info/Jeugd%20en%20media/download/Quality%20in%20children%27s%20television.pdf

Nikken, P., & van der Voort, T. H. A. (1997). Children's views on quality standards for children's television programs. *Journal of Educational Media*, *23*(2-3), 169–188.

Nikken, P., van der Voort, T. H. A., & van Bochove, E. (1996). Maternal quality standards for children's television programs. *Journal of Educational Media*, *22*(1), 41–54.

Nosek, B. A. (2005). Moderators of the relationship between implicit and explicit evaluation. *Journal of Experimental Psychology: General*, *134*(4), 565–584.

Nuszbaum, M., Voss, A., Klauer, K. C., & Betsch, T. (2010). Assessing individual differences in the use of haptic information using a German translation of the Need for Touch scale. *Social Psychology*, *41*(4), 263–274.

O'Hara, R. E., Walter, M. I., & Christopher, A. N. (2009). Need for cognition and conscientiousness as predictors of political interest and voting strategy. *Journal of Applied Social Psychology*, *39*(6), 1397–1416.

Ohler, P. (1994). *Kognitive Filmpsychologie: Verarbeitung und mentale Repräsentation narrativer Filme* [Cognitive film psychology: Information processing and mental representation of narrative films]. Münster: MAKS.

Oliver, M. B. (2008). Tender affective states as predictors of entertainment preference. *Journal of Communication*, *58*(1), 40–61.

Oliver, M. B. (2009). Entertainment. In R. L. Nabi & M. B. Oliver (Eds.), *The Sage handbook of media processes and effects* (pp. 161–175). Thousand Oaks, CA: Sage.

Oliver, M. B., & Bartsch, A. (2010). Appreciation as audience response: Exploring entertainment gratifications beyond hedonism. *Human Communication Research*, *36*(1), 53–81.

Oliver, M. B., & Hartmann, T. (2010). Exploring the role of meaningful experiences in users' appreciation of "good movies". *Projections: The Journal of Movies and Mind*, *4*(2), 128–150.

Oliver, M. B., & Krakowiak, K. M. (2009). Individual differences. In J. Bryant & M. B. Oliver (Eds.), *Media effects. Advances in theory and research* (3rd ed., pp. 517–531). New York, NY: Routledge.

Oliver, M. B., & Nabi, R. L. (Eds.). (2004). Exploring the concept of media enjoyment [Special issue]. *Communication Theory*, *14*(4).

Oliver, M. B., & Raney, A. A. (2011). Entertainment as pleasurable and meaningful: Identifying hedonic and eudaimonic motivations for entertainment consumption. *Journal of Communication*, *61*(5), 984–1004.

Oliver, M. B., Weaver, J. B., III, & Sargent, S. L. (2000). An examination of factors related to sex differences in enjoyment of sad films. *Journal of Broadcasting & Electronic Media*, *44*(2), 282–300.

Ostrom, T. M. (1989). Interdependence of attitude theory and measurement. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 11–36). Hillsdale, NJ: Erlbaum.

Otto, L., & Bacherle, P. (2011). Politisches Interesse Kurzskala (PIKS) – Entwicklung und Validierung [Short Scale Political Interest (SSPI) - Development and Validation]. *Politische Psychologie/Journal of Political Psychology*, *1*(1), 19–35.

Palmgreen, P. (1984). Uses and gratifications: A theoretical perspective. In R. N. Bostrom (Ed.), *Communication yearbook 8* (pp. 20–55). Beverly Hills, CA: Sage.

Palmgreen, P., Cook, P. L., Harvill, J. G., & Helm, D. M. (1988). The motivational framework for moviegoing: Uses and avoidances of theatrical films. In B. A. Austin (Ed.), *Current research in film. Audiences, economics and law* (Vol. 4, pp. 1–23). Norwood, NJ: Ablex.

Palmgreen, P., & Lawrence, P. A. (1991). Avoidances, gratifications, and consumption of theatrical movies: The rest of the story. In B. A. Austin (Ed.), *Current research in film. Audiences, economics and law* (Vol. 5, pp. 39–55). Norwood, NJ: Ablex.

Palmgreen, P., & Rayburn, J. D., II (1982). Gratifications sought and media exposure: An expectancy value model. *Communication Research*, *9*(4), 561–580.

Palmgreen, P., & Rayburn, J. D., II (1985). An expectancy-value approach to media gratifications. In K. E. Rosengren, L. A. Wenner, & P. Palmgreen (Eds.), *Media gratifications research. Current perspectives* (pp. 61–72). Beverly Hills, CA: Sage.

Pasek, J., & Krosnick, J. A. (2010). Optimizing survey questionnaire design in political science: Insights from psychology. In J. Leighley (Ed.), *The Oxford handbook of American elections and political behavior* (pp. 27–50). Oxford, UK: Oxford University Press.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, *46*(3), 598–609.

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic Press.

Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah, NJ: Erlbaum.

Paulhus, D. L., & Trapnell, P. D. (2008). Self-presentation of personality: An agency-communion framework. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality. Theory and research* (3rd ed., pp. 492–517). New York, NY: Guilford Press.

Payne, B. K., & Gawronski, B. (2010). A history of implicit social cognition: Where is it coming from? Where is it now? Where is it going? In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 1–15). New York, NY: Guilford Press.

Perloff, R. M. (1999). The third-person effect: A critical review and synthesis. *Media Psychology*, *1*(4), 353–378.

Peterson, R. A. (2001). On the use of college students in social science research: Insights from a second-order meta-analysis. *Journal of Consumer Research*, *28*(3), 450–461.

Petty, R. E., & Briñol, P. (2006). A metacognitive approach to 'implicit' and 'explicit' evaluations: Comment on Gawronski and Bodenhausen (2006). *Psychological Bulletin*, *132*(5), 740–744.

Petty, R. E., Briñol, P., & DeMarree, K. G. (2007). The Meta-Cognitive Model (MCM) of attitudes: Implications for attitude measurement, change, and strength. *Social Cognition*, *25*(5), 657–686.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change.* New York, NY: Springer.

Petty, R. E., & Krosnick, J. A. (Eds.). (1995). *Attitude strength: Antecedents and consequences*. Hillsdale, NJ: Erlbaum.

Piaget, J. (1926). *La représentation du monde chez l'enfant* [The child's conception of the world]. Paris: F. Alcan.

Piaget, J. (1937). *La construction du réel chez l'enfant* [The construction of reality in the child]. Neuchâtel, CH: Delachaux et Niestlé.

Plucker, J. A., Holden, J., & Neustadter, D. (2008). The criterion problem and creativity in film: Psychometric characteristics of various measures. *Psychology of Aesthetics, Creativity, and the Arts*, *2*(4), 190–196.

Plucker, J. A., Kaufman, J. C., Temple, J. S., & Qian, M. (2009). Do experts and novices evaluate movies the same way? *Psychology & Marketing*, *26*(5), 470–478.

Potter, W. J., Pashupati, K., Pekurny, R. G., Hoffman, E., & Davis, K. (2002). Perceptions of television: A schema. *Media Psychology*, *4*(1), 27–50.

Potts, R., Dedmon, A., & Halford, J. (1996). Sensation seeking, television viewing motives, and home television viewing patterns. *Personality and Individual Differences*, *21*(6), 1081–1084.

Prag, J., & Casavant, J. (1994). An empirical study of the determinants of revenues and marketing expenditures in the motion picture industry. *Journal of Cultural Economics*, *18*(3), 217–235.

Pratkanis, A. R. (1989). The cognitive representation of attitudes. In A. R. Pratkanis, S. J. Breckler, & A. G. Greenwald (Eds.), *Attitude structure and function* (pp. 71–98). Hillsdale, NJ: Erlbaum.

Pratkanis, A. R., & Greenwald, A. G. (1989). A sociocognitive model of attitude structure and function. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 22, pp. 245–285). San Diego, CA: Academic Press.

Prommer, E. (1999). *Kinobesuch im Lebenslauf: Eine historische und medienbiographische Studie* [Cinema-going in the life career: A historical and media-biographical study]. Konstanz: UVK Medien.

Raman, N. V., Chattopadhyay, P., & Hoyer, W. D. (1995). Do consumers seek emotional situations: The need for emotion scale. In F. R. Kardes & M. Sujan (Eds.), *Advances in consumer research* (Vol. 22, pp. 537–542). Provo, UT: Association for Consumer Research.

Raney, A. A. (2004). Expanding disposition theory: Reconsidering character liking, moral evaluations, and enjoyment. *Communication Theory*, *14*(4), 348–369.

Raney, A. A. (2006). The psychology of disposition-based theories of media enjoyment. In J. Bryant & P. Vorderer (Eds.), *Psychology of entertainment* (pp. 137–150). Mahwah, NJ: Erlbaum.

Raney, A. A., & Bryant, J. (2002). Moral judgment and crime drama: An integrated theory of enjoyment. *Journal of Communication*, *52*(2), 402–415.

Raykov, T. (2001). Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. *British Journal of Mathematical and Statistical Psychology*, *54*(2), 315–323.

Reid, S. A., & Hogg, M. A. (2005). A self-categorization explanation for the third-person effect. *Human Communication Research*, *31*(1), 129–161.

Reinstein, D. A., & Snyder, C. M. (2005). The influence of expert reviews on consumer demand for experience goods: A case study of movie critics. *Journal of Industrial Economics*, *53*(1), 27–51.

Rentfrow, P. J., Goldberg, L. R., & Zilca, R. (2011). Listening, watching, and reading: The structure and correlates of entertainment preferences. *Journal of Personality*, *79*(2), 223–258.

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology*, *84*(6), 1236–1256.

Robinson, D. C. (1975). A scale for measuring film analyticity. *Communication Research*, *2*(4), 396–411.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York, NY: Free Press.

Rosenberg, M. J. (1956). Cognitive structure and attitudinal affect. *The Journal of Abnormal and Social Psychology*, *53*(3), 367–372.

Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, & J. W. Brehm (Eds.), *Attitude organization and change. An analysis of consistency among attitude components* (pp. 1–14). New Haven, CT: Yale University Press.

Rössler, P. (1997a). Die Filmberichterstattung in der deutschen Tagespresse. FiT 94 – eine Inhaltsanalyse der Beiträge zum Thema Kino in den Printmedien [The coverage of movies in the German daily press: FiT 94 - a content analysis on the importance of the cinema in print media]. *Publizistik*, *42*(2), 181–204.

Rössler, P. (1997b). Filmkritiker und Publikum: Diskrepanzen und Übereinstimmungen. Ergebnisse einer Befragung von Filmrezensenten und Kinogängern [Film critics and the public: discrepancies and agreements: Results of a survey of film critics and moviegoers]. *Media Perspektiven*, *28*(3), 133–140.

Rössler, P. (1997c). Service statt Kritik? Die Zielgruppe Kinopublikum und die Filmberichterstattung in der deutschen Tagespresse [Service instead of critique? Film audience and film reviews in German daily newspapers]. In H. Scherer & H.-B. Brosius (Eds.), *Zielgruppen, Publikumssegmente, Nutzergruppen. Beiträge aus der Rezeptionsforschung* [Target groups, segments of the general public, user groups: Articles from reception research] (pp. 29–57). München: Reinhard Fischer.

Roth, M., & Herzberg, P. Y. (2004). A validation and psychometric examination of the Arnett Inventory of Sensation Seeking (AISS) in German adolescents. *European Journal of Psychological Assessment*, *20*(3), 205–214.

Rubin, A. M. (2009). Uses-and-gratifications perspective on media effects. In J. Bryant & M. B. Oliver (Eds.), *Media effects. Advances in theory and research* (3rd ed., pp. 165–184). New York, NY: Routledge.

Ruggiero, T. E. (2000). Uses and gratifications theory in the 21st century. *Mass Communication and Society*, *3*(1), 3–37.

Rusch, G. (1987). Cognition, media use, genres: Socio-psychological aspects of media and genres: TV and TV-genres in the Federal Republic of Germany. *Poetics*, *16*(5), 431–469.

Rydell, R. J., & McConnell, A. R. (2006). Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology*, *91*(6), 995–1008.

Sarason, I. G., Smith, R. E., & Diener, E. (1975). Personality research: Components of variance attributable to the person and the situation. *Journal of Personality and Social Psychology*, *32*(2), 199–204.

Sawhney, M. S., & Eliashberg, J. (1996). A parsimonious model for forecasting gross box-office revenues of motion pictures. *Marketing Science*, *15*(2), 113-131.

Scheele, B. (Ed.). (1992). *Struktur-Lege-Verfahren als Dialog-Konsens-Methodik: Ein Zwischenfazit zur Forschungsentwicklung bei der rekonstruktiven Erhebung subjektiver Theorien* [Structure-formation techniques as a dialogue–consensus methodology. Reconstruction of subjective theories: An interim balance of research]. Münster: Aschendorff.

Scheele, B., & Groeben, N. (1988). *Dialog-Konsens-Methoden zur Rekonstruktion Subjektiver Theorien: Die Heidelberger Struktur-Lege-Technik (SLT), konsensuale Ziel-Mittel-Argumentation und kommunikative Flussdiagramm-Beschreibung von Handlungen* [Dialogue-consensus methods for the reconstruction of subjective theories]. Tübingen: Francke.

Schenk, M., Büchner, B., & Rössler, P. (1986). TV-Programmvergleich. Ein Test neuer Ansätze in der Publikumsforschung [Comparison of television programs: A test of new approaches in public research]. *Rundfunk und Fernsehen*, *34*(1), 73–86.

Schenk, M., & Rössler, P. (1987). "Dallas" und "Schwarzwaldklinik". Ein Programmvergleich von Seifenopern im deutschen Fernsehen [*Dallas* and *Black Forest Hospital*. A comparison of soap operas in German TV]. *Rundfunk und Fernsehen*, *35*(2), 218–228.

Scherer, K. R. (1998). Emotionsprozesse im Medienkontext: Forschungsillustrationen und Zukunftsperspektiven [Emotion processes in media contexts: Research examples and future perspectives]. *Medienpsychologie*, *10*(4), 276–293.

Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.). (2001). *Appraisal processes in emotion: Theory, methods, research*. Oxford: Oxford University Press.

Schierman, M. J., & Rowland, G. L. (1985). Sensation-seeking and selection of entertainment. *Personality and Individual Differences*, *6*(5), 599–603.

Schmitt, M. (2004). Persönlichkeitspsychologische Grundlagen [Personality psychology foundations]. In R. Mangold, P. Vorderer, & G. Bente (Eds.), *Lehrbuch der Medienpsychologie* [Textbook of media psychology] (pp. 151–173). Göttingen: Hogrefe.

Schmitz, H., Knobloch, S., & Vorderer, P. (1999). Gutes Gefühl statt guter Geschmack: Die Bewertung von Spiel- und Fernsehfilmen durch die Zuschauer [Good feeling instead of good taste: Evaluation of feature films and television films by viewers]. In J. von Gottberg, L. Mikos, & D. Wiedemann (Eds.), *Mattscheibe oder Bildschirm. Ästhetik des Fernsehens* [Box or screen: Aesthetics of television] (pp. 157–179). Berlin: Vistas.

Schwan, S. (1995). "Love or crime or something else?" Schematische Wissensstrukturen und Filmrezeption ['Love or crime or something else?' Schematic knowledge structures and film recipience]. *Rundfunk und Fernsehen*, *43*(1), 26–40.

Schwan, S. (2001). *Filmverstehen und Alltagserfahrung: Grundzüge einer kognitiven Psychologie des Mediums Film* [Film comprehension and everday experience]. Wiesbaden: DUV.

Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition*, *25*(5), 638–656.

Schwarz, N., & Bohner, G. (2001). The construction of attitudes. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology. Intraindividual processes* (pp. 436–457). Oxford, UK: Blackwell.

Schweizer, K. (2010). Some guidelines concerning the modeling of traits and abilities in test construction. *European Journal of Psychological Assessment*, *26*(1), 1–2.

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, *51*(3), 515–530.

Shavelson, R. J., Webb, N. M., & Rowley, G. L. (1989). Generalizability theory. *American Psychologist*, *44*(6), 922–932.

Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change*. New Haven, CT: Yale University Press.

Silvia, P. J., & Berg, C. (2011). Finding movies interesting: How appraisals and expertise influence the aesthetic experience of film. *Empirical Studies of the Arts*, *29*(1), 73–88.

Simonton, D. K. (2002). Collaborative aesthetics in the feature film: Cinematic components predicting the differential impact of 2,323 Oscar-nominated movies. *Empirical Studies of the Arts*, *20*(2), 115–125.

Simonton, D. K. (2004). Group artistic creativity: Creative clusters and cinematic success in feature films. *Journal of Applied Social Psychology*, *34*(7), 1494–1520.

Simonton, D. K. (2005a). Cinematic creativity and production budgets: Does money make the movie? *Journal of Creative Behavior*, *39*(1), 1–15.

Simonton, D. K. (2005b). Film as art versus film as business: Differential correlates of screenplay characteristics. *Empirical Studies of the Arts*, *23*(2), 93–117.

Simonton, D. K. (2007a). Cinema composers: Career trajectories for creative productivity in film music. *Psychology of Aesthetics, Creativity, and the Arts*, *1*(3), 160–169.

Simonton, D. K. (2007b). Film music: Are award-winning scores and songs heard in successful motion pictures? *Psychology of Aesthetics, Creativity, and the Arts*, *1*(2), 53–60.

Simonton, D. K. (2007c). Is bad art the opposite of good art? Positive versus negative cinematic assessments of 877 feature films. *Empirical Studies of the Arts*, *25*(2), 143–161.

Simonton, D. K. (2009a). Cinematic success, aesthetics, and economics: An exploratory recursive model. *Psychology of Aesthetics, Creativity, and the Arts*, *3*(3), 128–138.

Simonton, D. K. (2009b). Controversial and volatile flicks: Concurrent dissension and temporal instability in film critic assessments. *Creativity Research Journal*, *21*(4), 311–318.

Slaney, K. L., & Maraun, M. D. (2008). A proposed framework for conducting data-based test analysis. *Psychological Methods*, *13*(4), 376–390.

Slater, M. D. (2003). Alienation, aggression, and sensation seeking as predictors of adolescent use of violent film, computer, and website content. *Journal of Communication*, *53*(1), 105–121.

Smith, E. R. (2009). Distributed connectionist models in social psychology. *Social and Personality Psychology Compass*, *3*(1), 64–76.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*(2), 108–131.

Smith, E. R., & Queller, S. (2001). Mental representations. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology. Intraindividual processes* (pp. 111–133). Oxford, UK: Blackwell.

Smith, M. B., Bruner, J. S., & White, R. W. (1956). *Opinions and personality*. Oxford, UK: John Wiley & Sons.

Smith, S. P., & Smith, V. K. (1986). Successful movies: A preliminary empirical analysis. *Applied Economics*, *18*(5), 501–507.

Snoek, J., & Bouwman, H. (1995). Wat kinderen leuk vinden aan tekenfilms [What children like about cartoons]. *Massacommunicatie*, *22*, 283–296.

Snyder, M. (1987). *Public appearances, private realities: The psychology of self-monitoring*. New York, NY: Freeman.

Sochay, S. (1994). Predicting the performance of motion pictures. *Journal of Media Economics*, *7*(4), 1–20.

Soto-Sanfiel, M. T., & Vorderer, P. (Eds.). (2011). Entertainment=Emotion [Special issue]. *Journal of Media Psychology: Theories, Methods, and Applications*, *23*(1).

Spitzenorganisation der Filmindustrie e. V. (2011). *Schlüsseldaten Filmwirtschaft 2010*. Retrieved from http://www.spio.de/index.asp?SeitID=3&TID=3

Starzyk, K. B., Fabrigar, L. R., Soryal, A. S., & Fanning, J. J. (2009). A painful reminder: The role of level and salience of attitude importance in cognitive dissonance. *Personality and Social Psychology Bulletin*, *35*(1), 126–137.

Stegert, G. (1993). *Filme rezensieren in Presse, Radio und Fernsehen* [Reviewing films in print, radio, and TV]. München: TR-Verlagsunion.

Stephenson, M. T., Velez, L. F., Chalela, P., Ramirez, A., & Hoyle, R. H. (2007). The reliability and validity of the Brief Sensation Seeking Scale (BSSS-8) with young adult Latino workers: Implications for tobacco and alcohol disparity research. *Addiction*, *102*(Suppl2), 79–91.

Steyer, R. (1987). Konsistenz und Spezifität: Definition zweier zentraler Begriffe der Differentiellen Psychologie und ein einfaches Modell zu ihrer Identifikation [Consistency and specificity: Definition of 2 central concepts in differential psychology, and a simple model for their identification]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *8*(4), 245–258.

Steyer, R., Ferring, D., & Schmitt, M. J. (1992). States and traits in psychological assessment. *European Journal of Psychological Assessment*, *8*(2), 79–98.

Steyer, R., & Schmitt, M. J. (1990). The effects of aggregation across and within occasions on consistency, specificity and reliability. *Methodika*, *4*, 58–94.

Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state–trait theory and research in personality and individual differences. *European Journal of Personality*, *13*(5), 389–408.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*(3), 220–247.

Strizhakova, Y., & Krcmar, M. (2007). Mood management and video rental choices. *Media Psychology*, *10*(1), 91–112.

Sturm, H. (1984). Wahrnehmung und Fernsehen: Die fehlende Halbsekunde. Plädoyer für eine zuschauerfreundliche Mediendramaturgie [Perception and television: The missing half second. Plea for a viewer-friendly media dramaturgy]. *Media Perspektiven*, *15*(1), 58–65.

Suckfüll, M. (1997). *Film erleben: Narrative Strukturen und physiologische Prozesse - "Das Piano" von Jane Campion* [Experiencing film: Narrative structures and physiological processes: The Piano by Jane Campion]. Berlin: edition sigma.

Suckfüll, M. (2004). *Rezeptionsmodalitäten: Ein integratives Konstrukt für die Medienwirkungsforschung* [Modes of reception. An integrative construct for media effects research]. München: Reinhard Fischer.

Suckfüll, M., & Scharkow, M. (2009). Modes of reception for fictional films. *Communications*, *34*(4), 361–384.

Swanson, D. L. (1987). Gratification seeking, media exposure, and audience interpretations: Some directions for research. *Journal of Broadcasting & Electronic Media*, *31*(3), 237–254.

Tamborini, R., & Stiff, J. (1987). Predictors of horror film attendance and appeal: An analysis of the audience for frightening films. *Communication Research*, *14*(4), 415–436.

Tamborini, R., Stiff, J., & Zillmann, D. (1987). Preference for graphic horror featuring male versus female victimization: Personality and past film viewing experiences. *Human Communication Research*, *13*(4), 529–552.

Tanaka, J. S., Panter, A. T., & Winborne, W. C. (1988). Dimensions of the need for cognition: Subscales and gender differences. *Multivariate Behavioral Research*, *23*(1), 35–50.

Tesser, A., Millar, K. U., & Wu, C.-H. (1988). On the perceived functions of movies. *Journal of Psychology: Interdisciplinary and Applied*, *122*(5), 441–449.

Thurstone, L. L. (1930). A scale for measuring attitude toward the movies. *Journal of Educational Research*, *22*, 89–94.

Tormala, Z. L., & Petty, R. E. (2001). On-line versus memory-based processing: The role of "Need to Evaluate" in person perception. *Personality and Social Psychology Bulletin*, *27*(12), 1599–1612.

Trepte, S. (1999). Forschungsstand der Medienpsychologie. *Zeitschrift für Medienpsychologie*, *11*(3), 200–218.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–69.

Valkenburg, P. M., & Cantor, J. (2000). Children's likes and dislikes of entertainment programs. In D. Zillmann & P. Vorderer (Eds.), *Media entertainment: The psychology of its appeal* (pp. 135–152). Mahwah, NJ: Erlbaum.

Valkenburg, P. M., & Janssen, S. C. (1999). What do children value in entertainment programs? A cross-cultural investigation. *Journal of Communication*, *49*(2), 3–21.

Van Deth, J. W. (1990). Interest in politics. In M. K. Jennings & J. W. van Deth (Eds.), *Continuities in political action. A longitudinal study of political orientations in three Western democracies* (Vol. 5, pp. 275–312). Berlin: de Gruyter.

Venkatraman, M. P. (1989). Opinion leaders, adopters, and communicative adopters: A role analysis. *Psychology & Marketing*, *6*(1), 51–68.

Visser, P. S., Bizer, G. Y., & Krosnick, J. A. (2006). Exploring the latent structure of strength-related attitude attributes. In M. P. Zanna (Ed.). *Advances in experimental social psychology, Vol. 38* (pp. 1–67). San Diego, CA: Elsevier Academic Press.

Visser, P. S., & Krosnick, J. A. (1998). Development of attitude strength over the life cycle: Surge and decline. *Journal of Personality and Social Psychology*, *75*(6), 1389–1410.

Vogel, I. C. (2007). *Das Sad-Film-Paradoxon: Ein theoretischer und empirischer Beitrag zum Anreiz trauriger Filme* [The sad film paradox. A theoretical and empirical contribution to the appeal of sad movies]. Aachen: Shaker.

Vorderer, P. (1992). *Fernsehen als Handlung: Fernsehfilmrezeption aus motivationspsychologischer Perspektive* [Television as action: TV film reception from a motivational psychological perspective]. Berlin: edition sigma.

Vorderer, P. (1994). Was macht die Rezeption von Filmen spannend? [What makes the reception of movies suspenseful?]. *Zeitschrift für Medienpsychologie*, *6*(2), 103–109.

Vorderer, P. (2011). What's next? Remarks on the current vitalization of entertainment theory. *Journal of Media Psychology: Theories, Methods, and Applications*, *23*(1), 60–63.

Vorderer, P., & Bube, H. (1996). Ende gut, alles gut? Eine empirische Studie über den Einfluß von emphatischem Streß und Filmausgang auf die Befindlichkeit von Rezipienten und deren Bewertung des Films. [All's well that ends well? An empirical study on the influence of empathic stress and film outcome on viewers' feelings and on their evaluation of the film]. *Medienpsychologie*, *8*(2), 128–143.

Vorderer, P., Knobloch, S., & Schramm, H. (2001). Does entertainment suffer from interactivity? The impact of watching an interactive TV movie on viewers' experience of entertainment. *Media Psychology*, *3*(4), 343–363.

Wallace, W. T., Seigerman, A., & Holbrook, M. B. (1993). The role of actors and actresses in the success of films: How much is a movie star worth? *Journal of Cultural Economics*, *17*(1), 1–27.

Wanderer, J. J. (1970). In defense of popular taste: Film ratings among professionals and lay audiences. *American Journal of Sociology*, *76*(2), 262–272.

Wartella, E., & Reeves, B. (1985). Historical trends in research on children and the media: 1900–1960. *Journal of Communication*, *35*(2), 118–133.

Waterman, A. S. (1993). Two conceptions of happiness: Contrasts of personal expressiveness (eudaimonia) and hedonic enjoyment. *Journal of Personality and Social Psychology*, *64*(4), 678–691.

Weaver, J. B., III. (1991). Exploring the links between personality and media preferences. *Personality and Individual Differences*, *12*(12), 1293–1299.

Weaver, J. B., III. Brosius, H.-B., & Mundorf, N. (1993). Personality and movie preferences: A comparison of American and German audiences. *Personality and Individual Differences*, *14*(2), 307–315.

Wegener, D. T., Downing, J., Krosnick, J. A., & Petty, R. E. (1995). Measures and manipulations of strength-related properties of attitudes: Current practice and future directions. In R. E. Petty & J. A. Krosnick (Eds.). *Attitude strength: Antecedents and consequences* (pp. 455–487). Hillsdale, NJ: Erlbaum.

Wegener, D. T., & Fabrigar, L. R. (2004). Constructing and evaluating quantitative measures for social psychological research: Conceptual challenges and methodological solutions. In C. Sansone, C. C. Morf, & A. T. Panter (Eds.), *The Sage handbook of methods in social psychology* (pp. 145–172). Thousand Oaks, CA: Sage.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). Thousand Oaks, CA: Sage.

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Erlbaum.

Wilson, T. D., & Hodges, S. D. (1992). Attitudes as temporary constructions. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 37–65). Hillsdale, NJ: Erlbaum.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*(1), 101–126.

Winkler, N., Kroh, M., & Spiess, M. (2006). *Entwicklung einer deutschen Kurzskala zur zweidimensionalen Messung von sozialer Erwünschtheit* [Development of a German short scale for measuring two dimensions of social desirability] (Discussion Papers No. 579). Berlin.

Wirth, W. (2006). Involvement. In J. Bryant & P. Vorderer (Eds.), *Psychology of entertainment* (pp. 199–213). Mahwah, NJ: Erlbaum.

Wolling, J. (2004). Qualitätserwartungen, Qualitätswahrnehmungen und die Nutzung von Fernsehserien: Ein Beitrag zur Theorie und Empirie der subjektiven Qualitätsauswahl von Medienangeboten [Quality expectations, quality perceptions and viewing of television serials: A contribution to the theoretical and empirical study of the subjective selection of media content]. *Publizistik*, *49*(2), 171–193.

Wolling, J. (2009). The effect of subjective quality assessments on media selection. In T. Hartmann (Ed.), *Media choice. A theoretical and empirical overview* (pp. 85–101). New York, NY: Routledge.

Wuss, P. (1992). Der Rote Faden der Filmgeschichten und seine unbewußten Komponenten: Topik-Reihen, Kausal-Ketten und Story-Schemata – drei Ebenen filmischer Narration [The red thread of film stories and its natural components: Topic series, causal chains and story schemes - three levels of film narration]. *montage/av*, *1*(1), 25–35.

Wuss, P. (1999). *Filmanalyse und Psychologie: Strukturen des Films im Wahrnehmungsprozess* [Film analysis and psychology: Structures of film within the process of perception] (2., durchges. und erw. Aufl.). Berlin: edition sigma.

Wuss, P. (2009). *Cinematic narration and its psychological impact: Functions of cognition, emotion and play*. Newcastle, UK: Cambridge Scholars.

Wyatt, R. O., & Badger, D. P. (1984). How reviews affect interest in and evaluation of films. *Journalism Quarterly*, *61*(4), 874–878.

Wyatt, R. O., & Badger, D. P. (1988). What newspaper film critics value in film and film criticism: A national survey. In B. A. Austin (Ed.), *Current research in film. Audiences, economics and law* (Vol. 4, pp. 54–71). Norwood, NJ: Ablex.

Wyatt, R. O., & Badger, D. P. (1990). Effects of information and evaluation in film criticism. *Journalism Quarterly*, *67*(2), 359–368.

Zanna, M. P., & Rempel, J. K. (1988). Attitudes: A new look at an old concept. In D. Bar-Tal & A. W. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 315–334). Cambridge, UK: Cambridge University Press.

Zillmann, D. (1988). Mood management through communication choices. *American Behavioral Scientist*, *31*(3), 327–340.

Zillmann, D. (2000). Mood management in the context of selective exposure theory. In M. E. Roloff (Ed.), *Communication yearbook* (Vol. 23, pp. 103–123). Thousand Oaks, CA: Sage.

Zillmann, D., & Cantor, J. R. (1977). Affective responses to the emotions of a protagonist. *Journal of Experimental Social Psychology*, *13*(2), 155–165.

Zillmann, D., Hay, T. A., & Bryant, J. (1975). The effect of suspense and its resolution on the appreciation of dramatic presentations. *Journal of Research in Personality*, *9*(4), 307–323.

Zillmann, D., & Vorderer, P. (Eds.). (2000). *Media entertainment: The psychology of its appeal*. Mahwah, NJ: Erlbaum.

Zuckerman, M. (1979). *Sensation seeking: Beyond the optimal level of arousal*. Hillsdale, NJ: Erlbaum.

Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking*. Cambridge, UK: Cambridge University Press.

Zuckerman, M., & Litle, P. (1986). Personality and curiosity about morbid and sexual events. *Personality and Individual Differences*, *7*(1), 49–56.

Zuwerink, J. R., & Devine, P. G. (1996). Attitude importance and resistance to persuasion: It's not just the thought that counts. *Journal of Personality and Social Psychology*, *70*(5), 931–944.

Appendix A

Table A1

Criteria Terms Used for Modified Structure Formation Technique

Criteria terms in alphabetical order (*N* = 274)			
abenteuerlich	emotionale Ansprache	Kernaussage(n)	Requisiten
Ablenkung	emotionale Momente	kindergeeignet	Romantik
Abstraktheit	emotionaler Effekt (berührt z. B.)	Klamauk	roter Faden
Abwechslung	Emotionalität	Klarheit (über Personen, Orte)	Schauspielerische Leistung
Abwechslung vom Alltag	Emotionen	Klasse	schauspielerisches Können
Abwechslungsreichtum	Entspannungsfaktor	Klassiker	schauspielerisches Talent
Action	ergreifend	Komik	Schnitt
Aktualität	Ernsthaftigkeit	Komplexität	Schöne Bilder
Altersspektrum der Zielgruppe	Erotik	Kosten	schöne Landschaften
Animation	erschütternd	Kosten-Nutzen	schwarz-weiß/ Farbfilm
Anregung	Existierende Buchvorlage	Kostüme	Setting
Anregung zum Nachdenken	Farbgebung	Kreativität (z. B. K. der Machart, K. der Geschichte, K. der Handlung)	Sexappeal
Anschaffungskosten	Filmfehler (z. B. F. in Story, F. in Plot)	Kritik	Sinn
Anspruch (z. B. künstlerischer A., intellektueller A.)	Filmische Ideen	kritische Auseinandersetzung mit bestimmten Themen	Situationskomik
Anteil an Spezialeffekten	Filmische Qualität	Kulisse	Soundtrack
Art der Geschichte	Gefühle	Kultur	sozialkritisch
Art der Verfilmung	Gefühlsbetonung	Kunst	Spannung
Art des Films	Gefühlsbreite	künstlerische Gestaltung	Spaß
Ästhetik	gefühlvoll	Kurzweiligkeit	Special-Effects
Attraktivität der Darsteller	Gehalt des Inhalts	Länge	Sprache
Aufklärung	geistreich	lehrreich	stimmige Abfolge von Szenen
Aufmachung	Gelungene Effekte	Licht	Stimmung
Aufnahmen	Genre	Liebe	Story
Aufwand	gesellschaftliche Relevanz	Liebesszenen	Synchronisation
aus dem Leben gegriffen	Gesellschaftskritik	Logik (z. B. Schlüssigkeit)	Szenenbild
Ausgang des Films	Gestaltung	lustige Dialoge	Technik
Aussage	Gewalt	Lustigkeit	Technische Gestaltung
Aussehen der Darsteller	Glaubwürdigkeit	Machart	technische Umsetzung

(table continues)

Table A1 (*continued*)

Außergewöhnlichkeit	Grafik	Maske	technischer Aufwand
Aussprache der Darsteller	große Gruppe ansprechen	mitfühlend	Tempo
Ausstattung	Grusel	mitreißend	Thema
Auszeichnungen (z. B. Oscars)	gute Kulisse/Kostüme etc.	Montage	Thema muss relevant sein
authentische Darsteller	gute Schauspieler	Moral	Themaauswahl
Authentizität (z. B. Story, Thema)	gute Story	Musik	Themenaufarbeitung
Bekanntheit	gute Umsetzung	Musikalische Untermalung	Tiefgang
Bekanntheit der Schauspieler	Handlung logisch nachvollziehbar	Nachfrage	tolle Szenen
Berühmtheit der Darsteller	Handlungsaufbau	Nachvollziehbare Gefühle	Ton (z. B. Qualität)
berührend	Handlungsstränge	nachvollziehbarer Humor	Tragik
beschreibend oder erklärend?	Handlungstiefe	nachvollziehbarer Szeneneinsatz	Trauerfaktor
Besetzung (z. B. zu den Rollen passende B., Stars)	Happy End	Neuheit	Traurigkeit
bewegend	Häufigkeit des Ortswechsels	Neuheit des Themas	Überraschung
bildend	Hauptdarsteller	optische Wirkung	Überschaubarkeit
Bilder	Herz	Originalität	Überzeugende Darsteller
Bildqualität	Hintergrund	passende Filmmusik	Überzeugende Darstellung
Bildsprache	historischer Hintergrund	persönliche Verbundenheit mit Thema	Überzeugende Schauspielleistung
biographisch - erfunden	Humor	persönlicher Bezug	Umsetzung
Botschaft	Ideen	Phantasie	Umsetzung der Geschichte
Budget	Identifikation mit den Protagonisten	Philosophischer Hintergrund	Umsetzung der Literaturvorlage
Charaktere	Individualität	Plot/Handlung	Umsetzung des Drehbuchs
Darsteller	individueller Geschmack	Pointe (Ende)	Unterhaltung
Darstellung	Informationsgehalt	präferierte Schauspieler	Verarbeitung des Thema
Der Buchvorlage entsprechend	Inhalt einer Moral	Produktion	Verpackung
detaillierte Darstellung von Personen, Räumen	inhaltliche Umsetzung	Produktionsaufwand	Verständlichkeit
Dialoge	Inhaltliches Interesse geweckt	Produktionskosten	vorbehaltlos
Dialogqualität	Innovation	Produktionsland	Vorhersagbarkeit des Ausgangs
Digitale Nachbearbeitung	Innovative Story	Produktionsort	Vorhersehbarkeit
Dramatik	Inspiration	Produktionszeitraum	wahrer Hintergrund
Dramaturgie	Intellekt	Qualität	Wahrheitsgehalt (z. B. historische W.)
Drehbuch (z. B. Stimmigkeit des D., Qualität des D.)	Intention	Qualität der Aufnahme	Wahrscheinlichkeit

(*table continues*)

Drehbuchauswahl	interessant	Qualität der Filmmusik	Wendungen
Drehdauer	interessante Handlung	Qualität der Special effects	Werbung
Drehkosten	interessante Story	Qualität des Materials	Wertevermittlung
Drehorte (z. B. passende D.)	Interesse	reale Szenen	Wissenserweiterung
Effekte	investierte Kosten	Realismus	Wissensvermittlung
Einfallsreichtum	Investiertes Geld	realistische Darstellungen	Witzigkeit
eingesetzte Technik (Echtheit)	Kamera	Realitätsnähe	Zeitlosigkeit
Eintauchen in eine fremde Welt	Kameraeinstellung	realitätsnahe Actionszenen	Zündstoffthemen
Einzigartigkeit	Kameraführung	Regie/Regieführung	
Eloquenz	keine einseitige Handlung	Regisseur (z. B. Bekanntheit)	

Figure A1. Excerpts from the photo-documentation of the modified structure formation technique: Examples for (A) how single cards were subsumed under a superior category; (B) subcategories were subsumed under main categories; and (C) final main categories and subcategories were visualized on a flip chart. Note that the three examples are from three different groups.
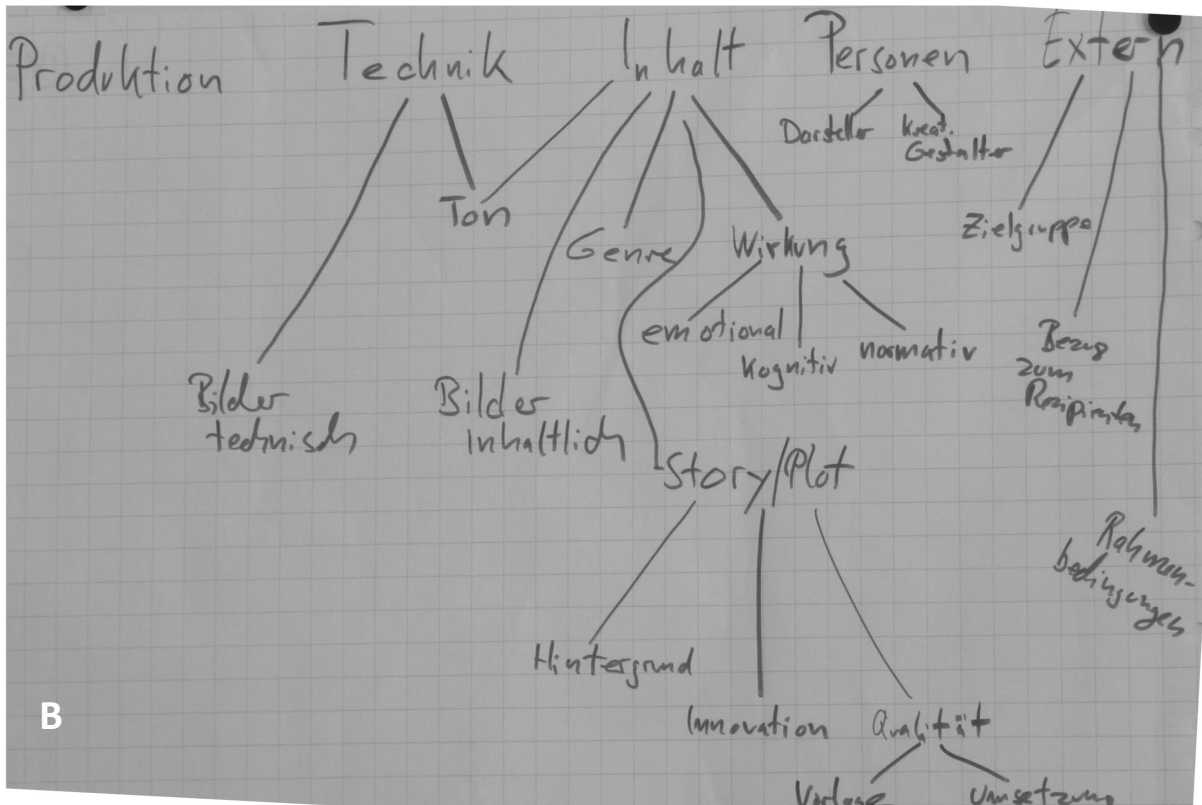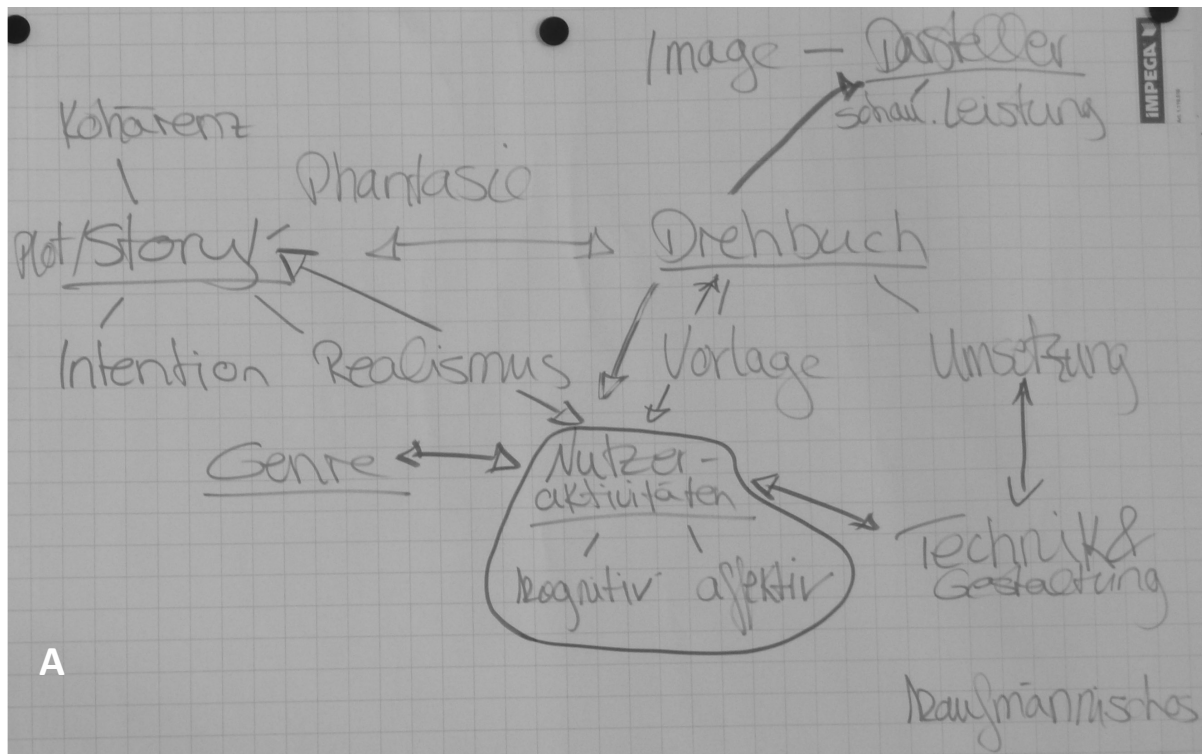
Figure A2. Excerpts from the photo-documentation of the modified structure formation technique: flip charts of (A) Group B and (B) Group C. See Figure A1 for flip chart of Group A.

Table A2

German Coding Scheme for Analyzing Results of the Modified Structure Formation Technique

Nr.	Kategorie	Code	Beschreibung
1.0	Genre	1	
1.1	Genre_Action	11	
2.0	Dauer des Films	2	Länge des Films
3.0	Regie	3	umfasst Regie, Regieführung, Regisseur/in etc.
4.0	Rezipient_Sonstige	4	wenn keiner oder mehreren Subkategorien zu zu ordnen (z.B. Unterhaltung, persönlicher Bezug, persönliche Relevanz kann sowohl affektiv, kognitiv als auch motivational sein)
4.1	Rezipient_affektiv	41	Emotionen, Gefühle, Stimmungen, etc.
4.2	Rezipient_kognitiv	42	Wissen, Lernen, Wahrnehmen, Erinnern, Nachdenken etc.
4.3	Rezipient_Motive_Sonstige	43	Erwartungen, Präferenzen (präf. Genres, Schauspieler, Regisseure etc., Motive etc.; wenn nicht zu 4.3.1 oder 4.3.2 zu zu ordnen
4.3.1	Rezipient_Motive_Ablenkung	431	Eskapismus, Ablenkung, Ablenkung vom Alltag
4.3.2	Rezipient_Motive_Interesse	432	Interesse
5.0	werkzentriert-inhaltlich_Sonstiges	5	nicht näher spezifiziert
5.1	werkzentriert-inhaltlich_Story/Plot	51	
5.1.1	werkzentriert-inhaltlich_Story/Plot_Umsetzung	511	
5.1.2	werkzentriert-inhaltlich_Story/Plot_Kohärenz	512	plausibel, logisch, fehlerfrei
5.1.3	werkzentriert-inhaltlich_Story/Plot_Botschaft	513	Sinn, Moral, Zweck, Intention, Botschaft, Absicht
5.1.4	werkzentriert-inhaltlich_Story/Plot_Innovation	514	Story/Plot sind kreativ, originell, einzigartig, innovativ, neu, einfallsreich, außergewöhnlich, individuell, phantasievoll etc.
5.1.5	werkzentriert-inhaltlich_Story/Plot_Ende	515	
5.2	werkzentriert-inhaltlich_Thema	52	
5.3	werkzentriert-inhaltlich_Anspruch	53	
5.5	werkzentriert-inhaltlich_Hintergrund	55	philosophischer, historischer, biografischer, detailgetreuer, aktueller, realistischer, realitätsnaher Hintergrund etc.
5.6	werkzentriert-inhaltlich_Vorlage	56	Entsprechung, Umsetzung etc. von Literaturvorlage/Drehbuch
5.7	werkzentriert-inhaltlich_Dialoge	57	
5.8	werkzentriert-inhaltlich_Dramaturgie	58	
5.8.1	werkzentriert-inhaltlich_Dramaturgie_Abwechslung	581	
5.9	werkzentriert-inhaltlich_Charaktere	59	
6.0	werkzentriert-formal_allgemein	6	z.B. Materialqualität

(*table continues*)

Nr.	Kategorie	Code	Beschreibung
6.1	werkzentriert-formal_Technik	61	
6.2	werkzentriert-formal_Innovation	62	
6.2.1	werkzentriert-formal_Technik_Umsetzung	621	
6.2.2	werkzentriert-formal_Technik_Ton	622	
6.2.2.1	werkzentriert-formal_Technik_Ton_Synchronisation	6221	
6.2.3	werkzentriert-formal_Technik_Gestaltung	623	
6.2.4	werkzentriert-formal_Technik_Schnitt/Montage	624	
6.2.4.1	werkzentriert-formal_Technik_Schnitt/Montage_Tempo	6241	
6.2.5	werkzentriert-formal_Technik_Kamera	625	
6.2.5.1	werkzentriert-formal_Technik_Kamera_Licht	6251	
6.2.6	werkzentriert-formal_Technik_Effekte	626	
6.2.7	werkzentriert-formal_Technik_Aufwand	627	
6.3	werkzentriert-formal_Ästhetik	63	
6.4	werkzentriert-formal_Ton/Musik	64	
6.5	werkzentriert-formal_Ausstattung	65	Ausstattung des Film, künstlerische Gestaltung
6.5.1	werkzentriert-formal_Ausstattung_Maske	651	
6.5.2	werkzentriert-formal_Ausstattung_Requisiten	652	
6.5.3	werkzentriert-formal_Ausstattung_Kulissen	653	
6.5.4	werkzentriert-formal_Ausstattung_Kostüme	654	
7.0	Darsteller_Sonstiges	7	
7.1	Darsteller_Attribute	71	
7.1.1	Darsteller_Attribute_Bekanntheit	711	
7.1.2	Darsteller_Attribute_Berühmtheit	712	
7.1.3	Darsteller_Attribute_Attraktivität	713	
7.1.4	Darsteller_Attribute_Aussehen	714	
7.2	Darsteller_Kompetenz	72	
7.2.1	Darsteller_Kompetenz_Talent	721	
7.2.2	Darsteller_Kompetenz_Können	722	
7.3	Darsteller_Performanz	73	falls nicht zu Subkategorien zu zu ordnen oder mehr als einer Subkategorie zu zu ordnen
7.3.1	Darsteller_Performanz_überzeugend	731	
7.3.2	Darsteller_Performanz_Aussprache	732	
7.3.3	Darsteller_Performanz_Leistung	733	
7.3.4	Darsteller_Performanz_Qualität	734	
7.3.5	Darsteller_Performanz_Authentizität	735	
7.4	Darsteller_Besetzung	74	
8.0	werkperipher_Produktion_Allgemein	8	
8.1	werkperipher_Produktion_Kosten	81	umfasst Drehkosten, Produktionskosten, Produktionsbudget
8.2	werkperipher_Produktion_Orte	82	umfasst Drehorte, Produktionsorte, Produktionsländer
8.3	werkperipher_Produktion_Aufwand	83	
8.4	werkperipher_Produktion_Zeit	84	umfasst Drehdauer, Produktionsdauer, Produktionszeitraum, Drehzeitraum

(table continues)

Table A2 (*continued*)

Nr.	Kategorie	Code	Beschreibung
9.0	werkperipher_Werbung/Marketing_Allgemein	9	umfasst Werbespots, Trailer, Bekanntheit des Films, Nachfrage/Angebot, Merchandising, crossmediale Verwertung (DVD, Downloads, TV etc.) u.ä.
9.1	werkperipher_Werbung/Marketing_Zielgruppe	91	umfasst alle Formen und Eigenschaften von Zielgruppen: Kinder, Senioren, Jugendlichen, Milieus, Männer, Frauen, Minoritäten, Majoritäten, Mainstream etc.
10.0	werkperipher_Reviews	10	umfasst alle Formen und Eigenschaften von Filmreviews: Filmkritik, Filmbesprechung, Filmtipp in Presse, Rundfunk, Internet etc.
11.0	werkperipher_Auszeichnungen	110	Preise, Auszeichnungen, Prädikate, Nominierungen etc.
12.0	werkperipher_Machart	12	Machart, Making of..., Umsetzung
13.0	Qualität	13	Qualität
14.0	Sonstiges	14	

Nr.	Kategorie	Code	Beschreibung

Appendix B

Hinweise zum Beantworten der Fragen

Die folgenden Fragen zielen darauf ab, zu erfahren, wie wichtig bestimmte Merkmale von Filmen für Sie persönlich sind, wenn Sie Filme bewerten.
Um die Wichtigkeit von Filmmerkmalen für Ihre eigene Bewertung abzustufen, stehen Ihnen fünf Antwortmöglichkeiten zur Auswahl: *überhaupt nicht wichtig, etwas wichtig, mäßig wichtig, sehr wichtig, äußerst wichtig.*

Für die Merkmale, die Ihnen am allerwichtigsten für Ihre Filmbewertungen erscheinen, wählen Sie bitte die höchste Ausprägung: *äußerst wichtig.*

Wenn Sie ein Merkmal nicht kennen oder kannten oder noch nicht auf die Idee kamen, anhand dieses Merkmals einen Film zu bewerten, geben Sie bitte *überhaupt nicht wichtig an.*

Um Ihnen das spontane Antworten zu erleichtern, sehen Sie sich bitte kurz das folgende Ausfüllbeispiel an anhand der folgenden Fragen.

Beispiel 1
Wenn Sie Autos bewerten, wie wichtig ist Ihnen dabei persönlich...

	überhaupt nicht wichtig	*etwas wichtig*	*mäßig wichtig*	*sehr wichtig*	*äußerst wichtig*
der Kaufpreis eines Autos?	☐	☐	☐	☐	☐

Um ein Auto positiv zu beurteilen, ist es **Person A** sehr wichtig, ein besonders teures Auto zu erwerben. Person A sollte die Antwortmöglichkeit *sehr wichtig* wählen.

Person B ist es sehr wichtig, ein Auto günstig zu erwerben. Auch Person B sollte also die Antwortmöglichkeit *sehr wichtig* wählen.

Person C der Kaufpreis egal. Sie sollte Antwortmöglichkeit *überhaupt nicht wichtig* wählen.

Wenn Sie Autos bewerten, wie wichtig ist Ihnen dabei persönlich...

	überhaupt nicht wichtig	etwas wichtig	mäßig wichtig	sehr wichtig	äußerst wichtig
das Material, aus dem die Sitzpolster eines Autos gemacht sind?	☐	☐	☐	☐	☐

Person A ist es nicht besonders wichtig, aus welchem Material die Sitzpolster sind. Ledersitzbezüge fände Person A gut.

Person A könnte mit *etwas wichtig* oder *mäßig wichtig* abstufen, wie stark sie die Wichtigkeit dieses Merkmals nun findet.

Person B hat sich noch keine Gedanken darüber gemacht haben, aus welchem Material Sitzpolster sein müssten, damit sie ein Auto positiv oder negativ bewertet.

Person B weiß nun durch diese Frage, dass das Material von Sitzpolstern wohl ein Bewertungskriterium sein kann.

Sie hat dieses Kriterium jedoch noch nicht bei der Bewertung von Autos herangezogen, deswegen sollte sie *überhaupt nicht wichtig* auswählen.

Person C ist der Meinung, dass das vom Autotyp abhänge. Bei einem Sportwagen fände Person C Ledersitze *äußerst wichtig*; für einen Kleinwagen wäre es egal, also *überhaupt nicht wichtig*.

Person C sollte nun aus ihren beiden Urteilen einen Mittelwert bilden und diesen angeben; in diesem Fall läge der mittlere Wert bei *mäßig wichtig*. Sie sollte also diese Antwortmöglichkeit auswählen.

Table B1

German Item Wordings for Phase II (After Pretesting)

	Wenn Sie Filme bewerten, wie wichtig ist bzw. sind Ihnen dabei persönlich(,)...
1.	das inhaltliche Thema eines Films?
2.	wie das Thema eines Films umgesetzt wird?
3.	die Geschichte/Handlung eines Films?
4.	dass ein Film Tiefgang hat?
5.	die Komplexität der Handlung eines Films?
6.	dass die Geschichte, die ein Film erzählt, einen philosophischen Hintergrund aufweist? (Philosophische Themen, vor deren Hintergrund Geschichten erzählt werden, können z.B. folgende Fragen beinhalten: Was ist gut? Was ist böse? Was ist Gerechtigkeit? Gibt es einen Gott? Was ist der Sinn des Lebens? Was ist Realität (Sein)? Was ist Traum (Schein)? Was ist die beste politische Ordnung? etc.)
7.	dass die Geschichte, die ein Film erzählt, einen wahren Hintergrund hat (z.B. ein historisches Ereignis aufgreift oder die Lebensgeschichte einer realen Person erzählt)?
8.	dass die Geschichte, die ein Film erzählt, realitätsnah ist?
9.	dass ein Film sich inhaltlich auf aktuelle Themen bezieht?
10.	dass ein Film auf einer literarischen Vorlage (z.B. Roman, Kurzgeschichte, Comic) basiert?
11.	dass ein Film mit seiner literarischen Vorlage (z.B. Roman, Kurzgeschichte, Comic) übereinstimmt?
12.	die Dialoge in einem Film?
13.	dass es Wendungen in der Handlung eines Films gibt?
14.	die Dramaturgie eines Films?
15.	die Charakterisierung der Figuren in einem Film?
16.	dass ein Film frei von gewalthaltigen Szenen ist?
17.	dass ein Film gewalthaltig ist?
18.	dass ein Film erotisch ist?
19.	dass ein Film actionhaltig ist?
20.	dass die Geschichte, die ein Film erzählt, schlüssig ist?
21.	die Glaubwürdigkeit der Geschichte, die ein Film erzählt?
22.	dass ein Film eine sozial-/gesellschaftskritische Botschaft hat?
23.	dass ein Film Erklärungen liefert (d.h. zu erklären versucht, aus welchen Gründen, Personen gehandelt haben bzw. was die Ursachen für die Ereignisse waren)?
24.	dass ein Film beschreibend ist (d.h. eine neutrale, beobachtende Position einnimmt und versucht, zu beschreiben/dokumentieren und nicht zu erklären)?
25.	dass ein Film sich kritisch mit seinem Thema auseinandersetzt?
26.	die Moral der Geschichte, die ein Film erzählt?
27.	dass ein Film eine Haltung einnimmt?
28.	dass ein Film eine Geschichte auf neue Art und Weise erzählt?
29.	dass die Geschichte, die ein Film erzählt, außergewöhnlich ist?
30.	dass die Geschichte, die ein Film erzählt, einzigartig ist?
31.	dass eine unvorhersehbare Wendung am Ende eines Films eintritt?
32.	dass ein Film ein „gutes" Ende (Happy End) hat?
33.	dass ein Film tragisch endet?
34.	dass das Ende eines Films offen bleibt?
35.	der Soundtrack eines Films?
36.	die Tonqualität eines Films?
37.	die Art und Weise wie ein Film geschnitten ist bzw. wie Filmszenen zusammengefügt sind?
38.	die Kameraführung und Kameraeinstellungen eines Films?
39.	die Ausleuchtung/die Lichtführung in einem Film?

(*table continues*)

Table B1 (*continued*)

40.	die Farbgebung in einem Film (z.B. Verwendung von schwarz/weiß, Rot als Signalfarbe etc.)?
41.	die Ausstattung eines Films (z.B. Szenenbild, Kulissen, Requisiten, Maske, Kostüme)?
42.	die Trickelemente in einem Film (z.B. Spezialeffekte wie Feuer, Explosionen, Stunts, Kampfszenen)?
43.	die digitale (Nach-)Bearbeitung eines Films?
44.	die „Echtheit" der eingesetzten Filmtechnik (z.B. Handkamera, Kamerafahrten)?
45.	dass ein Film aufwändig gemacht ist?
46.	die technische Gestaltung eines Films insgesamt?
47.	die schauspielerische Leistung der Darstellerinnen und Darsteller insgesamt?
48.	dass die Rollen/Charaktere in einem Film mit den passenden Darstellerinnen und Darstellern besetzt sind?
49.	dass die Rollen/Charaktere in einem Film untypisch besetzt sind (z.B. mit Actiondarstellern in Charakterrollen)?
50.	dass die Hauptdarstellerinnen und Hauptdarsteller attraktiv sind?
51.	die Bekanntheit der Hauptdarstellerinnen und Hauptdarsteller in einem Film?
52.	die Bekanntheit der Regisseurin oder des Regisseurs eines Films?
53.	dass ein Film sich an ein breites Publikum richtet?
54.	dass ein Film für Kinder geeignet ist?
55.	in welchem Land ein Film produziert wurde?
56.	die Besprechungen eines Films in den „Medien" (z.B. in der Presse, im Rundfunk, auf Filmseiten im Internet)?
57.	die Besprechungen eines Films mit Freunden, Bekannten o.ä.?
58.	die Auszeichnung(en) für einen Film (z.B. Oscar, Golden Globe, Goldene Palme, Goldener Bär, Deutscher Filmpreis, Grimme-Preis)?
59.	wie ein Film beworben wird (z.B. in TV- oder Kino-Trailern, auf Plakaten, in Zeitungs- oder Zeitschriftenanzeigen)?
60.	dass Sie einen Film romantisch finden?
61.	dass Sie einen Film ernsthaft finden?
62.	dass Sie einen Film lustig finden?
63.	dass ein Film unterschiedliche Gefühle bei Ihnen auslöst?
64.	dass ein Film Sie emotional berührt?
65.	dass Sie sich bei einem Film entspannen können?
66.	dass ein Film Sie gruselt?
67.	dass ein Film bei Ihnen eine fröhliche Stimmung auslöst?
68.	dass ein Film bei Ihnen eine traurige Stimmung auslöst?
69.	dass ein Film Sie überrascht?
70.	dass Sie einen Film spannend finden?
71.	dass Sie einen Film ästhetisch finden?
72.	dass Sie einen Film unterhaltsam finden?
73.	dass Sie einen Film mitreißend finden?
74.	dass Sie sich mit einer Hauptfigur eines Films identifizieren können?
75.	dass Sie sich in eine Hauptfigur eines Films einfühlen können?
76.	dass ein Film bei Ihnen Angst auslöst?
77.	dass ein Film frei von Szenen ist, die bei Ihnen Angst auslösen?
78.	dass ein Film bei Ihnen Ekel erregt?
79.	dass ein Film frei von Szenen ist, die bei Ihnen Ekel erregen?
80.	dass ein Film bei Ihnen Wut auslöst?
81.	dass ein Film frei von Szenen ist, die bei Ihnen Wut auslösen?

(*table continues*)

Table B1 (*continued*)

82.	dass ein Film Sie in eine fremde Welt eintauchen lässt?
83.	dass ein Film Sie vom Alltag ablenkt?
84.	dass ein Film Sie zum Handeln anregt (z.B. politisch, sozial, Informationssuche)?
85.	dass ein Film Ihre Fantasie anregt?
86.	dass ein Film Sie intellektuell fordert?
87.	dass ein Film Sie zum Nachdenken anregt?
88.	dass Ihnen ein Film Werte vermittelt?
89.	dass Sie durch das Anschauen eines Films Ihr Wissen erweitert wird?
90.	die Verständlichkeit der Aussage/Botschaft eines Films?
91.	dass Sie einen persönlichen Bezug zu den Inhalten eines Films haben (z.B. Film spielt in Heimatstadt; ähnliche Situation wie im Film dargestellt schon selbst erlebt)?
92.	dass ein Film sich eindeutig einem Filmgenre zuordnen lässt?
93.	dass ein Film als "Filmklassiker" gilt?

Note. German response categories were 0 (*überhaupt nicht wichtig*), 1 (*etwas wichtig*), 2 (*mäßig wichtig*), 3 (*sehr wichtig*), 4 (*äußerst wichtig*); only labels (not numbers) were displayed.

Appendix C

Table C1
Test-half Composition and Descriptive Statistics (N=275)

Test half	Item	t1 M	t1 SD	t2 M	t2 SD
SV1		.71	.95	.83	.96
	sv1	.71	.95	.83	.96
SV2		1.22	.86	1.24	.89
	sv2	1.57	1.14	1.48	1.10
	sv3	.88	.98	1.01	.94
SI1		2.37	1.11	2.32	1.05
	si1	2.37	1.11	2.32	1.05
SI2		1.95	1.02	2.01	1.00
	si2	2.10	1.10	2.13	1.07
	si3	1.80	1.18	1.90	1.12
CI1		2.55	1.13	2.39	1.01
	ci2	2.55	1.13	2.39	1.01
CI2		2.30	.99	2.23	.93
	ci1	2.55	1.10	2.47	1.02
	ci4	2.05	1.16	2.00	1.09
FX1		1.29	1.14	1.38	1.14
	fx3	1.29	1.14	1.38	1.14
FX2		1.13	.97	1.18	.96
	fx1	1.08	1.06	1.12	1.05
	fx2	1.17	1.14	1.25	1.06
RE1		1.28	.89	1.37	.85
	re2	1.53	1.12	1.57	1.00
	re4	1.04	1.00	1.16	1.00
RE2		1.23	.89	1.31	.87
	re1	1.19	1.05	1.32	1.03
	re5	1.28	1.08	1.31	1.03

Test half	Item	t1 M	t1 SD	t2 M	t2 SD
IN1		.97	1.09	1.08	1.13
	in4	.83	1.07	.93	1.12
	in3	1.12	1.33	1.23	1.32
IN2		1.05	1.04	1.24	1.13
	in1	1.23	1.28	1.47	1.32
	in2	.86	1.05	1.02	1.14
LH1		1.78	1.05	1.87	1.00
	lh2	1.61	1.14	1.66	1.07
	lh4	1.96	1.26	2.08	1.19
LH2		2.27	.98	2.26	.89
	lh1	1.56	1.25	1.63	1.13
	lh3	2.97	1.05	2.88	1.00
CS1		2.12	.87	2.04	.90
	cs5	1.49	1.04	1.50	1.09
	cs1	2.75	1.01	2.57	.96
CS2		1.89	.96	1.96	.92
	cs3	1.80	1.19	1.80	1.12
	cs4	1.99	1.06	2.11	.96

Note. For abbreviations and item wording see note to Table 4.

Appendix D

German Aesthetic Fluency in Film Scale

Instruction

Nun folgen einige Fragen zur Vertrautheit mit Begriffen/Namen aus dem Bereich 'Film'. Bitte beantworten Sie alle Fragen spontan und allein – ohne Verwendung von Hilfsmitteln (Lexika, Google, Freunde o. ä.). Es gibt keine richtigen oder falschen Antworten! Wenn Sie bei einer Frage nicht sicher sind, wählen Sie bitte die Antwortkategorie, die Ihrer Einschätzung am nächsten kommt.

Items: Item Stem and Response Scale

Response scale (before slash: version for terms / after slash: for persons; loosely based on Silvia & Berg, 2011).

0 = Ich habe noch nie davon gehört. / Ich habe noch nie von ihm gehört.

1 = Ich habe schon einmal davon gehört, weiß aber nicht wirklich etwas darüber. / Ich habe schon einmal von ihm gehört, weiß aber nicht wirklich etwas über ihn.

2 = Ich habe eine vage Idee davon, was es sein könnte. / Ich habe eine vage Idee davon, wer es sein könnte.

3 = Ich verstehe, was damit gemeint ist, wenn darüber geredet wird. / Ich weiß, wer gemeint ist, wenn über ihn geredet wird.

4 = Ich kann mit eigenen Worten erklären, was damit gemeint ist. / Ich kann mit eigenen Worten erklären, was sein Werk auszeichnet.

Item stem (Terms/persons in bold adapted from Silvia & Berg, 2011).

(1) Nouvelle Vague	(12) Geburt einer Nation
(2) Sergej Eisenstein	**(13) Das Cabinet des Dr. Caligari**
(3) McGuffin	(14) Autorenfilm
(4) Mise en scéne	**(15) Montage**
(5) Rainer Werner Fassbinder	(16) Amerikanische Nacht
(6) Rashomon	**(17) John Cassavetes**
(7) Eastmancolor	(18) Twin Peaks
(8) DEFA	(19) Todd-AO
(9) Academy Award	**(20) Continuity editing**
(10) Cinecittà	(21) SPIO
(11) Final Cut	(22) Dogma 95

Table D1
Psychometric Properties of the Final 10 Items at t1 (N = 587)

No.	*M*	*SD*	1	2	6	8	10	13	14	18	20	λ
1	1.47	1.52										.785
2	1.31	1.53	.59									.767
6	.92	1.49	.56	.53								.653
8	2.08	1.64	.48	.49	.35							.623
10	1.29	1.51	.52	.50	.42	.42						.643
13	2.13	1.57	.57	.58	.46	.50	.45					.748
14	2.73	1.28	.53	.52	.42	.47	.48	.56				.698
18	2.45	1.57	.48	.41	.38	.43	.37	.53	.41			.618
20	1.34	1.49	.48	.45	.43	.29	.42	.43	.44	.42		.604
22	1.70	1.64	.60	.61	.47	.45	.44	.54	.51	.49	.45	.747

Note. Rows x columns (1–22) = correlation matrix; λ = loadings on one factor.

Appendix E

German Film Expertise Scale

The initial item pool consisted of 10 items (items in bold adapted from Robinson, 1975):

(1)   Ich schaue mir Filme lieber im Kino an als zuhause.
(2)   Ich besuche mehrmals im Monat Filmvorführungen im Kino.
(3)   Ich besitze eine umfangreiche Filmsammlung.
(4)   Ich unterhalte mich gerne mit anderen über Filme.
(5)   Wenn in meinem Freundeskreis über einen Film gesprochen wird, kann ich meist mitreden.
(6)   Ich schaue mehrmals in der Woche Filme zuhause.
(7)   Ich sehe mir gerne Filme in der Originalsprache mit Untertiteln an.
(8)   Ich achte immer darauf, wer bei einem Film Regie geführt hat.
(9)   Ich habe bestimmte Lieblingskameraleute (Director of Photography).
(10)   Ich interessiere mich für Bücher über Filmtheorie.

The response scale is without middle category and ranges from 0 (*disagree completely*) to 3 (*agree completely*).

Table E1
Psychometric Properties of the Final Six Items at t2 (N = 273)

No.	M	SD	3	4	5	8	9	λ_1	λ_2
3	1.53	1.07						.56	
4	2.29	0.76	.44					.79	
5	2.09	0.77	.45	.67				.83	
8	1.26	1.07	.26	.30	.37				.81
9	0.68	0.92	.33	.34	.38	.66			.82
10	0.81	0.97	.26	.30	.25	.56	.55		.69

Note. Rows x columns (3–10) = correlation matrix; λ_i = loadings on two factors; φ_{12} = .54.

Appendix F

Table F1

Fit Indices for all Models Applied to Estimate the Latent Correlations Between the SMEC Scales and External Criterion Constructs (Table 11)

Construct	Study (*N*)	SB-χ^2/*df*	rRMSEA	rRMSEA 90% CI	SRMR	rCFI
FGP: Romance	2a (587)	2.42	.049	.045, .054	.062	.919
FGP: Action	2a (587)	2.56	.052	.047, .056	.063	.913
FGP: Animation	2a (587)	2.37	.048	.044, .053	.061	.921
FGP: Thriller	2a (587)	2.40	.049	.045, .053	.061	.920
FGP: Crime	2a (587)	2.45	.050	.045, .054	.062	.916
FGP: Adventure	2a (587)	2.44	.050	.045, .054	.062	.917
FGP: Science-Fiction/Fantasy	2a (587)	2.44	.050	.045, .054	.062	.918
FGP: Drama	2a (587)	2.44	.050	.045, .054	.062	.918
FGP: Tragedy	2a (587)	2.39	.049	.044, .053	.062	.921
FGP: Comedy	2a (587)	2.38	.048	.044, .053	.062	.922
FGP: Documentary	2a (587)	2.39	.049	.044, .053	.061	.920
FGP: Avant-garde	2a (587)	2.45	.050	.045, .054	.064	.919
FGP: Horror	2a (587)	2.45	.050	.045, .054	.062	.918
FGP: Erotic	2a (587)	2.33	.048	.043, .052	.061	.913
Aesthetic Fluency in Film	2a (587)	2.17	.044	.040, .047	.064	.922
Film Expertise	2b (273)	1.69	.050	.044, .056	.070	.922
Film-specific Evaluation scales	2b (273)	1.82	.055	.050, .060	.059	.918
Big Five	2a (587)	2.06	.042	.039, .045	.056	.909
Sensation Seeking	2a (587)	2.06	.043	.039, .046	.056	.920
Trendsetting	2a (587)	2.20	.045	.042, .049	.059	.913
Political Interest	2a (587)	2.13	.044	.040, .048	.060	.941
Need for Cognition	2a (587)	2.13	.044	.040, .048	.059	.931
Need to Evaluate	2a (587)	2.01	.042	.038, .045	.060	.923
Need for Affect	2a (587)	1.97	.041	.037, .044	.054	.920
Socially Desirable Responding	2a (587)	2.07	.043	.039, .047	.057	.917

Note. SB-χ^2/*df* = Satorra–Bentler χ^2/*df*; rRMSEA = robust root mean square error of approximation; SRMR = standardized root mean square residual; rCFI = robust comparative fit index; FGP = film genre preference.

Curriculum Vitae

Frank M. Schneider

Institut für Kommunikationspsychologie, Medienpädagogik und Sprechwissenschaft,
Institutsabteilung Kommunikationspsychologie/Medienpädagogik
Universität Koblenz-Landau, Campus Landau, Fachbereich 8: Psychologie
Fortstraße 7, 76829 Landau
E-Mail: schneiderfm@uni-landau.de

geboren am 21.09.1976 in Stuttgart-Bad Cannstatt; ledig, eine Tochter

BERUFLICHER WERDEGANG

Universität Koblenz-Landau, Fachbereich Psychologie, IKMS
Wissenschaftlicher Mitarbeiter **Seit 03/2008**

Universität Koblenz-Landau, Fachbereich Psychologie, IKMS
Geprüfte wissenschaftliche Hilfskraft **05/2005 – 02/2008**

Technische Universität Kaiserslautern, Fachbereich Sozialwissenschaften
Vertretung der Juniorprofessur für Methoden der empirischen Sozialforschung **11/2005 – 02/2006**

Universität Koblenz-Landau, Fachbereich Psychologie, IKMS
Studentische wissenschaftliche Hilfskraft **07/2001 – 04/2005**

Südwestrundfunk (SWR), Intendanz, Abt. Medienforschung/Programmstrategie, Baden-Baden
Freier Mitarbeiter **10/2001 – 03/2004**

AKADEMISCHE QUALIFIKATIONEN

Universität Koblenz-Landau, Campus Landau
Postgraduiertenstudium Kommunikationspsychologie/Medienpädagogik **2005 – 2010**

Universität Koblenz-Landau, Campus Landau
Diplom in Psychologie **10/2000 – 04/2005**

Universität Konstanz
Vordiplom in Psychologie **10/1998 – 09/2000**

Universität Erlangen-Nürnberg
Teilnehmer am Vorbereitungsprogramm für Studieninteressenten im Fach Psychologie, FIM **10/1997 – 09/1998**

ZIVILDIENST

KEKS e.V. (Kreis für Eltern von Kindern mit Speiseröhrenmissbildung), Stuttgart
Zivildienstleistender **09/1996 – 10/1997**

SCHULBILDUNG

Gottlieb-Daimler-Gymnasium, Stuttgart
Schulabschluss: Allgemeine Hochschulreife **1987 – 1996**

Sommerrain-Grundschule, Stuttgart
Grundschüler **1983 – 1987**