CrossMark

The effectiveness of celebrity endorsements: a meta-analysis

Johannes Knoll[1] · Jörg Matthes[1] (iD)

Abstract Celebrities frequently endorse products, brands, political candidates, or health campaigns. We investigated the effectiveness of such endorsements by meta-analyzing 46 studies published until April 2016 involving 10,357 participants. Applying multilevel meta-analysis, we analyzed celebrity endorsements in the context of for-profit and non-profit marketing. Findings revealed strong positive and negative effects when theoretically relevant moderators were included in the analysis. The most positive attitudinal effect appeared for male actors who match well with an implicitly endorsed object (d = .90). The most negative effect was found for female models not matching well with an explicitly endorsed object (d = −.96). Furthermore, celebrity endorsements performed worse compared to endorsements of quality seals, awards, or endorser brands. No publication bias was detected. The study has theoretical and practical implications, and provides an agenda for future research.

Keywords Meta-analysis · Celebrity endorsements · Advertising effects

Celebrity endorsements are a well-established marketing strategy used since the late nineteenth century (Erdogan 1999). While the strategy was first applied in traditional brand or product marketing (Erdogan 1999), it has spread to any form of marketing communication, including political marketing (Chou 2014, 2015), health communication, and the marketing of non-government organizations (NGOs; *Jackson 2008; *Wheeler 2009; *Young and Miller 2015). Current estimates indicate every fourth to fifth advertisement incorporates this strategy, though this varies across countries (USA: 19–25%, Elberse and Verleun 2012; Stephens and Rice 1998; UK: 21%, Pringle and Binet 2005; India: 24%, Crutchfield 2010; Japan: 70%, Kilburn 1998; Taiwan: 45%, Crutchfield 2010). In addition, longitudinal analyses show a steady increase over the past years (Erdogan 1999; Pringle and Binet 2005).

Hence, many studies have been conducted to test whether consumer attitudes and behavior are changed by celebrity endorsements. So far, results have been summarized in three narrative (Bergkvist and Zhou 2016; Erdogan 1999; Kaikati 1987) and one quantitative review (Amos et al. 2008; see the Appendix for a summary of the review). The quantitative review of Amos and colleagues focused on source effects of celebrity endorsers. In short, it asked which source variables (e.g., expertise, attractiveness) exert which influence on advertising effectiveness. However, it did not test whether the obtained effect sizes were significant, but solely tested whether they were significantly different from each other. Hence, up until now, there is no meta-analytic knowledge about whether celebrity endorsements actually influence consumers' responses, including the size of their influence. In addition, there is no knowledge about whether effects differ in terms of specific outcomes (e.g., cognitive, affective, behavioral). The reason is that Amos et al. (2008) applied only a combined measure of advertising effectiveness. Furthermore, frequently claimed propositions like the match-up hypothesis or the proposition of stronger effects in the case of unfamiliar brands have never been tested on a meta-analytic level. This seems particularly pressing considering the fact that practitioners frequently refer to such claims (Erdogan 1999). Besides, there

✉ Jörg Matthes
joerg.matthes@univie.ac.at

Johannes Knoll
johannes.knoll@univie.ac.at

[1] Department of Communication, University of Vienna, Währinger Straße 29, A-1090 Vienna, Austria

56

J. of the Acad. Mark. Sci. (2017) 45:55–75

are conflicting results of individual studies, for instance, when it comes to the endorser's sex or endorsement repetition (Bergkvist and Zhou 2016; Erdogan 1999). Last but not least, numerous studies have been conducted since the last quantitative review in 2004 (Bergkvist and Zhou 2016).

The present meta-analysis seeks to address these shortcomings by integrating research published through April 2016, by reporting average effect sizes according to various advertising outcomes including respective confidence intervals, and by performing moderator analyses testing the impact of various endorser and endorsed object variables. In terms of methodological advancement, we apply multilevel modeling accounting for the dependence of multiple effect sizes and we estimate publication bias, both important issues in meta-analysis (Borenstein et al. 2009). Finally, we provide practitioners with empirically derived implications for how to choose the right celebrity and offer researchers an agenda for future research.

Conceptual framework

Following McCracken (1989), celebrity endorsements are understood as a marketing technique in which an individual enjoying public recognition "uses this recognition on behalf of a consumer good by appearing with it in an advertisement" (p. 310). The effects of endorsements can well be explained within the advertising effectiveness model provided by Ladvidge and Steiner (Lavidge and Steiner 1961). Studies have mostly investigated celebrity endorsements according to one or more of the model's advertising functions (Bergkvist and Zhou 2016; Erdogan 1999; Kaikati 1987). Furthermore, the model has revealed itself as fruitful in a similar meta-analysis (Grewal et al. 1997). It enables a

systematic organization of the analyzed dependent variables and moderators, and specifies their relationships (see Fig. 1). According to the model, advertising serves to influence three basic psychological dimensions: the cognitive, the affective, and the conative. "Advertising's cognitive function provides information and facts for the purpose of making consumers aware and knowledgeable about the sponsored brand. Advertising's affective function creates liking and preference for the sponsored brand – preference presumably refers to more favorable attitudes. Advertising's affective function, therefore, is to persuade. Finally, advertising's conative function is to stimulate desire and cause consumers to buy the sponsored brand" (Grewal et al. 1997, p. 2). Important to note, we do not suggest that these outcomes necessarily take place in a particular sequence (i.e., cognition => affect => behavior). Following more recent advancements in the conceptualization of advertising effects, we propose that each of the outcomes may be independently influenced by celebrity endorsements. In addition, all outcomes are assumed to be interrelated. They possibly influence or interact with each other (Vakratsas and Ambler 1999). This is indicated by the double-headed arrows in Fig. 1.

Cognitive effects

Cognitive effects include awareness and knowledge about an endorsed object. Establishing awareness starts from creating attention and interest (Lavidge and Steiner 1961). Directing one's attention involves controlled as well as automatic processes (Kahneman 1973). Both processes can be influenced by celebrity endorsements. First, people who are interested in a particular celebrity are assumed to purposefully direct their
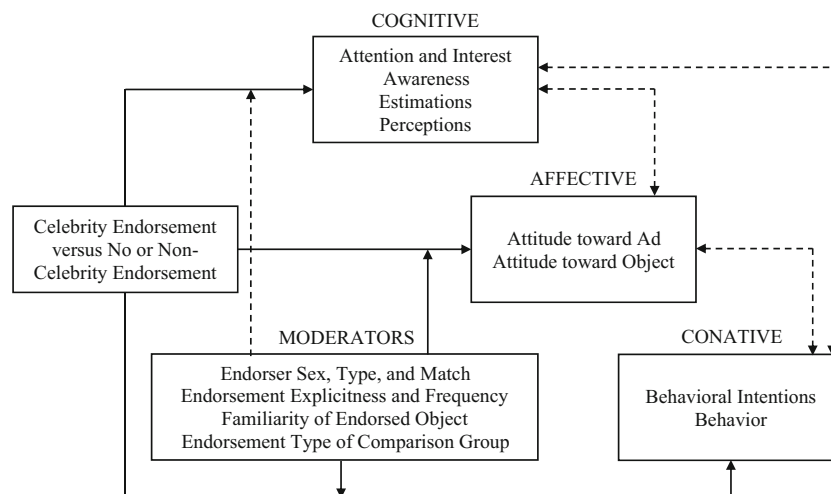
Fig. 1 Celebrity endorsement effectiveness model adapted from Grewal et al. (1997)

attention to this celebrity's ad (*Wei and Lu 2013). Second, people's attention is automatically directed. Humans tend to give preferential treatment to stimuli that are related to their goals (Lang 2000). In addition, celebrities are well-known, resulting in more accessible representations in memory (Erfgen et al. 2015). This should foster automatic attention, too (Bargh and Pratto 1986).

Once a celebrity endorsement grabs their attention, consumers are assumed to become more interested in the advertised object as compared with a non-endorsed or other-endorsed object. This due to the fact that celebrities possess inherent news value caused by their celebrity status (Corbett and Mori 1999). Since celebrities are generally liked, consumers also tend to be more motivated to assess what kind of object a celebrity is endorsing. As a result, object recall and recognition is assumed to be enhanced due to greater message elaboration (*Petty et al. 1983).

In terms of knowledge, celebrity endorsements are assumed to influence the meaning of the endorsed object (*Miller and Allen 2012) as well as perceptions about its price, its taste level, the risk of buying it, or the perceived information value of the endorsement (*Biswas et al. 2006; *Dean and Biswas 2001; *Freiden 1982; *Friedman et al. 1976; *Young and Miller 2015). Based on the mechanism behind these effects, consumers are assumed to conclude that an object has a specific attribute when they perceive this object as paired with a celebrity known for this attribute (e.g., premium price with a high-class celebrity; *Miller and Allen 2012). The process can be conceptualized as propositional learning (De Houwer 2009). Consumers have experienced in their past that people frequently present themselves with objects they share similarities with (Elliot and Wattanasuwan 1998). "Once a relation between two events has been discovered in the past, it is likely that this knowledge is used to generate propositions about similar events in the present" (De Houwer 2009, p. 8). As a result, celebrity attributes created through celebrities' role in society transfer to associated objects (McCracken 1989). In conclusion, it is proposed that celebrity endorsements influence consumers' cognitions including attention and interest, awareness, as well as perceptions.

H1:   As compared with non-celebrity endorsements or no endorsements, celebrity endorsements evoke greater attention, interest, and awareness as well as perceptions more in line with the respective endorser.

Affective effects

Affective effects pertain to attitudes toward the ad and attitudes toward the advertised object. This influence may best be explained with regard to balance theory (Heider 1946, 1958;

see also Mowen and Brown 1981). The theory explains a person's desire to maintain consistency among a triad of linked cognitions. It follows that people generally strive for a consistent organization of their cognitive structures, experiencing this state as most tension-free. In the case of celebrity endorsements, the cognitive triad consists of the consumer, the celebrity, and the endorsed object or the endorsed ad, respectively. A consistent state is achieved if the consumer perceives the celebrity and the endorsed object/ad as equally valenced (i.e., as both positive or both negative) because celebrities endorsing an object are usually seen as positively related to that object or the respective ad (Erdogan 1999). Starting from the premise that researchers and practitioners usually employ likeable celebrities, it can be hypothesized that celebrity endorsements positively impact consumers' attitudes toward the ad and attitudes toward the endorsed object. Only then are consumers' attitudes and their liking for the respective celebrity of the same valence (i.e., both positive; Heider 1946, 1958). Although there may be similar effects for likeable non-celebrity endorsers, these are assumed to be notably weaker. This is due to the fact that consumers are familiar with celebrities by definition. As a result, relationships with celebrities are more affectional as compared with unknown non-celebrity endorsers (Dibble et al. 2016).

H2:   As compared with non-celebrity endorsements or no endorsements, celebrity endorsements evoke more positive attitudes toward the ad and the endorsed object.

Behavioral effects

Behavioral effects include purchasing or using an object (e.g., *Freiden 1982; *Kamins 1989; *Kamins and Gupta 1994; *Roozen and Claeys 2010; *Siemens et al. 2008), sharing object information, volunteering, supporting a charitable cause, or voting for a political candidate (Myrick and Evans 2014; *Pease and Brewer 2008; *Wei and Lu 2013; *Wheeler 2009). Such effects are frequently explained with regard to the theory of planned behavior (Ajzen 1991). According to the theory, behavior is strongly determined by behavioral intentions. These are, in turn, influenced by consumers' attitudes, the perceived subjective norm, and the perceived behavioral control. As long as consumers are able to exert the respective behavior (behavioral control), and as long as consumers do not feel social pressure to avoid the behavior (subjective norm), attitudes largely predict behavioral intentions. The assumptions have been supported by various meta-analyses (Armitage and Conner 2001; Kim and Hunter 1993). Accordingly, the more positive attitudes assumed in H2 should lead to stronger behavioral intentions and respective

58

J. of the Acad. Mark. Sci. (2017) 45:55–75

behavior. Corresponding effects were, for instance, found by Fleck et al. (2012) and Mishra and Mishra (2014). We consequently hypothesize:

H3:   As compared with non-celebrity endorsements or no endorsements, celebrity endorsements evoke stronger behavioral intentions and behavior.

Moderators

Studies investigating the applied advertising effectiveness framework have consistently found that people respond differently to advertisements depending on characteristics of the ad, the advertised object, and individual characteristics (Vakratsas and Ambler 1999). This is frequently intended by the advertiser, tailoring advertisements to specific consumers and their needs (Lavidge and Steiner 1961). In line with this reasoning, we included various moderators within our framework accounting for the fact that consumers do not respond uniformly to advertising (cf. Figure 1; Lavidge and Steiner 1961). Following Grewal et al. (1997), our analysis of moderators is limited to those that (1) are theoretically relevant, (2) provide a sufficient number of effect sizes, (3) show sufficient variance to test the moderation, and (4) are important to advertisers. In terms of number of effect sizes, Higgins and Green (2011) suggest considering moderator analysis only if there are ten or more studies incorporating the moderators. Seven moderators met the criteria: endorser sex, endorser type, endorser match, endorsement explicitness, endorsement frequency, familiarity of the endorsed object, and endorsement type of the comparison group.

Endorser sex Though endorser sex has generally been viewed as influential (e.g., Erdogan 1999; McCracken 1989), hardly any study explicitly addressed this variable in empirical research. Most studies have investigated either female or male endorsers (for the only exception, see Freiden 1984). "The dearth of research on endorser gender effects is somewhat surprising as persuasion research shows that men and women respond differently to male and female communicators" (Bergkvist and Zhou 2016, p. 11). Hence, meta-analysis seems especially valuable (Lipsey and Wilson 2001). Assumptions about possible effects may be derived from studies on non-celebrity spokespersons. According to Kenton (1989), the credibility and persuasiveness of a spokesperson depends on four dimensions: goodwill and fairness (e.g., unselfishness), prestige (e.g., power, status), expertise (e.g., competence), and self-presentation (e.g., confidence). Research has revealed women to be higher ranking on goodwill and fairness, whereas men outperform women on the remaining dimensions (Kenton 1989). As a result, male

spokespersons were frequently more persuasive than female ones (e.g., Cabalero et al. 1989; Whittaker 1965). Transferring this to the present context, consumers may perceive male celebrity endorsers as more credible due to higher levels of expertise and prestige (Cabalero et al. 1989). As a result, male celebrities are assumed to evoke stronger endorsements effects when compared to female ones.

H4:   Male celebrity endorsers evoke stronger endorsement effects when compared to female ones.

Endorser type No study has explicitly investigated different types of celebrity endorsers. Instead, studies have typically focused on only one type. For instance, studies have explored actors, models, musicians, athletes, or TV hosts (e.g., *Dean and Biswas 2001; *Frizzell 2011; *Pease and Brewer 2008; *Wheeler 2009; *Wei and Lu 2013). By joining the results of several studies, meta-analysis can provide information whether certain endorser types perform better than others do (Lipsey and Wilson 2001).

Starting from the premise that endorsement effects depend on the strength of the relationship a consumer shares with a celebrity (McCracken 1989), research on parasocial interaction can provide insights. Specifically, studies have revealed that people tend to develop relationships with celebrities, merely known from the media, just as they would do with real life persons (Dibble et al. 2016): Upon encountering a celebrity on television, radio, or the Internet, consumers may parasocially interact with the celebrity, storing this experience in a relationship schema (Klimmt et al. 2006). The more frequently a celebrity is encountered and the more intense each interaction experience is, the more likely a strong consumer–celebrity relationship is formed (Klimmt et al. 2006). Looking at different kinds of celebrities, consumers are particularly likely to form a strong relationship with actors. First, consumers are audiovisually exposed to actors, creating a particularly rich experience, and second, experience is usually based upon multiple encounters over a longer period: "Over time, viewers become familiar with characters and performers on continuing series and often feel as though they know these individuals as well as they know their friends and neighbors. The importance of characters to viewers frequently extends beyond the viewing situation to include the sense of having personal relationships with the characters, deep concern about what happens in their 'lives,' and/or a desire to become like them in significant ways" (Hoffner and Buchanan 2005, p. 326).

According to McCracken (1989), this exact type of relationship causes consumers to accept celebrities' influence more readily. The following hypothesis is proposed:

H5:   Actors elicit stronger celebrity endorsement effects when compared to other types of celebrities such as models, musicians, athletes, or TV hosts.

Endorser match Several studies have investigated the so-called product match-up hypothesis that assumes the effectiveness of celebrity endorsements is partially dependent on the degree of perceived fit between an endorsed object and the respective celebrity (Erdogan 1999). A good match may be an attractive model presenting cosmetics, whereas a bad match may be an athlete trying to sell a guitar. The process underlying the product match-up hypothesis can be explained with regard to Social Adaptation Theory (Kahle and Homer 1985; Kamins 1990) or Schema Theory (Lynch and Schuler 1994). Social Adaptation Theory assumes that people use information sources as long as they facilitate adaptation to their environment. If a match exists between a spokesperson and a product on some relevant attribute, the spokesperson becomes an information source of adaptive significance on which people may rely (Kamins 1990). Schema Theory posits that attributes of celebrities can be integrated more easily with existing product schemas if the celebrity schemas match the product schemas (Lynch and Schuler 1994). Both theories assume enhanced effects in the case of congruence. Accordingly, several studies have supported the product match-up hypothesis (Erdogan 1999). We thus hypothesize larger effects for object–endorser congruence compared to incongruence.

H6:   Congruent celebrity endorsers evoke stronger endorsements effects when compared to incongruent ones.

Endorsement explicitness Explicitness can broadly be categorized into two modes: implicit and explicit endorsements. Whereas implicit endorsements refer to situations where celebrities simply use an object or merely appear jointly without overtly announcing their support ("I use this object"; *Miller and Allen 2012), explicit endorsements refer to situations where celebrities overtly express their support for an object ("I endorse this object"; *Miller and Allen 2012). To the best of our knowledge, no study has ever compared both modes; instead, they have researched either implicit or explicit endorsements. Though effects have been found with both modes, one mode may be more effective than the other (implicit: e.g., *Miller and Allen 2012; explicit: *Dean and Biswas 2001; *Friedman and Friedman 1979). According to Russell and Stern (2006), consumers infer the celebrity–object association to be of greater strength if celebrities explicitly express their support, signaling commitment and reliability. In addition, consumers may not even realize that an object is endorsed if the endorsement is too subtle. We consequently propose that explicit endorsements are more effective that implicit ones.

H7:   Explicit endorsements evoke stronger effects than implicit ones.

Endorsement frequency Celebrities may also vary in their endorsement frequency. Consumers are highly likely to encounter celebrity endorsements multiple times via various media channels, including TV, billboards, print advertising, radio, and the Internet. Research on classical conditioning suggests that effects may occur as early as a single pairing of a celebrity with an endorsed object (e.g., Ambroise et al. 2014; Gorn 1982). However, other research suggests that effects tend to be greater the greater the number of pairings. For instance, Stuart et al. (1987) increased the number of pairings from one to three, to ten, and eventually to twenty, revealing a steady increase in effectiveness. Although these results do not directly refer to celebrity endorsements, similar effects can be assumed because celebrity endorsements are often seen as a certain type of classical conditioning (e.g., *Chen et al. 2012). The following hypothesis is proposed:

H8:   Celebrity endorsement effects increase with increased endorsement exposure.

Familiarity of the endorsed object Next to the celebrity, the endorsed object itself may impact endorsement effectiveness (*Friedman and Friedman 1979). For instance, researchers assume stronger effects, with decreasing familiarity with an endorsed object (*Miller and Allen 2012). Object familiarity can be understood as the number of object-related experiences accumulated by a consumer (Alba and Hutchinson 1987). These experiences can be obtained directly and indirectly, such as through celebrity endorsements (Kent and Allen 1994). The more familiar a person is with an object, the more comprehensive his or her knowledge structures can become (Keller 2012). Given that consumers already possess a rich network of associations representing an object, attitudes, and behavior appear more difficult to change (Cacioppo et al. 1992). Accordingly, Ambroise et al. (2014) reported stronger celebrity endorsement effects with unfamiliar compared to familiar brands. Similarly, Shimp et al. (1991) showed the likelihood of conditioning effects for unknown or moderately known objects, but not for well-known ones. We consequently propose stronger celebrity endorsement effects for unfamiliar objects when compared to familiar ones.

H9:   Celebrity endorsement effects are stronger for unfamiliar objects when compared to familiar ones.

Endorsement type of the comparison group Investigating the effectiveness of celebrity endorsements through experiments, researchers have chosen various control groups. Frequently, celebrities are compared with a non-endorsed condition (e.g., *Martín-Santana and Beerli-Palacio 2013), an expert (e.g., Biswas et al. 2006), or an ordinary consumer (e.g., *Dong 2015). Less frequently, celebrities are compared with an unknown model or athlete (e.g., *Roozen and Claeys 2010), an employee of the selling company (*Maronick 2005), a quality seal or an award (*Dean and Biswas 2001), or an endorser brand from the same product category (*Sengupta et al. 1997). Studies typically apply one or two of these comparison groups. Thus, they enable assertions about whether celebrity endorsements outperform a single kind of endorsement or no endorsement. By contrast, meta-analysis enables comparisons across all types of endorsements simultaneously. Therefore, we can see whether celebrity endorsements outperform any other kind of endorsement. We can also test whether specific differences in performance (e.g., celebrity vs. expert) are significantly different from other performance differences (e.g., celebrity vs. ordinary consumer). Marketing managers can thus gain valuable knowledge when deciding on celebrity endorsers, any other kind of endorsement, or no endorsement at all. We consequently ask:

RQ1: Do celebrity endorsements differ in their effectiveness depending on the control group applied?

A concise summary of the existing knowledge on celebrity endorsement effects can be found in Table 1. Looking at the main results, celebrity endorsements are shown to affect cognitive, affective, and conative outcomes. Furthermore, most studies have looked at endorsements of for-profit causes. Results frequently appear to be mixed. In addition, some studies show no effects at all. This meta-analysis will shed light on these mixed results by calculating an overall effect. In addition, mixed results can be clarified by adding potential moderators to the analysis. Furthermore, the meta-analysis will close gaps in the literature by investigating between-study differences that cannot be explored with single studies (e.g., endorser sex, endorser type, endorsement explicitness). Looking at the investigated outcomes, most studies have investigated affective reactions followed by cognitive and conative ones. Meta-analysis will provide insights about whether there are any differences in terms of effectiveness per outcome type.

Methods

Study retrieval

Literature search Studies were collected from three major databases (Business Source Premier, PsychINFO, Communication and Mass Media Complete). The search included all peer-

Table 1 Summary of research on the effectiveness of celebrity endorsements

Investigated effects	Number of studies	Dependent variables	Investigated ad format	Cause of endorsed object	Main results
Cognitive	20 (43% of all investigated studies)	Attention, interest, recall, recognition, meaning, estimated price, estimated taste, perceived risk, perceived information value, perceived credibility	Print ad (85%) Radio spot (5 %) Direct mailing (5%) Press text (5%)	For-profit (85%) Non-profit (15%)	Celebrity endorsements increase consumers' attention and sometimes their interest. Results for recall and recognition are mixed. The same pertains to estimations and perceptions. There are strong meaning transfer effects.
Affective	35 (76% of all investigated studies)	Attitude toward the ad, attitudes toward the endorsed object, evoked feelings	Print ad (80%) Direct mailing (3%) Press text (14%) Packaging (3%)	For-profit (69%) Non-profit (31%)	Celebrity endorsements increase attitudes toward the ad and toward the endorsed object. This is particularly true when compared to no endorsement and when the celebrity matches the endorsed object.
Conative	19 (41% of all investigated studies)	Intention to purchase, intention to use, intention to vote, intention to share information or inform oneself, brand choice, intention to volunteer, intention to support a charitable cause	Print ad (80%) TV spot (5%) Direct mailing (5%) Press text (5%) Packaging (5%)	For-profit (79%) Non-profit (21%)	Results for purchase intention and using a product appear to be mixed. Furthermore, celebrity endorsements mostly enhance one's intention to support a charitable cause and to volunteer. The same applies to voting. There are no effects on brand choice or intention to share information or to inform oneself.

reviewed articles written in English and published through April 2016. The databases were examined using the term celebrit* in combination with endors*, spokes*, or advert* in any available search field. The search resulted in 1025 articles. About 300 of them were quantitative studies, including content analyses, surveys, and experimental studies.

Inclusion criteria These quantitative studies were narrowed down based on the impact of celebrity endorsements on endorsed objects. Three criteria had to be met. First, only experimental studies were included because only they enable causal assertions (Shaughnessy and Zechmeister 1997). The studies had to compare an experimental group to a control group. While the experimental group had to feature a celebrity endorsing an object, the control group had to include the same object, either non-endorsed or endorsed by a non-celebrity spokesperson. Studies that compared various types of celebrity endorsements but did not feature a non-celebrity control group were excluded (e.g., Ambroise et al. 2014; Kamins 1990). Second, the celebrities had to be actually existing celebrities, thus excluding studies that investigated the impact of fictitious and imagined celebrities, as their validity is arguably limited. Third, the studies also had to report effect measures related to the endorsed object, excluding studies that solely reported measures related to the endorser (e.g., Cho 2010) or the general acceptance of celebrity endorsements (e.g., Becker 2013). In addition, a measure was considered only if it was possible to obtain at least two effect sizes. Otherwise, the meta-analyzed effect size would equal the sole obtained effect size, rendering meta-analysis useless. This resulted in 15 eligible measures: attention to and interest in an ad, awareness of an endorsed object (recognition and recall), attitude toward an ad, attitude toward the endorsed object, perceived credibility of the ad and advertiser, meaning transfer (in the sense of transferring a celebrity's meaning to a brand), evoked feelings, estimated price of a product, taste of a product, estimated information value of an ad, planning to inform oneself more about an endorsed object, perceived increase of knowledge, perceived risk when buying or using a product, brand choice, and behavioral intentions (intention to purchase or use an object, intention to volunteer, intention to support a charitable cause by spending time or money, and intention to share an endorsed object online; cf. Motyka et al. 2014). No limitations were placed regarding the endorsed object encompassing any kind of object, such as product, brand, organization, behavior, or charitable cause.

Results Based on these criteria, 44 manuscripts remained (the majority of the 300 quantitative studies were content analyses, surveys, or experimental studies comparing celebrities with celebrities). Eight of these (Chou 2014; Fireworker and Friedman 1977; Freiden 1984; Jain et al. 2011; Ross et al. 1984; Sanbonmatsu and Kardes 1988;

Veer et al. 2010) had to be excluded, as they lacked appropriate statistical information to calculate effect sizes with the formulas suggested by Lipsey and Wilson (2001). Beforehand, all authors had been contacted and asked to provide missing statistical information if possible. According to Eisend (2009), about 18% exclusion is not uncommon in meta-analysis, and it matches other meta-analyses in marketing (Brown and Stayman 1992; Szymanski et al. 1995; Tellis 1988). The remaining 36 manuscripts yielded 46 independent studies, coming to 10,357 participants.

Meta-analytic procedures

Effect size calculation The standardized mean difference (d) was used as the effect size estimate according to the formulas provided by Lipsey and Wilson (2001). All available statistical information was incorporated (e.g., means, standard deviations, t- and F-statistics, and frequencies). Since this effect size estimate has been shown to be upwardly biased when calculated from small sample sizes (Lipsey and Wilson 2001), all estimates were corrected for sample size bias (Hedges 1981). Positive d-values indicated a stronger effect of a celebrity endorsement compared to a non-endorsed or non-celebrity endorsed message, whereas negative d-values indicated a stronger effect of the non-endorsed or non-celebrity endorsed message. In total, 367 effect sizes were obtained. The ratio of effect sizes (367) to the number of studies (46) is the rule rather than the exception when analyzing various dependent variables (Eisend 2006, 2009; Szymanski et al. 1995).

Effect size integration and meta-analysis Estimates were based on random-effects models. Fixed-effects models assume that all studies included in the meta-analysis are practically identical, having the same true effect size. In contrast, random-effects models assume differing true effect sizes varying, for instance, because of different participants or treatments. Specifically, true effect sizes are assumed to be distributed around some mean whereby the studies included in the analysis are assumed to represent a random sample (Borenstein et al. 2009). This model was much more realistic, as participants and study settings certainly differed across studies. In addition, results may be generalized because the investigated studies are treated as a random subset of a larger study population (Hedges and Vevea 1998). Several studies reported results that enabled obtaining more than one effect size per dependent variable. Performing a meta-analysis on these studies would violate the assumption of independence of effect sizes and assign more weight to the studies producing more than one effect size. Previous studies mostly ignored these problems, aggregated effect sizes into a single effect size (or chose only one effect size

per study), or performed the so-called shifting the unit of analysis approach (Cheung 2014). This approach averages effect sizes within differing units depending on the current research question (e.g., study as a unit or study characteristics, such as gender of participants as a unit).

While ignoring these problems is clearly not satisfactory, the latter two approaches are rather broadly accepted (Borenstein et al. 2009; Cooper 2010). However, aggregating effect sizes or choosing only one effect size per study may strongly reduce the number of effect sizes, thus lowering the power of statistical tests. In addition, statistical information is lost, resulting in less precise estimates. The same deficits apply to shifting the unit of analysis (Cheung 2014). Researchers recently suggested treating meta-analysis as a multilevel model to address these drawbacks (e.g., Cheung 2014; Field 2015; Konstantopoulos 2011). The basic idea nests the effect sizes (first level) within the studies (second level; Konstantopoulos 2011). The resulting model then looks like Eqs. (1) and (2):

$$\gamma_i = \lambda_i + e_i \text{(first level or within−study model)} \quad (1)$$

$$\lambda_i = \beta_0 + u_i \text{(second level or between−study model)}. \quad (2)$$

"Effect sizes (γ) in the ith study are predicted from the 'true' effect size for that study (λ_i) and some error (e_i) (note that the variance of e_i is the sampling variance of that study). The true effect size for a study is made up of the average population effect (β_0) – which is the thing we usually want to estimate in meta-analysis – and some between-study error (u_i) (note that the variance of this between-study error is the heterogeneity of effect sizes across studies, which in traditional meta-analysis is denoted as τ^2)" (Field 2015, p. 18).

Writing the model in a single-level notation results in Eq. (3):

$$\gamma_i = \beta_0 + u_i + e_i. \quad (3)$$

In this equation, it becomes evident that the variance of an observed effect size (γ_i) is decomposed into a sampling variance component (e_i) and the between-study error or random effect (u_i), as in traditional meta-analysis. However, since u_i denotes a study-specific random effect of an ith study, the same random effect can be assigned to effect sizes stemming from the same study while effect sizes stemming from different studies receive different random effects (Konstantopoulos 2011; Viechtbauer 2015). Consequently, all effect sizes can be taken into account without aggregation and loss of information. The dependence or independence of the effect sizes is explicitly modeled by assigning the correct random effect. Furthermore, a third level may be introduced when estimating an overall effect size composed of effect sizes (first level) nested within different types of effect sizes, that is, dependent variables (second level), which are, in turn, nested within different studies (third level). The variance would then be decomposed into sampling

variance, between-type of effect size variability, and between-study variability. This analysis is necessary for testing whether it makes sense to analyze different types of effect sizes separately (between-type of effect size variability) and whether it makes sense to analyze our moderators at all (between-study variability; Konstantopoulos 2011).

Following these recommendations, all analyses were carried out using the rma.mv() function of the R metafor package (Viechtbauer 2010). A maximum likelihood estimator, the typical method to estimate multilevel models, was applied (see Konstantopoulos 2011; van den Noortgate et al. 2014). Average effect sizes were estimated taking the random-effects perspective, and moderator analyses were performed applying the mixed-effects models (meta-regression). As the studies showed considerable variance in sample size and some studies produced multiple effect size estimates, effect sizes were weighted by sample size and the number of effect sizes per study. Specifically, effects sizes were weighted by the ratio of their study's sample size to the number of effect sizes measuring the same dependent variable within the study (Eisend 2009). As a result, studies reporting only one effect size received the same weight as studies reporting multiple effect sizes if their sample size was equal.

Moderators

The moderators can be grouped as endorser variables, endorsed object variables, and endorsement type of the comparison group. The variables were coded by two independent coders based on the information available in the manuscripts and complemented by the English Wikipedia pages of celebrities where necessary. Agreement was perfect except for endorser match, which yielded an acceptable Krippendorff's alpha of .74 (Krippendorff 2004). Discrepancies were resolved by discussion after a review of the article.

Endorser variables The endorser's sex was coded as female (0) or male (1), according to the description of the study authors. Typical descriptions were male/female, Mr./Mrs., or he/she. If the manuscript provided no gender information, the endorser's English Wikipedia page was consulted. The endorser type was coded as actor (0), model (1), athlete (2), musician (3), or TV host (4), according to the description of the authors. The authors typically described their celebrities by using one of the aforementioned professions. If the manuscript provided no related information, the endorser's English Wikipedia page was consulted. If the Wikipedia page presented various professions, the first was chosen.

The endorser match was coded as incongruent (0) or congruent (1), according to the description of the authors. Frequently used descriptions were not matching/matching, not fitting/fitting, or being incongruent/congruent to an endorsed object.

Furthermore, some studies reported pretests explicitly testing the congruence of the endorser and endorsed object. The moderator was then coded accordingly. If the authors used existing advertising, the endorser was coded as congruent because advertisers usually put considerable time and effort into finding a matching endorser (Erdogan 1999). The endorsement explicitness was coded as implicit (0) or explicit (1) according to the description of the authors. Following *Miller and Allen (2012), the endorsement was coded as implicit when the endorser and the endorsed object appeared merely as paired without the endorser explicitly announcing his or her endorsement (e.g., classical conditioning procedure or ad merely displaying object and celebrity). An explicit endorsement was coded if the endorser's support for an object could be explicitly read or heard by the study participants (e.g., "I think XY is …" or "I love XY"). In addition, signatures were coded as explicit endorsements.

The endorsement frequency was coded continuously starting from one (1) and stretching—theoretically—infinitely, though the maximum number of endorsements was 10.

Endorsed object variables The familiarity of the endorsed object was coded as unfamiliar (0) or familiar (1) according to the description of the authors. Familiarity refers to whether the endorsed object was known by the participants. Typical descriptions in the articles were unknown/known, fictitious, or having a strong reputation. Furthermore, some studies reported pretests assessing object familiarity. The moderator was then coded accordingly.

Endorsement type of the comparison group We coded whether the comparison group perceived the object as non-endorsed (0; appearing without any support) or as endorsed by a non-celebrity spokesperson or organization. The endorsed categories were expert (1), an employee of the selling company (2), an ordinary consumer (3), an unknown model or athlete (4), a quality seal or award (5), a government employee[1] (6), or an endorser brand from the same product category (7). The authors typically described the comparison groups by using one of the aforementioned category names.

Results

Overall analysis

Testing first whether it makes sense to analyze different types of effect sizes separately and whether it makes sense to analyze the moderators at all, we calculated the three-level model

(Konstantopoulos 2011). Effect sizes (first level) were nested within different types of effect sizes (second level), which were, in turn, nested within different studies (third level). We observed no significant overall effect of the celebrity endorsements on participants' responses (d = .04, 95% CI (−.09, .17), ns). However, highly significant heterogeneity was found among effect sizes (Q (366) = 1095.77, $p < .001$). This suggests that effect sizes vary considerably due to the type of effect size differences (second level) and/or study differences (third level). The I^2 statistic—the amount of total variability (sampling variance + heterogeneity) that can be attributed to the heterogeneity among the true effects (Higgins and Thompson 2002)—provided more details. About half the total variability could be attributed to the between-study heterogeneity ($I^2 = 51.53\%$, third level) and about 11% could be attributed to the between-type of effect size within study heterogeneity ($I^2 = 11.45\%$, second level). Hence, it seemed reasonable to explain the between-study heterogeneity by moderator analysis and to examine the different types of effect sizes separately, given their heterogeneity.

Table 2 shows the meta-analytic results for the most frequently investigated dependent variables.[2] The first column presents the average effect size for a specific dependent variable. The second and third column display subgroup results of this average effect size. They differentiate studies that featured a comparison group receiving no endorsement (second column) from studies that featured a comparison group receiving an object endorsed by a non-celebrity spokesperson (third column). In addition, the third column specifies the particular types of endorsements received by the comparison groups. There were almost no average effects for any of the dependent variables. Only one significant effect size emerged. Celebrity endorsements positively affected consumers attitudes toward the endorsed object when celebrity endorsements were compared to a non-endorsed condition (d = .24; 95% CI (.04, .43), $p < .05$). As opposed to this affective effect, there were neither cognitive nor conative effects on average.

Moderator analysis

Moderator analyses were conducted for the dependent variables of attitude toward the endorsed object, attitude toward the ad, and behavioral intention, all featuring a sufficient number of effect sizes to conduct the analyses (Higgins and Green 2011). Tests for heterogeneity revealed significant heterogeneity among all three types of effect sizes (attitude toward the endorsed object: Q (116) = 372.67, $I^2 = 68.87\%$, $p < .001$; attitude toward the ad: Q (44) = 200.81, $I^2 = 78.11\%$, $p < .001$; behavioral intention: Q (92) = 188.33, $I^2 = 51.15\%$, $p < .001$). As indicated by the I^2 statistic, the level of heterogeneity was medium to high,

---

[1] Endorsement by a government employee refers to a study by Frizzell (2011) in which the effectiveness of a celebrity endorsement was compared against the endorsement of a State Department representative (political issue).

[2] Meta-analytic results for less frequently investigated dependent variables (k < 10) can be obtained from the authors upon request.

Table 2 Meta-analytic results (Random-Effects Model) for dependent variables arranged according to average effect size, effect size based on comparisons with no endorsement, and effect size based on comparisons with other endorsements

Dependent variable	k	N	Average			No endorsement			Other endorsement			Type of endorsm.
			95% CI			95% CI			95% CI			
			d	LL	UL	d	LL	UL	d	LL	UL	
Awareness (recognition and recall)	24	3351	−.04	−.30	.22	−.05	−.47	.37	−.04	−.32	.23	C, M
Attitude toward ad	45	1930	.06	−.32	.44	.30	−.32	.93	.01	−.40	.41	E, EM, C, M
Attitude toward object	117	6389	.05	−.10	.21	.24*	.04	.43	−.07	−.24	.10	E, C, M, Q, G, B
Behavioral intentions	93	4241	.02	−.12	.16	.01	−.21	.24	.02	−.13	.17	E, EM, C, M

k: number of effect sizes; N: combined sample size; d: standardized mean difference; CI: confidence interval with lower (LL) and upper limit (UL); E: expert; EM: employee of a selling company; C: ordinary consumer; M: unknown model/athlete; Q: quality seal/award; G: government employee; B: endorser brand; * $p < .05$

suggesting great between-study variability that may be explained by the moderators (Huedo-Medina et al. 2006).

Table 3 presents the results of the first meta-regression testing the influence of the comparison group's endorsement on the effect size for attitude toward the object. The analysis was conducted only for this dependent variable. Only this measure allowed for the simultaneous analysis of almost all the moderator categories. The category of employee of the selling company was missing because no study included both attitude toward the endorsed object and employee of the selling company. The comparison group's endorsement is a categorical variable, so the moderator was dummy coded with no endorsement (0) representing the reference category (Field 2013). As a result, the intercept's coefficient (Table 3), being significant and positive, represents the average effect size of celebrity endorsements compared to a non-endorsed group. The remaining regression coefficients represent the change in this effect size when the comparison group features some kind of endorsement instead of no endorsement. As seen in Table 3, the effect size decreased when the comparison group featured some other kind of endorsement, as indicated by all the regression coefficients being negative. However, the decrease appeared to be significant only when celebrity endorsements were compared to endorsements of an unknown model or athlete, quality seal or award, government employee, or endorser brand. Subtracting these decreases from the intercept (.24), it became evident that the respective effect size was then negative in all cases. A separate test for significance revealed these negative effect sizes to be (marginally) significant for a quality seal or award endorsement, and a endorser brand (unknown model: d = −.32, 95% CI (−.73, .08), ns; quality seal: d = −.44, 95% CI (−.81, −.07), $p < .05$; government employee: d = −.45, 95% CI (−1.01, .12), ns; endorser brand: d = −.58, 95% CI (−1.16, .01), $p = .05$). In terms of the research question, it can thus be concluded that celebrity endorsements perform worse compared to quality seals, award endorsements, or endorser

Table 3 Meta-regression results for testing the influence of comparison group's endorsement type on effect size (attitude toward the endorsed object; k = 117)

Predictor	95% CI			
	B	SE	LL	UL
Intercept (= average effect size of celebrity endorsements compared to no endorsement)	.24**	.09	.07	.41
Change in effect size if comparison group features not "no endorsement" (intercept), but includes a(n)				
Expert endorsement	−.04	.24	−.52	.44
Ordinary consumer endorsement	−.19	.13	−.45	.06
Unknown model/athlete endorsement	−.56*	.22	−.99	−.12
Quality seal/award endorsement	−.68***	.18	−1.04	−.32
Government employee endorsement	−.69*	.30	−1.27	−.10
Endorser brand from same product category	−.81**	.32	−1.43	−.20

The coefficient of −.56* indicates, for instance, that the average effect size of celebrity endorsements significantly decreases by −.56 when the comparison group is exposed to an object endorsed by an unknown model or athlete instead of exposure to an non-endorsed object. Comparison to a non-endorsed object is the reference category

B: unstandardized regression coefficient; SE: standard error; CI: confidence interval with lower (LL) and upper limit (UL); * $p < .05$; ** $p < .01$; *** $p < .001$

brands, but they perform better when compared to no endorsement.

Further meta-regressions were conducted to integrate all assumed moderators (Table 4). The moderators were entered hierarchically starting with comparison group's endorsement (Model 1), followed by endorser variables (Model 2), and eventually complemented by an endorsed object variable (Model 3). For the moderator comparison group's endorsement, the categories were merged into two groups (no endorsement (0) vs. non-celebrity endorsement (1), as in Table 2. The remaining moderators were coded as displayed in the methods section. The endorser type was dummy coded with actor (0) representing the reference category. The results can be seen in Table 4. As can be seen, the impact of celebrity endorsements decreased when compared to an endorsed comparison group instead of a non-endorsed one (Model 1, first column). Looking at the endorser

variables (Model 2, first column), we see that male endorsers performed substantially better than female endorsers, supporting H4. Looking at endorser type, effect size decreased when an object was endorsed by a model, athlete, musician, or TV host instead of an actor (the reference category). However, only a model, musician, or TV host performed significantly less well. Hypothesis 5 was thus partially confirmed. The product match-up hypothesis (H6) was supported, indicated by enhanced effects for congruent endorsers when compared to incongruent ones. By contrast, the seventh hypothesis was rejected, as implicit endorsements performed substantially better that explicit ones. This is the opposite of what we predicted. Likewise, the eighth hypothesis was rejected as no impact was seen for endorsement frequency.

As seen at the bottom of Table 4, R^2 indicates that more than 91% of the heterogeneity could be explained by both blocks of moderators. This explanation was significant, as indicated by the

Table 4 Meta-regression results for testing the impact of moderators on effect size of various dependent variables

Predictor (Moderator)	Attitude toward endorsed object			Attitude toward ad		Behavioral intention	
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 1	Model 2
Intercept (= average effect size of celebrity endorsements when all moderators included in the model equal zero).	.28* (13)	−.03 (.13)	−.02 (.11)	.35 (.34)	−.16 (.10)	.09 (.14)	−.26** (.08)
Change in effect size if							
Comparison group endorsed instead of non-endorsed	−.50*** (.14)	−.18* (.09)	−.20* (.09)	−.41 (.30)	−.05 (.10)	−.001 (.11)	.07 (.07)
Endorser type model instead of actor		−.38*** (.11)	−.39*** (.10)		−.23* (.12)		.00 (.11)
Endorser type athlete instead of actor		−.20 (.14)	−.12 (.13)		.11 (.18)		−.25 (.16)
Endorser type musician instead of actor		−.36* (.17)	−.46** (.15)		.13 (.14)		.31 (.17)
Endorser type TV host instead of actor		−.36* (.17)	−.21 (.17)		−		−
Endorser sex male instead of female		.50* (.22)	.46* (.20)		1.47*** (.29)		1.08*** (.26)
Endorser match congruent instead of incongruent		.44*** (.11)	.47*** (.11)		.20 (.13)		.18* (.07)
Endorsement explicit instead of implicit		−.75*** (.19)	−.55** (.19)		−1.43*** (.28)		−.81** (.26)
Endorsement frequency increases by one unit		−.01 (.05)	−.01 (.05)		−		−
Familiarity of endorsed object known instead of unknown			−.30* (.14)		−		
Q_M	13.71***	82.11***	103.32***	1.87	85.36***	0	50.98***
Q	207.66***	119.08***	113.27***	119.34***	27.09	121.30***	57.78
R^2	21.19	91.69	95.72	2.42	100	0	100
k	63	63	63	37	37	69	69

The coefficient of −.50* indicates, for instance, that the average effect size of celebrity endorsements significantly decreases by −.50 when the comparison group is exposed to an endorsed object instead of exposure to a non-endorsed object. In case of categorical predictors, the reference category presents the category which is introduced as "instead of", e.g., instead of actor

QM: Q-statistic for test of moderators: significant values indicate that all moderators taken together explain a significant amount of heterogeneity in the respective effect size; Q: Q-statistic for test for residual heterogeneity: significant values indicate that the remaining heterogeneity may be explained by other moderators not included in the model; R^2: amount of heterogeneity accounted for; k: number of effect sizes; * $p < .05$; ** $p < .01$; *** $p < .001$

test of moderators ($Q_M = 82.11$, $p < .001$). The test for residual heterogeneity indicated a significant amount of heterogeneity remaining, possibly explained by the last moderator ($Q = 119.08$, $p < .001$). Indeed, familiarity affected effect size as assumed in H9 (Model 3 first column). Celebrity endorsement effects were stronger for unfamiliar objects when compared to familiar ones. For attitude toward the ad and behavioral intention, the results appeared quite similar. The only differences were that endorser match had no influence on attitude toward the ad and that the influence of endorser type did partially differ in both cases. The moderators were able to explain 100 % of the heterogeneity. The tests for residual heterogeneity were accordingly insignificant. The impact of familiarity and endorsement frequency could not be tested as all effect sizes pertained to unknown objects endorsed once.

The combined impact of all moderators can be seen in Table 5. It displays the predicted d values and confidence intervals for an actor endorsing an unfamiliar object a single time compared to no endorsement of the same object at varying levels of the endorser's sex, endorser match, and endorsement mode. Values were calculated with the predict() function (Viechtbauer 2010). The d values for other endorser or object types (e.g., model or familiar object) may be obtained by adding or subtracting the respective regression coefficient from Table 4. Based on the results for attitude toward the endorsed object, the highest effect size can be expected for male actors, matching the object, and endorsing it implicitly ($d = .90$; 95% CI (.54, 1.25)). In contrast, the lowest effect size appears for female actors not matching an explicitly endorsed object ($d = −.58$; 95% CI (−1.02, −.13)). It is important to note that female celebrities can likewise have a positive impact, for instance, endorsing a congruent object implicitly ($d = .44$; 95% CI (.20, .68)). Again, effect size patterns were similar for attitude toward the ad and behavioral intention.

Publication bias

Before starting the bias analysis, the effect sizes were aggregated within the studies. Publication bias analysis checks whether certain studies are more likely to be published than others. Publication bias was assessed applying funnel plots and Egger's regression test for funnel plot asymmetry (Egger et al. 1997). Figure 2 displays the funnel plots for attitude toward the endorsed object, attitude toward the ad, and behavioral intention. The x-axis indicates the observed effect size whereas the y-axis displays the standard error as an indicator of sample size (Borenstein et al. 2009). The funnel plots showed no evidence of publication bias in terms of smaller studies missing at the bottom left corner (i.e., no evidence that smaller studies with minor effect sizes failed to be published). This is further confirmed by Egger's regression tests being insignificant in all cases (attitude toward endorsed object: $t(28) = 1.86$, $p = .07$; attitude toward ad: $t(8) = .01$, $p = .99$; behavioral intentions: $t(15) = 1.45$, $p = .17$).

Discussion

Main findings and contributions

The main findings are summarized in Table 6. The analysis revealed a zero overall effect of celebrity endorsements on consumers' responses. Yet there were strong effects on some dependent measures and under some conditions. The main contribution of our study, therefore, lies in understanding the possible causes of this variability rather than focusing on a summary effect (Thompson and Sharp

Table 5 Predicted d values for actors endorsing an unfamiliar object once compared to no endorsement for various dependent variables at levels of the moderators sex, match, and endorsement explicitness

Levels of moderators			Attitude toward endorsed object[a]				Attitude toward ad[a]				Behavioral intention[a]			
					95% CI				95% CI				95% CI	
Sex	Match	Endorsement explicitness	Predicted d	SE	LL	UL	Predicted d	SE	LL	UL	Predicted d	SE	LL	UL
Female	Incongruent	Implicit	−.03	.10	−.22	.17	−.16	.10	−.36	.04	−.26	.08	−.42	−.10
Female	Incongruent	Explicit	−.58	.23	−1.02	−.13	−1.59	.29	−2.15	−1.03	−1.07	.26	−1.58	−.56
Female	Congruent	Implicit	.44	.12	.20	.68	.04	.12	−.18	.27	−.08	.09	−.26	.10
Female	Congruent	Explicit	−.11	.24	−.58	.36	−1.39	.32	−2.01	−.76	−.89	.28	−1.44	−.34
Male	Incongruent	Implicit	.43	.19	.05	.81	1.30	.32	.68	1.92	.82	.26	.30	1.34
Male	Incongruent	Explicit	−.12	.15	−.41	.16	−.12	.18	−.49	.24	.01	.07	−.13	.15
Male	Congruent	Implicit	.90	.18	.54	1.25	1.51	.26	1.00	2.02	1.00	.25	.51	1.49
Male	Congruent	Explicit	.35	.13	.09	.60	.08	.14	−.20	.36	.19	.07	.05	.32

[a] Frequency was set to one. All other categorical moderators included in the regression model (Table 4) were set to zero. That is, predicted d values refer to actors endorsing an unfamiliar object once compared to no endorsement of the same object; d: standardized mean difference; SE: standard error; CI: confidence interval with lower (LL) and upper limit (UL)
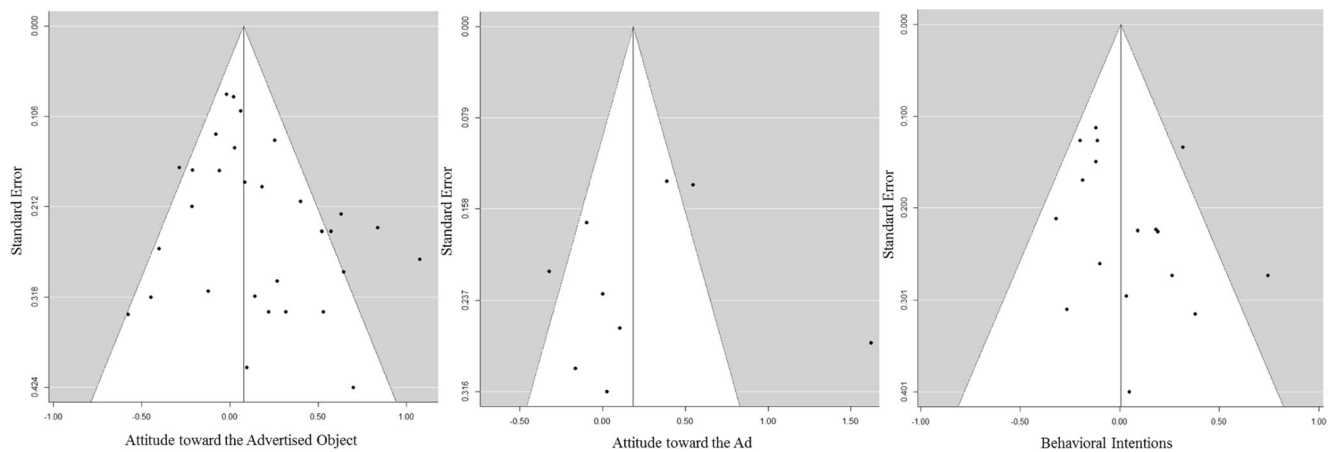
Fig. 2 Funnel plots of the studies in the meta-analysis for various dependent variables

1999). This fits nicely with the notion of Lipsey and Wilson (2001) saying that "contemporary meta-analysis is increasingly attending to the variance of effect size distributions rather than the means of those distributions. That is, the primary question of interest often has to do with identifying the sources of differences in study findings" (p. 8f).

Different dependent measures Surprisingly, almost no average effects were observed for standard measures, such as awareness, attitude toward the ad, or purchase intention. A significant and positive average effect size emerged only for attitude toward the endorsed object. Furthermore, this was the case only when the respective comparison group was not endorsed. This is consistent with the meta-regression results

Table 6 Key findings from the analysis

Hypothesis/research question	Result	Specific finding
H1 (cognitive)	rejected	There were no average effects of celebrity endorsements on consumers' awareness.
H2 (affective)	partially confirmed	Celebrity endorsements positively affect consumers' attitude toward the endorsed object when compared to no endorsement. There were no average effects for attitude toward the ad.
H3 (behavioral)	rejected	There were no average effects of celebrity endorsements on consumers' behavioral intentions.
H4 (endorser sex)	confirmed	Male endorsers perform better than female endorsers do.
H5 (endorser type)	partially confirmed	Actors perform better than models, musicians, and TV hosts do.
H6 (endorser match)	confirmed	Congruent endorsers perform better than incongruent endorsers do.
H7 (endorsement explicitness)	rejected	Implicit endorsements perform better than explicit endorsements do.
H8 (endorsement frequency)	rejected	Endorsement frequency has no impact on consumer responses.
H9 (familiarity)	confirmed	Endorsements of unfamiliar objects perform better than endorsements of familiar objects do.
RQ1 (type of comparison group)	difference	Celebrity endorsements perform best when compared to no endorsement. They perform less well compared to quality seals, awards, or endorser brands.

H: hypothesis; RQ: research question

68

J. of the Acad. Mark. Sci. (2017) 45:55–75

assessing the influence of comparison group's endorsement type on attitude toward the advertised object (Table 3). Celebrity endorsements positively affected consumers' attitudes compared to no endorsement, and this effect was significantly lower and negative when celebrity endorsements were compared to an unknown model or athlete, a quality seal or award, a government employee, or an endorser brand. Hence, several low or zero average effect sizes were at least partially due to being based on comparison groups producing negative effect sizes and on comparison groups producing positive effect sizes. As a result, there were small or no effects on average. Interestingly, although effect sizes lowered significantly in the aforementioned cases (unknown model or athlete, quality seal or award, government employee, endorser brand), celebrity endorsements performed worse only when compared to a quality seal or award, or an endorser brand. In these cases, the effect sizes were negative and significant themselves rather than just significantly lower (Viechtbauer 2010).

Moderators We integrated several endorser variables and the properties of the endorsed object. Overall, the moderators appeared very effective in explaining the between-study heterogeneity accounting for 95–100% of the variability with respect to three key dependent variables (attitude toward the endorsed object, attitude toward the ad, behavioral intention). Moreover, all moderated effects came into existence when controlling for the other moderators (Field 2013). For the endorser variables, the impact of the endorsement frequency on effect size was investigated. Contrary to our assumption, there was no effect. This result contradicts conventional thinking that endorsements' effectiveness is enhanced with increasing repetitions (*Till et al. 2008). Although endorsement effects have been found at single exposures (Ambroise et al. 2014), advertisements are assumed to be learned more thoroughly when exposed multiple times (wearin), as long as consumers are not bored or annoyed by too much repetition (wearout; Campbell and Keller 2003). Wearout seems rather unlikely in this study, as the maximum number of exposures included in the specific analysis was five, and as "five pairings is not so many as to cause subject boredom but is likely to lead to conditioning effects" (*Till et al. 2008). With the maximum number of five repetitions, the amount of variance needed to produce an effect of repetition was most likely too small. For instance, Stuart et al. (1987) included up to 20 repetitions in their conditioning study to test for frequency effects. In addition, it was impossible to control for exposure time since most studies failed to report it. Hence, some studies might have exposed their participants once but for a rather long time, while others may have exposed their participants multiple times in short durations. Both manipulations, repetition and duration, may have resulted in similar effects. Given these limitations, it seems premature to reject the frequency hypothesis. Instead, future studies should account for exposure time,

too. These studies should also look for a possible suppression effect (Koch and Zerback 2013). Specifically, repeated celebrity endorsements may lead to enhanced advertising outcomes while at the same increasing the likelihood of evoking reactance due to too much (forced) exposure. As result, increased reactance may reflect negatively on advertising outcomes and suppress the positive effect of endorsement repetition. That is, the effect of repetition on advertising outcomes may be positive and negative (mediated through reactance) at the same time leading to a zero total effect.

For endorser type, actors performed best followed by athletes and TV hosts, which were followed by models and musicians. The results were quite similar across the three dependent variables. As initially outlined, the enhanced effects of actors may best be explained by consumers being exposed to them audiovisually as well as multiple times over the years. As a result, consumers are likely to develop stronger consumer–celebrity relationships (Klimmt et al. 2006). This is particularly true with a TV series (Hoffner and Buchanan 2005). In addition, actors may generally be more famous, at least compared to models, which explains the comparatively weak effects of the latter.

Relatively strong effects appeared regarding the endorser's sex. Male celebrities evoked substantially stronger effects compared to female celebrities. We attribute the stronger effects to the male spokespersons' greater prestige and expertise resulting in stronger credibility (Kenton 1989; Whittaker 1965). Compared to all previous research (see Erdogan 1999), we are now able to provide practitioners with a clear effect direction. In addition to the sex differences, this study is the first to confirm the product match-up hypothesis on a meta-analytic level. Congruent endorsers produce significantly greater effect sizes compared to incongruent ones. Interestingly, the hypothesis could only be confirmed with regard to attitude toward the endorsed object and behavioral intention. Both refer to evaluations or behavior directly related to the endorsed object. In contrast, no effect was found for attitude toward the ad. Hence, celebrity–object match matters when it comes to attitudes toward the advertised object and purchasing the object. However, celebrity–object match does not necessarily matter when it comes to attitudes toward the ad. This is due to the fact that an object does not necessarily equal the style or imagery of its ad. Both may be completely different. Just think about a company intending to change the image of its brand. Employed ads will likely be quite different from the existing brand image in order to change the image. Hence, a match with an endorsed object is evidently of less importance when evaluating the ad because object and ad do not necessarily equal. In addition, marketers and researchers usually test the match with an endorsed object and not the match with the advertisement itself (e.g., *Chen et al. 2012; Kamins 1990; *Kamins and Gupta 1994).

Effects may arise when explicitly manipulating the match with an ad.

For endorsement explicitness, strong effects appeared, yet opposite the direction we expected. Celebrities implicitly endorsing an object enhanced consumers' attitudes and behavioral intentions substantially more compared to explicit endorsements. The Persuasion Knowledge Model presents a possible explanation (Friestad and Wright 1994). According to this model, consumers develop persuasion knowledge throughout a lifetime of being exposed to persuasive communication. This knowledge is likely to be activated when consumers recognize a persuasive attempt. For instance, consumers tend to recognize persuasive attempts when messages "create the perception that the source was trying 'too hard' to sell his case" (Smith 1977, p. 198). Explicit endorsements then act as some kind of forewarning of the persuasive intent of the endorser. Consumers are more motivated to counterargue the endorsement in order to reassert their freedom (Petty and Cacioppo 1979). This is particularly likely if the endorsed object is of high relevance (Petty and Cacioppo 1979).

Another explanation for this counterintuitive finding may be the fact that implicit endorsement are mostly perceived as merely conveying a celebrity's personal object experience (e.g., by using a brand), whereas explicit endorsements are perceived as conveying a clear recommendation to buy or use an object. As a result, implicit endorsements try to persuade their audience to a lesser extent compared to explicit ones. According to Pornpitakpan (2004), this should enhance endorsers' trustworthiness, leading to stronger effects (see Pornpitakpan 2004). Future studies should test whether implicit endorsers are perceived as more trustworthy and less likely to activate persuasion knowledge.

In terms of familiarity, celebrity endorsements appeared more effective in the case of unfamiliar objects when compared to familiar ones. This result was expected and is in line with past research. Given that consumers already possess a rich network of associations representing an object, attitudes and behavior appear more difficult to change (Cacioppo et al. 1992). That is not to say that outcomes related to familiar objects cannot be influenced at all. The influence is just weaker or more difficult to accomplish.

Theoretical and managerial implications

The hierarchy of advertising effects model by Lavidge and Steiner (1961), which was adapted from Grewal et al. (1997), provided a fruitful theoretical framework. It enabled a systematic organization of all relevant advertising outcomes as well as the integration of relevant moderators (Lipsey and Wilson 2001). In addition, readers can grasp at a glance which relationships have been tested and which relationships have been impossible to test, although they might be theoretically relevant. Future studies can precisely look at these relationships. Furthermore, the model—being a broad overarching framework—enabled the integration of various sub-theories to explain specific effects.

The match-up hypothesis could be supported on a meta-analytic level. The proposed congruency effect can be regarded as quite robust. However, it was not possible to test which of the theoretical explanations is more accurate. In fact, both may be accurate: a matching celebrity being an information source of adaptive significance (Social Adaptation Theory) is also more likely to be easily integrated with existing brand schemas (Schema Theory). Instead of continuing to prove the effect, future research should rather look for its underlying psychological mechanism as well as boundary conditions. It still remains difficult to know which dimensions should be matched between an endorsed object and a celebrity (Amos et al. 2008).

The meta-analysis was also able to support the familiarity proposition. It suggests that attitudes toward unfamiliar objects are easier to change when compared to familiar objects (Cacioppo et al. 1992). That is due to the fact that attitudes are based on attitude-relevant information. If relatively little attitude-relevant information is available in memory, attitudes are primarily based on the information provided by the celebrity endorsement, leading to stronger effects. By contrast, if consumers have access to a relatively large body of attitude-relevant information in memory, attitudes are based on the endorsement *and* information in memory. As a result, endorsement effects are comparably smaller (Cacioppo et al. 1992). Since ads typically feature familiar objects, future research is advised to look for factors boosting celebrity endorsement effects in the case of high familiarity (Kent and Allen 1994). Endorsement repetition or a particular strong consumer–celebrity relationship may present such factors.

In terms of managerial implications, marketers should consider the following findings when choosing their endorser. The findings are accompanied by an assessment of their impact magnitude. Magnitude was assessed according to an effect size investigation from 300 meta-analyses ($d < .30$: small: $d = .50$: medium; $d > .67$: large; Lipsey and Wilson 2001), providing marketers with guidance when deciding on which of the findings to focus more.

1. Matching endorsers elicit more favorable attitudes and stronger behavioral intentions when compared to non-matching ones (impact magnitude: medium).
2. Male endorsers elicit more favorable attitudes and stronger behavioral intentions when compared to female ones (impact magnitude: medium to large).
3. Implicit endorsements elicit more favorable attitudes and stronger behavioral intentions when compared to explicit ones (impact magnitude: medium to large).

4. Actors elicit more favorable attitudes when compared to models, musicians, and TV hosts (impact magnitude: small to medium).
5. Endorsements of unfamiliar objects elicit more favorable attitudes when compared to endorsements of familiar objects (impact magnitude: small).
6. Celebrity endorsements elicit less favorable attitudes when compared to endorsements by quality seals, awards, and endorser brands (impact magnitude: medium).

In general, celebrity endorsements are undoubtedly an effective way of marketing communication. They enhance attitudes and reinforce behavioral intentions, provided marketers choose the right endorser. Marketers have to make these choices carefully, as celebrity endorsements can evoke strong negative outcomes, too. Incongruent male or female celebrities may very likely result in negative effects when endorsing an object explicitly. Marketers are, therefore, advised to back these decisions with market research (Agrawal and Kamakura 1995).

Limitations and agenda for future research

A great advantage of meta-analysis lies in identifying significant gaps in the literature. Such gaps become visible when the meta-analysis cannot report findings on a particular dependent variable or topic. In addition, a meta-analysis may leave blind spots because theoretically relevant moderators cannot be analyzed, as too few studies have investigated these. We discuss these gaps in the following sections.

Understudied dependent variables The most important understudied variables pertain to recognition and recall, meaning transfer, and behavioral measures in general. Beginning with recognition and recall, only about ten effect sizes could be obtained from the literature regarding each measure. This deficit is particularly relevant as marketers are frequently interested in favorable object recognition and recall employing celebrity endorsements (Erfgen et al. 2015). Moreover, the revealed average effect was close to zero (cf. awareness Table 2). This suggests no impact or a moderated impact of celebrity endorsements. Having robust knowledge of one or the other is vital to marketers. "A common concern is that consumers will focus their attention on the celebrity and fail to notice the brand being promoted" (Erdogan 1999, p. 296). Researchers have recently started to investigate this concern, concluding that celebrities might indeed overshadow an endorsed object (Erfgen et al. 2015).

A similar pattern appears when looking at meaning transfer. Marketers frequently seek to transfer celebrity meaning to a brand to build or reposition its image (Keller 2012). However, research investigating celebrity meaning transfer is relatively scarce. It only started a few years ago, resulting in a limited number of effect sizes that could be obtained (k = 9; Galli and Gorn 2011). Though the existing studies suggest strong transfer effects, further research is needed, also specifying boundary conditions (e.g., brand familiarity, *Miller and Allen 2012). Finally, researchers should dedicate more resources to measuring behavior. As seen in Table 2, scholars only measured behavioral intentions. While closely related, behavioral intentions do not fully explain true behavior (Kim and Hunter 1993).

Understudied moderators For the theoretically relevant moderators, our study could not integrate several moderators due to missing coding information, no variability within the moderators, or simply a lack of studies. The advertising vehicles in which the celebrity endorsements were integrated present such a factor (cf. Table 1). Almost all the studies employed print advertisements or similar stimuli. In contrast, very few studies looked at radio, television, or online advertising, rendering moderator analysis impossible (e.g., *Myrick and Evans 2014; *Toncar et al. 2007; *Wei and Lu 2013). This seems even more serious considering that more than 60% of global advertising can be attributed to television and the Internet (Bergkvist and Zhou 2016). Also, exposure time could not be included. Increasing exposure time enhances processing capacity, which may lead to stronger and more durable effects (Petty and Cacioppo 1986). Similarly, few studies provided information about the endorser's valence (positive vs. negative), trustworthiness, attractiveness, or expertise.

Longer-term effects Thus far, almost all studies measured effects immediately after exposure. This is particularly problematic as advertisers are mostly interested in longer-term effects (Eisend and Langner 2010). In addition, various advertising studies have shown that the effectiveness of advertisements varies across time (Bergkvist and Zhou 2016). For instance, Eisend and Langner (2010) were able to show that a celebrity's attractiveness exerts its main impact right after exposure while expertise exhibits its main influence in a delayed situation. They conclude that "effects on attitude toward the brand can considerably differ depending on whether the measurement occurs immediately after ad exposure or with a delay" and that studies would strongly "benefit from including delayed measures in ad testing, particularly when they deal with celebrity endorsers in the advertisements" (Eisend and Langner 2010, p. 543).

Underlying psychological processes In addition, hardly any study looked at the underlying psychological mechanisms of celebrity persuasion (Bergkvist and Zhou 2016). Instead, research was mostly focused on showing main effects or on looking for possible moderators. As a result, theoretical depth is rather low (Bergkvist and Zhou 2016). For instance, there is

no clear knowledge about whether endorsement effects vary depending on high or low effort processing (Petty and Cacioppo 1986). The few studies that integrated processing style suggested stronger effects in low effort processing (*Dong 2015; *Petty et al. 1983; *Sengupta et al. 1997). However, several studies investigating other issues in celebrity endorsements found strong effects when participants fully concentrated on the endorsements (e.g., *Friedman and Friedman 1979; *Kamins and Gupta 1994; *La Ferle and Choi 2005). Further research is needed that clarifies this matter and looks for other possible mechanisms besides message elaboration.

Non-profit advertising Non-profit advertising certainly deserves more attention because it has been steadily growing over the past years including politics, the health sector, or any kind of NGO communication (*Wheeler 2009). Yet research is still very limited. For instance, whether celebrity endorsements can change voting behavior is still understudied (*Pease and Brewer 2008). Accordingly, van Steenburg (2015) recently concluded a review by asking: "Are voters consumers? Can the two be treated similarly when it comes to marketing strategy and the marketing mix? Is selling a candidate the same as selling a car? Do theoretical foundations of consumer behaviour hold in voter behaviour? […] As of yet, all of these represent untapped discoveries" (p. 216). The same applies to the effectiveness of celebrity endorsements in any kind of health or health-related communication as well as environmental communication (Boyland et al. 2013; *Myrick and Evans 2014; *Wu et al. 2012).

Cross-cultural differences So far, the majority of studies has been conducted in the U.S. However, there is undoubtedly a strong interest in celebrity endorsements in such emerging countries as India or China (Chou 2014; Mishra and Mishra 2014). And even though more and more studies are conducted in Asia, it has yet to be tested whether the same mechanisms that apply to Western celebrity endorsements apply to Asian cultures as well. Because a lot of aspects of consumer behavior are culture-bound, culture-adequate methods are urgently needed (de Mooij and Hofstede 2011).

Side effects Finally, future research may also dedicate itself more strongly to side effects or unintended effects. This pertains particularly to vulnerable audiences like, for instance, children or adolescents. It is widely accepted that childhood and adolescence is *the* developmental period during which human beings complete the process of identity formation (Lloyd 2002). Next to their family and friends, they frequently refer to mass media when looking for role models (Hoffner and Buchanan 2005). Celebrities depicted in the media and advertising can serve as such models given that they are considered relevant by consumers (Lockwood and Kunda 1997).

This may pose a problem because children's and adolescents' understanding of advertising may not be as advanced when compared to adults. Specifically, past research has revealed that children do not necessarily comprehend the persuasive intent of advertising and constitute a vulnerable audience, deserving of special protection (Kunkel 2001).

Conclusion

The study sought to quantify the effectiveness of celebrity endorsements on a meta-analytic level across a variety of measures. The results showed a zero effect when averaging across all studies. However, we found strong attitudinal and behavioral effects when including theoretically relevant moderator variables. In particular, effects on attitudes and behavior were found to be strongest when choosing a male actor that matches the endorsed object and expresses his endorsement implicitly. Given the continuing growth of celebrity endorsements in product marketing, politics, and health communication, the study provides essential knowledge to researchers and marketers.

Appendix

Summary of the meta-analysis by Amos et al. (2008)

This meta-analysis focused on source effects of celebrity endorsers on advertising effectiveness. Analyzed source variables were negative information, expertise, attractiveness, credibility, trustworthiness, likeability, familiarity, and performance. Advertising effectiveness was understood rather broadly, combining various measures of effectiveness into one effect size variable (purchase intention, brand attitude, attitude toward advertisement, believability, recall, and recognition). The actual analysis focused on the comparison of effect sizes according to the source variables. In addition, it was tested whether effect sizes significantly differ according to four methodological dimensions: surveys vs. experiments; student vs. non-student samples; U.S. vs. non-U.S. studies; main vs. interaction effects. Thirty two studies published through 2004 were part of the analysis. It included surveys and experiments whereby the majority of effect sizes were obtained from surveys.

Compared to the present analysis, Amos et al. (2008) did not test whether the obtained effect sizes were significant, but solely whether they were significant *different* from each other (according to various source variables). Hence, no results were provided that enable the assertion that celebrity endorsements exert a significant influence as well as an assertion about its size. In addition, there were

no results for individual measures of advertising effectiveness (e.g., cognitive, affective, conative) because all measures were combined into one variable. Furthermore, Amos et al. (2008) neither accounted for dependency among the obtained effect sizes nor provided results in terms of the performance of celebrity endorsements when compared to other kinds of endorsements. Last but not least, results are less clear in terms of a causal interpretation since surveys were included predominantly. We want to point out that we do not perceive our analysis as more valuable, but rather as focusing on completely different aspects as Amos et al. (2008).

References

References marked with an asterisk indicate studies included in the meta-analysis.

Agrawal, J., & Kamakura, W. A. (1995). The economic worth of celebrity endorsers: an event study analysis. *Journal of Marketing, 59*(3), 56–62.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211.

Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of consumer expertise. *Journal of Consumer Research, 13*(4), 411–454.

Ambroise, L., Pantin-Sohier, G., Valette-Florence, P., & Albert, N. (2014). From endorsement to celebrity co-branding: personality transfer. *Journal of Brand Management, 21*(4), 273–285.

Amos, C., Holmes, G., & Strutton, D. (2008). Exploring the relationship between celebrity endorser effects and advertising effectiveness. A quantitative synthesis of effect size. *International Journal of Advertising, 27*(2), 209–234.

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: a meta-analytic review. *British Journal of Social Psychology, 40*(4), 471–499.

*Aureliano-Silva, L., Lopes, E. L. de Lamonica Freire, O. B., & da Silva, D. (2016). The brand's effect on the evaluation of advertising endorsed by celebrities: an experimental study. *Brazilian Business Review, 12*(4), 57–78.

Bargh, J. A., & Pratto, F. (1986). Individual construct accessibility and perceptual selection. *Journal of Experimental Social Psychology, 22*(4), 293–311.

Becker, A. B. (2013). Star power? Advocacy, receptivity, and viewpoints on celebrity involvement in issue politics. *Atlantic. Journal of Communication, 21*(1), 1–16.

Bergkvist, L., & Zhou, K. Q. (2016). Celebrity endorsements: a literature review and research agenda. *International Journal of Advertising, 35*, 642–663.

*Biswas, D., Biswas, A., & Das, N. (2006). The differential effect of celebrity and expert endorsements on consumer risk perceptions. The role of consumer knowledge, perceived congruency, and product technology orientation. *Journal of Advertising, 35*(2), 17–31.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis. Statistics in practice.* Chichester, U.K.: John Wiley & Sons.

Boyland, E. J., Harrold, J. A., Dovey, T. M., Allison, M., Dobson, S., Jacobs, M. C., & Halford, J. C. G. (2013). Food choice and overconsumption: effect of a premium sports celebrity endorser. *The Journal of Pediatrics, 163*(2), 339–343.

Brown, S. P., & Stayman, D. M. (1992). Antecedents and consequences of attitude toward the ad: a meta-analysis. *Journal of Consumer Research, 19*(1), 34–51.

Cabalero, M., Lumpkin, J. R., & Madden, C. S. (1989). Using physical attractiveness as an advertising tool: An empirical test of attraction phenomenon. *Journal of Advertising, 29*(4), 16–22.

Cacioppo, J. T., Marshall-Goodell, B. S., Tassinary, L. G., & Petty, R. E. (1992). Rudimentary determinants of attitudes: classical conditioning is more effective when prior knowledge about the attitude stimulus is low than high. *Journal of Experimental Social Psychology, 28*(3), 207–233.

Campbell, M. C., & Keller, K. L. (2003). Brand familiarity and advertising repetition effects. *Journal of Consumer Research, 30*(2), 292–304.

*Chao, P., Wührer, G., & Werani, T. (2005). Celebrity and foreign brand name as moderators of country-of-origin effects. *International Journal of Advertising, 24*(2), 173–192.

*Chen, C. Y., Lin, Y. H., & Hsiao, C. L. (2012). Celebrity endorsement for sporting events using classical conditioning. *International Journal of Sports Marketing & Sponsorship, 13*(3), 209–219.

Cheung, M. W.-L. (2014). Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. *Psychological Methods, 19*(2), 211–229.

Cho, V. (2010). The endorser's persuasiveness on the purchase intention of high-involvement products: a comparison between a newly launched product and a mature one. *Journal of Global Marketing, 23*(3), 226–242.

Chou, H. Y. (2014). Effects of endorser types in political endorsement advertising. *International Journal of Advertising, 33*(2), 391–414.

Chou, H. Y. (2015). Celebrity political endorsement effects: a perspective on the social distance of political parties. *International Journal of Communication, 9*, 523–546.

Cooper, H. M. (2010). *Research synthesis and meta-analysis. a step-by-step approach. Applied social research methods series* (Vol. Vol. 2). Los Angeles: Sage.

Corbett, J. B., & Mori, M. (1999). Medicine, media, and celebrities: news coverage of breast cancer, 1960–1995. *Journalism and Mass Communication Quarterly, 76*(2), 229–249.

Crutchfield, D. (2010). Celebrity endorsements still push product. Why, in the era of social media, the rewards continue to outweigh the risks. *Advertising Age*, (accessed September 8, 2015), [available at http://adage.com/article/cmo-strategy/marketing-celebrity-endorsements-push-product/146023].

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior, 37*(1), 1–20.

De Mooij, M., & Hofstede, G. (2011). Cross-cultural consumer behavior: a review of research findings. *Journal of International Consumer Marketing, 23*(3), 181–192.

*Dean, D. H., & Biswas, A. (2001). Third-party organization endorsement of products: an advertising cue affecting consumer prepurchase evaluation of goods and services. *Journal of Advertising, 30*(4), 41–57.

Dibble, J. L., Hartmann, T., & Rosaen, S. F. (2016). Parasocial interaction and parasocial relationship: conceptual clarification and a critical assessment of measure. *Human Communication Research, 42*(1), 21–44.

*Dixon, H., Scully, M., Kelly, B., Donovan, R., Chapman, K., & Wakefield, M. (2014). Counter-advertising may reduce parent's susceptibility to front-of-package promotions on unhealthy foods. *Journal of Nutrition Education and Behavior, 46*(6), 467–474.

*Dong, Z. (2015). How to persuade adolescents to use nutrition labels: effects of health consciousness, argument quality, and source credibility. *Asian Journal of Communication, 25*(1), 84–101.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629–634.

Eisend, M. (2006). Two-sided advertising: a meta-analysis. *International Journal of Research in Marketing, 23*(2), 187–198.

Eisend, M. (2009). A meta-analysis of humor in advertising. *Journal of the Academy of Marketing Science, 37*(2), 191–203.

Eisend, M., & Langner, T. (2010). Immediate and delayed advertising effects of celebrity endorsers' attractiveness and competence fit. *International Journal of Advertising, 29*(4), 527–546.

Elberse, A., & Verleun, J. (2012). The economic value of celebrity endorsements. *Journal of Advertising Research, 52*(2), 149–165.

Elliot, R., & Wattanasuwan, K. (1998). Brands as symbolic resources for the construction of identity. *International Journal of Advertising, 17*(2), 131–144.

Erdogan, B. Z. (1999). Celebrity endorsement: a literature review. *Journal of Marketing Management, 15*(4), 291–314.

*Erfgen, C., Zenker, S., & Sattler, H. (2015). The vampire effect: when do celebrity endorsers harm brand recall? *International Journal of Research in Marketing, 32*(2), 155–163.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics. And sex and drugs and rock 'n' roll*. Los Angeles: SAGE.

Field, A. (2015). Dread returns to mega-silly one. *Health Psychology Review, 9*(1), 15–20.

Fireworker, R. B., & Friedman, H. H. (1977). The effects of endorsements on product evaluation. *Decision Sciences, 8*(3), 576–583.

Fleck, N., Korchia, M., & Roy, I. L. (2012). Celebrities in advertising: looking for congruence or likability. *Psychology and Marketing, 29*(9), 651–662.

*Freiden, J. B. (1982). An evaluation of spokesperson and vehicle source effects in advertising. *Journal of Current Issues and Research in Advertising, 5*(1), 77–88.

Freiden, J. B. (1984). Advertising spokesperson effects: an examination of endorser type and gender on two audiences. *Journal of Advertising Research, 24*(5), 33–41.

*Friedman, H. H., & Friedman, L. (1979). Endorser effectiveness by product type. *Journal of Advertising Research, 19*(5), 63–71.

*Friedman, H. H., Termini, S., & Washington, R. (1976). The effectiveness of advertisements utilizing four types of endorsers. *Journal of Advertising, 5*(3), 22–24.

Friestad, M., & Wright, P. (1994). The persuasion knowledge model: how people cope with persuasion attempts. *Journal of Consumer Research, 21*(1), 1–31.

*Frizzell, C. (2011). Public opinion and foreign policy: the effects of celebrity endorsements. *The Social Science Journal, 48*(2), 314–323.

Galli, M., & Gorn, G. (2011). Unconscious transfer of meaning to brands. *Journal of Consumer Psychology, 21*(3), 215–225.

Gorn, G. J. (1982). The effects of music in advertising on choice behavior: a classical conditioning approach. *Journal of Marketing, 46*(1), 94–101.

Grewal, D., Kavanoor, S., Fern, E. F., Costley, C., & Barnes, J. (1997). Comparative versus noncomparative advertising: a meta-analysis. *Journal of Marketing, 61*(4), 1–15.

Hedges, L. V. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*(2), 107–128.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3*(4), 486–504.

Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology, 21*(1), 107–112.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley and Sons.

Higgins, J. P. T., & Green, S. (2011). Cochrane handbook for systematic reviews of interventions version 5.1.0. (accessed September 8, 2015), [available at www.cochrane-handbook.org].

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*(11), 1539–1558.

Hoffner, C., & Buchanan, M. (2005). Young adults' wishful identification with television characters: the role of perceived similarity and character attributes. *Media Psychology, 7*(4), 325–351.

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological Methods, 11*(2), 193–206.

*Jackson, D. J. (2008). Selling politics. *Journal of Political Marketing, 6*(4), 67–83.

*Jackson, D. J., & Darrow, T. I. A. (2005). The influence of celebrity endorsements on young adults' political opinions. *The Harvard International Journal of Press/Politics, 10*(3), 80–98.

Jain, V., Roy, S., Daswani, A., & Sudha, M. (2011). What really works for teenagers: human or fictional celebrity? *Young Consumers, 12*(2), 171–183.

Kahle, L. R., & Homer, P. M. (1985). Physical attractiveness of the celebrity endorser: a social adaptation perspective. *Journal of Consumer Research, 11*(4), 954–961.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.

Kaikati, J. G. (1987). Celebrity advertising: a review and synthesis. *International Journal of Advertising, 6*(2), 93–105.

*Kamins, M. A. (1989). Celebrity and noncelebrity advertising in a two-sided context. *Journal of Advertising Research, 29*(3), 34–43.

Kamins, M. A. (1990). An investigation into the "match-up" hypothesis in celebrity advertising: when beauty may be only skin deep. *Journal of Advertising, 19*(1), 4–13.

*Kamins, M. A., & Gupta, K. (1994). Congruence between spokesperson and product type: a matchup hypothesis perspective. *Psychology and Marketing, 11*(6), 569–586.

Keller, K. L. (2012). *Strategic brand management. Building, measuring, and managing brand equity*. Boston: Pearson.

Kent, R. J., & Allen, C. T. (1994). Competitive interference effects in consumer memory for advertising: the role of brand familiarity. *Journal of Marketing, 58*(3), 97–105.

Kenton, S. B. (1989). Speaker credibility in persuasive business communication: a model which explains gender differences. *Journal of Business Communication, 26*(2), 143–157.

Kilburn, D. (1998). Star power. *Adweek, 39*(2), 20–21.

Kim, M. S., & Hunter, J. E. (1993). Attitude-behavior relations: A meta-analysis of attitudinal relevance and topic. *Journal of Communication, 43*(1), 101–142.

Klimmt, C., Hartmann, T., & Schramm, H. (2006). Parasocial interactions and relationships. In J. Bryant & P. Vorderer (Eds.), *Psychology of entertainment* (pp. S291–S313). Mahwah: Lawrence Erlbaum.

Koch, T., & Zerback, T. (2013). Helpful or harmful? How frequent repetition affects perceived statement credibility. *Journal of Communication, 63*(6), 993–1010.

Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level meta-analysis. *Research Synthesis Methods, 2*(1), 61–76.

Krippendorff, K. (2004). *Content analysis: an introduction to its methodology*. Thousand Oaks, CA: Sage.

Kunkel, D. (2001). Children and television advertising. In D. G. Singer & J. L. Singer (Eds.), *Handbook of children and the media* (pp. 375–393). Thousand Oaks: Sage.

*La Ferle, C., & Choi, S. M. (2005). The importance of perceived endorser credibility in South Korean advertising. *Journal of Current Issues and Research in Advertising, 27*(2), 67–81.

Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication, 50*(1), 46–70.

Lavidge, R. J., & Steiner, G. A. (1961). A model for predictive measurements of advertising effectiveness. *Journal of Marketing, 25*(6), 59–62.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis. Applied social research methods series* (Vol. 49). SAGE: Thousand Oaks.

Lloyd, B. T. (2002). A conceptual framework for examining adolescent identity, media influence, and social development. *Review of General Psychology, 6*(1), 73–91.

Lockwood, P., & Kunda, Z. (1997). Superstars and me: predicting the impact of role models on the self. *Journal of Personality and Social Psychology, 73*(1), 91–103.

Lynch, J., & Schuler, D. (1994). The matchup effect of spokesperson and product congruency: a schema theory interpretation. *Psychology and Marketing, 11*(5), 417–445.

*Maronick, T. J. (2005). Celebrity versus company president as endorsers of high risk products for elderly consumers. *Journal of Promotion Management, 11*(4), 63–80.

*Martín-Santana, J. D., & Beerli-Palacio, A. (2013). Magazine advertising: factors influencing the effectiveness of celebrity advertising. *Journal of Promotion Management, 19*(2), 139–166.

McCracken, G. (1989). Who is the celebrity endorser? Cultural foundations of the endorsement process. *Journal of Consumer Research, 16*(3), 310–321.

*Miller, F. M., & Allen, C. T. (2012). How does celebrity meaning transfer? Investigating the process of meaning transfer with celebrity affiliates and mature brands. *Journal of Consumer Psychology, 22*(3), 443–452.

Mishra, A. A., & Mishra, A. (2014). National vs. local celebrity endorsement and politics. *International Journal of Politics, Culture, and Society, 27*(4), 409–425.

Motyka, S., Grewal, D., Puccinelli, N. M., Roggeveen, A. L., Avnet, T., Daryanto, A., de Ruyter, K., & Wetzels, M. (2014). Regulatory fit: a meta-analytic synthesis. *Journal of Consumer Psychology, 24*(3), 394–410.

Mowen, J. C., & Brown, S. W. (1981). On explaining and predicting the effectiveness of celebrity endorsers. In K. B. Monroe (ed), *Advances in consumer research* (pp. 437–441). Ann Abor, MI: Association for Consumer Research.

*Myrick, J. G., & Evans, S. D. (2014). Do PSAs take a bite out of shark week? The effects of juxtaposing environmental messages with violent images of shark attacks. *Science Communication, 36*(5), 544–569.

*Park, M., Turner, B. A., & Pastore, D. L. (2008). Effective public service advertisements to attract volunteers for the Special Olympics: an elaboration likelihood perspective. *Sport Management Review, 11*, 165–192.

*Pease, A., & Brewer, P. R. (2008). The Oprah factor: the effects of a celebrity endorsement in a presidential primary campaign. *The International Journal of Press/Politics, 13*(4), 386–400.

Petty, R. E., & Cacioppo, J. T. (1979). Effects of forwarning of persuasive intent and involvement on cognitive responses and persuasion. *Personality and Social Psychology, 5*(2), 173–176.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood of persuasion. *Advances in Experimental Social Psychology, 19*, 123–205.

*Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: the moderating role of involvement. *Journal of Consumer Research, 10*(2), 135–146.

Pornpitakpan, C. (2004). The persuasiveness of source credibility: a critical review of five decades' evidence. *Journal of Applied Social Psychology, 34*(2), 243–281.

Pringle, H., & Binet, L. (2005). How marketers can use celebrities to sell more effectively. *Journal of Consumer Behaviour, 4*(3), 201–214.

*Roozen, I., & Claeys, C. (2010). The relative effectiveness of celebrity endorsement for print advertisement. *Review of Business and Economics, 1*, 76–89.

Ross, R. P., Campbell, T., Wright, J. C., Huston, A. C., Rice, M. L., & Turk, P. (1984). When celebrities talk, children listen: an experimental analysis of children's responses to TV ads with celebrity endorsement. *Journal of Applied Developmental Psychology, 5*(3), 185–202.

*Roy, S. (2012). To use the obvious choice: investigating the relative effectiveness of an overexposed celebrity. *Journal of Research for Consumers, 22*, 41–69.

Russell, C. A., & Stern, B. B. (2006). Consumers, characters, and products: a balance model of sitcom product placement effects. *Journal of Advertising, 35*(1), 7–21.

Sanbonmatsu, D. M., & Kardes, F. R. (1988). The effects of physiological arousal on information processing and persuasion. *Journal of Consumer Research, 15*(3), 379–385.

*Sengupta, J., Goodstein, R. C., & Boninger, D. S. (1997). All cues are not created equal: obtaining attitude persistence under low- involvement conditions. *Journal of Consumer Research, 23*(4), 351–361.

Shaughnessy, J. J., & Zechmeister, E. B. (1997). *Research methods in psychology*. New York: McGraw-Hill.

Shimp, T. A., Stuart, E. W., & Engle, R. W. (1991). A program of classical conditioning Experiments testing variations in the conditioned stimulus and context. *Journal of Consumer Research, 18*(1), 1–12.

*Siemens, J. C., Smith, S., Fisher, D., & Jensen, T. D. (2008). Product expertise versus professional expertise: Congruency between an endorser's chosen profession and the endorsed product. *Journal of Targeting, Measurement and Analysis for Marketing, 16*(3), 159–168.

Smith, M. J. (1977). The effects of threats to attitudinal freedom as a function of message quality and initial receiver attitude. *Communication Monographs, 44*(August), 196–206.

*Stafford, M. R., Stafford, T. F., & Day, E. (2002). A contingency approach: the effects of spokesperson type and service type on service advertising perceptions. *Journal of Advertising, 31*(2), 17–35.

Stephens, A., & Rice, A. (1998). Spicing up the message. *Finance Week, 76*(26), 46–47.

Stuart, E. W., Shimp, T. A., & Engle, R. W. (1987). Classical conditioning of consumer attitudes: four experiments in an advertising context. *Journal of Consumer Research, 14*(3), 334–349.

Szymanski, D. M., Troy, L. C., & Bharadwaj, S. G. (1995). Order of entry and business performance: an empirical synthesis and reexamination. *Journal of Marketing, 59*(4), 17–33.

Tellis, G. J. (1988). The price elasticity of selective demand: a meta-analysis of econometric models of sales. *Journal of Marketing Research, 25*(4), 331–341.

Thompson, S. G., & Sharp, S. J. (1999). Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics in Medicine, 18*(20), 2693–2708.

*Till, B. D., Stanley, S. M., & Priluck, R. (2008). Classical conditioning and celebrity endorsers: an examination of belongingness and resistance to extinction. *Psychology and Marketing, 25*(2), 179–196.

*Toncar, M., Reid, J. S., & Anderson, C. E. (2007). Effective spokespersons in a public service announcement. *Journal of Communication Management, 11*(3), 258–275.

Vakratsas, D., & Ambler, T. (1999). How advertising works: what do we really know? *Journal of Marketing, 63*(1), 26–43.

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2014). Meta-analysis of multiple outcomes: a multilevel approach. *Behavior Research Methods, 47*, 1274–1294.

Van Steenburg, E. (2015). Areas of research in political advertising: a review and research agenda. *International Journal of Advertising, 34*(2), 195–231.

Veer, E., Becirovic, I., & Martin, B. A. (2010). If Kate voted Conservative, would you? *European Journal of Marketing, 44*(3/4), 436–450.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the meta-phor package. *Journal of Statistical Software, 36*(3), 1–48.

Viechtbauer, W. (2015). Package metafor. accessed September 8, 2015, [available at https://cran.r-project.org/web/packages/metafor/metafor.pdf].

*Wei, P. S., & Lu, H. P. (2013). An examination of the celebrity endorsements and online customer reviews influence female consumers' shopping behavior. *Computers in Human Behavior, 29*(1), 193–201.

*Wheeler, R. T. (2009). Nonprofit advertising: impact of celebrity connection, involvement and gender on source credibility and intention to volunteer time or donate money. *Journal of Nonprofit & Public Sector Marketing, 21*(1), 80–107.

Whittaker, J. O. (1965). Sex differences and susceptibility to interpersonal persuasion. *The Journal of Social Psychology, 66*, 91–94.

*Wu, W.-Y., Linn, C. T., Fu, C. S., & Sukoco, B. M. (2012). The role of endorsers, framing, and rewards on the effectiveness of dietary supplement advertisements. *Journal of Health Communication, 17*(1), 54–75.

*Yoon, D., & Kim, Y. K. (2016). Effects of self-congruity and source credibility on consumer responses to coffeehouse advertising. *Journal of Hospitality Marketing & Management, 25*(2), 167–196.

*Young, K., & Miller, Y. D. (2015). Keeping it natural: does persuasive magazine content have an effect on young women's intentions for birth? *Women & Health, 55*(4), 447–466.