MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„The Hostile Media Effect: A Preregistered Meta-Analytic Investigation"

verfasst von / submitted by

Julian Krismer, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 840
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Psychologie UG2002
Betreut von / Supervisor:	Mag. Dr. Ulrich Tran, Privatdoz.

Danksagung

An dieser Stelle möchte ich die die Gelegenheit ergreifen, mich bei all jenen Menschen zu bedanken, die mich in meinem Studium und um mein Studium herum begleiteten. Allen voran natürlich meiner Familie und insbesondere meinen Eltern und meinem Bruder, die mir diese Möglichkeit gaben und mich auf vielen Ebenen dabei unterstützten.

Ein großes Dankeschön ist auch an meine FreundInnen aus dem Office gerichtet, die mir zwar auch inhaltlich enorm weiterhelfen konnten, aber, noch viel wichtiger, vor allem auch darüber hinaus. Auch an meine FreundInnen außerhalb des Studiums ist dieser Dank gerichtet, mit denen diese Zeit eine so angenehme werden hat können.

Zusätzlich möchte ich mich noch bei Mag. Dr. Tran bedanken, der diese Arbeit betreute und mir dadurch die Möglichkeit gab, mich inhaltlich weiterzuentwickeln.

The Hostile Media Effect: A Preregistered Meta-Analytic Investigation

Krismer Julian, BSc

University of Vienna

Abstract

The *hostile media effect* (HME)–the tendency that individuals with a strong pre-existing opinion on an issue perceive neutral, even handed news coverage on that issue as biased against their side and in favour of the opposite point of view–has been the subject of intensive research across many different disciplines. First termed by Vallone, Ross, & Lepper in 1985, looking back at 35 years of HME research reveals this phenomenon to be quite robust and observable in many different contexts. The present multilevel meta-analysis aimed at summarizing already existing research and quantifying an overall effect size, investigating possible moderator variables, and applying state-of-the-art analyses for publication bias. Forty-six effect sizes from 23 different studies were synthesized into an overall effect of $g = 0.67$, 95% CI = [0.42; 0.93]. No significant moderating effects could be found for type of HME (*relative* or *absolute*), reach, involvement, and study quality. Year of publication, on the other hand, turned out to be a significant moderator of HME. No indication of publication bias was detected in *p*-curve, *p*-uniform, *p*-uniform*, or Egger's regression test. Although the effect appears to be robust, further research is needed to address differences in study-design and methodological approaches, as well as theoretical conceptualizations of moderating variables and practical implications of this media bias.

Keywords: Hostile Media Effect, meta-analysis, multilevel, media bias, bias.

Abstract (German)

Der Hostile Media Effect (HME) beschreibt ein oft untersuchtes und nachgewiesenes psychologischen Phänomen, wonach Personen, die mit Medienberichten über ein bestimmtes Thema, zu dem sie bereits eine Vormeinung haben, konfrontiert werden, diese Medienberichte als voreingenommen und tendenziös wahrnehmen. Genauer gesagt, wird diesen Berichten unterstellt, die jeweils gegenteilige Meinung zu vertreten. Der HME ist seit der Einführung des Begriffes 1985 im Fokus interdisziplinärer Forschungsansätze, die den HME in verschiedenen Kontexten auf die Probe stellten. Diese Studie ist ein Versuch, die existierende, oft sehr heterogene Literatur in einer Multilevel Metaanalyse zu synthetisieren und einen Gesamteffekt zusammenzufassen, mögliche Moderatorvariablen zu identifizieren und hinsichtlich eines möglichen Publication Bias zu untersuchen. Aus 23 Studien wurden 46 Effektsärken identifiziert, die einen Gesamteffekt von $g = 0.67$, 95% KI = [0.42; 0.93] ergaben. Das Publikationsjahr erwies sich als der einzige signifikanter Moderator, im Gegensatz zum HME-Typus, Involvierung, Reichweite und Studienqualität. Dem neuesten Stand der Technik entsprechen wurden Publication Bias Tests durchgeführt (*p*-curve, *p*-uniform, *p*-uniform* und Egger's Regression Test), die allesamt keine Hinweise auf Publication Bias lieferten. Somit erwies sich der HME als robuster Effekt, der in verschiedenen Kontexten zu beobachten ist. Allerdings ist ein Resümee für zukünftige Forschung zu ziehen, die sich insbesondere den methodischen Unterschieden und der theoretischen Konzeptualisierung widmen sollte.

Schlagwörter: Hostile Media Effect, Metaanalyse, Multilevel, Medien-bias, bias.

Table of Contents

1. Introduction

As accusations of media bias rise in political and public discourse, so does the amount of literature in psychological and communicational sciences investigating the underlying phenomena and the origin of these accusations. Of particular interest is the perception of a media bias against one's own beliefs. In 1985, Vallone, Ross, and Lepper introduced the term *hostile media effect* (HME) in their study on partisans' differences in perceptions of the same media story on the middle east conflict. Their findings suggest that participants either supporting or opposing the two parties involved in this conflict perceived the same, ostensibly even-handed news story to be biased against their side and favouring the opposing side of the isle.

These findings are particularly interesting considering an apparently contradictory study by Lord, Ross, and Lepper (1979), in which they already found that partisans with a strong opinion on a certain issue process information in a biased manner. Specifically, their study showed that these partisans were likely to draw support for their own beliefs from contradictory or indifferent evidence for the issue at hand, a phenomenon they then called *biased assimilation*.

Gunther and Schmitt (2004) tried to resolve this apparent contradiction by arguing that the nature of the medium plays a crucial role in how partisans perceive the message. According to them, the difference between a low-reaching scientific study and a high-reaching mass media report invokes the difference of perceiving a bias in favour or against one's own opinion. This argument is in line with the so-called *third-person effect* (TPE) introduced by Davison (1983), which will be explained in more detail when discussing possible influences of high or low reaching stimuli.

In communication science, perceptions of message biases were most commonly examined in the context of (mass) media and led to what is now also known as the *hostile media bias,* the *hostile media perception,* or the *hostile media phenomenon*.

1.1. The HME in the Past

When searching for what is meant by the HME as put forth by Vallone et al. (1985), we need to take a closer look at their initial study. Their methodological approach included three groups of participants, one being explicitly “pro-israeli”, one being explicitly “pro-arab”, and one neutral group, in the context of the so-called Beirut massacre. The neutral group can be neglected for the moment, as the bias between the two partisan groups is relevant for this phenomenon. The participants had to watch the same, ostensibly neutral news coverage on the Beirut massacre, before rating their perceived bias of the news report on a 9-point Likert-scale (ranging from 1 to 9, with 1 indicating a bias against Israel, 9 indicating a bias in favour of Israel, and 5 indicating no bias at all). Interestingly, the pro-israeli group perceived the news report to biased against their side (*mean* = 2.9), and the pro-arab group perceived the news report to be biased against their side as well (*mean* = 6.7). This significant difference in directional bias of perception was then termed as the HME.

In 2011, Hansen and Kim published a meta-analysis to quantify the summary effect of the HME literature. They found an overall effect of $r = .296$, 95% CI = [.223; .367] after synthesizing reported effect sizes from 34 different studies. A closer look at this particular article reveals significant methodological weaknesses that threaten the robustness of their findings, for example using ordinary least squares (OLS) instead of weighted least squares (WLS) approach for their meta-regression or failing to report (or conduct) the recommended Fisher’s *z*-transformation for correlational effect sizes (Borenstein, Hedges, Higgins, & Rothstein, 2009).

Additionally, looking directly at methodological approaches used in the primary studies included in this analysis showed that different approaches were used to investigate the HME, or even different operationalizations of what was measured.

Similar problems arise when looking at Perloff's (2015) systematic review on the HME literature. Different studies on the HME used different measurements or operationalizations, which necessarily brought up different interpretations of the underlying effect.

Considering these limitations and the diverse body of literature on the HME, a re-evaluation of the overall effect is necessary. But first we have to take a look at how and why the theoretical underpinnings of the HME differ across studies, before carefully operationalizing it in the context of the current study.

1.2. Methodological Differences in HME Research

A lot of follow-up studies on the HME differ slightly or substantially from Vallone et al.'s (1985) specific approach. For example, Gunther and Christen (2002) used a different method to evaluate the HME. In their study they drew from a national probability sample, asked for partisanship on a specific issue and then whether these partisans perceived news coverage (unspecified and without using a controlled stimulus material) on this issue as biased in one direction or the other. The obvious difference in their methodological approach compared to Vallone et al. (1985) is the absence of a controlled stimulus for which the HME is measured. In other words, instead of an experimental investigation, they used a survey without controlling the setting in which the HME was measured.

Another different approach in the context of the HME was used for example by Coe et al. (2008). They investigated perceptions of biases in news coverage and late-night talk shows, namely in *Fox News*, *CNN* and *The Daily Show*. After measuring partisanship via selfcategorization into two groups (democratic or republican), the HME was measured as a kind of general bias perceived in stimuli from these three different sources, but without specific direction. Therefore, the two partisan groups differed in their perceptions of the strength of the bias, but no indication could be drawn whether this bias was perceived in favour of democrats or republicans. Without this direction in the measured bias, it is hard to tell whether this methodological approach really measures bias against one's own point of view, and if it therefore can be called the HME or rather some kind of perceived source bias (Dunn, 2012).

These two examples, along with Vallone et al.'s approach from 1985, make up the most important differences in methodological approaches in the HME literature. Although all of these approaches are measuring at least related phenomena, it is important to consider these distinctions when comparing different studies.

These distinctions are often not considered, especially in Hansen and Kim (2011) and Perloff (2015), two recent studies that quantitatively and qualitatively summarize theoretical considerations and empirical findings on the HME. Although Perloff briefly addresses the difference between experimental study and national survey, he does not articulate whether we are dealing with the same phenomena, even after defining the HME as "the tendency for individuals with a strong pre-existing attitude on an issue to perceive that ostensibly neutral, even-handed media coverage of the topic is biased against their side and in favour of their antagonists' point of view" (Perloff 2015, p. 707). Consequently, according to this definition, Coe et al. (2008) did not really measure HME (as their perceived bias did not have such direction), and neither did Gunther and Christen (2002), because we cannot know whether the media coverage that their participants had in mind while reporting their perceived bias really was "neutral" and "even-handed".

But what about non-neutral news coverage? Apart from the fact that neutrality of news reports is, strictly speaking, very rare and difficult to measure, why would the HME only be observed for neutral stimuli? Surely, we are dealing with the same underlying phenomenon when partisans differ in perception of strength (and direction) of bias in non-neutral news reports instead of neutral ones. Therefore, Gunther, Christen, Liebhart, and Chia (2001) introduced the term *relative* HME in order to deal with this certain kind of bias. Instead of the classical understanding of the HME in the context of neutral news coverage (which became known as the *absolute* HME), the *relative* HME can be observed when the news coverage is not neutral, but biased in any way, and partisans consider the news to be biased in the same direction, but still *relatively* hostile to their own point of view. This distinction is important, because we can now measure the same underlying phenomenon, but without having to assure neutrality of the stimulus. This was not considered in any way by Hansen and Kim (2011).

1.3. The HME in the Present Context

So, for the present analysis, specifications were used to guarantee comparability between different studies. First, only experimental designs were included. Specifically, only studies in which a controlled stimulus (neutral or non-neutral) was used to measure the HME were considered, national surveys in which general bias in media coverage on an issue were excluded.

Secondly, only studies in which the measured bias was directional (implying a bias either in favour or against ones' own point of view) were included. Studies like Coe et al. (2008), in which the measured bias was not direction-specific, were excluded, because a non-directional bias is hardly comparable with what Perfloff (2015) defined the HME to be.

In addition to quantifying the overall effect size, additional moderating variables were investigated in order to explain the expected heterogeneity in the literature, as well as state-of-the-art methods on publication bias.

2. Methods

In order to synthesize the literature on the HME and estimate an overall effect size as well as possible moderating variables, a preregistered meta-analysis was conducted, focussing on distinguishing different methodological approaches in HME studies from one another and on improving past attempts to synthesize effects of media bias. In contrast to Hansen and Kim (2011), the difference between two group means with respective standard deviations was used to calculate the effect size, which was then converted from Cohen's *d* into Hedges's *g* in order to correct for possible overestimated effects of Cohen's *d* in small-sample studies (Borenstein et al., 2009).

2.1. Inclusion Criteria and Eligibility

Studies had to fulfil certain methodological criteria in order to be included in this meta-analysis, following theoretical considerations and to assure comparability between studies. (1) Studies had to use a specific, controlled stimulus for which the HME was tested in both groups. Surveys that did not control for a certain stimulus but investigated bias for general perception of the media on a certain issue were excluded; (2) the scale used for measuring the perceived bias had to be bidirectional, with one end representing bias in favour of one side, and the other end representing bias in favour of the other side; and (3) two partisan group (in favour and against an issue at hand) had to be used in order two compare between-group differences. The effect size was then calculated using group means and standard deviations from the two partisan groups.

Possible third, neutral groups were not considered in the present study (for example in Vallone et al. 1985).

Studies that met the inclusion criteria but lacked crucial information to calculate the effect size, for example failing to report respective standard deviations for mean scores or *F* – Values for two-group-comparisons (e.g. Giner-Sorolla & Chaiken, 1994), had to be excluded.

Grey literature (e.g. master or doctoral dissertations, preprints, and unpublished studies in general) was included alongside published literature to avoid over-representing statistically significant findings, to avoid overestimating effect-sizes and to aim for a more precise estimate of the summary effect in comparison to meta-analyses without grey literature (Conn, Valentine, Cooper, & Rantz, 2003).

2.2. Data Collection

The data collection was carried out in August 2019, finishing on the 19th of August. The following search terms were used to collect studies on the HME, all were applied on Google Scholar and PsycINFO: "Hostile Media Effect" OR "Hostile Media Bias" OR "Hostile Media Perception" OR "Hostile Media Phenomenon". Additionally, a cited reference search was carried out in Google Scholar for all studies that cited Vallone et al. (1985). An originally planned search through EBSCOhost could not be carried out in the end due to lack of accessibility.

The data collection and other guidelines of data analysis and reporting followed the PRISMA statement (Moher et al., 2009).

2.3. Coding Procedure

The following information was extracted from all primary studies: study name, authors, year of publication, numbers of observations per study, topic of interest, the type of HME measured (*relative* or *absolute*), the total sample size across the two groups, mean scores, standard deviations or standard errors (if reported) and sample sizes in each group, *t*-values or *F*-values for two groups if standard deviations were not reported, reach, involvement, and study quality, measured via the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies (National Heart, Lung, & Blood Institute, 2014).

2.4. Methods of Analysis

2.4.1. Summary effect. The overall effect size of the HME was estimated and reported in Hedges' *g.* First, group means and standard deviations (or *F* - values for $df = 1$) were used for each individual observation to estimate Cohen's *d,* which was then converted into Hedges' *g.* As multiple effect sizes were reported in most of the primary studies, a three-level mixed effects model was applied to account for effect size dependency (Cheung, 2014). In this model, the sampling variation for each effect size is on level 1, the variation for effect sizes within each study on level 2, and the variation of effect sizes across studies on level 3.

2.4.2. Model-fit. In order to verify whether the three-level model is the best fit for the dataset, a test for significance for within-study variance (level 2) and between-study variance (level 3) was conducted via two separate log-likelihood-ratio tests. For each of these two, the fit for the original model with freely estimated variances over all three levels was compared to the fit of the model in which the variance on the second or third level was set to zero (Assink, & Wibbelink, 2016).

2.4.3. Heterogeneity and level-specific variance. In addition to testing for significance of variance components on each individual level, the amount of variance distributed on each level out of the total variance can be quantified in multi-level approaches, using formula 14 on page 215 from Cheung (2014). This can be used as an addition to log-likelihood-ratio tests in multilevel analyses, which can lack statistical power without a high number of primary studies or effect sizes (Assink, & Wibbelink, 2016).

2.4.4. Moderator analysis. A common method for investigating sources of heterogeneity in meta-analyses is to examine the influence of moderator variables (Borenstein et al., 2009). In other words, in addition to an estimation of the overall effect size it is common to investigate why the effect size differs from study to study. In the context of the HME, Perloff (2015) addressed three specific moderator variables: reach, involvement and social identity. Additionally, Hansen and Kim (2011) quantified the moderating effects of involvement, medium (for which the HME was examined), and method (the context of the study being an experiment or a national survey), although their results remain questionable, as they used OLS approach in their meta-regression instead of WLS, which should be applied in a meta-analytic design. In the present study, a number of possible moderator variables were investigated.

2.4.4.1. Absolute or relative HME. Although *relative* and *absolute* HME can be attributed to the same underlying phenomenon, it is possible that they differ in their effect size. This difference will therefore be investigated as a possible moderating variable.

The difference between *absolute* and *relative* HME is usually made along the neutrality of the stimulus (Gunther et al., 2001). In the present study, the neutrality of the news report was not considered. Instead, the difference between the *absolute* and the *relative* HME was made by investigating whether the two group means are placed on the same side of the HME-scale (indicating that the bias is perceived in the same direction, but still relatively more hostile to one's own position), or on different sides (indicating an *absolute* HME). The *absolute* and the *relative* HME were divided into two categories, and their possible difference in effect size and the significance of this categorization was tested in a subgroup analysis (Borenstein et al., 2009).

2.4.4.2. Involvement. The question whether HME is present in the entire population or only among highly involved individuals can be traced back to Vallone et al. (1985). This line of reasoning was followed for example by Giner-Sorolla and Chaiken (1994) and by Gunther et al. (2001), the latter addressing the issue of investigating whether HME is only observable in highly involved partisans. Therefore, a lot of following studies only considered highly involved partisans and excluded participants which did not hold strong pre-existing attitudes. This problem has then been addressed as a focal point of HME research (Gunther et al., 2001) and Hansen and Kim (2011) investigated differences in HME of high-level involvement, medium-level involvement, and low-level involvement methodological approaches. Although involvement resulted in their only significant moderating variable (b = .176, p = .045), this finding remains questionable due to inaccurate model usage and lack of transparency of their involvement rating. To code for the three levels of involvement, they regarded samples that were drawn from specific groups of interest as highly involved and randomly selected participants as lowly involved, with a sample that lies “between these two extremes” (Hansen & Kim, 2011, page. 173) being assigned a medium-level involvement.

For the present study, a different approach to classifying participants’ involvement in primary studies was used, following several theoretical considerations: (1) the differences in sampling procedure (interest groups or “random” sample) might not be the best way to classify low-level involvement, as “randomly” selected participants might very well be highly involved in prominent topics (like general issues surrounding political topics, etc.); (2) although they gave examples for studies that were coded as high or low involvement of the respective participants, their coding procedure, especially their judgement of connectedness between subjects and issues, remains unclear and difficult to replicate (Perloff, 2015); and (3) they completely ignored pre-test evaluations of partisanship that were applied in virtually all primary studies and used bidirectional Likert-scales, which also measures at least a similar construct to involvement. After all, a subject's personal involvement with an issue might be multi-dimensional (considering emotional or intellectual involvement, etc.) and therefore remains difficult to evaluate, especially in a meta-analytic context.

Nevertheless, available information in primary studies (the pre-test of evaluating strength of partisanship) opens up the possibility of judging some kind of connectedness between subjects and issues. So, for the present study, this information was used to code for what is usually called "involvement" in the context of the HME (might also be considered as "strength of partisanship"). The two partisan groups were therefore coded as high in involvement when (1) a pre-test of partisanship was implemented via a bidirectional Likert-scale; and (2) only those participants scoring very high and very low on this scale were included in the analysis, with those scoring at the middle point (and close to the middle) being considered as neutral and excluded from the analysis. Observations that did not follow this specific pre-test procedures were coded low in involvement due to lack of possible definitions on how low and medium involvement might differ from one another. Possible differences between these two groups were then investigated via a subgroup analysis.

In addition to already addressed problems with operationalizing what is called involvement, this approach was used in most of the HME investigations at hand (34 out of 46 observations) and brings another problem with it, namely being an arbitrary categorization (in this case: a dichotomization) of a continuous variable (Altman & Royston, 2006). Although a convenient simplification is gained both in statistical analysis and interpretation, it has severe drawbacks, for example reducing statistical power in already small sampled studies and underestimating variation in the relationship of predictor and outcome.

2.4.4.3. Reach. Perloff (2015) mentions reach as a possible moderating variable in the context of HME. This interpretation can be traced back to 1983, when Davison introduced the already mentioned TPE into the communicational science research, which is used to describe the tendency of individuals to perceive other people (third persons) to be more likely to be influenced or persuaded by mass media, or media in general. Gunther and Mundy (1993) suggested that this effect might be bigger in the context of harmful stimuli, in contrast to stimuli that is in line with individuals' beliefs. In HME studies, this work can serve as possible explaining factor of the HME, when individuals perceive like-minded "third-persons" to be persuaded by news coverage that is not in line with their opinion and that is therefore viewed as hostile news.

Theoretical frameworks like the TPE assume that the HME depends on the stimulus reaching a broad and heterogeneous audience. Similar to involvement, it is difficult to judge whether participants assume that a stimulus material reaches a high number of people. Gunther and Schmitt (2004) tried manipulating the stimulus materials context in order to test whether the HME differs when high or low reach is assumed by the participants. Similar approaches were used by Gunther and Liebhart (2006), both indicating lower HME scores in the low-reach conditions.

For the present study, differences in reach were included as a possible moderating variable. Therefore, stimulus material was classified as high reaching when it was presented to imply reaching a broad audience (being framed as a newspaper article, either verbally or visually). If the stimulus was framed as not reaching a broader audience (for example being explicitly presented as a student essay), study reach was coded as low.

2.4.4.4. Study quality. The problem of varying quality in primary studies can pose a serious threat for the validity of meta-analyses and systematic review (Moher et al., 1998) and has therefore been of special interest in the literature on synthesizing findings from different studies (Cooper, Hedges, & Valentine, 2009; Card, 2011). In a replication project for 11 randomly selected meta-analyses, Moher et al. (1998) found that studies being rated lower in quality tend to overestimate the effect size, while studies with higher quality tend to find smaller effects. In order to assess whether a systematic difference in estimating the effect exists in HME literature, the moderating effect on quality in primary studies on the HME was also investigated in the current meta-analysis.

Quality of primary studies can often be difficult to assess for meta-analyses and systematic reviews (Card, 2011). Although a lot of different checklists or manuals for reported items and study quality exist (Page, McKenzie, & Higgins, 2017), most of them are best suited for clinical, intervention type methodological approaches. With the current meta-analysis dealing with non-randomized, observational primary studies, their quality was assessed via the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies developed by the National Heart, Lung, and Blood Institute (2014). This checklist consists of 14 items that address different indicators of study quality, like reported sample size justification or methodological soundness. Because this quality assessment tool was originally designed for clinical studies, it was slightly modified to fit the present study. Although difficulties still remain on assessing study quality unidimensionally (Card, 2011), the current quality rating seems best fit for the primary studies involved in this synthesis but has to be interpreted with caution.

2.4.4.5. Year of publication. The often-observable change of effect sizes in meta-analyses over time has increasingly drawn attention in different fields and is also known as the Proteus phenomenon (Ioannidis & Trikalonis, 2005). The Proteus phenomenon is used to describe meta-analytic data, in which extreme and often contradictory results can be observed in early stages. Pietschnig, Siegel, Eder, and Gittler (2019) found that effect sizes are more likely to decline than to increase over time at a ratio of 2:1, at least for research on intelligence. In order to investigate whether changes of the effects size are observable in HME literature, the primary studies' year of publication was included as a possible moderating variable and analysed in a meta-regression.

2.4.5. Publication bias. Publication bias, the tendency that studies are more likely to be published based on the strength or direction of their findings, can be observed across many disciplines in academic research (Dickersin, 1990). So far, the HME literature has not been examined for possible publication bias.

For the current study, examination of *p*-values in order to test for publication bias were carried out with state-of-the-art methods, including (1) *p*-curve (Simonsohn, Nelson, & Simmons, 2014), which is used to look at the distribution of significant *p*-values for a certain effect size for suspicious shapes; (2) *p*-uniform (van Assen, van Aert, & Wicherts, 2015), estimating a summary effect size under account for possible biases for significant *p*-values; and (3) *p*-uniform* (van Aert, & van Assen, 2018), correcting for possible overestimation of effect sizes given by *p*-uniform and also including nonsignificant *p*-values.

In addition to methods based on *p*-values, Sterne & Egger (2005) proposed a meta-regression to examine systematic connections between effect size reported in primary studies and their respective standard errors. A possible connection of this kind was also examined in the present study.

2.4.6. Software. For statistical analysis, R version 3.6.0 was used as well as the R package *metafor* version 2.1-0 (Viechtbauer, 2010) for summary effect size estimation, plotting and moderator analysis. For publication bias analyses, the *puniform* package (van Aert, 2019) version 0.1.2 and the *meta* package (Balduzzi, Rücker, & Schwarzer, 2019) version 4.9-9 were used (see Appendix for R-code).

2.4.7. Family-wise error correction. Because multiple hypothesis tests were carried out in the present study, it is important to consider family-wise error rates (FWER). Therefore, adjusted *p*-values using Holm's sequential Bonferroni procedure (Holm, 1979) will be reported rather than original *p*-values in an attempt to avoid alpha-error inflation.

3. Results

3.1. Study Selection

The search process resulted in 1160 studies, after removing duplicates, that were screened for eligibility. One-hundred and thirty-three full texts were then analysed for eligibility and resulted in a final sample of 23 studies (see PRISMA flow diagram, Figure 1), from which 46 observed effect sizes were estimated (see Table A1, with other important study characteristics). The total number of included studies is substantially smaller than the previous meta-analysis on the HME by Hansen and Kim (2011), who included 34 primary studies in their analysis. This is the result of the rather strict inclusion criteria implied in the present study.

The final sample consisted of 12 studies that were also included in Hansen and Kim's (2011) final sample and 11 studies that are not found in the previous meta-analysis. The 22 remaining studies in Hansen and Kim (2011) were excluded from the current investigation for different reasons.

(1) The first reason for exclusion was if the bias was measured without a specific direction. Studies were excluded that only measured differences in strength of perceived bias between two partisan groups without any information on the direction in which this bias was perceived (Coe et al., 2008; D'Alessio, 2005; Feldman, 2008).

(2) The second exclusion criterion was if the studies did not report enough information needed to calculate an effect size. Often, important values like mean scores, standard deviations, or *F*-values were not reported (Choi, Yang, & Chang, 2009; Giner-Sorolla & Chaiken, 1994).

(3) The lack of a controlled stimulus was the third exclusion criteria. If the bias was measured in the context of general news coverage, in opposition to a controlled stimulus (like a newspaper article), the study had to be excluded (Dalton, Beck, & Huckfeldt, 1998; Gunther & Chia, 2001; Huge & Glynn, 2010; Tsfati & Cohen, 2005; Tsfati, 2007).

(4) The measurement of partisanship was also a necessary requirement. Studies that measured a generally perceived bias without a difference between two partisan groups were excluded (Arpan & Raney, 2003; Gunther, Miller, & Liebhart, 2009).

(5) Finally, the last group of exluded studies, were either irretrievable, not found during the search procedure or measured a different construct, that is not comparable in a meta-analysis (Arpan, Bae, Chen, & Greene, 2009; Choi & Chang, 2007; Gunther, 1992; Hoffner & Toohey, 2007; Jeon, 2009; Kinnally, 2008; Peng, 2005; Perloff, 1989; Schmitt, Gunther, & Liebhart, 2004; Yan et al., 2007).

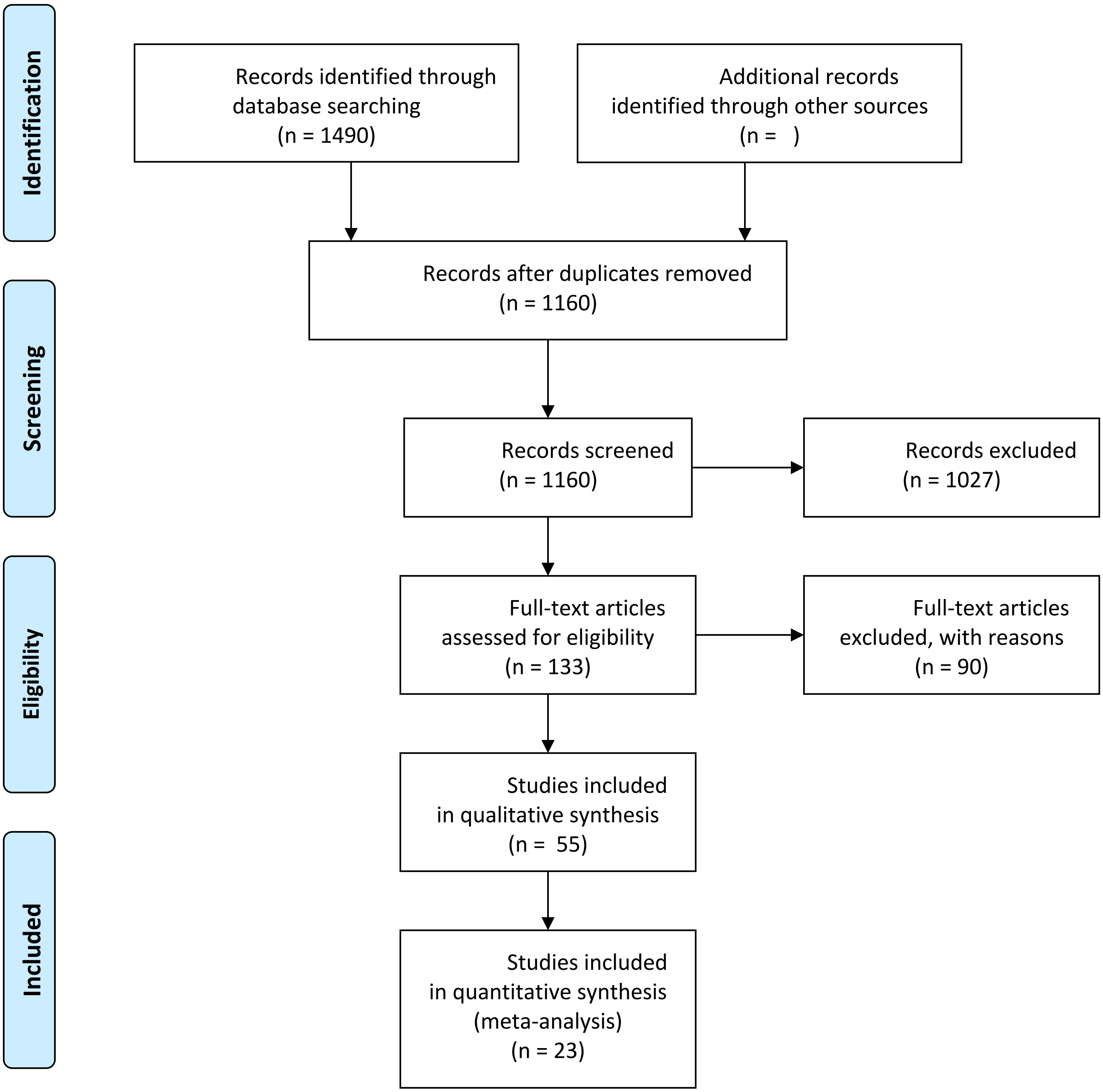

Figure 1: Prisma flow diagram containing information on number of studies included and excluded.

3.2. Summary Effect

The results for the summary effect estimated by the three-level model are summarized in Table 1, with an overall effect of $g = 0.67$, 95% CI = [0.42; 0.93], $p < .0001$. Figure 2 depicts all effects sizes from different studies and observations as well as an estimated summary effect in a forest plot.

Table 1

Results from the three-level model for the overall effect (k = 46)

	g	**SE**	**CI.lb**	**CI.ub**	***p*-value**	**QE**
Model results	0.67	0.13	0.42	0.93	< .0001*	414.41

Note: Results from the estimation of the overall effect including corresponding standard error (SE), 95% confidence interval (CI.lb = lower boundary, CI.ub = upper boundary), *p*-value (significant) and test for heterogeneity (QE).

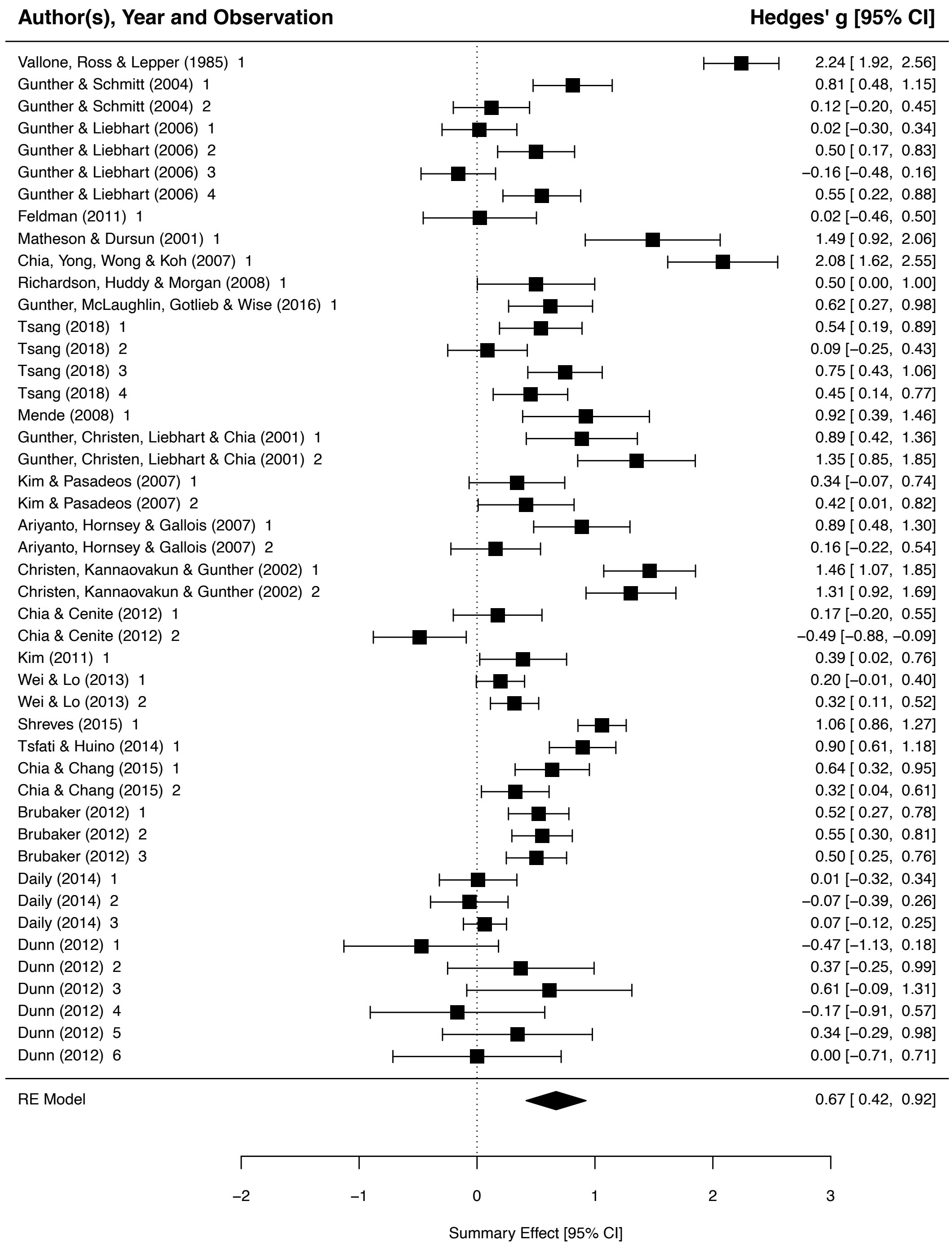

Figure 2. Forest plot of the summary effect estimated by the three-level model. *Observations* refers to the number of observations within one primary study. Primary studies' Hedges' *g* with respective CI's are mentioned on the right side of the plot. 46 observations from 23 studies were included in the final model.

3.3. Model Fit

The fit of the three-level model was conducted via two separate log-likelihood-ratio tests, showing significant results for variance on level two (within-study variance) and on level three (between-study variance), indicating that the three-level model seems to be the best fit for the data in the present study.

The results for the within-study variance are summarized in Table 2, indicating a significantly better fit for the three-level model than for a two-level model with the variance on level two set to zero (p = .0009).

Table 2

Log-likelihood-ratio test for within-study variance (level two)

	df	**LRT**	**AIC**	**BIC**	***p*-value**	**QE**
Full	3		69.35	74.77		414.41
Reduced	2	11.08	78.43	82.04	.0009*	414.41

Note: Test for within-study variance. The full model represents the three-level model, whereas the reduced model refers to an estimated two-level model. df refers to the degrees of freedom, the reduced model having one degree of freedom less, because the within-study variance is not represented in this model. LRT = result from the log-likelihood-ratio test statistic. *p*-value indicating significant within-study variance. QE = heterogeneity for all effect size in the data.

Results for the test statistic for between-study variance (level three) are summarized in Table 3, again indicating a better fit for the three-level model than for a two-level model with between-study variance set to zero (p = .0002).

Table 3

Log-likelihood-ratio test for between-study variance (level three)

	df	LRT	AIC	BIC	*p*-value	QE
Full	3		69.35	74.77		414.41
Reduced	2	14.21	81.57	85.18	.0002*	414.41

Note: Test for within-study variance. The full model represents the three-level model, whereas the reduced model refers to an estimated two-level model. df refers to the degrees of freedom, the reduced model having one degree of freedom less, because the within-study variance is not represented in this model. LRT = result from the log-likelihood-ratio test statistic. *p*-value indicating significant within-study variance. QE = heterogeneity for all effect size in the data.

3.4. Heterogeneity and Level-Specific Variance

In addition to examining whether significant variance can be observed on all three levels, it is possible to quantify the amount of variance on each level out of the total variance in the data (Assink, & Wibbelink, 2016). Results from the test statistic recommended by Assink and Wibbelink (2016) estimated the typical sampling variance (level one) to amount to 7.59%, the within-study variance (level two) to 12.66%, and the between-study variance (level three) to 79.74% out of the total variance in the dataset.

3.5. Moderator Analysis

An overview of results from the moderator analysis is given in Table 4. The meta regression for year of publication was the only moderator variable for which a significant interaction could be found.

Table 4

Results from the moderator analysis

Variable	***p***	***intercept***	***slope***
abs. or rel. HME	.34	0.79 [0.52; 1.05]	-0.26 [-0.54; 0.007]
reach	.07	0.71 [0.45; 0.96]	-0.41 [-0.72; -0.10]
involvement	1.0	0.62 [0.32; 0.92]	0.21 [-0.38; 0.8]
year	.002*	113.12 [55.31; 170.92]	-0.06 [-0.09; -0.03]
study quality	1.0	0.32 [-0.73; 1.38]	0.06 [-0.11; 0.22]

Note: Results from the moderator analyses. Values of intercepts and slopes are given with 95% CI in parentheses. The first 3 moderators were examined via subgroup analyses, the last two via meta regressions. The reference group at the intercept for *absolute* and *relative* HME is the *absolute* HME, and for reach and involvement it is the group of studies coded as high.

3.5.1. Absolute or relative HME. Following the pre-registered plan to investigate possible differences in effect size of the *absolute* and the *relative* HME, a subgroup analysis was conducted for these effects. The results from the present study indicate no moderating effect for the type of HME (*relative* or *absolute*) measured (p = .34).

Although not significantly differing from each other, differences in the size of the *relative* and the *absolute* HME are summarized in Figure 3 and Figure 4.

Note that the last observation in Dunn (2012) was excluded for this moderator analysis, because this specific effect size turned out to be zero (g = 0, 95% CI = [-0.71; 0.71]) and could therefore not be categorized as *relative* or *absolute*.

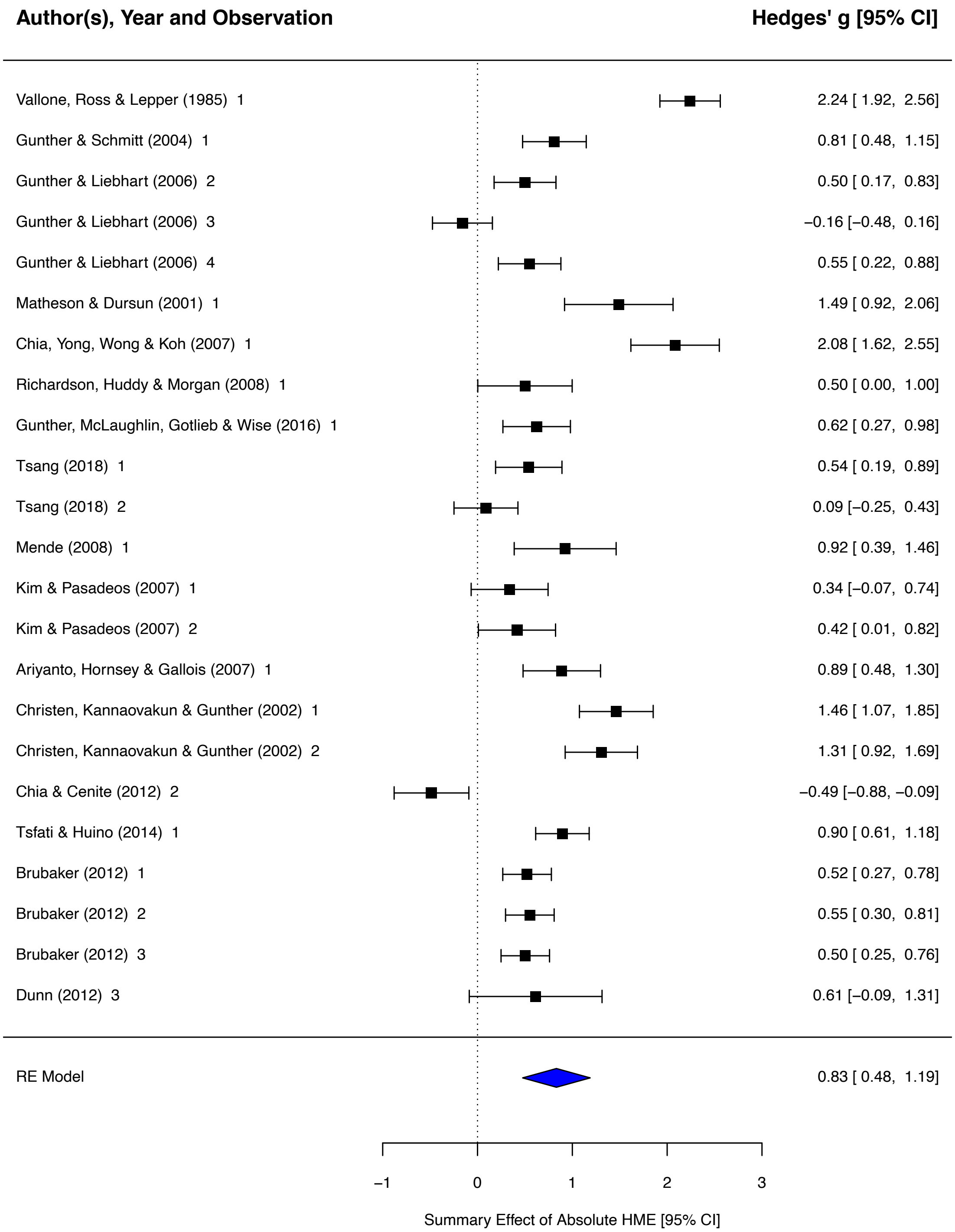

Figure 3. Summary effect of the *absolute* HME. 23 observations out of 16 studies were coded as measuring the *absolute* HME, resulting in an overall *absolute* HME of $g = 0.83$, 95% CI = [0.48; 1.19].

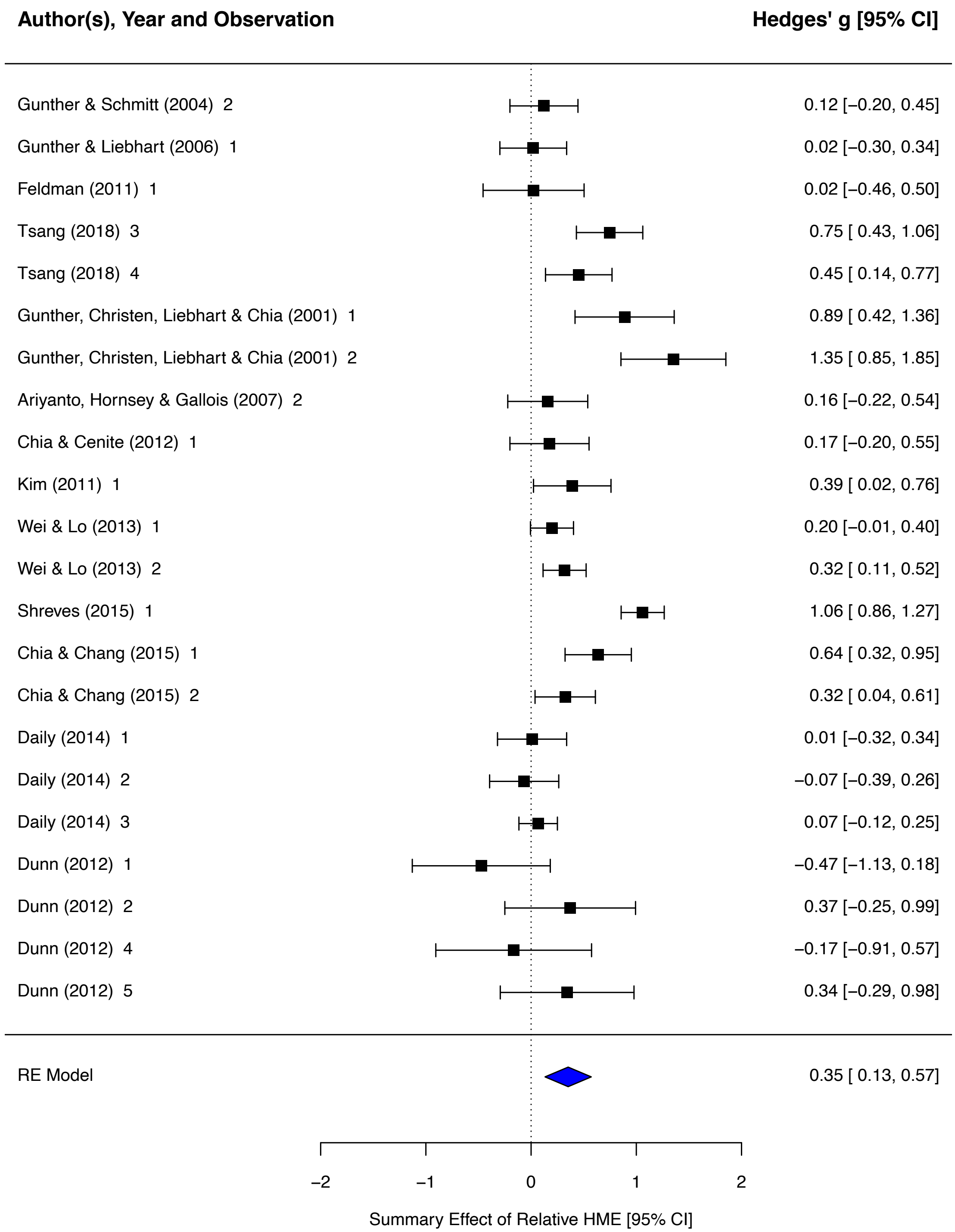

Figure 4. Summary effect of the *relative* HME. 22 observations out of 13 studies were coded as measuring the *relative* HME, resulting in an overall *relative* HME of $g = 0.35$, 95% CI = [0.13; 0.57].

3.5.2. Involvement. Like the type of HME being measured, the results from the subgroup analysis for involvement do not support involvement being a moderating variable for the HME ($p = 1$). Differences in the overall effect size for studies coded as high or low in involvement are visualized in Figure 5 and Figure 6.

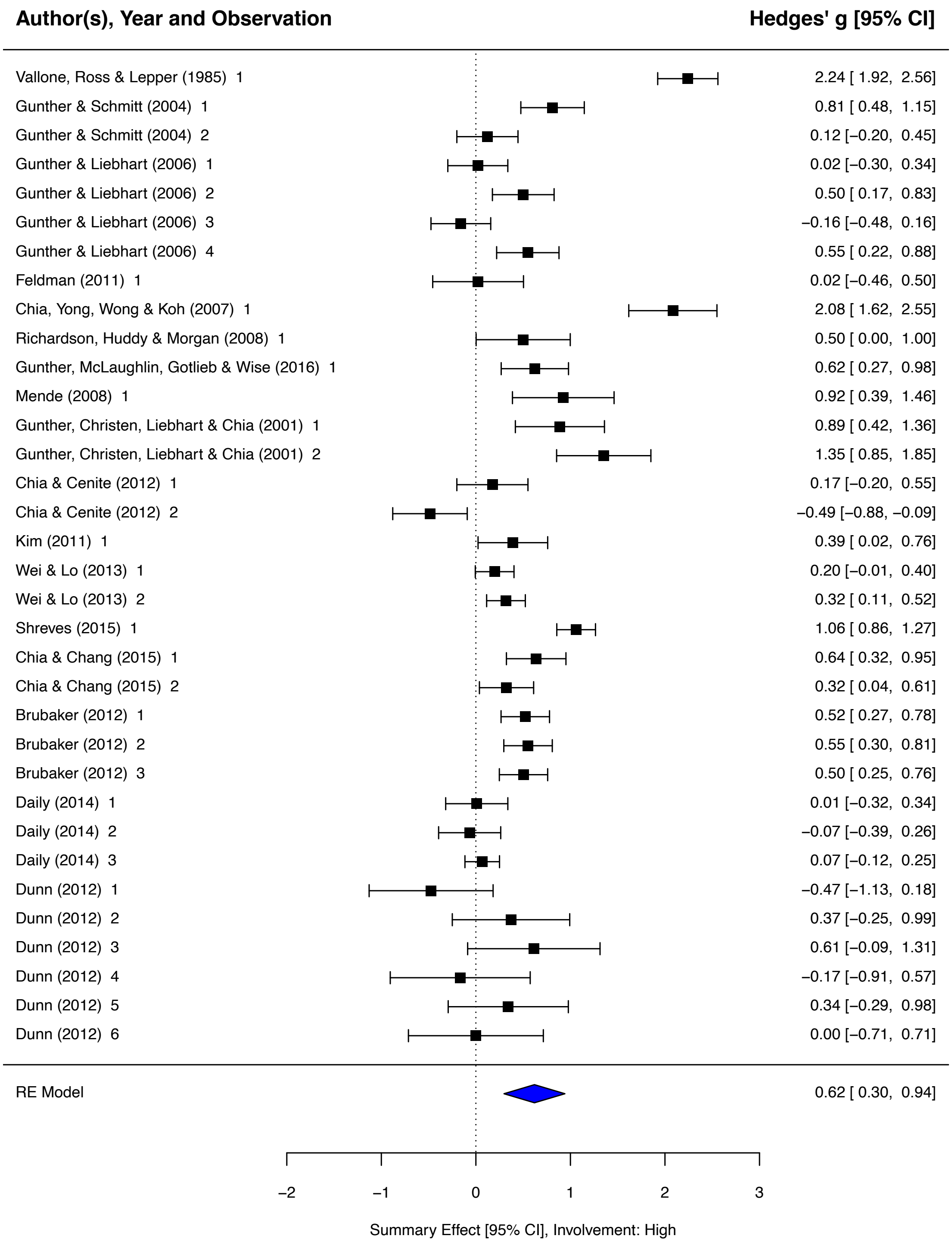

Figure 5. Summary Effect of studies coded as high in involvement. 34 observations out of 17 studies were included in the analysis, resulting in an overall effect of $g = 0.62$, 95% CI = [0.3; 0.94].

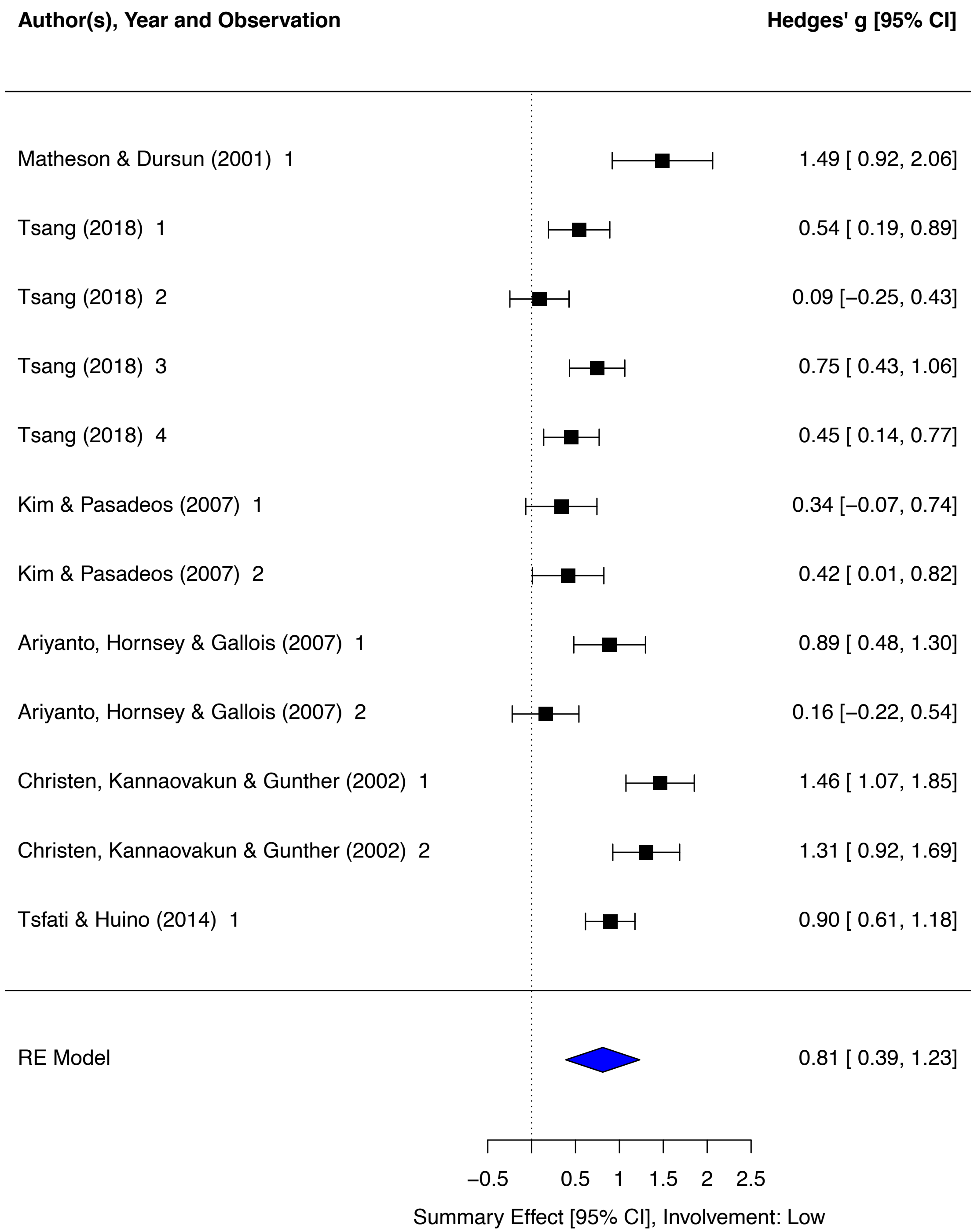

Figure 6. Summary Effect of studies coded as high in involvement. 12 observations out of 6 studies were included in the analysis, resulting in an overall effect of g = 0.81, 95% CI = [0.39; 1.23].

3.5.3. Reach. The results from the subgroup analysis for reach indicate no significant influence on the HME (p = .07). The overall effects for studies coded high or low in reach are summarized in Figure 7 and Figure 8.

Notice that only 5 studies were coded as low in reach, calling for caution in interpreting the results and the robustness of this effect size.

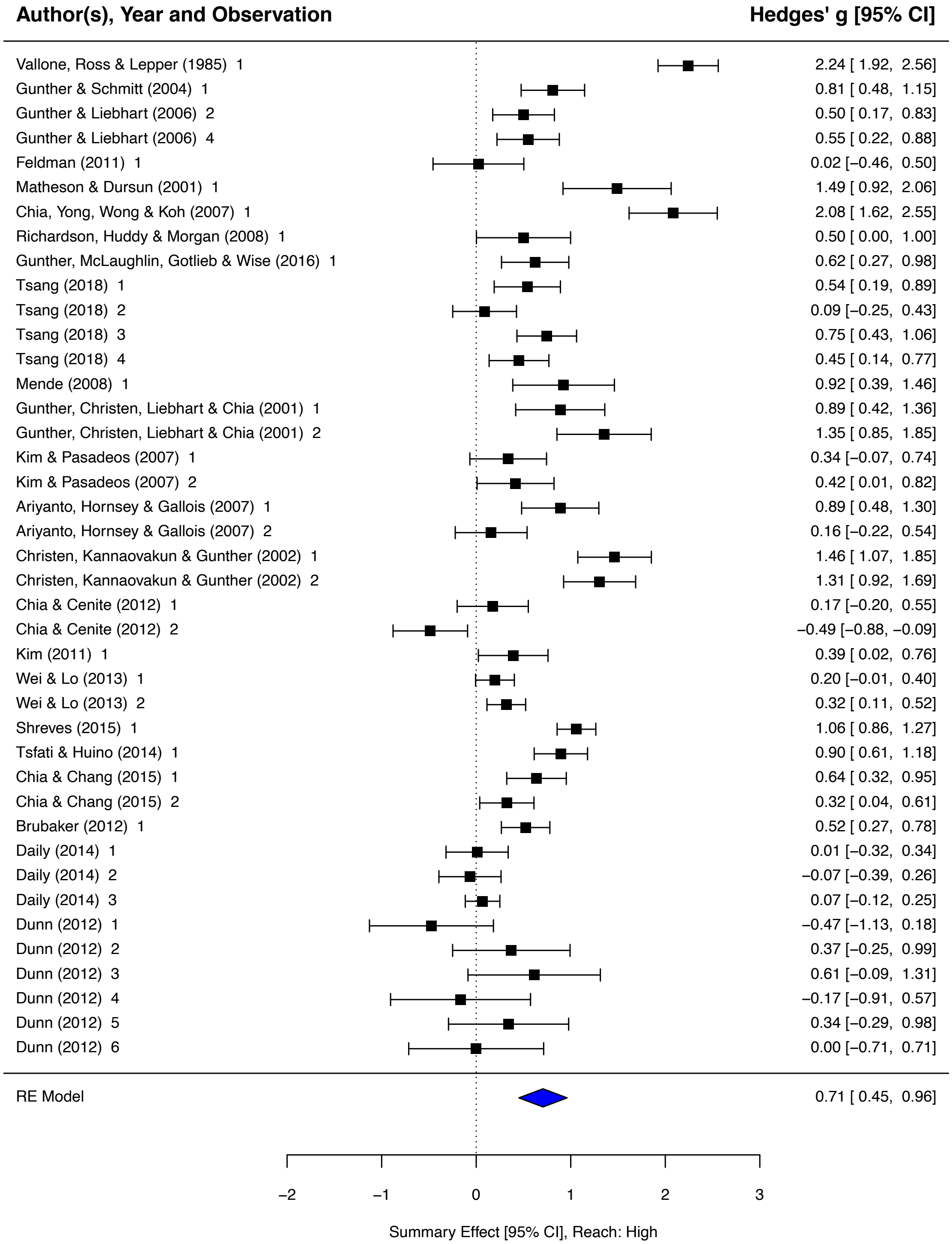

Figure 7. Summary Effect of studies coded as high in reach. 41 observations out of 22 studies were included in the analysis, resulting in an overall effect of g = 0.71, 95% CI = [0.45; 0.96].

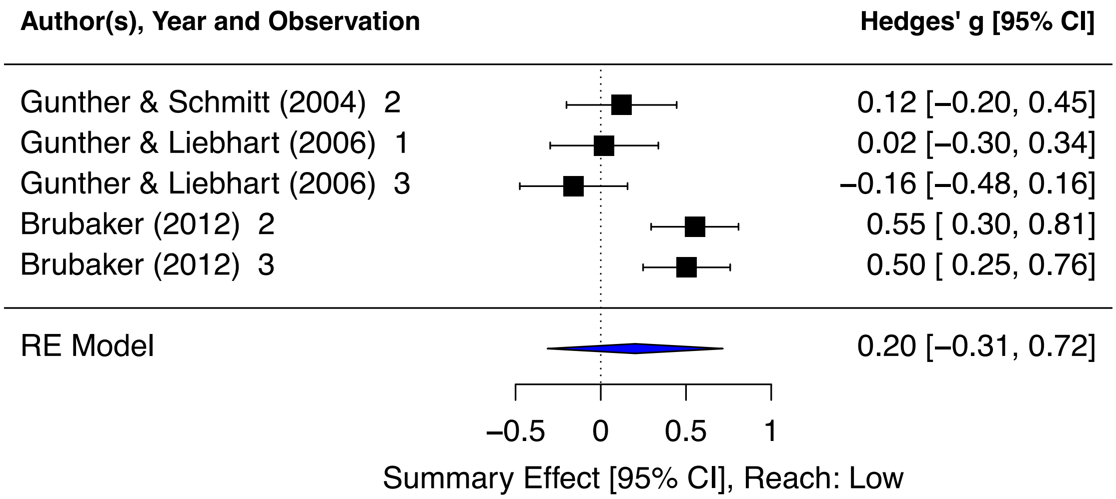

Figure 8. Summary Effect of studies coded as low in reach. 5 observations out of 3 studies were included in the analysis, resulting in an overall effect of $g = 0.2$, 95% CI = [-0.31; 0.72].

3.5.4. Study quality. Again, results from the meta-regression for study quality revealed no significant moderating effect of study quality on the size of the HME ($b = 0.06$, $p = 1$). The results from the meta-regression are visualized in Figure 9.

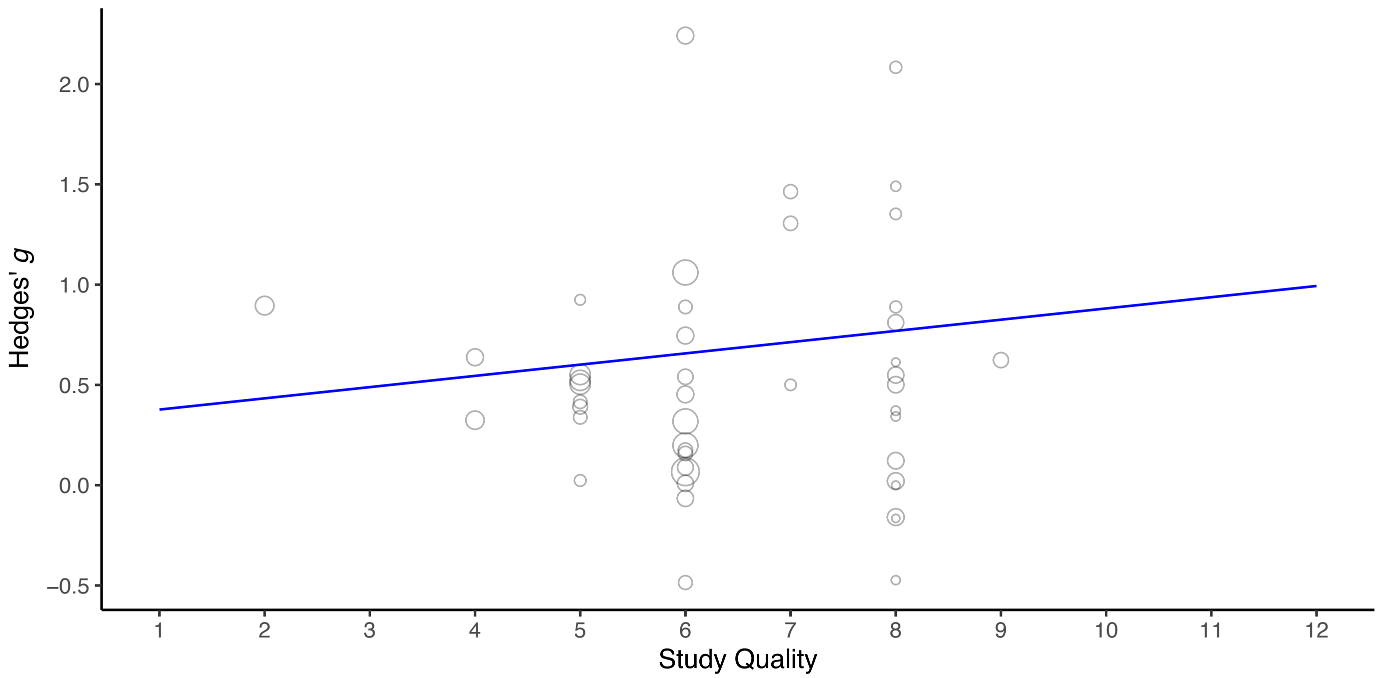

Figure 9. Results from the meta-regression for study quality: $b = 0.06$, $p = 1$.

3.5.5. Year of publication. The test for year of publication indicated a significant moderating effect ($b = -0.06$, $p = 0.002$, see Figure 10), showing that the size of the effect decreased over time.

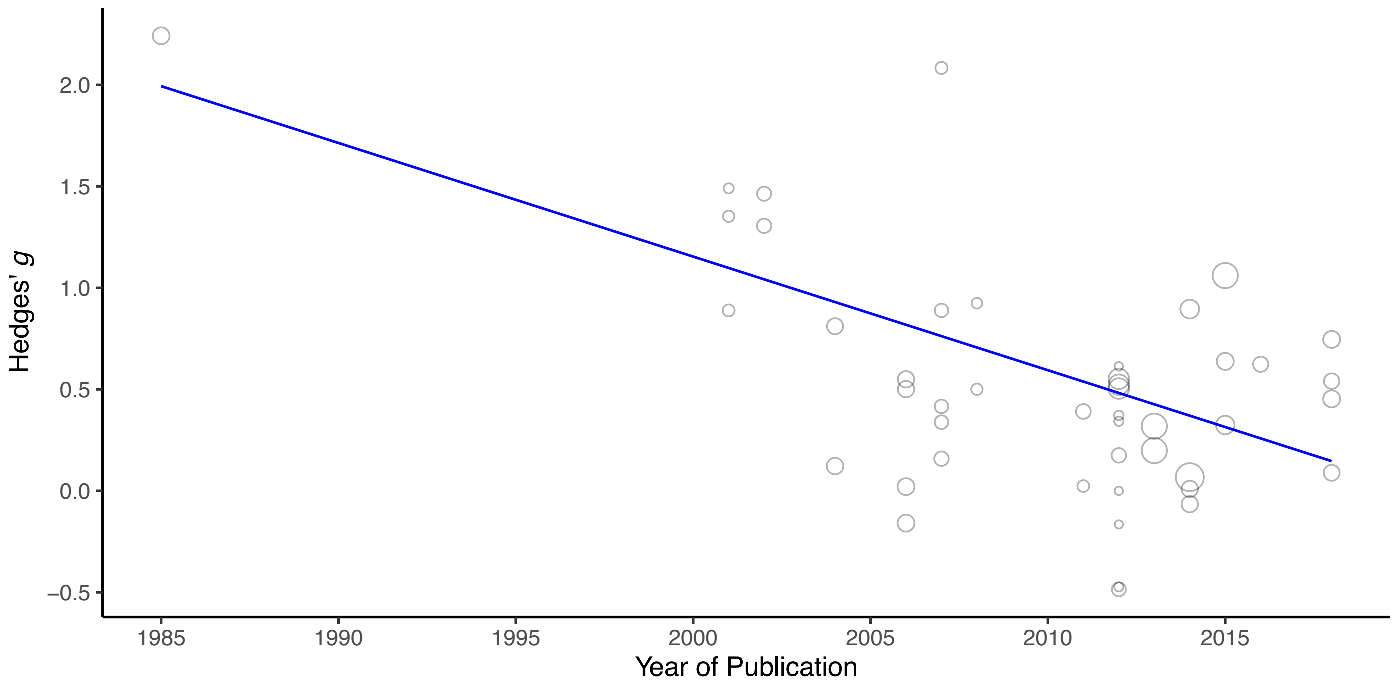

Figure 10. Results from the meta-regression for year of publication: b = -0.06, p = 0.002.

3.6. Publication Bias

Important to note is that for all statistical analyses based on *p*-values (*p*-curve, *p*-uniform and *p*-uniform*) the original three-level model could not be used due to lack of compatibility with the R-functions used to carry out these analyses and all observations were treated as individual studies. Therefore, the effect size in *p*-uniform was estimated by a mixed-effects model and the *p*-uniform* effect size was estimated via a normal fixed-effects model.

3.6.1. p-curve. Results from the *p*-curve analysis revealed no indication of publication bias in the present dataset (see Figure 11). Twenty-eight significant results were included in the *p*-curve analysis and depict a perfectly right-skewed distribution of observed *p*-values.

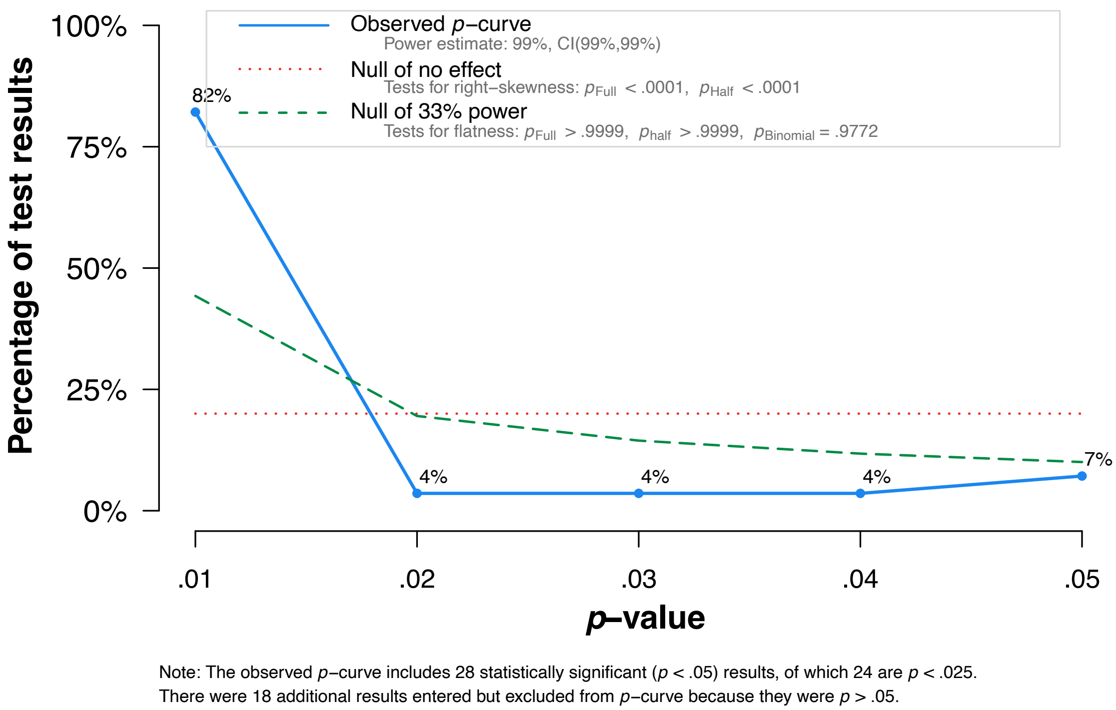

Figure 11. Results from the *p*-curve analysis. 28 significant observations (treated as individual studies) were included. No indication of publication bias was observed.

3.6.2. p-uniform. The *p*-uniform analysis estimated the effect size of the mixed-effects model as $g = 0.68$, 95% CI = [0.55; 0.83], $p < .001$. Twenty-seven significant observations were included in the analysis, and no indication of publication bias could be found ($p = 1$, see Figure 12).

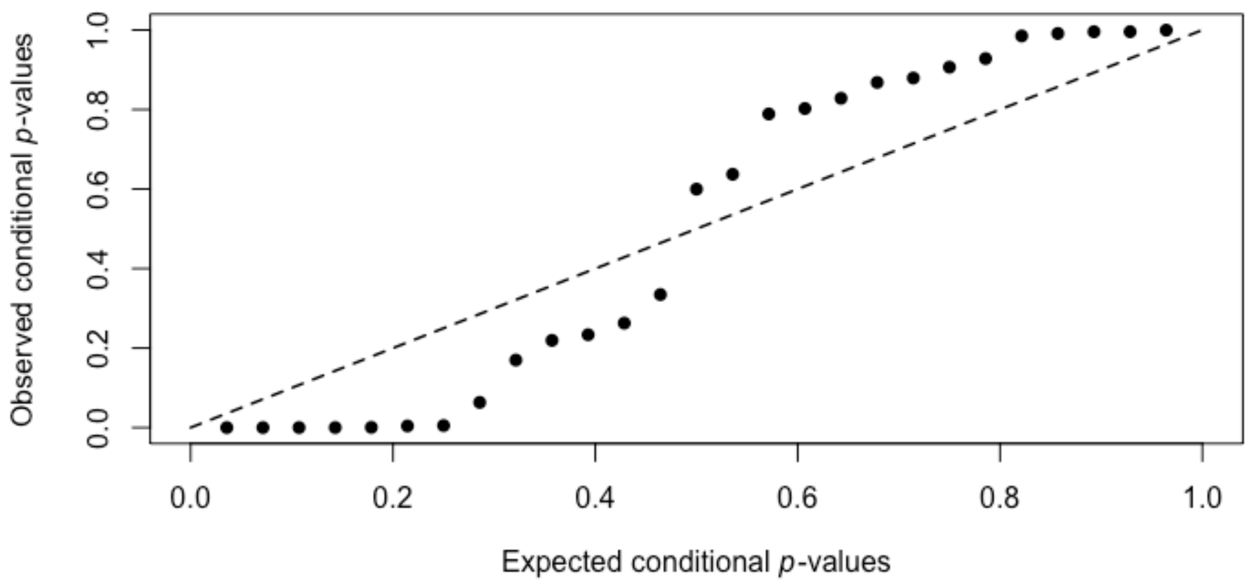

Figure 12. Results from the *p*-uniform analysis. 27 significant observations (treated as individual studies) were included. No indication of publication bias was observed.

3.6.3. p-uniform*. As well as the previous tests for publication bias, *p*-uniform* revealed no indication of publication bias ($p = 1$). The overall effect was estimated in a fixed-effects meta-analytic model and resulted in an effect size of $g = 0.55$, 95% CI = [0.28; 0.8], $p < .001$.

3.6.4. Egger's regression test. Results from the Egger's regression line (Sterne & Egger, 2005) also showed no indication of a present publication bias. Although the slope parameter for the regression line for effect sizes and standard errors turned out to be positive, it remained highly non-significant ($b = 0.55$, $p = 1$). Figure 13 depicts the funnel plot of effect sizes with respect to their standard errors.

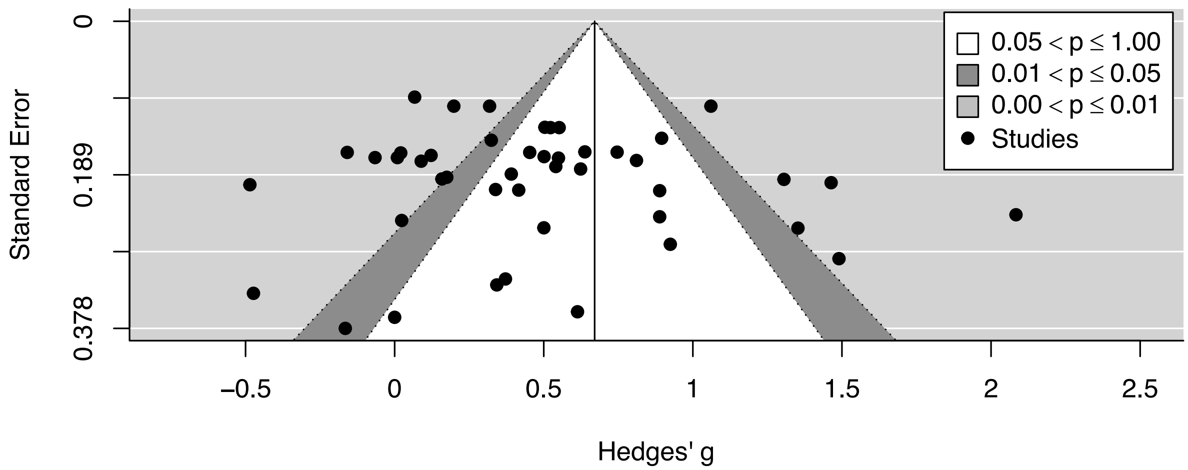

Figure 13: Funnel plot of effect sizes and standard errors in primary studies. Egger's regression line showed no significant indication of asymmetry (b = 0.55, p = 1).

4. Discussion

The present meta-analysis systematically examined the evidence for the HME within a certain, pre-specified context. Forty-six effect sizes out of 23 primary studies were synthesized into a summary effect of g = 0.67. This effect is of comparable size with the summary effect of the previous meta-analysis by Hansen and Kim (2011), who estimated the effect size to be r = .296. So, after differences in operationalization of the HME and leaving methodological weaknesses aside, their overall effect size seems to have been replicated in the present study. Thus, our findings fall in line with results from an extensive amount of research that has been done on the HME over the last three decades.

Again, it has to be mentioned that the present study only included studies that followed a carefully drawn methodological approach, including distinctions that have not been made as careful as they should have been in the previous meta-analytic attempts. The problem of mixing apples and oranges is a dangerous threat to validity in psychological research, one that is also extended into meta-analyses (Sharpe, 1997). In in the context of the HME this is a serious problem that needs to be addressed. The purpose of the current study was to do exactly that: By carefully compiling evidence for one specific methodological approach, the distinction to other approaches that are used to investigate the HME should become more and more obvious. This paper should be considered as a starting point for laying the pathway for future research on the HME.

4.1. Moderator Variables and the HME

The influence of moderator variables has been a focal point of the discussion in the HME context. Hansen and Kim (2011) identified involvement, experimental condition, and the medium (newspaper or others) as potentially moderating variables, whereas Perloff (2015) threw reach, and social identity into the mix, as well as a few mediators (selective recall, selective categorization, different standards, and prior beliefs about media).

Apart from their theoretical considerations, the present study included moderator variables following common meta-analytic practices, like the evolution of the effect over time, possible effect size changes due to study quality and making the distinction of thy type of HME that was measured (*absolute* or *relative*). Although only the year of publication turned out to have significant moderating effects on the HME, these results have to be interpreted with caution.

Some limitations have to be considered when looking at the remaining results of the present examination of moderating factors. For example, problems with operationalizations of the term involvement are problems that cannot be addressed in meta-analytic contexts, as systematic variations cannot be applied. Specific research that primarily address the issue of involvement within a carefully drawn methodological framework can lead the HME literature into the future and help to develop a solid foundation of the underlying psychological phenomenon. Especially distinguishing the HME from similar effects like biased assimilation (Lord et al., 1979) can clarify situations in which information is processed in a biased manner, and the specification of what involvement really is in the present context is needed to support the underlying theoretical framework.

4.2. Publication Bias

Results from tests for publication bias (*p*-curve, *p*-uniform, *p*-uniform*, and Egger's regression test) showed no indication of such a bias existing in the dataset. But one important limitation needs attention when we try to interpret these findings in the present context. The underlying assumption of publication bias refers to the tendency, that the likelihood of studies being published is a result of the strength or the significance of their main results (Dickersin, 1990). But in the present study, a lot of effect sizes were calculated and included based on whether they fit the conceptualized theory, not on what the authors of the primary studies reported as their main findings. This resulted in effects sizes being taken from pre-tests, being taken out of a number of items that were originally investigated, or HME-scales being recoded (if two biases were being examined separately, they were converted into one bidirectional item) in order to make them compatible with the theoretical concept and to guarantee comparability between them. This approach therefore neutralizes the fundamental assumptions that need to be made for publication bias analyses, as publications are less likely to be dependent on these effect sizes.

Furthermore, this design also resulted in the low number of included studies. If a meta-analysis published in 2011 (Hansen & Kim) identifies and includes 34 studies, it looks odd when a meta-analysis on the same topic carried out 8 years later only finds 23 studies to be available. But considering the widely heterogeneous literature and the many different methodological approaches, this unfathomable body of research justifies a strict methodological design to be able to conceptually distinguish (and compare) the certain types of phenomena that are being measured.

4.3. Limitations and Implications for Future Research

Even though necessary boundaries were set in an attempt to guarantee the comparability between primary studies, a few possible blurring differences in methodological approaches could not be eradicated. Apart from possible influencing factors that were analysed in the present study (with mixed results), a few obstacles for future research need to be addressed.

First, the number of items used to measure the HME varies substantially within the literature. Vallone, Ross and Lepper (1985) used 7 items to assess perceived bias from the stimulus, ranging from *"Overall treatment of Israeli in programs"* to *"Perceptions of the personal views of editors of these programs"*, even including percentages for *"[…] favourable and unfavourable references to Israel"*. In general, the number of items used to measure the HME falls within the range of 1-10 items in the dataset. Although a lot of these approaches surely resemble items that measure conceptually related (or even the same) construct, predictive validity is often under threat by the use of single items (Diamantopoulos, Sarstedt, Fuchs, Wilczynski, & Kaiser, 2012). Future research is needed to address this issue in the context of the HME.

A similar issue appears when looking at the scale used to measure the HME. Besides being measured with two separate unidirectional items (one covering bias in favour of on side, one covering bias in favour of the opposite side) or with one bidirectional item, the Likert-type scale differs substantially in size (5-11 points for the present studies). Leung (2011) found that although 4-, 5-, 6-, and 11-point Likert scales apparently don not differ on important values like mean or standard deviations, increasing the number of scale points might reduce skewness and help approaching a normal distribution.

The population from which participants were drawn also differs from study to study. Two main approaches could be observed when analysing the literature: one approach is to sample from a "random" population, like all students at a certain university (for example Kim & Pasadeos, 2007), and the second main approach is to actively sample from a targeted group that appears to be involved in the issue at hand (for example by Mende, 2008). The idea behind the second approach is to actively sample from a highly involved group (although this idea can be questioned, see *methods*), the strength of partisanship is then still measured via a single Likert scale. One needs to be careful when comparing results from participants sampled by these two different approaches.

4.4. Conclusion

Results from the present meta-analysis support the idea of a perceived bias in the media against one's own point of view when holding a strong opinion on a certain topic. The effect is robust with respect to type of HME, reach, involvement and study quality (and can be observed across different topics and within different contexts). Although limitations need consideration and interpretation needs caution, a new approach of distinguishing different methodological approaches from one another and to synthesize findings within a certain framework can help future research digging into the causes and effects of perceived biases in the media. Media usage is getting more and more fragmented—especially for younger people (Geraci & Nagy, 2004)—

therefore, underlying perceptions and misperceptions need to be addressed. Hopefully, this analysis yields an essential contribution to this body of research.

References

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ*, *332*(7549), 1080. https://doi.org/10.1136/bmj.332.7549.1080

Ariyanto, A., Hornsey, M. J., & Gallois, C. (2007). Group allegiances and perceptions of media bias. *Group Processes & Intergroup Relations*, *10*(2), 266–279. https://doi.org/10.1177/1368430207074733

Arpan, L. M., & Raney, A. A. (2003). An experimental investigation of news source and the Hostile Media Effect. *Journalism & Mass Communication Quarterly*, *80*(2), 265–281. https://doi.org/10.1177/107769900308000203

Arpan, L.M., Bae, B. J., & Greene, G. H. (2009). *Does humor attenuate hostility? A comparison of hostile media perceptions of news and late-night comedy*. Presented at the International Communication Association annual conference, Chicago. Retrieved from https://www.semanticscholar.org/paper/Does-Humor-Attenuate-Hostility-A-Comparison-of-of-Arpan-Bae/0e60b0202bbdad36b1667e7dba29d4dc5061ed22

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology*, *12*(3), 154–174. https://doi.org/10.20982/tqmp.12.3.p154

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: a practical tutorial. *Evidence Based Mental Health*, *22*(4), 153–160. https://doi.org/10.1136/ebmental-2019-300117

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis* (1st ed.). Hoboken, NJ, United States: Wiley.

Brubaker, P. J. (2012). *Do you see what I see? An examination of hostile media perceptions online* (Dissertation). Retrieved from https://etda.libraries.psu.edu/catalog/14170

Card, N. A. (2015). *Applied Meta-Analysis for Social Science Research*. New York: Guilford Publications.

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. https://doi.org/10.1037/a0032968

Chia, S. C., & Cenite, M. (2012). Biased news or biased public? *Journalism Studies*, *13*(1), 124–140. https://doi.org/10.1080/1461670x.2011.601957

Chia, S. C., & Chang, T.-K. (2015). Not my horse: Voter preferences, media sources, and hostile poll reports in election campaigns. *International Journal of Public Opinion Research*, *29*(1), 23–45. https://doi.org/10.1093/ijpor/edv046

Chia, S. C., Yong, S. Y. J., Wong, Z. W. D., & Koh, W. L. (2007). Personal bias or government bias? Testing the Hostile Media Effect in a Regulated Press System. *International Journal of Public Opinion Research*, *19*(3), 313–330. https://doi.org/10.1093/ijpor/edm011

Choi, J., & Chang, J. (2007). *Exploring behavioral consequence of hostile media perceptions: Its impact on advocacy evaluation and strategies.* Presented at the International Communication Association annual conference, San Francisco.

Choi, J., Yang, M., & Chang, J. J. (2009). Elaboration of the Hostile Media Phenomenon. *Communication Research*, *36*(1), 54–75. https://doi.org/10.1177/0093650208326462

Christen, C. T., Kannaovakun, P., & Gunther, A. C. (2002). Hostile Media Perceptions: Partisan assessments of press and public during the 1997 United Parcel Service Strike. *Political Communication*, *19*(4), 423–436. https://doi.org/10.1080/10584600290109988

Coe, K., Tewksbury, D., Bond, B. J., Drogos, K. L., Porter, R. W., Yahn, A., & Zhang, Y. (2008). Hostile news: Partisan use and perceptions of cable news programming. *Journal of Communication*, *58*(2), 201–219. https://doi.org/10.1111/j.1460-2466.2008.00381.x

Conn, V. S., Valentine, J. C., Cooper, H. M., & Rantz, M. J. (2003). Grey literature in meta-analyses. *Nursing Research*, *52*(4), 256–261. https://doi.org/10.1097/00006199-200307000-00008

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The Handbook of Research Synthesis and Meta-Analysis* (2nd ed.). New York: Russell Sage Foundation.

D'Alessio, D. (2003). An experimental examination of readers' perceptions of media bias. *Journalism & Mass Communication Quarterly*, *80*(2), 282–294. https://doi.org/10.1177/107769900308000204

Daily, K. (2014). *Explicating the hostile media perception: How source credibility influences partisans' responses to balanced news coverage of health policies.* (Dissertation). Retrieved from http://hdl.handle.net/1903/16179

Dalton, R. J., Beck, P. A., & Huckfeldt, R. (1998). Partisan cues and the media: Information flows in the 1992 presidential election. *American Political Science Review*, *92*(1), 111–126. https://doi.org/10.2307/2585932

Davison, W. P. (1983). The Third-Person Effect in communication. *Public Opinion Quarterly*, *47*(1), 1–15. https://doi.org/10.1086/268763

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective. *Journal of the Academy of Marketing Science*, *40*(3), 434–449. https://doi.org/10.1007/s11747-011-0300-3

Dickersin, K. (1990). The existence of Publication Bias and risk factors for its occurrence. *JAMA: The Journal of the American Medical Association*, *263*(10), 1385–1389. https://doi.org/10.1001/jama.1990.03440100097014

Dunn, S. W. (2011). *Perceptions of bias in a changing media environment: The hostile media effect for objective and ideological media.* (Dissertation). Retrieved from https://doi.org/10.17615/w4v8-3z72

Feldman, L. (2008). *To opine or not to opine: The consequences of opinionated news for political information processing, attitudes, and knowledge* (Dissertation). Retrieved from https://search.proquest.com/openview/e3947f3b483304a181b6808f402ee3f2/1?pq-origsite=gscholar&cbl=18750&diss=y

Feldman, L. (2010). Partisan differences in opinionated news perceptions: A test of the Hostile Media Effect. *Political Behavior*, *33*(3), 407–432. https://doi.org/10.1007/s11109-010-9139-4

Geraci, J. C., & Nagy, J. (2004). Millennials - the new media generation. *Young Consumers: Insight and Ideas for Responsible Marketers*, *5*(2), 17–24. Retrieved from https://www.ingentaconnect.com/content/mcb/yc/2004/00000005/00000002/art00002

Giner-Sorolla, R., & Chaiken, S. (1994). The Causes of hostile media judgments. *Journal of Experimental Social Psychology*, *30*(2), 165–180. https://doi.org/10.1006/jesp.1994.1008

Gunther, A. C. (1992). Biased press or biased public? Attitudes toward media coverage of social groups. *Public Opinion Quarterly*, *56*(2), 147. https://doi.org/10.1086/269308

Gunther, A. C., Christen, C. T., Liebhart, J. L., & Chia, S. C.-Y. (2001). Congenial public, contrary press, and biased estimates of the climate of opinion. *Public Opinion Quarterly*, *65*(3), 295–320. https://doi.org/10.1086/322846

Gunther, A. C., & Liebhart, J. L. (2006). Broad reach or biased source? Decomposing the Hostile Media Effect. *Journal of Communication*, *56*(3), 449–466. https://doi.org/10.1111/j.1460-2466.2006.00295.x

Gunther, A. C., McLaughlin, B., Gotlieb, M. R., & Wise, D. (2016). Who says what to whom: content versus source in the Hostile Media Effect. *International Journal of Public Opinion Research*, edw009. https://doi.org/10.1093/ijpor/edw009

Gunther, A. C., Miller, N., & Liebhart, J. L. (2009). Assimilation and contrast in a test of the Hostile Media Effect. *Communication Research*, *36*(6), 747–764. https://doi.org/10.1177/0093650209346804

Gunther, A. C., & Mundy, P. (1993). Biased optimism and the Third-Person Effect. *Journalism Quarterly*, *70*(1), 58–67. https://doi.org/10.1177/107769909307000107

Gunther, A. C., & Chia, S. C.-Y. (2001). Predicting pluralistic ignorance: The Hostile Media Perception and its consequences. *Journalism & Mass Communication Quarterly*, *78*(4), 688–701. https://doi.org/10.1177/107769900107800405

Gunther, A. C., & Schmitt, K. (2004). Mapping boundaries of the Hostile Media Effect. *Journal of Communication*, *54*(1), 55–70. https://doi.org/10.1111/j.1460-2466.2004.tb02613.x

Hansen, G. J., & Kim, H. (2011). Is the media biased against me? A meta-analysis of the Hostile Media Effect research. *Communication Research Reports*, *28*(2), 169–179. https://doi.org/10.1080/08824096.2011.565280

Hoffner, M., & Toohey, R. (2007). *The Hostile Media Effect in the 2004 U.S. presidential election.* Presented at the International Communication Association annual conference, San Francisco.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70. Retrieved from https://www.jstor.org/stable/4615733?seq=1#metadata_info_tab_contents

Huge, M., & Glynn, C. J. (2010). Hostile media and the campaign trail: Perceived media bias in the race for governor. *Journal of Communication*, *60*(1), 165–181. https://doi.org/10.1111/j.1460-2466.2009.01473.x

Ioannidis, J. P. A., & Trikalinos, T. A. (2005). Early extreme contradictory estimates may appear in published research: The Proteus phenomenon in molecular genetics research and randomized trials. *Journal of Clinical Epidemiology*, *58*(6), 543–549. https://doi.org/10.1016/j.jclinepi.2004.10.019

Jeon, J. (2009). *The influence of prior issue attitudes on perception bias and perceived message credibility in participatory websites* (Master's dissertation). Retrieved from https://www.semanticscholar.org/paper/The-influence-of-prior-issue-attitudes-on-bias-and-Jeon/b7538b6c199fdea93ee25f12466a54b63344b786

Kim, K. S. (2010). Public understanding of the politics of global warming in the news media: the hostile media approach. *Public Understanding of Science*, *20*(5), 690–705. https://doi.org/10.1177/0963662510372313

Kim, K. S., & Pasadeos, Y. (2007). Study of partisan news readers reveals Hostile Media Perceptions of balanced stories. *Newspaper Research Journal*, *28*(2), 99–106. https://doi.org/10.1177/073953290702800207

Kinnally, W. (2008). *Reactance and the hostile media effect: Placing the effect within the theory.* (Dissertation). Retrieved from https://diginole.lib.fsu.edu/islandora/object/fsu%3A254109

Leung, S.-O. (2011). A comparison of psychometric properties and normality in 4-, 5-, 6-, and 11-point Likert scales. *Journal of Social Service Research*, *37*(4), 412–421. https://doi.org/10.1080/01488376.2011.580697

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*(11), 2098–2109. https://doi.org/10.1037/0022-3514.37.11.2098

Matheson, K., & Dursun, S. (2001). Social identity precursors to the Hostile Media Phenomenon: Partisan perceptions of coverage of the Bosnian Conflict. *Group Processes & Intergroup Relations*, *4*(2), 116–125. https://doi.org/10.1177/1368430201004002003

Mende, A. M. (2008). *Testing the hostile media effect under selective exposure* (Bachelor's Dissertation). Retrieved from http://citation.allacademic.com/meta/p232794_index.html

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., ... Tugwell, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*. https://doi.org/10.1371/journal.pmed.1000097

Moher, D., Pham, B., Jones, A., Cook, D. J., Jadad, A. R., Moher, M., … Klassen, T. P. (1998). Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *The Lancet*, *352*(9128), 609–613. https://doi.org/10.1016/s0140-6736(98)01085-x

National Heart, Lung, & Blood Institute, 2014. (2014). Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies. Retrieved from https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools

Page, M. J., McKenzie, J. E., & Higgins, J. P. T. (2018). Tools for assessing risk of reporting biases in studies and syntheses of studies: a systematic review. *BMJ Open*, *8*(3), e019703. https://doi.org/10.1136/bmjopen-2017-019703

Peng, Z. (2005). *Ideology, source, news content, and perception of media bias and credibility: An empirical study on hostile media effect* (Dissertation). Retrieved from https://search.proquest.com/openview/0b9db8fb4218d77217abc66841af573a/1?pq-origsite=gscholar&cbl=18750&diss=y

Perloff, R. M. (1989). Ego-Involvement and the Third Person Effect of televised news coverage. *Communication Research*, *16*(2), 236–262. https://doi.org/10.1177/009365089016002004

Perloff, R. M. (2015). A three-decade retrospective on the Hostile Media Effect. *Mass Communication and Society*, *18*(6), 701–729. https://doi.org/10.1080/15205436.2015.1051234

Pietschnig, J., Siegel, M., Eder, J. S. N., & Gittler, G. (2019). Effect declines are systematic, strong, and ubiquitous: A meta-meta-analysis of the Decline Effect in intelligence research. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02874

Richardson, J. D., Huddy, W. P., & Morgan, S. M. (2008). The Hostile Media Effect, Biased Assimilation, and perceptions of a presidential debate. *Journal of Applied Social Psychology*, *38*(5), 1255–1270. https://doi.org/10.1111/j.1559-1816.2008.00347.x

Schmitt, K. M., Gunther, A. C., & Liebhart, J. L. (2004). Why partisans see mass media as biased. *Communication Research*, *31*(6), 623–641. https://doi.org/10.1177/0093650204269390

Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review*, *17*(8), 881–901. https://doi.org/10.1016/s0272-7358(97)00056-1

Shreves, W. B. (2015). *Disarming the hostile media: an intervention to reduce perceptions of media bias by reducing naïve realism* (Dissertation). Retrieved from https://ir.ua.edu/handle/123456789/2533

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, *143*(2), 534–547. https://doi.org/10.1037/a0033242

Sterne, J. A. C., & Egger, M. (2006). Regression methods to detect publication and other bias in meta-analysis. *Publication Bias in Meta-Analysis*, 99–110. https://doi.org/10.1002/0470870168.ch6

Tsang, S. J. (2018). Empathy and the Hostile Media Phenomenon. *Journal of Communication*, *68*(4), 809–829. https://doi.org/10.1093/joc/jqy031

Tsfati, Y. (2007). Hostile Media Perceptions, presumed media influence, and minority alienation: The case of Arabs in Israel. *Journal of Communication*, *57*(4), 632–651. https://doi.org/10.1111/j.1460-2466.2007.00361.x

Tsfati, Y., & Cohen, J. (2005). Democratic consequences of Hostile Media Perceptions. *Harvard International Journal of Press/Politics*, *10*(4), 28–51. https://doi.org/10.1177/1081180x05280776

Tsfati, Y., & Huino, H. (2014). Research findings weaken perceptions of media bias. *Newspaper Research Journal*, *35*(1), 82–95. https://doi.org/10.1177/073953291403500107

Vallone, R. P., Ross, L., & Lepper, M. R. (1985). The hostile media phenomenon: Biased perception and perceptions of media bias in coverage of the Beirut massacre. *Journal of Personality and Social Psychology*, *49*(3), 577–585. https://doi.org/10.1037/0022-3514.49.3.577

van Aert, R. C. M. (2019). puniform: Meta-analysis methods correcting for publication bias. R package version 0.2.1. https://CRAN.R-project.org/package=puniform

van Aert, R. C. M., & van Assen, M. A. L. M. (2018). *Correcting for publication bias in a meta-analysis with the p-uniform* method*. https://doi.org/10.31222/osf.io/zqjr9

van Assen, M. A. L. M., van Aert, R. C. M., & Wicherts, J. M. (2015). Meta-analysis using effect size distributions of only statistically significant studies. *Psychological Methods*, *20*(3), 293–309. https://doi.org/10.1037/met0000025

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3). Retrieved from https://lirias.kuleuven.be/1059637?limo=0

Yan, L., Tay, M., Yap, P., Hao, X., Chia, S., Detenber, B., & Lee, W. (2007). *Personal involvement and perception of press credibility.* Presented at the International Communication Association annual conference, San Francisco.

Wei, R., & Lo, V.-H. (2013). Perceived poll bias and inference of biased poll impact in the 2010 U.S. midterm elections: Media performance in the eyes of the beholder. *The Open Communication Journal*, *7*(1), 12–20. https://doi.org/10.2174/1874916x20130509002

Appendix

Table A1

Summary of studies

Author(s)	n	*g*	Effect	Invol.	Reach	Qual.
Vallone, Ross & Lepper (1985)	95	2,242	A	H	H	6
Gunther & Schmitt (2004)	151	0,811	A	H	H	8
Gunther & Schmitt (2004)_2	151	0,122	R	H	L	8
Gunther & Liebhart (2006)	160	0,021	R	H	L	8
Gunther & Liebhart(2006)_2	154	0,501	A	H	H	8
Gunther & Liebhart (2006)_3	164	-0,159	A	H	L	8
Gunther & Liebhart (2006)_4	150	0,549	A	H	H	8
Feldman (2011)	102	0,024	R	H	H	5
Matheson & Dursun (2001)	60	1,490	A	L	H	8
Chia, Yong, Wong, & Koh (2007)	109	2,084	A	H	H	8
Richardson, Huddy, & Morgan (2008)	69	0,500	A	H	H	7
Gunther, McLaughlin, Gotlieb, & Wise (2016)	141	0,624	A	H	H	9
Tsang (2018)	130	0,541	A	L	H	6
Tsang (2018)_2	134	0,089	A	L	H	6
Tsang (2018)_3	165	0,746	R	L	H	6
Tsang (2018)_4	156	0,453	R	L	H	6
Mende (2008)	58	0,924	A	H	H	5
Gunther, Christen, Liebhart, & Chia (2001)	76	0,889	R	H	H	8
Gunther, Christen, Liebhart, & Chia (2001)_2	76	1,353	R	H	H	8
Kim & Pasadeos (2007)	95	0,338	A	L	H	5
Kim & Pasadeos (2007)_2	95	0,416	A	L	H	5

Ariyanto, Hornsey, & Gallois (2007)	104	0,889	A	L	H	6
Ariyanto, Hornsey, & Gallois (2007)_2	108	0,158	R	L	H	6
Christen, Kannaovakun, & Gunther (2002)	135	1,464	A	L	H	7
Christen, Kannaovakun, & Gunther (2002)_2	135	1,306	A	L	H	7
Chia & Cenite (2012)	109	0,175	R	H	H	6
Chia & Cenite (2012)_2	102	-0,485	A	H	H	6
Kim (2011)	138	0,391	R	H	H	5
Wei & Lo (2013)	422	0,199	R	H	H	6
Wei & Lo (2013)_2	422	0,318	R	H	H	6
Shreves (2015)	420	1,061	R	H	H	6
Tsfati & Huino (2014)	213	0,895	A	L	H	2
Chia & Chang (2015)	163	0,638	R	H	H	4
Chia & Chang (2015)_2	189	0,324	R	H	H	4
Brubaker (2012)	252	0,523	A	H	H	5
Brubaker (2012)_2	252	0,552	A	H	L	5
Brubaker (2012)_3	252	0,504	A	H	L	5
Daily (2014)	154	0,009	R	H	H	6
Daily (2014)_2	154	-0,066	R	H	H	6
Daily (2014)_3	530	0,067	R	H	H	6
Dunn (2012)	50	-0,473	R	H	H	8
Dunn (2012)_2	46	0,372	R	H	H	8
Dunn (2012)_3	42	0,613	A	H	H	8
Dunn (2012)_4	29	-0,166	R	H	H	8
Dunn (2012)_5	42	0,343	R	H	H	8
Dunn (2012)_6	31	0,000	0	H	H	8

Note: Table of all studies included in the analysis with values used for summary effect and moderator analysis.

Author(s) and year of publication are depicted in the most left columns. If more than one observation was reported, this is indicated by the number next to the year of publication. n = sample size for each observation. *g* is the reported (or calculated) effect size for each observation. Effect is either A (*absolute*) or R (*relative*), which was then used for the moderator analysis. Involvement and Reach was coded either as H (high) or L (low) for the subgroup analyses. Each study was assigned a specific value of study quality, ranging from 0 to 13. Higher scores indicate higher study quality.