Linking survey and media content data: Opportunities, considerations, and pitfalls

Claes H. De Vreese, Mark Boukes, Andreas Schuck, Rens Vliegenthart, Linda Bos & Yph Lelkes

 View supplementary material

 Published online: 24 Oct 2017.

 Submit your article to this journal

 View related articles

 View Crossmark data

Routledge
Taylor & Francis Group

Linking survey and media content data: Opportunities, considerations, and pitfalls

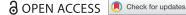Claes H. De Vreese[a], Mark Boukes[a], Andreas Schuck[a], Rens Vliegenthart[a], Linda Bos[a], and Yph Lelkes[b]

[a]Amsterdam School of Communication Research ASCoR, University of Amsterdam; [b]Annenberg School of Communication, University of Pennsylvania

ABSTRACT

In media effects research a fundamental choice is often made between (field) experiments or observational studies that rely on survey data in combination with data about the information environment or media coverage. Such studies linking survey data and media content data are often dubbed "linkage studies." On the one hand, such designs are the state of the art in our field and on the other hand, they come with a long list of challenges and choices. This article reviews the *rationales* for linkage studies, outlines different *types* of linkage studies, reviews the *state-of-the-art* in this area, discusses which survey and content *items* to use in an analysis, reviews different types of analyses, outlines considerations for alternative specifications, and provides a step-by-step example.

Introduction

There is a rich tradition of "linkage studies" that combine survey data with content analysis data in communication science and, in part, public opinion and political science research. Even though the example used here and several of the published studies are in the realm of political communication, many of the observations, considerations, and challenges also pertain to linkage studies in, for example, campaign assessments and health communication (e.g., Nagelhout et al., 2012). Some of the original studies date back to the 1970s (e.g., Erbring, Goldenberg, & Miller, 1980). Shoemaker and Reese (1990) and more recent publications provide a good introduction to the topic (De Vreese, 2014), an explication of the idea (Schuck, Vliegenthart, & De Vreese, 2016b), and identification of problems and solutions with respect to measurement error (Scharkow & Bachl, 2016). These are good publications to consult for additional information.

Why combine content and survey data? The advantages are self-evident. On the one hand, content analyses are revealing about content patterns, but nothing can be said about the implications or even *effects*. On the other hand, survey data can say a lot about the use of media or about attitudes, "but these attitudes are hard to meaningfully relate to the media without knowing about the contents of the media" (De Vreese, 2014, p. 339).

The underlying principle is simple. The answer given to a question whether media usage is related to or even affecting a given dependent variable is potentially improved by combining content data and survey data. For example, if there is interest in measuring political knowledge, and one of the knowledge items is about the contents of an ongoing tax reform, a survey-only based study that would find a correlation between reading quality newspaper X and the dependent variable may actually be misleading. What if, in fact, newspaper X did not report about the tax reform? In that case the

positive relation between exposure to newspaper X and being informed about the reform is a spurious one and probably only indicative of the fact that knowledgeable persons tend to read paper X.

With a content analysis at their disposal, scholars would be in a much better position to make inferences about the nature of the relationship. If a content analysis would show that paper X indeed reported extensively about the tax reform, we are in a stronger position to make claims about the positive coefficient. If, moreover, paper Y did not report about the reform and being exposed to that paper does not show a relationship with this knowledge measure, this adds confidence to the assumption that it is not exposure to paper X per se that drives this relationship, it is because the content in paper X provided the opportunity to learn this fact. However, this relationship is only substantiated by eyeballing the survey coefficients and the content data.

Some studies rely on exposure measures in surveys to build an argument for specific media effects. Eveland and Scheufele (2000), for example, related survey-based measures of media exposure from the American National Election Study to tap political knowledge. While theoretically interesting and innovative, such designs say little about the actual impact of the media *content* and can thus be dubbed "mere exposure studies", i.e., they show a plausible correlation between media usage and an outcome variable. However, as Graber (2004, p. 516) concluded, most scholars do not pay much attention to the actual content of the information that they assume is driving the effect.

An overview of types of linkage studies

Linkages between survey data and media content data can take different forms, each offering advantages and limitations.

Descriptive content and exposure studies

When combining media content data, there are different models and levels to consider. One might describe media content at the *aggregate* during a longer time period and link this to the development in public opinion during this period. Such studies are either exploratory in nature (e.g., Statham & Tumber, 2013) or aggregate level times series analyses (e.g., Hester & Gibson, 2003; Vliegenthart & Boomgaarden, 2007). Another option is to *describe* media content during a period while media exposure is measured in a survey. This will allow the researcher to offer conclusions about the type of media content that was available. This description of media content can either relate to the period before a single cross-sectional survey or it can pertain to the period in between two panel survey waves so as to get a sense of the media coverage between the two survey periods. Such linkages can be described as correlational linkages.

Media content and cross-sectional survey data

A more compelling linkage is provided if a researcher has individual outlet content data *and* individual level media exposure measures in the survey. The survey items can then be linked, at the individual level, to the contents of the individual media outlets. This idea is not novel (see Erbring et al., 1980; Noelle-Neumann, 1973; Rössler, 1999), but its application is still limited. Such studies all use media content analyses to enrich the individual level survey exposure measures. This approach is more informed and obviously contains more details about the media content than relying on "empty media exposure variables" specifications only. However, at the end of the day, such an approach still suffers from the general limitations of the cross-sectional survey approach.

Media content and rolling cross-sectional survey data

The rolling cross sectional (RCS) design is a survey design that provides detailed observations of, for example, campaign dynamics. As noted by Johnson and Brady (2002, p. 283), its essence is to

conduct a cross-sectional survey but then use the timing of the interview systematically across the data collection period. It typically involves daily interviews with a sub-set of the sample during a specific period (e.g., an election campaign) and has specific requirements and recommendations for re-contacting respondents. The design was explored as add-ons to the 1984 and 1988 U.S. national election studies and fully developed in the 1988 and subsequent Canadian national election studies (Johnson & Brady, 2002). It later became a hallmark of the Annenberg National Elections Studies (Romer, Kenski, Winneg, Adasiewicz, & Jamieson, 2006) and has been applied in, for example, the European Parliament Campaign Study in 2009 (de Vreese et al., 2010), and the longitudinal German elections study (Schmitt-Beck, Bytzek, Rittinger, Roßteutscher, & Weßels, 2009).

The combination of RCS survey data and media data can take different forms. Aggregate comparisons between developments in public opinion and media content are possible, formal time series comparisons are possible, as are individual level dynamic analyses of the relationship between exposure to media content and a specific dependent variable (see, e.g., Elenbaas, Boomgaarden, Schuck, & De Vreese, 2013). However, ultimately, the RCS design suffers the same limitation as the cross sectional design in terms of making causal inferences because it is still only a *relationship* that is established between exposure and an outcome variable rather than an *effect* (such as within-individual change).

Media content and panel survey data

A panel survey has the major advantage that it enables the research to study intra-individual change in addition to the change in the sample overall. Panel designs have long played a prominent role in the study of political communication and electoral behavior. A prominent example is Lazarsfeld, Berelson, and Gaudet's *The People's Choice* (1944) which relied on a seven-wave, monthly panel survey of residents of Erie County, Ohio. The American national election study has included a panel component since 1948. Panel surveys have the distinct advantage of offering more leverage with regard to *change* and *causality*.

The use of panel surveys is on the rise in political communication research. Eveland and Morey (2011) report that in the 1990s an average of two articles were published per year using panel data, but between 2002 and 2006, no less than nine studies in political communication using panel data were published, per year. Panel data can be combined with all sorts of contextual data (such as campaign efforts of political candidates or parties, developments in the economy etc.), but they can also very well be combined with media content data.

An example comes from De Vreese and Semetko (2004) analysis of vote choice in the Danish 2000 euro referendum. They measured media exposure in the panel survey and then weighted the exposure to specific media outlets with information about that outlet's coverage of the referendum in terms of amount and tone of news. This was generated from a systematic content analysis of the major newspapers and television news shows. While controlling for the vote intention at wave 1 (including a lagged measure, see Yanovitzky & Cappella, 2001), they could thus assess the impact of exposure to media content on the actual vote (assessed at wave 2). In recent years, several studies have relied on panel survey data integrated with media content data, e.g., Bos, Van Der Brug, and De Vreese (2011), Matthes (2012), Schuck and De Vreese (2009), and Schuck, Vliegenthart, and De Vreese (2016a). Ultimately, the purpose of this data integration is to improve the estimation of media effects on the one hand, but also to improve the relationship between theory and data on the other (see also Schuck et al., 2016b).

What do we know? The state of affairs

As evidenced above the basic idea of combining (panel) survey and media data has been around for more than half a century. In a prominent example, Erbring et al. (1980) combined the media exposure measures in the U.S. national election study with a content analysis of daily newspapers.

Front-page articles were coded for the main issue and merged with the survey data by matching each respondent with information about the content in the particular paper the respondent had read. This study also included contextual information on unemployment rates, thereby allowing the researchers to disentangle the contribution of media content and real world variables. As aptly summarized by Erbring and colleagues (1980, p. 21) their study had the benefits of being national in scope, they had actual media content data, a linkage to real-world data, and the ability to link the media and real-world data to the individual level data in the survey. The study also highlighted the importance of spatial variation in content by looking at local news, a point also emphasized later by Schuck et al. (2016a) and Slater (2016).

Other examples include the extensive study by Kepplinger, Brosius, and Staab (1991) that analyzes public opinion formation about three conflicts in German media. This was done with a sophisticated model that captured both media selection and media effects. They relied on public opinion survey data and a media content analysis in the six months prior to the survey. In assessing the relationship between exposure and opinions they merged the survey data and the content data to create a "specific information index", summing up the information across the media.

As noted by Scharkow and Bachl (2016), a recent wave of studies has further developed this tradition, often in the combination of panel survey with content analysis.[1] Collectively, this new generation of linkage studies has *substantively* yielded new insights such as the impact of news coverage on, e.g., public attitudes, knowledge, economic evaluations, and electoral behavior. The studies have shown the conditionality of effects, by either weighing in contextual or individual level factors, and their interactions with media exposure. *Methodologically*, the studies have developed the linking of survey and content data by bringing in specific content features (e.g., visibility, topics, evaluations of the economy, and political actors), by combining or contrasting content features, and working with different weights (e.g., weighing up recent content).

These advances do not imply that a linkage study is—per se—superior to another approach. This always depends on the question to be answered. Moreover, also a weighted exposure measure may not—per se—perform better empirically. Schuck et al. (2016b) argue why that may not be the case. If a study would assess the impact of newspaper reading on knowledge about a specific political topic, but a content analysis of that newspaper would show virtually no coverage of that specific topic, it would be substantively hard to interpret a positive survey data-based coefficient for a respondent reading this paper. A weighted, potentially non-significant, exposure measure that might reduce the empirical importance of reading the newspaper because of the absence of relevant content would be more meaningful in this case. In short, it is important to measure exposure to specific media *content* that is congruent with the dependent variable studied.

A couple of developments affect this renewed interest in the linkage approach: (1) the collection of (dynamic) survey data has become better feasible (both in terms of online technologies and web and more recently mobile based data collection). Some of these developments come at the cost of concerns about recruitment, sample bias etc., but it is generally easier to obtain public opinion data that are more dynamic; (2) the development in (automated) content analysis is rapid. Only a decade ago the majority of studies relied fully on human coders, today automated content analyses are gaining ground. Together, the ability to collect more dynamic opinion data and more rapid and large-scale content data, also increase the interest in and enthusiasm for linkage studies.

Getting specific I: What survey items/variables to use?

It is the survey-based exposure measure which is the starting point for thinking about the actual linkage with media content. However, these self-reported exposure measures are often criticized as unreliable. In a nutshell the argument has been that they are so poor that we should rather use knowledge measures as a proxy. Price and Zaller (1993) frame it this way: "If one's goal is to predict reliably whether or not a person has received any particular item of news, we conclude, general political knowledge is clearly the most useful predictor". Indeed it has been convincingly and

repeatedly demonstrated that self-reports over-estimate news consumption (e.g., Prior, 2009). This is problematic, but it also seems that those who report the most frequent exposure are the ones most exposed, albeit at a too high self-reported level. Self-reports can also be optimized (or made less problematic): The American National Election study assesses days of week and type of medium (Althaus & Tewksbury, 2007). Even better is asking about specific news outlets (De Vreese & Semetko, 2004; Dilliplane, Goldman, & Mutz, 2013; Slater, 2004), preferably in combination with an assessment of frequency (Andersen, De Vreese, & Albæk, 2016).[2] This discussion is still ongoing and also touches on the ambition to include information processing variables and attention measures.[3] The challenge is additionally augmented by current changes in the media landscape. This pertains both the emergence of new information sources (e.g., social media; Gil De Zúñiga, Jung, & Valenzuela, 2012) as well as changing patterns (time/place/platform/device) of consumption. Ceteris paribus, *our advice is to include exposure measures in the survey in the most granular and detailed way possible (time and resources permitting) to allow for multiple approaches in subsequent linkages (see discussion below). The exposure measures in the survey should match the selection of outlets that are subsequently subject to the content analysis. Additional survey items measuring e.g. attention to news and/or individual information processing style can be considered as well since they can further extend modelling options in the subsequent linkage.*

Getting specific II: What content indicators to use?

On the media content side, there are obviously endless features that could be included. However, "…theory is the guiding light to making the actual combination between the two types of data sources" (De Vreese, 2014, p. 339; translated from Danish). To illustrate we might consider which content features to include in studies relying on concepts such as agenda-setting, priming, framing and evaluations. In all cases, some indicator of *visibility* (of topics or actors) is a likely candidate. After all the visibility of different topics or actors in the news relates to both agenda-setting and priming (salience of topics and salience when assessing actors like the government; see Hopmann, Van Aelst, & Legnante, 2011). Furthermore, the *tone* of the news is important to, for instance, understand candidate evaluations and vote intentions/choice. And *frames* may matter for issue understanding, attitudes, and electoral behavior. The crucial observation is that there is no *a priori* content feature which is "correct" to include. This really depends on the theoretical question at stake. The theoretical context may also indicate whether it makes sense to consider adjacent or conceptually related features while being cautious if these features correlate, making it hard to disentangle which feature exactly is causing the effect.

In addition to the *substantive features* (such as issue or actor visibility, tone, frames, etc.; see Table 1), one can also consider a range of *formal features*. These include aspects such as prominence, size/length, or position of news in an outlet. It is a reasonable assumption that prominent or large news stories might have a bigger effect because they are more likely to catch attention. As will become evident below, there is not (yet) a well-established standard for how to consider these formal features (see Geiss, 2016; Wolling, 2002). *At this stage, and at the absence of more empirical evidence or substantive theoretical reasoning, we would recommend focusing on substantive content features* first *in any linkage analysis and adding formal features only later in the analysis and report the outcome as part of additional analyses and/or robustness checks, which can also further inform this discussion.*

Table 1. Examples of substantive content features to be used in linkage studies.

Concept	Content feature
Agenda setting	Issue salience
Framing	Frames (e.g.,strategy frame -> cynicism; conflict frame -> mobilization; etc.)
Priming	Issue salience and evaluations
Evaluations	Tone
…	…

How to combine exposure and content?

There are, of course, many options to consider when actually constructing the linkage. Again, the guiding light is the *theoretical* puzzle at hand.

If the endeavor is to test for the effects of content in a specific *medium* (e.g., TV news vs. newspapers), it makes sense to organize the exposure measures per medium (e.g., Andersen et al., 2016). If the interest is to test for the effects of content in a specific *genre*, it may make sense to distinguish between quality outlets (e.g., public service news and broadsheet papers vs. commercial news and tabloids). In cross-national or other comparative research, such distinctions might need to be based on functional equivalence in the absence of fully comparable outlets. For example, a tabloid newspaper in one country might not be comparable to that in another country, but they might still be classified together (De Vreese, Esser, & Hopmann, 2016).

If the interest is to test for the effects of content exposure all together, it may make sense to create a single exposure measure. The same is true when there are no significant content differences across outlets (see e.g., Schuck & De Vreese, 2008). Of course, if content features differ substantially between outlets this obviously does provide a good rationale to consider a linkage approach. If summed or additive exposure measures are used, it is recommended to run genre/medium/outlet analyses to see if general exposure effects are (in particular) driven by specific outlets. If a person is exposed to positive news in one outlet and negative in another, this would lead to exposure to balanced news in a combined measure which might be of interest, pending the question. Of course, if the contents are basically identical across different outlets, a weighted measure will correlate strongly with an unweighted "mere" exposure measure. But it could still be of interest to "deflate" a high (unweighted) exposure score if the outcome of the content analysis shows virtually no relevant content, or vice versa in the case of a high degree of relevant homogenous coverage. At any rate, knowing about the content even if it does not affect the actual analysis is of course superior to mere *assumptions* about content.

In the absence of a specific (theoretical) question about medium, genre, or total exposure, the choice for specifying the exposure measures can also be made on the basis of the content features. If, say, all newspapers report in a similar vein about a topic, it may make less sense to differentiate reading these papers. However, as Bachl and Scharkow (2017) convincingly demonstrate, content analyses typically convolute and "suppress" variation due to coding misclassifications so it is wise to consider additional correction of content data before concluding a lack of, e.g., between-outlet variance (see Bachl & Scharkow, 2017, for recommendations on correction).

In the absence of actual variation, it might rather be a question about how much content is consumed, regardless of the actual outlet. As an example, if all news is reporting negatively about the economy, it should matter more whether a respondent is exposed to less or a lot of news rather than which specific outlet s/he is exposed to. Conversely, if newspapers report very differently about a topic (such as the economy), this can be made visible by keeping the papers separately in the analysis. In general, *it is recommended to evaluate the relevant content features* first *and in order to inform the choice of how to combine different outlets in the analysis and to provide information about the robustness of the findings if alternative specifications were used.*

Let us now turn to an actual example of linking. We here look at constructing a weighted measure of exposure to the visibility of a specific topic. The weighted exposure measure (X_1) is calculated by weighting the media exposure measure using the following equation:

$$X_{1i,t} = \sum_{j} exposure_{i,j,t}^{*}visibility_{j,t}.$$

With *exposure*$_{i,j,t}$ the number of days respondent *i* reports to use outlet *j* in a typical week at time point *t*, and *visibility*$_{j,t}$ the mean visibility of topic X in outlet *j* in the period preceding time point *t*.

This example is relatively straightforward. Media effects theory might suggest different additional steps to consider in the analysis (e.g., Geers & Bos, 2016). Whether or not these are pertinent

depends on the research topic and how prescriptive theories and conventions are. For example, if the theoretical question is about the impact of the absolute visibility of, say, the economy, maybe there is no need to consider other topics in the exposure measure. However, if the theoretical question is about the relative importance of the economy (vis-à-vis other topics), one might consider a relative rather than an absolute measure.

Some research suggests a primacy effect, where information is most impactful when it is first encountered, while other research puts a premium on recent information, and yet other research areas are agnostic or silent about this. Depending on the expectations, one might include for example a higher weight to more recent information (recent as in closer to the date of interviewing for example). Again, as a general guideline, and if such specification is used, we recommend reporting alternative outcomes of in-/excluding such a weight as additional robustness check and to be fully transparent of how such analytical choices affect the outcome. Accordingly, we will pick up this aspect again in the section about robustness and additional analyses below.

Another consideration could be whether to consider the mere visibility of a topic or also the valence (tone) of the coverage. We will return to the question if these should be considered to weigh in evenly in terms of relative importance. Again, theory might prescribe that visibility is the key factor to look at, or tone, or a combination, and extant research might be ambivalent. The same goes for, for example, frames, which can have a separate effect or one in conjunction with the visibility of the framed topic. In such cases it can be worthwhile specifying a model with visibility separately, one with tone or frame separately, and one combined to try and disentangle this question. Likewise, if one is interested in the impact of media candidate evaluations on sympathy for candidates in a campaign, one might want to control for the visibility of these candidates in the news. For example, when selecting content indicators, one needs to think about confounding factors.

Again, the general recommendation is to present one or more analyses, while evaluating the impact and *robustness of the findings if alternative specifications were used. Such alternative specifications can be reported in a separate robustness section, notes or appendices, depending on the prescribed format or conventions.*

Analysis steps

Once a weighted content score (combining data on individual exposure and content analysis) has been assigned to a respondent, this score is used as an independent variable to explain for example attitudes, behavioural intentions, or knowledge (i.e., dependent variable). The analytical strategy depends on the substantial research question and the research design. In cross-sectional designs the analysis specifies a relationship rather than a causal mechanism.

Panel studies offer (a) the opportunity to focus on *change* in the dependent variable between the current wave and the previous wave and (b) analyze and use specifically the content between those panel waves to account for changes (or stability) in the dependent variable. Below, we provide a description of some of the more basic statistical models by which the data can be analyzed. In many instances, it is recommendable to run different specifications to test the robustness of the findings. Angrist and Pischke (2008) recommend running a lagged dependent variable model and a fixed effect model to bracket the true effect.

Statistically, several considerations play a role in determining an appropriate model. In many instances, there is not a single "right" solution and different specifications can be used to test the robustness of the findings. Important considerations relate to the *temporal* dependencies of observations, as well as *variation across individuals* that need to be properly modeled. The most basic model that is often chosen is a lagged dependent variable model. In that situation the score on the dependent variable in the previous wave is added as an independent variable. If the survey has more than two waves, standard errors can be corrected for the fact that observations are not completely independent—each respondent has multiple observations in the dataset. Additionally,

time fixed effects (wave dummies) can be added to account for aggregate (overall) changes in the level of the dependent variable between waves. By controlling for this previous value, the aim is to achieve two goals: (a) the focus is on how the score on the dependent variable has changed compared to the previous measurement; and (b) unit level heterogeneity (i.e., differences across individuals that cannot be accounted for by the independent variables in the model (i.e., the media variable(s)) are captured by the lagged dependent variable. While these goals are sometimes achieved, this is not always the case.

First, if the correlation between the current value of the dependent variable and the previous dependent variable is very high (say above .9), estimation of effects of other variables will be problematic (see Achen, 2000, for a statistical discussion)—a situation that is comparable with the issue of non-stationarity in regular time series. In that situation, it is recommended to subtract the previous value from the current value (*difference*) of the dependent variable and explain *changes* in this variable. Using a change-variable usually also solves the issue of heterogeneity: changes often do not differ systematically across individuals with different characteristics.

Second, the lagged dependent variable might not capture heterogeneity in an appropriate manner: there still exists variation across individuals that is not captured by variables in the model, which can cause omitted variable bias. One option to deal with this is to add individual-level control variables that can account for this variation—usually these are background characteristics, or variables such as political interest. In general, comparable to the context of experimental research, adding control variables can in all instances function as a robustness check for the findings from a simple lagged-dependent variable model and in some instances, they might contain relevant information as well. However, if there is no a priori *theoretical* reason to expect an effect from a certain control variable, it might be better not to include it in the model for reasons of parsimony.

An alternative approach to dealing with respondent-level heterogeneity is to rely on a fixed effects model (Baltagi, 2001). In that situation, dummy variables are added for each respondent (minus one) as independent variables to account for inter-respondent variation (least squares dummy variables fixed effects).[1] In that situation, the focus of the analysis is on variation *within* the individual, since all variance across individuals is absorbed by the dummy variables (and adding control variables does not make sense either). Unlike the lagged dependent variable model, after including the fixed effects, there is assumed not to be any autocorrelation left between the waves. Fixed effects can usually not be combined with a lagged-dependent variable because a combination will result in estimations that are not consistent, unless the number of waves is considerable (i.e., more than 5—depending on the stability of the dependent variable and the level of respondent-level heterogeneity). Table 2 below summarizes the discussion above and the characteristics of each of the models.

A more recent application of panel modeling in communication science is the random intercept cross-lagged panel model. Baumgartner, Van Der Schur, Lemmens, and Te Poel (2017) point out that regular cross-lagged panel models do not distinguish between- and within-person changes well over time. They propose an adapted version of the cross-lagged panel model based on a multilevel logic—the random intercept cross-lagged panel model (RI-CLPM) (Hamaker, Kuiper, & Grasman, 2015). The RI-CLPM allows distinguishing between within-person and between-person variation. This approach is beneficial to consider when modeling panel data and looking for effects of exposure to actual content as well. *In general,*

Table 2. Different specifications.

Model	# waves	Effect of LDV	Heterogeneity w/o control variables
Lagged-dependent variable model	2- +/- 5	~<.9	no
Lagged-dependent variable model with control variables	2- +/- 5	~<.9	yes
Change model	≥2	~>.9	(almost) per definition absent
Fixed effects model	2- +/- 5	not included	per definition absent
Fixed effects model with lagged dependent variable	~ >5	~<.9	per definition absent

the model specification is a function of both the substantive question and the design and measures available in a study whereby specific models have specific (dis)advantages as outlined above.

Robustness and additional analyses

We have touched upon the need to check for the robustness of findings above in the respective sections that deal with the different approaches and substantive content (and formal) features to consider in a linkage approach. In more general terms, it is important to consider (and discuss) plausible alternatives to the specification a researcher chooses to present, to develop a sense of how robust or dependent on specific assumptions results are. When it comes to operational decisions researchers take in their analytic approach, Gelman and Loken (2013), for example, have warned of "the garden of forking parks," i.e. researchers taking decisions which are contingent on their data and performing analyses which might be plausible as such but for which equally plausible alternative analyses would have been possible as well (yielding other results). Also, recent more general discussions about reproducibility in the social sciences and concerns regarding "p-hacking" all point to the importance to demonstrate the robustness of findings (e.g., Gelman, 2015; Vermeulen et al., 2016).

Regarding the linking of media content and survey data this means researchers should consider reasonable alternatives to their own operational decisions in order to demonstrate how sensitive their own analysis is to such decisions. Generally, for researchers using this approach this implies to consider different specifications of the above formula and document how findings change when different alternative specifications are employed. This can not only substantiate and strengthen the presentation of findings but also make a valid contribution to theory-building and inform future empirical research. Thus, a general advice is to assess different options and summarize these in footnotes or in a robustness/alternative specification section and/or in a technical appendix.

To avoid misunderstandings, and as discussed elsewhere (Schuck et al., 2016b), the decision of the final model specification should not be driven by empirical considerations (alone) and non-consistent findings with alternative specifications do *not* refute the core findings derived from a theory-driven model. Rather, such differences can well be explained by the different emphasis given to certain content features in a given model, for example, or by different weights that are applied. This is not problematic as such, rather *we recommend to report the outcomes of alternative specifications to (a) assure full replicability and transparency and to demonstrate the robustness of findings and/or their sensitivity to alternative operationalizations, and (b) to inform future research and further advance the theoretical and empirical discussion regarding the role and impact of such alternative specifications.*

Factors to routinely check involve the *visibility* and *prominence* of a given content feature. Most approaches weigh in and match mean scores of a certain content feature with exposure (see, e.g., Lengauer & Höller, 2012). However, such mean scores can be based on just a few observations or a large number. If *visibility* is considered, i.e., in terms of total volume or frequency of a certain content feature, there are different ways to weigh these in. One approach is to multiply the content score with the visibility score before combining it with exposure. However, if an analysis includes for example print and TV coverage the mere visibility of relevant items (i.e., content features) will typically be higher for print coverage because of media format differences—but likelihood of actual exposure to a *particular* news item might arguably be higher for TV news items.

The question how print and TV news items compare in terms of their likely influence also extends to the concept of *prominence*. Prominence is different from visibility in that it not only considers mere *volume* but also the *placement* of an item. The underlying assumption being that prominent placement increases the likelihood of actual exposure and processing of such news items (see, e.g., Vliegenthart, Schuck, Boomgaarden, & De Vreese, 2008). Here the problem of how different placement within the same medium compared to another needs to be considered. For example, is a news item on the title page of a newspaper twice as important (i.e., prominent in terms of likely impact and attached importance as well as likely attention devoted to it) as an item on page 2 or further back in the same newspaper? Furthermore, as for visibility, the same problem of how different media compare to one another must be

considered with regard to prominence as well. For example, if a news item on a newspaper title page is considered to be more important, how does this translate to relative prominence assessments for TV news items? Would one then, for example, also need to weigh the first item of the televised news more?

In general, more research is needed tapping into the question of how different placement in different media affects attention and likelihood of exposure (and thus relative prominence and, consequently, likely impact), but in absence of such studies researchers need to take reasonable choices and be aware that these can impact their findings significantly (Geiss, 2016; Wolling, 2002). Additional formal news features researchers might want to consider include the *size* or *length* of news items, as well as the use of *visuals* or other visual features (such as headlines vs. text, etc.), for example.

In a recent study, Geiss (2016) takes a stab at some of these questions. He concludes that outlet-specific and fine-grained exposure measures are superior to binary measures when tapping media use in surveys. More specifically on the content features, he suggests that "news stories should be weighted with indicators of their salience [...] e.g. by their positioning or their headline size" (Geiss, 2016, p. 5). He does not find conclusive evidence with respect to memory curves. He also does not find much difference as to whether news items should be weighted jointly by their length/duration and their salience or by their salience only. Obviously work on such content feature combinations is in progress and the field is yet too premature to make definitive recommendations.

Aside from considering news attention and the visibility and/or prominence of certain content features we also recommend paying special attention to the use of different possible *weights* to content features, the consideration of *alternative content features* as well as the simultaneous presence of *multiple content features* with theoretical relevance. If the same results can be obtained regardless of which content feature is weighed in, or differences are marginal, it needs a strong theoretical argument for why such analysis is still meaningful. Justifying and explicitly addressing ones' choices in this respect is much akin to discussions about the use of time lags in (pooled) time-series analyses which are sometimes justified and sometimes not (see, e.g., Wilson & Butler, 2007).

To be clear, we do believe that such theoretical reasoning is of central importance to any linking analysis and can indeed provide strong reasons for why results based on specific modeling choices are relevant and valid. Obtaining the same or similar results with other empirical operationalizations does not per se refute a theoretically driven model. The same is true if other operationalizations which seem equally plausible at first hand lead to different findings. In any scenario, sound theoretical reasoning should instruct and guide modeling choices. However, such alternative operationalizations should be tested and reported to assure transparency and replicability.

Additionally, the same is true for alternative operationalizations considering *multiple content features* and their relative importance and relation to another. Weighting in one specific content feature when there are several other relevant ones, or even possibly opposing ones, can be detrimental to the conceptual strength of linkage approaches (i.e., the increased precision of effect estimates). For example, if there are opposing content features (e.g., one that is assumed to be mobilizing and one that is supposed to be demobilizing), weighing in only one without considering the other can be conceptually problematic. Focusing on one without the other might be less precise than considering both or their (outlet-specific) relation to another. Attempting to build in both does pose the question how such different content features relate to one another (both in terms of the direction of the assumed impact but also in terms of their relative contribution and/or relation to another, etc.). This could require subtracting one content score from another, for example in a study testing the impact of strategy framing on cynicism, which could also consider the presence of more substantial issue focus in the same coverage and not just that of strategy framing alone. There is no clear-cut answer how to deal with this and it again depends on the theoretical puzzle at stake, but when more concrete empirical evidence is absent, and *as a general rule of thumb, it is recommendable to document the results of different alternative specifications and to discuss their respective underlying*

assumptions as well as their conceptual implications. Reporting the outcomes of different approaches can also inform other studies and contribute to future theory-building and an emerging picture of how such content features might relate to another.

Finally, as touched upon above, researchers should also consider weighing in *time*. For example, in campaign studies or any other design tracking the impact of news coverage over time and with reference to a specific event in time, it seems plausible to attach greater weight to news items published and thus processed closer to that point in time (see Vliegenthart et al., 2008). Again, we are missing empirical evidence what a reasonable weight would be, e.g., with regard to the assumed difference in the impact of the same news item published 30 days, 5 days, or 1 day before an event. Thus, there are different options to account for such differences in timing, e.g., adopting a more linear approach or one (e.g., logarithmic) that attached even greater weight to the most recent coverage. Again researchers are recommended to test and report reasonable alternatives, not only to demonstrate the robustness or volatility of their findings to a certain operational decision but also to inform other researchers who can use such knowledge for their own operational decision or further theory development.

In addition to the alternative specifications, such as outlined above, researchers may also consider extrapolating from their content and/or survey data to create *counterfactuals*. For example, it may be estimated what effects would be if individuals were exposed more or less by using one or more media differently, holding the level of media content constant. Or what would happen if media use patterns were held constant, but if media content in one or more media looked differently. Such an approach was for example used in De Vreese, Azrout, and Möller (2016) and can help to illustrate findings and the magnitude of effects.

To conclude with, *one of the key promises and advantages of any linkage approach is an increase in precision of the estimates and a closer link between theory and data.* However, such precision is necessarily and inevitably compromised not only by the reliability of survey measures but also by the reliability of the other part of such data combination, i.e. the media content measures. As recent research has shown, low reliability of these components usually is connected to an *under*estimation rather than an exaggeration of effects (see Scharkow & Bachl, 2016). Therefore, *we recommend that researchers should reflect on the reliability of their measures and consider available correction methods and see this as further encouragement to develop or employ more reliable measures as otherwise they risk that the exact benefit of such linkage approaches, i.e., increased precision of their estimates, is compromised by low reliability of their survey and content analytic measures.*

An example

The section below provides a straightforward example of how a content analysis can be combined with data from a survey to analyze how specific content features influence a certain dependent variable. The data originate from Vliegenthart's NWO-VIDI (2015–2019) project "Media coverage as a catalyst for economic crises? Causes, content and consequences of economic news coverage." A content analysis of Dutch print media and a three-wave panel survey among a sample of the Dutch population in the period February to June 2015 were conducted.[4]

Please notice that the analyses are purely a *statistical example* of how such data can be used in a linkage analysis; therefore, we do not go into further detail about any theoretical rationale that underlies the models we present down here or of any plausible alternatives which as we stressed above should also be reported and discussed in an actual linkage study. The analyses predict how exposure to economic news itself (*visibility*, which is a count of the number of news items) as well as an additional content feature (*tone*) predict what respondents expect with regards to the state of the national economy in the coming 12 months. This outcome variable is presumably affected by news coverage as aggregate-level analyses have shown (Damstra & Boukes,

forthcoming). Additionally, we analyze whether the strength of this effect depends on the *prominence* in the newspaper of the articles or the *recency* by which they are published. To come to the analysis, we followed a number of steps. The datasets as well as an annotated STATA code that describe the procedure step-by-step are available online as supplemental materials.

Within the content analysis dataset

Step 1. Generating a variable indicating the wave: First, a variable is created which identifies whether an article is published *before* Wave 1, between Wave 1 and Wave 2, between Wave 2 and Wave 3, or after Wave 3.

Step 2. Generating a variable indicating the week (recency moderator): Analyses will include a measurement of tone that is weighted by *recency* (more recent articles arguably would have a stronger effect). A variable is created that indicates the particular week in which an article was published. Following the approach of Vliegenthart et al. (2008), the strongest weight is given to the most recent week and this weight gradually decreases.

Because there was a time period of eight weeks between the waves of the survey, the last (most recent) week before a survey wave commenced was weighted by 8, the week before by 7, etc., until the last week which received the weight of 1. To generate results that are comparable to the unweighted measurement, these weights are divided by 4.5, which is the average weight, i.e., $(8 + 7 + \ldots + 2 + 1)/8 = 4.5$.

As an alternative to this recency moderator, we also generated a variable indicating whether a news item was published in the two weeks before the survey wave or the weeks further away in time. The impact of recent articles (i.e., within two weeks of the survey launch) was weighted by $\frac{4}{3}$ whereas for the others this was set to $\frac{2}{3}$. Accordingly, recent articles can later be given a double weight of those of non-recent articles, while altogether the non-standardized effect coefficients remain comparable (i.e., $[\frac{4}{3} + \frac{2}{3}]/2 = 1$).

Step 3. Selecting the news items of interest: The dependent variable is citizens' expectation of how the Dutch economy will develop in the coming year. For the purpose of this example, we therefore *only* analyze the articles that contained an evaluation of the Dutch economy; articles that did not carry an evaluation of the economy were discarded. In our article sample, 6,671 articles did not evaluate the state of the Dutch economy, and were therefore dropped from the analysis just as all other articles that did not deal with economic news. 1,211 articles remained that are used in the analyses.

Step 4. Constructing the independent variables: We construct a variable in the content analysis dataset that measures the tone of an article. Tone was measured using a scale that runs from −2 (completely negative) via 0 (mixed or neutral) to +2 (completely positive). The stories with an evaluation are used in the analyses. The tone varied over time and on average became more positive toward the end of the survey period (i.e., with every additional day later in 2015), $b = .002$, $p = .020$.

Step 5. Constructing the moderating variables: Additional variables have to be created to analyze whether effects of a news item's tone was stronger for items that were published more *recently* before a survey wave or more *prominently* in a newspaper. Respectively, the tone variable was multiplied with the recency factor or with a factor indicative of the length (i.e., [tone × length]/average length). Including these newly created variables in a linear model and comparing results with a model that includes the unweighted variable, we can test whether more recent and more prominent articles have stronger effects.

Step 6. Aggregating dataset to outlet/wave level: Now the necessary variables on the level of the individual news item have been created, the dataset has to be aggregated into observations per outlet per time period. The dataset is therefore aggregated on two variables: wave and outlet.

This means we get scores for every variable per outlet per wave. In practice, this means that we create a new dataset that has observations for the number of articles, the sum of the tone in those articles, and the weighted tone of those articles (i.e., tone × article recency or length) for every newspaper and for every time period.

While aggregating the dataset, the researcher can choose to calculate the sum of all observations on one variable or the mean score. In terms of tone, this would involve (a) calculating the tone by adding up all the individual tone scores in a summed score or (b) calculating the average of all the individual tone scores in a mean score. In the following analyses, we used the sum scores. The advantage of sum scores is that the interpretation of effect coefficients is more straightforward. From this sum score, mean scores can still easily be calculated by simply dividing this variable by the number of articles per outlet/wave-combination (i.e., visibility).

Within the survey dataset

Step 7. Preparing the media exposure independent variable: The survey dataset needs to include exposure variables that measure how often an individual respondent is exposed to the media outlets available in the content analysis dataset. First, variables are created that indicate for every individual outlet on a scale from 0–1 how often a respondent is exposed to it. This means that in the next step in which media exposure data are combined with data of the content analysis results, values of zero will be assigned to respondents who were not exposed to a certain media outlet (i.e., multiplying by 0), and a score that is *equal to* the observation in the media content analysis for respondents who always consume a certain outlet (i.e., multiplying by 1).

Second, a variable is created that indicates how often respondents are exposed to the combination of media outlets that are analyzed. This variable is generated by summing the variables that were created for the individual outlets. It functions as the "empty" media exposure variable in which media content is *not* taken into account.

Step 8a. Preparing the dependent variable: This is a straightforward step that is part of any analysis; also those not linking survey and content analysis data. Create the dependent variable that one is interested in; when using latent scales, reliability needs to be satisfactory in all the waves. For practical reasons, one gives these newly created variables a consistent name that ends with the wave number (e.g., EconomicExpectationNL_Wave1).

To improve the assessment of causality, analyses can control for the lagged dependent variable. This means that when the economic expectation in Wave 2 is estimated, analyses include the expectation respondents held in Wave 1 as a control variable.

Step 8b. Preparing the dependent variable (change scores): An alternative way to control for existing values is by using the calculated *change* in the dependent variable as variable to be predicted in the analysis (see above). This is particularly recommended when the correlation between the dependent variables of the different waves is strong (i.e., above .9). This dependent variable is calculated by the following calculation for every survey wave (except for the first): Dependent variable $_t$ – Dependent variable $_{t-1}$

Combining the content analysis and survey datasets

Step 9. Creating media content weighted exposure variables: The actual linkage of content analysis and survey measurement is done in the survey dataset; after all, that is also where one finds the dependent variable. New variables need to be computed in the survey dataset. For every individual i per wave j, we multiply the individual exposure to media outlets k with the presence of the content feature of interest in the specific outlet k. Hence, the formula that we follow is:

$$\text{Independent variable }_{ij} = \sum_i \text{ exposure to }[\text{outlet } k_i] \times \text{content feature of }[\text{outlet } k_j].$$

The self-reported exposure variables are already available in the survey dataset; the content feature scores need to be retrieved from the aggregated content analysis dataset. Therefore, the content feature scores are manually added to the syntax (i.e., do-file) that calculates the individual exposure to a content feature in a certain media outlet (Step 5 in the Appendix generated a spreadsheet with these results that can be used for copy-pasting).

This needs to be done for every outlet/wave combination. In our case, we have three waves that we are interested in and seven national newspapers. So, three variables have to be created per independent variable, which consists of the summation of the exposure to the seven outlets multiplied by their respective content feature score (i.e., visibility, tone, or weighted tone). The reason to sum the scores of the different outlets is that, for this example but also in many other cases, no substantially differential effects for different outlet types were expected. In other cases, one may decide to calculate separate scores and variables for different outlet types (e.g., popular vs. quality newspapers; television vs. newspapers; etc.).

Now, we have created a dataset with on every row a respondent's scores on the independent variables (i.e., how much exposure to a certain content feature) as well as on the dependent variable (i.e., expectation regarding the Dutch economy in Wave 1, 2, and 3).

Step 10. Reshape the dataset into a long format: If, *a priori*, one has no theoretical expectations that relationships between variables should differ for the different waves of data collection, these are most efficiently analyzed in one pooled model. To do so, the dataset needs to be organized in such a way that each observation represents a combination of respondent and wave. After reshaping the dataset, every respondent (*n*) fills a number of lines in the dataset that is equal to the number of survey waves (*j*). So, in total the dataset will be filled with $n \times j$ cases.

Step 10b. Creating z-standardized variables (standardized): Because many of the statistical models necessary to analyze panel data (e.g., fixed and random effects models, regressions with clustered standard errors) cannot generate standardized effects, it is necessary to *manually* create standardized versions of the variables (i.e., dependent and independent variables) that will be used in these analyses. Variables are z-standardized by subtracting the mean from the original scores and then dividing the difference by the standard deviation. While the unstandardized variables (Step 9) allow for more straightforward interpretation of effect coefficients, standardized variables give insight in the effect strength of independent variables and allow for comparison across variables.

Analyses

Using the dataset created above, we can analyze the causal relationship between the content to which people are exposed and their subsequent economic expectations. We start with an analysis of how news coverage affected the perceptions of the national economy in Wave 2 and Wave 3; Wave 1 is only used for the lagged dependent variable and to calculate the change score for the dependent variable in Wave 2.

Step 11. Analyze correlation between dependent variable in different waves: First, it is important to assess the *correlation* between the dependent variables of the different waves. In this example's case, expectations of the future national economy in one wave correlate on average at $r = .65$ with the expectation in the subsequent wave. Hence, analyses using the lagged dependent variable are most appropriate, because $r < .90$. However, checking the robustness of these findings with change models can further substantiate the confidence in our results. Accordingly, we present these as well for our final model (Model 4d, in Table 4).

Table 3. Correlations between independent variables: i.e., self-reported exposure, actual exposure to articles about economy (visibility), the tone to which respondents are exposed, and the weighted tone.

	Self-reported exposure	Visibility	Tone	Weighted tone (recency I)	Weighted tone (recency II)	Weighted tone (prominence)
Self-reported exposure	1.00					
Visibility	0.85	1.00				
Tone	0.23	0.51	1.00			
Weighted tone (recency I)	0.18	0.46	0.98	1.00		
Weighted tone (recency II)	0.17	0.44	0.98	1.00	1.00	
Weighted tone (prominence)	−0.03	0.19	0.76	0.67	0.67	1.00

Step 12. Check for issues of multicollinearity: Next, it is important to check how strongly the different independent variables correlate with one another to avoid or at least get insight in potential issues of multicollinearity. Table 3 shows the correlation table between the six independent variables of interest.

With strong correlations between, especially, the self-reported exposure variable and the visibility variable ($r = .85$) we have to be careful about potential issues of multicollinearity and preferably do not include these simultaneously in one model. The strong correlations with self-reported exposure indicates that there was limited variation in visibility between the different outlets; so, the more one is exposed to any of the outlets, the more likely one is to pick up a certain visibility of the economic issue, irrespective of which outlet a respondent consumes. Although there is a moderately strong correlation between the visibility and tone variable ($r = .51$), this level of correlation does not cause problems when including both in the same model.

Because the variable "self-reported exposure" correlates strongly with visibility in particular and also with tone, potential multi-collinearity issues could arise when included in the same model. Therefore, analyses have been done with exposure and visibility simultaneously in one model as well as separately; so, potential differences in findings can be observed.

The correlation table also shows that the tone variable that is weighted by recency in both cases (recency I: decreasing weight over time; recency II: double value to items in the 2 weeks before survey wave) very strongly correlates with the original tone variable. This implies that the tone did not differ much in the different weeks before the survey waves. Because of the strong correlation between the original tone variable and the weighted tone scores, it is not possible to include these in *one* regression model to assess whether there is an interaction effect between tone and recency that amplifies the effect of tone (i.e., such a model would include variables for tone, recency, and for tone weighted by recency). Consequently, we add these in separate models and compare the effects (and effect strengths).

Step 13. Step-by-step OLS regression analysis: Analyses are gradually built up. Having a dependent variable that is measured on an interval scale (0–10) and an independent variable on ratio level (i.e., number of articles, the sum of tone), OLS regression technique is the appropriate way to analyze this relationship. Because we have repeated measurement for individuals, we use robust clustered standard errors to correct for intra-individual correlation. Additionally, we control in the analyses for the wave of the analysis to account for time fixed effects.

We estimate five different regression models that built on the previous model: Model 1—Analysis with only the lagged dependent variable; Model 2—Analysis which includes self-reported exposure to newspapers; Model 3—Analysis which includes the actual exposure to economic news items; Model 4—Analysis which includes the tone of news to which the respondent is actually exposed; Model 5—Analysis with weighted measurements of tone.

Model 1. Table 4a shows the results of the analysis with economic expectations using the lagged dependent variable (Model 1). The models show that the lagged dependent variable

Table 4. Regression models predicting economic expectations in wave 2 and 3.

Independent Variables	Model 1					Model 2					Model 3a					Model 3b				
	Only lagged DV					Self-reported exposure					Self-reported exposure and visibility					Visibility only				
	b	s.e.	b^*	t	p	b	s.e.	b^*	t	p	b	s.e.	b^*	t	p	b	s.e.	b^*	t	p
Intercept	2.10	0.10		21.96	0.000	2.09	0.10		21.72	0.000	2.09	0.10		21.70	0.000	2.10	0.10		21.84	0.000
Wave	−0.02	0.03	−0.01	−0.81	0.418	−0.02	0.03	−0.01	−0.81	0.419	−0.02	0.03	−0.01	−0.79	0.432	−0.02	0.03	−0.01	−0.81	0.418
Lagged DV: Expectations Wave$_{t-1}$	0.64	0.01	0.64	52.03	0.000	0.64	0.01	0.64	51.97	0.000	0.64	0.01	0.64	51.94	0.000	0.64	0.01	0.64	52.00	0.000
Self-reported exposure						0.01	0.02	0.01	0.62	0.535	0.09	0.05	0.04	1.87	0.061					
Visibility											−0.00	0.00	−0.04	−1.74	0.082	−0.00	0.00	−0.00	−0.12	0.904
Tone																				
R^2			0.43					0.43					0.43					0.43		

Independent Variables	Model 4a					Model 4b					Model 4c					Model 4d				
	Visibility and tone					Tone					Self-reported exposure, visibility, tone					Change model (DV$_t$ − DV$_{t-1}$)				
	b	s.e.	b^*	t	p	b	s.e.	b^*	t	p	b	s.e.	b^*	t	p	b	s.e.	b^*	t	p
Intercept	1.97	0.10		20.08	0.000	2.01	0.10		20.59	0.000	1.97	0.10		20.10	0.000	0.41	0.10		4.32	0.000
Wave	−0.06	0.03	−0.03	−1.94	0.053	−0.04	0.03	−0.03	−1.36	0.173	−0.06	0.03	−0.03	−1.88	0.060	−0.15	0.04	−.10	−3.88	0.000
Lagged DV: Expectations Wave$_{t-1}$	0.64	0.01	0.64	52.09	0.000	0.64	0.01	0.64	52.00	0.000	0.64	0.01	0.64	52.05	0.000					
Self-reported exposure											0.02	0.05	0.01	0.46	0.644					
Visibility	−0.00	0.00	−0.04	−3.31	0.001						−0.00	0.00	−0.05	−2.11	0.035	−0.00	0.00	−0.05	−3.76	0.000
Tone	0.01	0.00	0.11	4.65	0.000	0.00	0.00	0.05	3.26	0.001	0.01	0.00	0.11	4.27	0.000	0.01	0.00	0.11	3.54	0.000
R^2			0.43					0.43					0.43					0.00		

Note. Cells contain unstandardized (b) coefficients with robust clustered standard errors (s.e.) in parentheses, standardized coefficients (b^*) and probabilities (p; two-tailed).

has a strong impact on the dependent variable with a standardized regression coefficient above .60.

This model without any other independent variables already explains 43% of the variance in economic expectations. In the next models (2, 3, and 4), the R^2 did not increase. This implies that the inclusion of media content did not improve the overall prediction of economic expectations, but it still adds precision in the measurement as such.

Model 2. Adding the self-reported exposure measurement to the model did not yield an effect of media exposure: Self-reported exposure had no significant effect.

Model 3. In Model 3a and 3b, we include the visibility variable which indicates the number of articles to which a respondent has been exposed as estimated by the combination of survey (i.e., self-reported exposure to outlets) and content analysis data (i.e., count of news items per outlet). Separate models are run for the combination of this visibility variable with the self-reported exposure as well as with the visibility variable alone; the reason being that the two variables strongly correlate ($r = .85$, see Table 3) and the analysis may thus have to cope with issues of multicollinearity, even though Variance Inflation Factors (VIF) are yielded below the rule-of-thumb cut-off point of 10, VIF < 7.4.

The analysis shows that substantially different results are yielded when both variables are included in one model vis-a-vis a model that only estimates the effect of visibility: Visibility has no effect when included alone in the model (3b), but visibility *negatively* predicts the economic expectations when analysis controls for self-reported exposure (Model 3a).

Comparing Model 3a and Model 3b for Wave 3, thus, showed that inclusion of highly correlated factors can significantly influence the findings. In this case, a marginally significant effect was only yielded when both the self-reported exposure *and* visibility variable were included in one model. Admittedly, the unstandardized effects are rather small. These unstandardized coefficients guide the interpretation of substantive interpretation of findings (i.e., the effect of being confronted with one extra newspaper article about the economy); yet, for the interpretation of effect strengths, one should examine the standardized effects (Beta) that are between small and moderate.

Model 4. In the next step, we included the tone as independent variable in the model (Model 4a/b, Table 4). Because self-reported exposure and visibility correlate strongly ($r = .85$; Table 3), we do not include the exposure measurement in basic model (4a). The correlation between visibility and tone is moderately strong ($r = .51$; Table 3). We analyze the data with visibility and tone simultaneously (Model 4a) as well as with tone alone (Model 4b) to show the differences between both models. Model 4a yields VIF values that do not exceed 2.2; accordingly, no statistical problems occur when both variables are included simultaneously.

Results show that visibility has a negative effect on perceptions whereas tone has a positive effect (Model 4a). Thus, when people are exposed to more positive news, they will develop more optimistic expectations of the national economy. When visibility is dropped from the model (Model 4b), the effect of tone remains positive and significant. Similarly, tone's effect remains positive and significant in a model that additionally *includes* self-reported exposure (Model 4c). In all these models, the standardized effect of tone remains stable and considerable.

Finally, we test a model with *change scores* as dependent variable to examine the robustness of tone's effect on economic expectations. Model 4d demonstrates that the effect of tone is significant in such an equation as well; visibility has a negative effect on changes in economic expectations. In this example, the change model, thus, replicates the findings of the lagged dependent variable model (4a), which demonstrates the *robustness* of the relationships that are established.

Table 5. Regression models predicting economic expectations with different operationalizations of (weighted) tone.

Independent Variables	Model 5a Tone without weight					Model 5b Tone weighted by recency					Model 5c Tone weighted by recency of 2 weeks					Model 5d Tone weighted by prominence				
	b	s.e.	b*	t	p	b	s.e.	b*	t	p	b	s.e.	b*	t	p	b	s.e.	b*	t	p
Intercept	1.97	0.10		20.08	0.000	2.12	0.10		21.97	0.000	2.13	0.10		22.07	0.000	1.78	0.12		14.89	0.000
Wave	−0.06	0.03	−0.03	−1.94	0.053	−0.11	0.04	−0.06	−2.77	0.006	−0.13	0.04	−0.07	−3.18	0.001	0.07	0.04	0.04	2.02	0.044
Lagged DV: Expectations Wave$_{t-1}$	0.64	0.01	0.64	52.09	0.000	0.64	0.01	0.64	52.09	0.000	0.64	0.01	0.64	52.10	0.000	0.64	0.01	0.64	52.15	0.000
Self-reported exposure																				
Visibility	−0.00	0.00	−0.04	−3.31	0.001	−0.00	0.00	−0.03	−2.75	0.006	−0.00	0.00	−0.03	−3.00	0.003	−0.00	0.00	−0.00	−0.38	0.701
Tone	0.01	0.00	0.11	4.65	0.000	0.01	0.00	0.11	3.67	0.000	0.01	0.00	0.12	4.38	0.000	0.00	0.00	0.06	4.12	0.000
R^2			0.43					0.43					0.43					0.43		

Note. Cells contain unstandardized (b) coefficients with robust clustered standard errors (s.e.) in parentheses, standardized coefficients (b*) and probabilities (p; two-tailed).

Model 5.Next, we can examine whether the effects of tone vary when these are weighted for how recently news items were published or how prominently those were presented in a newspaper (i.e., longer length). Because the models that include both visibility and tone did not cause problems with multicollinearity, we start with Model 4a that was already presented in Table 6, and then follow with a variety of weighted tone measurements.

Model 5a is this basic model that provides the estimates for an unweighted measurement of tone. In Model 5b, tone is weighted by the recency of the article, whereas model 5c weights articles in the two weeks *before* the survey double as compared to the articles published in the weeks more distant in time (i.e., 3–8 weeks). Model 5d presents the results for a tone measurement that is weighted by prominence.

The findings in Table 5 demonstrate that the effect of tone does not become stronger for weighted measurements. The standardized effects of tone weighted by recency (Model 5b/c) are very similar to those of the unweighted measurement (Model 5a; $b^\star$ = .11). When compared to the tone measurement weighted by prominence, the effect strength even decreases significantly ($b^\star$ = .06; 95% CI [0.03, 0.09]. So, in this case the findings are robust for different operationalizations of tone; it has a consistent positive effect, but this does not strengthen with more sophisticated weighting procedures.

Please note that the R^2 barely changes for any of the models when compared to the model that only includes the lagged dependent variable (Model 1). This means that the media content to which people are exposed does not explain much extra in terms of how economic expectations developed. So, while we find significant effects of tone and visibility, these are rather weak media effects. One can also infer this from the Beta coefficients that only marginally pass the .10 level, and hence point toward relatively weak effects. One explanation for this can be found in the measurement error that accompanies any (large-scale) content analysis and self-reports of media exposure (Scharkow & Bachl, 2016), making it unlikely to yield strong effects.

Step 13. Fixed effect models: Alternatively, one could analyze the data using fixed effect models. The advantage of this kind of models is that they make use of repeated measurements; therefore, such kinds of models can analyze the relationship between the content to which one is exposed and the dependent variable at t_1, t_2, t_i in one model. The disadvantage is that such models do not allow the inclusion of lagged dependent variables nor time-invariant variables (e.g., age, gender, but also media exposure if only measured once).

Table 6 shows the results of the fixed effect models. Model 6a shows that visibility has a positive effect on the economic expectations. This outcome is inconsistent with the findings of Tables 4 and 5 (Models 3, 4, and 5) where insignificant as well as negative effects were yielded for visibility. Hence, one can conclude that the effects found for models with visibility as *only* predictor are not robust across different estimation methods.

Model 6b (see Table 6) shows the effects for a model that combines visibility and tone as independent variables in one model. Both have a strong effect: Visibility a negative effect, whereas tone has a positive effect. This model, therefore, replicates the findings of Models 4a/b and Models

Table 6. Fixed effect models predicting economic expectations with visibility and tone.

Independent Variables	Model 6a Visibility					Model 6b Visibility and tone				
	B	Std. Error	b*	t	Sig	B	Std. Error	b*	t	Sig
Intercept	4.94	0.02		220.97	0.000	5.11	0.02		206.17	0.000
Visibility	0.01	0.00	0.17	10.75	0.000	−0.01	0.00	−0.16	−5.75	0.000
Tone						0.01	0.00	0.16	14.25	0.000
R^2 (within)	0.02					0.05				
R^2 (overall)	0.00					0.00				

5a/d that were presented before. The positive effect of tone thus proves to be robust across models and the negative effect of visibility as well *if* tested simultaneously with tone in one model.

Probably the positive effect of visibility in Model 6a turns negative in Model 6b that controls for tone, because both variables are correlated ($r = .51$, see Table 3). So, visibility in a model that does not control for tone (6a), may partly function as a proxy for the tone. Because economic news coverage in the period of study was relatively positive, higher values of visibility thus do not only indicate reading more newspaper articles but *also* reading more *positive* news. Once a model controls for tone (6b), the effects of visibility and tone can be disentangled.

Step 14. Random effect models (when possible): Instead of fixed effects models, scholars may conduct random effect models *but only* after a Hausman specification test shows that the coefficients are not significantly different from those in a fixed effect model. In our case we find a large and significant Hausman statistic, which means we should not use a random effects model, $\chi^2(2) = 28.95$, $p < .001$. Accordingly, we do not present those results.

Conclusion

This article revisited some of the rationales and requirements for conducting linkage studies. In a nutshell we believe that while survey data can say a lot about the use of media or about attitudes, survey data alone are hard to interpret in the absence of knowledge about the content of the media. The article provided an overview of different types of linkage studies and how these are currently conducted. In a more practical section, the article discussed which survey and content items to use in an analysis, what types of analyses are available, what alternative specifications could be considered, and it provided a step-by-step example. Throughout the article we tried to provide insights on which considerations and choices are available when conducting linkage analyses. We have explicitly refrained from prescribing a single approach, and offered mostly advice on what to consider, test, and report. That said, as we summarized throughout the article we consider this advice essential to linkage analyses: (1) *include exposure measures in the survey in the most granular and detailed way possible (time and resources permitting) to allow for multiple approaches in subsequent linkages*; (2) *focus on substantive content features* first *in any linkage analysis and add formal features only later in the analysis (and report the outcome as part of additional analyses and/or robustness checks)*; (3) *evaluate the robustness of the findings by testing alternative specifications (and report these)*; (4) *be attentive to the choice of model specification which should be a function of both the substantive question and the design and measures available in a study*; and (5) *reflect on the reliability of the content measures and consider available correction methods*.

In general, to advance linkage studies documenting alternative specifications and explicating the underlying assumptions can inform other studies and contribute to future theory-building and modeling choices. We hope this provides a useful reference in the continued pursuit of optimal design, measurement, and analysis of linkage studies in communication science.

Notes

1. These include, for example, Adriaansen, Van Praag, and De Vreese (2010), Arendt and Brantner (2015), Banducci, Giebler, and Kritzinger (2015), Boomgaarden, Van Spanje, Vliegenthart, and De Vreese (2011), Boomgaarden, Vliegenthart, and De Vreese (2012), Bos et al. (2011), Burscher, Van Spanje, and De Vreese (2015), Desmet, Van Spanje, and De Vreese (2015), De Vreese and Semetko (2004), De Vreese et al. (2016), Druckman (2005), Elenbaas et al. (2013), Garz (2012), Geers and Bos (2016); Hopmann, Vliegenthart, De Vreese, and Albæk (2010), Hopmann, Shehata, and Strömbäck (2015), Jebril, De Vreese, Van Dalen, and Albæk (2013), Kalogeropoulos, Albæk, De Vreese, and Van Dalen (2016), Kepplinger, Geiss, and Siebert (2012), Kiousis (2011), Lengauer and Höller (2012), Matthes (2012), Moon (2013), Nir and Druckman (2008), Peter (2004), Schuck and De Vreese (2009), Schuck, Boomgaarden, and De Vreese (2013), Schuck et al. (2016a), Shehata (2014), and Van Spanje and De Vreese (2014).
2. For a recent overview of exposure measures, see De Vreese and Neijens (2016). They summarize "typical examples of self-report exposure measures" ranging from unaided recall, to aided recall, recognition, etc. Typically, scholars have relied on some measure of exposure to a medium (e.g., television), and/or an outlet (e.g., a news show),

sometimes augmented with an exposure frequency estimate. The free and open access website www.mediaexposuremeasures.org offers an overview of existing studies and their measures of media exposure.

3. Many surveys include measures of news attention and some matching approaches include news attention (either in an *additive* or *multiplicative* approach) (see, e.g., Hobolt & Tilley, 2014; Möller & De Vreese, 2015; Schuck & De Vreese, 2008) whereas others focus on exposure alone or even warn against building in attention measures (e.g., Slater, 2004).

4. The content analysis has been conducted with a team of 22 coders (see Boukes & Vliegenthart, 2017, for further details). For the sake of simplicity, the current example only employs the data of national newspapers (i.e., not regional newspapers or other media types). The only variable manually coded and used for the analysis presented in the current study is the article's tone with which the Dutch economy is normatively evaluated in the news item. Inter-coder reliability was assessed with 170 articles of which 89 qualified as economic news. The standardized Lotus (see Fretwurst, 2015) coefficient was satisfactory (std. λ = .73).

Funding

References

Achen, C. H. (2000, July). *Why lagged dependent variables can suppress the explanatory power of other independent variables*. Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago.

Adriaansen, M. L., Van Praag, P., & De Vreese, C. H. (2010). Substance matters: How news content can reduce political cynicism. *International Journal of Public Opinion Research*, 22(4), 433–457. doi:10.1093/ijpor/edq033

Althaus, S. L., & Tewksbury, D. H. (2007). *Toward a new generation of media use measures for the ANES: Report to the Board of Overseers* (ANES Pilot Study Report No. nes011903). Retrieved from: http://www.electionstudies.org/Library/papers/Pilot2006/nes011903.pdf

Andersen, K., De Vreese, C. H., & Albæk, E. (2016). Measuring media diet in a high-choice environment: Testing the list-frequency technique. *Communication Methods and Measures*, 10(23), 81–98. doi:10.1080/19312458.2016.1150973

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Arendt, F., & Brantner, C. (2015). Toward an implicit cognition account of attribute agenda setting. *International Journal of Communication*. 9, 2735–2756. Retrieved from http://ijoc.org/index.php/ijoc/article/view/4001

Bachl, M., & Scharkow, M. (2017). Correcting measurement error in content analysis. *Communication Methods & Measures*, 11, 87–104. forthcoming 10.1080/19312458.2017.1305103

Banducci, S., Giebler, H., & Kritzinger, S. (2015). Knowing more from less: How the information environment increases knowledge of party positions. *British Journal of Political Science*, 1–18. doi:10.1017/S0007123415000204

Baltagi, B. H. (2001). Econometric Analysis of Panel Data (second ed.) John Wiley & Sons.

Baumgartner, S., Van Der Schur, W., Lemmens, J., & Te Poel, F. (2017). *The relationship between media multitasking and attention problems in adolescents: Results of two longitudinal studies. Human Communication Research*, forthcoming, doi:10.1111/hcre.12111

Boomgaarden, H. G., Van Spanje, J., Vliegenthart, R., & De Vreese, C. H. (2011). Covering the crisis: Media coverage of the economic crisis and citizens' economic expectations. *Acta Politica*, 46(4), 353–379. doi:10.1057/ap.2011.18

Boomgaarden, H. G., Vliegenthart, R., & De Vreese, C. H. (2012). A worldwide presidential election: The impact of the media on candidate and campaign evaluations. *International Journal of Public Opinion Research*, 24(1), 42–61. doi:10.1093/ijpor/edr041

Bos, L., Van Der Brug, W., & De Vreese, C. H. (2011). How the media shape perceptions of right-wing populist leaders. *Political Communication*, 28(2), 182–206. doi:10.1080/10584609.2011.564605

Boukes, M., & Vliegenthart, R. (2017). A general pattern of newsworthiness? Analyzing news values in popular, quality, regional and financial newspapers. *Journalism*. Advance online publication. doi:10.1177/1464884917725989

Burscher, B., Van Spanje, J., & De Vreese, C. H. (2015). Owning the issues of crime and immigration: The relation between immigration and crime news and anti-immigrant voting in 11 countries. *Electoral Studies*, 38, 59–69. doi:10.1016/j.electstud.2015.03.001

Damstra, A., & Boukes, M. (forthcoming). The economy, the news and the public: A longitudinal study of the impact of economic news on economic evaluations and expectations. *Communication Research*.

De Vreese, C. H., Van Spanje, J. H. P., Schuck, A. R. T., Boomgaarden, H. G., Vliegenthart, R., Azrout, R., & Elenbaas, M. (2010). 2009 European election campaign study, scenario I: Data and documentation. University of Amsterdam.

De Vreese, C. H. (2014). Kombination af indholdsanalyse og spørgeskemadata. In D. N. Hopmann, & M. Skovsgaard (Eds.), *Forskningsmetoder i journalistik og politisk kommunikation* (pp. 339–356). Copenhagen: Hans Reitzels Forlag.

De Vreese, C. H., Azrout, R., & Möller, J. (2016). Cross road elections: Change in EU performance evaluations during the European Parliament elections 2014. *Politics and Governance*, 4(1), 69–82. doi:10.17645/pag.v4i1.462

De Vreese, C. H., Esser, F., & Hopmann, D. N. (2016). *Comparing political journalism*. New York: Routledge.

De Vreese, C. H., & Neijens, P. (2016). Measuring media exposure in a changing communications environment. *Communication Methods and Measures*, 10(2–3), 69–80. doi:10.1080/19312458.2016.1150441

De Vreese, C. H., & Semetko, H. A. (2004). News matters: Influences on the vote in the Danish 2000 Euro referendum campaign. *European Journal of Political Research*, 43(5), 699–722. doi:10.1111/j.0304-4130.2004.00171.x

Desmet, P., Van Spanje, J., & De Vreese, C. H. (2015). Discussing the democratic deficit: Effects of media and interpersonal communication on satisfaction with democracy in the European Union. *International Journal of Communication*, 9, 3177–3198.

Dilliplane, S., Goldman, S. K., & Mutz, D. C. (2013). Televised exposure to politics: New measures for a fragmented media environment. *American Journal of Political Science*, 57(1), 236–248. doi:10.1111/j.1540-5907.2012.00600.x

Druckman, J. N. (2005). Media matter: How newspapers and television news cover campaigns and influence voters. *Political Communication*, 22(4), 463–481. doi:10.1080/10584600500311394

Elenbaas, M., Boomgaarden, H. G., Schuck, A. R. T., & De Vreese, C. H. (2013). The impact of media coverage and motivation on performance-relevant information. *Political Communication*, 30(1), 1–16. doi:10.1080/10584609.2012.737411

Erbring, L., Goldenberg, E. N., & Miller, A. H. (1980). Front-page news and real-world cues: A new look at agenda-setting by the media. *American Journal of Political Science*, 24(1), 16–49. doi:10.2307/2110923

Eveland, J. W., & Scheufele, D. A. (2000). Connecting news media use with gaps in knowledge and participation. *Political Communication*, 17(3), 215–237. doi:10.1080/105846000414250

Eveland, W. P., Jr, & Morey, A. C. (2011). Challenges and opportunities of panel designs. In E. P. Bucy, & R. L. Holbert (Eds.), *The sourcebook for political communication research: Methods, measures, and analytical techniques* (pp. 19–33). New York, NY: Routledge.

Fretwurst, B. (2015). *Lotus Manual: Reliability and accuracy with SPSS*. Zürich, Switzerland: Institut für angewandten Kommunikationsforschung.

Garz, M. (2012). Job insecurity perceptions and media coverage of labor market policy. *Journal of Labor Research*, 33 (4), 528–544. doi:10.1007/s12122-012-9146-9

Geers, S., & Bos, L. (2016). Priming issues, party visibility, and party evaluations: The impact on vote switching. *Political Communication*, 1–23. doi:10.1080/10584609.2016.1201179

Geiss, S. (2016). *Media effects or method artifacts? Comparing strategies for linking content analysis and survey data*. Paper presented to the Annual conference of the section Methods of Communication Research, DGPuK, Amsterdam.

Gelman, A. (2015). The connection between varying treatment effects and the crisis of unreplicable research: A Bayesian perspective. *Journal of Management*, 41(2), 632–643. doi:10.1177/0149206314525208

Gelman, A., & Loken, E. (2013). *The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time* (Unpublished manuscript). Department of Statistics, Columbia University, New York.

Gil De Zúñiga, H., Jung, N., & Valenzuela, S. (2012). Social media use for news and individuals' social capital, civic engagement and political participation. *Journal of Computer-Mediated Communication*, 17(3), 319–336. doi:10.1111/j.1083-6101.2012.01574.x

Graber, D. (2004). Mediated politics and citizenship in the twenty-first century. *Annual Review of Psychology*, 55, 545–571. doi:10.1146/annurev.psych.55.090902.141550

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, 20(1), 102–116. doi:10.1037/a0038889

Hester, J. B., & Gibson, R. (2003). The economy and second-level agenda setting: A time-series analysis of economic news and public opinion about the economy. *Journalism & Mass Communication Quarterly*, 80(1), 73–90. doi:10.1177/107769900308000106

Hobolt, S., & Tilley, J. (2014). *Responsibility without accountability in the European Union*. Oxford: Oxford University Press.

Hopmann, D. N., Shehata, A., & Strömbäck, J. (2015). Contagious media effects: How media use and exposure to game-framed news influence media trust. *Mass Communication and Society*, 18(6), 776–798. doi:10.1080/15205436.2015.1022190

Hopmann, D. N., Van Aelst, P., & Legnante, G. (2011, May). *Partisan bias in the news*. Paper presented at the Annual Conference of the International Communication Association, Boston.

Hopmann, D. N., Vliegenthart, R., De Vreese, C. H., & Albæk, E. (2010). Effects of election news coverage: How visibility and tone influence party choice. *Political Communication*, 27(4), 389–405. doi:10.1080/10584609.2010.516798

Jebril, N., De Vreese, C. H., Van Dalen, A., & Albæk, E. (2013). The effects of human interest and conflict news frames on the dynamics of political knowledge gains. *Scandinavian Political Studies*, *36*(3), 201–226. doi:10.1111/1467-9477.12003

Johnson, R., & Brady, H. E. (2002). The rolling cross-section design. *Electoral Studies*, *21*(2), 283–295. doi:10.1016/S0261-3794(01)00022-1

Kalogeropoulos, A., Albæk, E., De Vreese, C. H., & Van Dalen, A. (2016). News priming and the changing economy: How economic news influences government evaluations. *International Journal of Public Opinion Research*. doi:10.1093/ijpor/edv048

Kepplinger, H. M., Brosius, H.-B., & Staab, J. F. (1991). Opinion formation in mediated conflicts and crises: A theory of cognitive-affective media effects. *International Journal of Public Opinion Research*, *3*(2), 132–156. doi:10.1093/ijpor/3.2.132

Kepplinger, H. M., Geiss, S., & Siebert, S. (2012). Framing scandals: Cognitive and emotional media effects. *Journal of Communication*, *62*(4), 659–681. doi:10.1111/j.14602466.2012.01653.x

Kiousis, S. (2011). Agenda-setting and attitudes: Exploring the impact of media salience on perceived salience and public attitude strength of US presidential candidates from 1984 to 2004. *Journalism Studies*, *12*(3), 359–374. doi:10.1080/1461670X.2010.501149

Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1944). *The people's choice: How the voter makes up his mind in a presidential election*. New York, NY: Duell, Sloan and Pearce.

Lengauer, G., & Höller, I. (2012). Contest framing and its effects on voter (de)mobilisation: News exposure and its impact on voting turnout in the 2008 Austrian elections. *Javnost - the Public: Journal of the European Institute for Communication and Culture*, *19*(4), 73–91. doi:10.1080/13183222.2012.11009097

Matthes, J. (2012). Exposure to counterattitudinal news coverage and the timing of voting decisions. *Communication Research*, *39*(2), 147–169. doi:10.1177/0093650211402322

Möller, J., & De Vreese, C. (2015). Spiral of political learning: The reciprocal relationship of news media use and political knowledge among adolescents. *Communication Research*, 1–17. doi:10.1177/0093650215605148

Moon, S. J. (2013). Attention, attitude, and behavior second-level agenda-setting effects as a mediator of media use and political participation. *Communication Research*, *40*(5), 698–719. doi:10.1177/0093650211423021

Nagelhout, G. E., Van Den Putte, B., De Vries, H., Crone, M., Fong, G. T., & Willemsen, M. C. (2012). The influence of newspaper coverage and a media campaign on smokers' support for smoke-free bares and restaurants and on second-hand smoke harm awareness. *Tobacco Control*, *21*(1), 24–29. doi:10.1136/tc.2010.040477

Nir, L., & Druckman, J. N. (2008). Campaign mixed-message flows and timing of vote decision. *International Journal of Public Opinion Research*, *20*(3), 326–346. doi:10.1093/ijpor/edn035

Noelle-Neumann, E. (1973). Return to the concept of the powerful mass media. *Studies in Broadcasting*, *9*, 67–112.

Peter, J. (2004). Our long 'return to the concept of powerful mass media': A cross-national comparative investigation of the effects of consonant media coverage. *International Journal of Public Opinion Research*, *16*(2), 144–168. doi:10.1093/ijpor/16.2.144

Price, V., & Zaller, J. (1993). Who gets the news? Alternative measures of news reception and their implications for research. *Public Opinion Quarterly*, *57*(2), 133–164. doi:10.1086/269363

Prior, M. (2009). The immensely inflated news audience: Assessing bias in self-reported news exposure. *Public Opinion Quarterly*, *73*(1), 130–143. doi:10.1093/poq/nfp002

Romer, D., Kenski, K., Winneg, K., Adasiewicz, C., & Jamieson, K. H. (2006). *Capturing campaign dynamics, 2000 and 2004: The national Annenberg election survey*. Philadelphia, PA: University of Pennsylvania Press.

Rössler, P. (1999). The individual agenda-designing process: How interpersonal communication, egocentric networks, and mass media shape the perception of political issues by individuals. *Communication Research*, *26*(6), 666–700. doi:10.1177/00936509902600600

Scharkow, M., & Bachl, M. (2016). How measurement error in content analysis and self-reported media use leads to minimal media effect findings in linkage analysis: A simulation study. *Political Communication*, Advance online publication. doi:10.1080/10584609.2016.1235640

Schmitt-Beck, R., Bytzek, E., Rittinger, H., Roßteutscher, S., & Weßels, B. (2009, May 21–25). *The German longitudinal election study (GLES)*. Paper prepared for delivery at the 2009 Annual Conference of the International Communication Association (ICA), New Opportunities and New Choices for Campaign and Election Studies, Chicago.

Schuck, A. R. T., Boomgaarden, H. G., & De Vreese, C. H. (2013). Cynics all around? The impact of election news on political cynicism in comparative perspective. *Journal of Communication*, *63*(2), 287–311. doi:10.1111/jcom.12023

Schuck, A. R. T., & De Vreese, C. H. (2008). The Dutch no to the EU constitution: Assessing the role of EU skepticism and the campaign. *Journal of Elections, Public Opinion and Parties*, *18*(1), 101–128. doi:10.1080/17457280701858656

Schuck, A. R. T., & De Vreese, C. H. (2009). Reversed mobilization in referendum campaigns: How positive news framing can mobilize the skeptics. *The International Journal of Press/Politics*, *14*(1), 40–66. doi:10.1177/1940161208326926

Schuck, A. R. T., Vliegenthart, R., & De Vreese, C. H. (2016a). Who's afraid of conflict? The mobilizing effect of conflict framing in campaign news. *British Journal of Political Science*, *46*(1), 177–194. doi:10.1017/S0007123413000525

Schuck, A. R. T., Vliegenthart, R., & De Vreese, C. H. (2016b). Matching theory and data: Why combining media content with survey data matters. *British Journal of Political Science*, 46(1), 205–213. doi:10.1017/S0007123415000228

Shehata, A. (2014). Game frames, issue frames, and mobilization: Disentangling the effects of frame exposure and motivated news attention on political cynicism and engagement. *International Journal of Public Opinion Research*, 26(2), 157–177. doi:10.1093/ijpor/edt034

Shoemaker, P., & Reese, S. D. (1990). Exposure to what? Integrating media content and effects studies. *Journalism & Mass Communication Quarterly*, 67(4), 649–652.

Slater, M. D. (2004). Operationalizing and analyzing exposure: The foundation of media effects research. *Journalism and Mass Communication Quarterly*, 81(1), 168–183. doi:10.1177/107769900408100112

Slater, M. D. (2016). Combining content analysis and assessment of exposure through self-report, spatial, or temporal variation in media effects research. *Communication Methods & Measures*, 10(2–3), 173–175. doi:10.1080/19312458.2016.1150969

Statham, P., & Tumber, H. (2013). Relating news analysis and public opinion: Applying a communications method as a 'tool' to aid interpretation of survey results. *Journalism*, 14(6), 737–753. doi:10.1177/1464884913491044

Van Spanje, J., & De Vreese, C. (2014). Europhile media and eurosceptic voting: Effects of news media coverage on eurosceptic voting in the 2009 European parliamentary elections. *Political Communication*, 31(2), 325–354. doi:10.1080/10584609.2013.828137

Vermeulen, I., Beukeboom, C. Batenburg, A., Avramiea, A., Stoyanov, D., van de Velde, B., & Oegema, D. (2016). Blinded by the Light: how a focus on statistical "significance" may cause p-value misreporting and an excess of p-values just below .05 in communication science. *Communication Methods & Measures*, 9(4), 253–279.

Vliegenthart, R., & Boomgaarden, H. G. (2007). Real-world indicators and the coverage of immigration and the integration of minorities in Dutch newspapers. *European Journal of Communication*, 22(3), 293–314. doi:10.1177/0267323107079676

Vliegenthart, R., Schuck, A. R. T., Boomgaarden, H. G., & De Vreese, C. H. (2008). News coverage and support for European integration, 1990–2006. *International Journal of Public Opinion Research*, 20(4), 415–439. doi:10.1093/ijpor/edn044

Wilson, S. E., & Butler, D. M. (2007). A lot more to do: The sensitivity of time-series cross-section analyses to simple alternative specifications. *Political Analysis*, 15(2), 101–123. doi:10.1093/pan/mpl012

Wolling, J. (2002). Methodenkombination in der Medienwirkungsforschung: Der Entscheidungsprozess bei der Verknüpfung von Umfrage- und Inhaltsanalysedaten. *ZUMA Nachrichten*. 50, 54–85. Retrieved from http://nbn-resolving.de/urn:nbn:de:0168-ssoar-207905

Yanovitzky, I., & Cappella, J. N. (2001). Effect of call-in political talk radio shows on their audiences: Evidence from a multi-wave panel analysis. *International Journal of Public Opinion Research*, 13(4), 377–397. doi:10.1093/ijpor/13.4.377