

This is an Accepted Manuscript of a book chapter in a book edited by Claes de Vreese, Frank Esser and David Nicolas Hopmann published by Routledge as *Comparing Political Journalism* in 2016. The book is available online at: <https://www.routledge.com/Comparing-Political-Journalism/de-Vreese-Esser-Hopmann/p/book/9781138655867>

Chapter 2

How we did it. Approach and methods

David Nicolas Hopmann, Frank Esser, and Claes de Vreese with Toril Aalberg, Peter Van Aelst, Rosa Berganza, Nicolas Hubé, Guido Legnante, Jörg Matthes, Stylianos Papathanassopoulos, Carsten Reinemann, Susana Salgado, Tamir Sheafer, James Stanyer, and Jesper Strömbäck

The analyses in this book are based on a dataset covering information on more than 7,500 news items and more than 28,000 sources that appear in the news items. How did we gather these news items? In the process of designing a comparative study, numerous decisions have to be made – many of which are not straightforward (Rössler 2012). These decisions relate to the sampling procedure, the construction of the codebook, inter-coder reliability testing, and the strategy of analysis. The goal of this chapter is to provide a technical overview of how we created the main data source for this book and to present the methods applied. In so doing, we intend to provide enough background information for the subsequent analyses that are needed to evaluate and contextualize the data and the results. This chapter does not aim to provide an in-depth discussion or analysis of the methods and strategies applied. Numerous books and articles with extended methodological discussions on cross-country media content analysis have been previously published (e.g., Hopmann and Skovsgaard 2014; Krippendorff 2004; Neuendorf 2002; Peter and Lauf 2002; Riffe, Lacy, and Fico 2005; Rössler 2012). This chapter, by contrast, describes the application of methods.

The chapter proceeds as follows. In the first part, the period of sampling and its various steps are presented, involving countries, news outlets, and news articles. In the second part, we provide a short overview of the codebook. The third part presents our testing of inter-coder reliability, both across countries and within countries. We then briefly explain how our analyses in the subsequent chapters were conducted and why. The chapter's concluding section reflects on some of the advantages and challenges of the data sources and the methods that we have chosen.

Sampling strategies

The more than 7,500 news items analyzed in this book were gathered in 16 different countries. Heading from north to south, the countries are Norway, Sweden, the United Kingdom, Denmark, Germany, the Netherlands, Belgium, France, Switzerland, Austria, Italy, Spain, Portugal, Greece, Israel, and the United States. In short, the country sample covers all larger Western European countries (barring Finland and Ireland) as well as Israel and the United States.

Behind the choice of countries were several considerations. First, we opted for a design that has a set of comparable countries (established Western democracies). Second, the sample of countries captures variance in several dimensions on the media and political systems level, which is crucial for the analyses to be presented in the book (i.e., we decided to cover countries that vary on a number of relevant independent variables, including their media markets and journalistic professionalism). Third, and related to the second consideration, we wanted to include countries that cover the different models of media and politics identified in previous research (i.e., the liberal model, the polarized pluralist model, and the democratic corporatist model, see Hallin and Mancini 2004). The

presentation of the key independent variables in Chapter 4 will show how these countries vary with respect to their media markets and political systems. By investigating these 16 countries, we are able to study how these factors correlate with news content. Fourth, we wanted to include newspaper- as well as television-centric countries (Norris 2000; Shehata and Strömbäck 2011). Fifth, the United States was included because of its prominent role in previous news research. In addition, it is seen as a prime example of a liberal media system (Hallin and Mancini 2004). Both Israel and the United States are two countries where public service broadcasting (PSB) has only small market shares (comparable in Europe only to Greece), which is another important reason for their inclusion. Finally, there were practical considerations. The scholars behind this study are based in 14 different European countries plus Israel.

Sampling units

With regard to the choice of news outlets to be included in this study, our goal was twofold. On the one hand, we intended to cover the variation in each country's media market as comprehensively as possible. On the other hand, we aimed for sample equivalency across countries; that is, we wanted to sample comparable units (Rössler 2012, pp. 461–462). A list of all included news outlets is shown in Table 2.1.

With respect to newspapers, we therefore decided to sample the two most popular upmarket outlets in each country and, where possible, one politically left of center and one politically right of center. In most European countries, upmarket (sometimes also called broadsheet) newspapers have a long tradition of being affiliated with specific political camps (see, e.g., the discussions in Hallin and Mancini 2004; Seymour-Ure 1974). Where the leading national upmarket newspapers include newspapers with different political leanings, our sampling strategy included the leading left-of-center and right-of-center newspapers. In addition, we sampled one mass-market (i.e., tabloid) newspaper from each country. Previous research has indicated that mass-market newspapers differ in their news reporting from upmarket newspapers (Albæk et al. 2014). By including both upmarket and mass-market newspapers across all 16 countries, we are able to study the implications of editorial missions and organizational goals for news content. Differences in news content as a consequence of editorial missions or organizational goals are often assumed; the question is whether they can be empirically confirmed.

For a majority of citizens, the most important source for political news continues to be television news (Shehata and Strömbäck 2014). For all countries, we included the most widely watched public service broadcasts and commercial news broadcasts. Except for a few outliers, public service broadcasting (PSB) is widely watched across the sampled countries, and in many countries, public service broadcasters are leading players in the media market (Esser, de Vreese et al. 2012). Previous research has indicated that there are systematic differences between the style of news reporting on public service broadcasters and commercial broadcasters (Aalberg and Curran 2012; Cushion 2012; de Vreese et al. 2006). By including both types of broadcasters, we are able (as was the case with the two types of newspapers) to study whether they differ systematically across countries, and how.

Increasingly, citizens consume news through websites. This trend poses a challenge to comparative research, given the blurred boundaries of website production and the vast differences in consumption at the individual and media-system levels. Since the major news sites in nearly all countries are mostly the online affiliates of established newspapers or television broadcasters (Mutz and Young 2011; Shehata and Strömbäck 2014, p. 107), in Table 2.1, we included the websites of the newspapers and broadcasters chosen. This approach of choosing online sources also has the advantage that they are fairly comparable across media systems. These websites were downloaded on the sampling days to make their content available for subsequent content analysis.

More generally, it is important to note that many of the media outlets analyzed in this study are not centrally archived. We therefore had to tape or download newscasts, buy newspapers at the newsstand, copy websites for offline use, and so on.¹

Unit of analysis

After having decided which news outlets to include in the content analyses, the next step was to define the unit of analysis – in our case, the news item. Everyone with experience in media content analysis knows that it is anything but easy to define a news item; in particular, a common challenge is deciding when a televised news story ends and a new one begins. The challenge is only magnified in comparative content analysis (Rössler 2012, p. 465). The details of our approach for each type of medium are explained later.

As a general rule, we included only domestically relevant political news items. So, for instance, a short report on some remote country's election result without any reference to domestic politics was not included. In practical terms, this strategy implied searching for news items that verbally or visually referred to at least one domestic political actor (including political parties or political institutions).

Beginning with television, we followed the approach for defining a news item that has been used in previous international comparative studies (e.g., de Vreese et al. 2006; Schuck, Boomgaarden, and de Vreese 2013). A television news item may have several different components, including an introduction by an anchor, a field report, a so-called two-way with a journalist, an interview with a politician, and more. To begin with, we defined a television news item as having one topic. If the topic changes, by definition, a new news item begins. 'Packages' dealing with the same topic (e.g., a field report followed by a two-way with the correspondent) were also treated as two or more different news items. Short teasers or very short news items were not included in the content analysis.

With newspapers, defining a news item is more straightforward since each item is usually clearly graphically separated from the next. The criterion of reference to a domestic political actor was limited to headlines, subheadlines, first paragraphs, and visuals. We sampled only news items appearing in the international or domestic political sections of newspapers; that is, we did not sample news items appearing in sections with a focus on such things as regional, local, cultural, and sports issues. We excluded news articles that were shorter than five lines of text, unless they were major front-page headlines, as is sometimes the case in mass-market newspapers.

Finally, for websites, we treated news items as consisting of text, text with a visual, or text with a video, generally following the same rules as for newspaper article sampling. We did not include very short news items (i.e., less than two sentences) but did include the text introducing a video (the video itself was not content analyzed in any detail). Side stories embedded within a news item that had their own headlines were coded as separate news items, as was also the case with newspaper items.

In a first step, coders were asked to count the number of news items published on a specific day by a specific outlet and to numerate all eligible news items. In a second step, if a specific news outlet published more than five (for websites, three) news items on a specific day, coders had to choose five (for websites, three) random news items that were then included in the content analysis. A randomizer (offered by the website Random.org) drew five random news items from all eligible news items.

Sampling period

We sampled news items during a constructed fortnight (i.e., 14 days in total), stretching from 15 April to 15 July 2012. The main advantage of constructing two weeks from a total period of three months over choosing two consecutive weeks is that special events – even if occurring in only one country – distort the sample (on constructed weeks in content analyses, see Riffe et al. 2005). Special events include major accidents or natural disasters. A special event could still be captured by our sampling procedure if it occurred on one of the days included in the sample, but our strategy ensured that it does not color the entire sample. In other words, in our content analysis, it would remain what it is – a special event.

There were three exceptions to the sampling period: France, Greece, and the Netherlands. In these cases, elections were held in the period mentioned earlier. But, as explained in Chapter 1, one of the goals of this study was to analyze political news coverage in routine times. The sampling was therefore postponed in these countries and took place in the period from 15 September to 15 December 2012.²

Codebook

Apart from reviewing the current state of research on a number of key concepts in the political communication literature, the second goal of our 2012 *Journalism* special issue was “to contribute to increasing standardization of how key concepts are conceptualized and, perhaps most important, operationalized and investigated empirically” by suggesting “how each of the selected key concepts should be conceptualized and operationalized” (Esser, Strömbäck, and de Vreese. 2012, pp. 140–141). Each review article included in the special issue presented an overview of how a given concept can be operationalized.

As explained in the Introduction, the goal of the present study is to turn these suggestions into practice. The detailed codebook that was used during the data gathering for the present book was based on this previous review. As shown in Table 2.2, the included variables covered six dimensions on the news item level: formal characteristics (e.g., type of media outlet), strategy versus game framing, interpretive journalism, negativity and style, policy substance, and issue of the news item. For each news item, up to five actors could be coded. On the actor level, we coded formal characteristics (e.g., gender) and the favorability of their appearances.

The attentive reader will notice that we modified and updated some of the measures that were originally suggested in our *Journalism* review of concepts. The changes are hardly surprising. Suggesting a measure in a review is one thing; actually implementing it in a large-scale comparative content analysis is another. Details on the measures used and their derivations will be described in each of the chapters presenting our findings. What is most important is that we applied the same definitions of the variables across all countries. Although one can always discuss whether the operationalization of a concept is too broad or too narrow, the major advantage of this approach is that we can compare levels across countries; that is, our results will inform us about the *differences* between countries.

Inter-coder reliability across countries

Ensuring inter-coder reliability in comparative research is a major challenge – even more so in our study, given the many different languages (see, e.g., Peter and Lauf 2002; Rössler 2012). In this study, we have taken several steps to ensure inter-coder reliability. In a first step, we tested the codebook using English-language material among all authors of the book. The goal was obviously to ensure a common understanding of how to apply the codebook across countries. As Rössler (2012) noted, ensuring a common understanding of the constructs to be measured is especially challenging in comparative research. Based on the results of this initial coding, some variable descriptions and definitions were revised and updated.

In a second step, local coders were recruited and trained. Following suggestions from methodological research on comparative content analyses (Peter and Lauf 2002, p. 826), we recruited native speakers as local coders but whose English proficiency was sufficient to read the codebook in English and to code the first set of coder-training material in English. To ensure a common understanding of concepts across countries, the coder training began with one English-language set of testing material that was used in all countries. In the subsequent third step, the local coders performed the coding of the sampled news items (details are available upon request).³

In a fourth and final step, we formally tested the inter-coder reliability based on English-language material after the country-specific coding had been completed. Using five news examples, this test

was performed by coders who had participated in the country-specific content analyses.⁴ The summarized results of this final test are reported in Table 2.2.

As one can see in Table 2.2, we report two versions of Fretwurst's *lotus* (Fretwurst 2015a, 2015b). The *lotus* coefficient can be applied to categorical, ordinal, or metrical scales and can be reported as both unstandardized and standardized measures of inter-coder reliability. This measure of inter-coder reliability has a number of advantages. The unstandardized *lotus* is directly interpretable, representing the percentage agreement of coders with the most used category by all coders. This approach ignores coder agreement on other categories other than the most used category (Fretwurst 2015a, 2015b). The standardized *lotus* measure is a chance-corrected version; that is, the computation of the *lotus* also takes into account the number of categories used by coders. The reasoning is that the more categories, the more difficult it is to reach an agreement. Finally, comparing countries that have provided information on how they contribute to the overall *lotus* coding results (reported in Table 2.2) is a straightforward task.

What do the results reported in Table 2.2 tell us? Looking at the unstandardized *lotus* first, we see that the coding of formal characteristics achieved good inter-coder reliability scores. With respect to the substantive variables, the results for interpretive journalism, 'policy substance and issue,' and 'strategy and issue framing' are also acceptable. The results for 'negativity and style' are somewhat lower. Variables such as 'incapability' or 'negative tonality' are notoriously difficult to code given their evaluative character. On the actor level, we also find good scores for formal characteristics of actors, but favorability of actor appearances was more difficult to code, as one would expect. The chance-corrected *lotus* scores are generally somewhat lower but acceptable. Again, as one would expect, evaluative variables such as 'incapability,' 'negative tonality,' and 'favorability' towards actors have the lowest scores.

In Table 2.3, we also report the details of the international inter-coder reliability test per country. That is, these results tell us the inter-coder reliability of each country. Low overall scores tell us that a country has had a negative impact on the summarized results reported in Table 2.2. Overall, we see no major differences between countries. Countries such as Israel, Italy, and Switzerland score somewhat lower than other countries. The crucial story to tell from Table 2.3, however, is that despite minor differences between countries, average inter-coder reliability scores across countries are acceptable.

We do want to be explicit that – as in most cases – the inter-coder reliability of our study is open for further improvement. We stress that the test was conducted on identical stories in the project language English, but that the actual coding of the 16-country material was done in the coders' native language. This difference is important as project language tests (such as ours in English) typically yield lower reliability scores and may thus underestimate the quality of the actual coding (see Rössler 2012).

Analytical strategy

After collecting data, the next important question arises – how to analyze the collected data. Our goal was to choose an adequate analytical strategy that readers would find easily accessible. Therefore, the standard analytical strategy in this book's analyses was to use media outlets ($N = 160$) as the unit of analysis. The rationale behind this strategy was our interest in explaining the *outlets'* news coverage across our sample of countries. As will be explained in Chapter 3, the main independent variables that we were interested in are at the media outlet level (i.e., type of medium) and at the country level (i.e., information on the political and media systems of a given country). In those cases where it was necessary, the unit of analysis was actors, of which we coded more than 28,000. As mentioned previously, up to five actors could be coded per news story. Using actors as a unit of analysis was partly relevant for the chapters on personalization and political balance. Finally, overviews across countries are based on simple country-level means across all news stories (or actors), unless another approach is specified in the presentation of the results.

In short, the six concept chapters and the cross-country analysis in Chapter 7 are mainly based on regression analyses on the news outlet level as well as country means. In the case of regression analysis, we computed robust standard errors (Rogers 1994). By doing so, we aimed to take into account the grouping of cases into country clusters. We did not opt for a multilevel analysis, given the frequent low *Ns*, for example, on a country level, where we included ten news outlets per country. When graphs report confidence intervals, they are based on the standard formula for confidence intervals.⁵

Discussion

The starting point of this book was the goal to describe and explain differences in political journalism across established Western democracies. Analyses of news content can take many different forms and approaches, depending on the key questions at stake. Conducting media content analyses in 16 countries simultaneously is anything but easy. Given these comparative ambitions, we had a trade-off by focusing on certain aspects while leaving out more in-depth aspects of the coverage. This lacuna is often the price of working systematically and comparatively. We accomplished our content analysis task through local, native-speaking coders. While the inter-coder reliability across countries could surely have been higher, we are confident that the coding across countries was of a quality that ensured the comparability of the results across countries. Moreover, given the nature of our endeavor – testing multi-item measures for 6 key concepts across 16 countries for the first time – we were willing to relax more stringent requirements in the hope that the work will spark more research that can also improve, in a more focused fashion, certain subfacets. We fully acknowledge that our analysis follows the logic of systematic, comparative content analysis. Using other approaches would have enabled us to go more into depth with particular cases or aspects, and we hope that subsequent research, from a variety of perspectives and approaches, will engage with this work as they further develop their research agendas.

In our previous *Journalism* special issue (Esser et al. 2012), we described the state of the art for a number of concepts that are important in the fields of political communication research and journalism studies. Now the time had come to implement the measures. For this undertaking, we needed additional data from other sources. The next chapter presents the explanatory logic of our analyses and the additional data sources that we drew on (besides the media content analysis described in this chapter). What we found and what we did not find in the 16 media systems included in our study, and how the differences between media systems can be explained, will be presented in the Chapters 4–9.

Authors' acknowledgements

We are indebted to Benjamin Fretwurst for providing us with the necessary tools to compute the *lotus* inter-coder reliability coefficients reported in this chapter. For more details, see www.iakom.ch/lotus.html (last accessed 10 June 2015). Our thanks also go to Sven Engesser for his helpful input.

Notes

- 1 In a few cases, data were missing or not available: for Denmark, pictures were not included on some websites as the program that was used to download the website was unable to perform certain operations (for example, it could not handle Danish characters in links and references). For Italy, on May 20, it was not possible to download some news from websites; therefore, we performed a new sampling on relevant news of the day. For the U.S., April 26 was replaced with May 4 due to missing data.
- 2 In the Netherlands, parliamentary elections took place September 12, 2012, implying that the Dutch data covered an immediate postelection period. As in some other European countries, Dutch elections are typically announced only with very short notice, making it difficult to plan accordingly.
- 3 The content analysis of the U.S. data was conducted by native English speakers residing in the United Kingdom.
- 4 No Norwegian and Spanish coder was available at the time of the test; these two countries are therefore missing from the results reported in Tables 2.2 and 2.3. In most other cases, one coder per country was asked to complete the test coding.
- 5 The formula used to compute confidence intervals of proportions was $1.96 \times \sqrt{p \times (1 - p) / N}$; in the case of country means, it was $1.96 \times se$. Given the large N s in most cases, using a t -value taking into account the available degrees of freedom instead of the z -value 1.96 would unlikely yield substantially different conclusions.

Table 1: News outlets included in the content analyses, by country and type of outlet.

Country	Newspaper			Television, news bulletin		Upmarket		Websites		
	Left-of- centre	Right-of- centre	Mass- market	Public- service	Commerci- al	Left-of- centre	Right-of- centre	Mass- market	Public- service	Commerci- al
Austria	Der Standard	Die Presse	Kronenzeitung	ORF1, ZiB	ATV, Aktuell	derstandar d.at	diepress.co m	krone.at	news.orf.at	kuurier.at, <i>see note A</i>
Belgium	De Morgen	De Standaard	Het Laatste Nieuws	VRT, Het Journaal	VTM, Het Nieuws	demorgen. be	standaard. be	hln.be	deredactie. be	nieuws.vt m.be
Denmark	Politiken	Jyllands-Posten	Ekstra Bladet	DR, TV-Avisen	TV2, Nyhederne	politiken.d k	jp.dk	ekstrablade t.dk	dr.dk/nyhe derne	nyhederne. tv2.dk
France	Le Monde	Le Figaro	Le Parisien	France 2, Journal de 20h	TF1, Journal de 20h	lemonde.fr	lefigaro.fr	leparisien.f r	info.france 2.fr	lei.tfl.fr
Germany	Süddeutsche Zeitung	Frankfurter Allgemeine Zeitung	Bild	ARD, Tagesschau	RTL, Aktuell	sueddeutsc he.de	faz.net	bild.de	tagesschau. de	rfl.de/rflakt uell
Greece	Tanea	Kathimerini	Espresso	NET, News	Mega, News	tanea.gr	kathimerin i.gr	espressone ws.gr	ert.gr	megatv.vo m
Israel	Haaretz	Ydiot Aharonot, <i>see note B</i>	Israel Hayom	Channel 1, Evening News	Channel 2, Evening News	haaretz.co. il	Ydiot Aharonot, <i>see note C</i>	israelhayo m.com	ynet.co.il, <i>see note C</i>	mako.co.il
Italy	La Repubblica	Il Giornale	Altri Mondi, <i>see note D</i>	RAI1, TG1	Canale5, TG5	repubblica. it	ilgiornale.i t	quotidiano. net	tg1.rai.it	tgcom24.m ediaset.it
Netherlands, the	De Volkskrant	NRC Handelsblad	De Telegraaf	NOS, Journaal	RTL, Nieuws	volkskrant. nl	nrc.nl	telegraaf.nl	nos.nl	rtnieuws.n l
Norway	Dagsavisen	Aftenposten	VG	NRK, Dagsrevyen	TV2, Nyhetene	aftenposten .no	dagsavisen .no	vg.no	nrk.no	tv2.no/nyh eter
Portugal <i>see note H</i>	Jornal de Notícias	Público	Correio da Manhã	RTP1, Telejornal	TVI, Jornal das 8	publico.pt	jo.pt	cmjornal.x lpt	rtp.pt	tvi.iol.pt
Spain	El País	El Mundo	20Minutos	TVE, Telediario de la Noche	Tele5, Informativos	elpais.com	elmundo.e s	20minutos. es	rtve.es/tele diario	telecinco.e s/informati vos
Sweden	Dagens Nyheter, <i>see note E</i>	Svenska Dagbladet	Aftonbladet	SVT, Rapport	TV2, Nyheterna	dn.se (centre)	svd.se	aftonbladet .se	svt.se	tv2.se/nyh eter
Switzerland	Tages-Anzeiger	Neue Zürcher Zeitung	Blick	SF, Tagesschau	TeleZüri, ZüriNews	nzz.ch	tagesanzei ger.ch	blick.ch	tagesschau. .sf.tv	radio24.ch, <i>see note F</i>
United Kingdom	Guardian	Daily Telegraph	The Sun	BBC, News at Ten	ITV, News at Ten	guardian.c o.uk	telegraph.c o.uk	thesun.co.u k	bbc.co.uk/ news	itv.com/ne ws
United States	New York Times	Los Angeles Times	USA Today	NBC, Nightly News	PBS, News Hour	nytimes.co m	latimes.co m	usatoday.c om	msnbc.msn .com	npr.org, <i>see note G</i>

Notes:

A: Kurier was taken because the private broadcaster does not offer an online news site.

B: Yediot Aharonot is considered a popular newspaper and its ideology is rather centre-based.

C: Radio was taken since the website of the Israeli public television provides only video and not text. Reshet Bet's news is part of the Israeli Public Broadcast Authority, as is Israeli Public Television.

D: Given the absence of a mass vs. upmarket distinction in Italy, we took the Altri Mondì ("Other Worlds") section of the sport newspaper La Gazzetta Dello Sport, which reports to its mass audience the news of the day and the main political news. This section hasn't a news website, thus we took the website of quotidiano.net, the national network of locally relevant mass newspapers.

E: In Sweden there is no major left-of-center upmarket newspaper on the national level. Dagens Nyheter is considered as liberal and slightly right-of-centre, but compared to Svenska Dagbladet it is more to the centre.

F: Website of Radio 24 was taken because TeleZüri does not have a news website.

G: NPR was taken because it is a fairer comparison to the much used MSNBC website.

H: For Portugal, we chose the largest news outlets, but the categorization into "left-of-centre" or "right-of-centre" does not apply here.

Table 2: International inter-coder reliability, summarized results.

Variable name (coding categories)	Variable group	Fretwurst's lotus	Fretwurst's lotus (std.)
<i>News item level</i>			
Type of medium (1-4)	Formal characteristics	.94	.92
Type of news item (1-6)	Formal characteristics	.94	.93
Commentary by journalist (0/1)	Interpretative journalism	.80	.60
Consequence speculation by journalist (0/1)	Interpretative journalism	.83	.66
Reasons provided by journalist (0/1)	Interpretative journalism	.73	.46
Conflict (0-3)	Negativity and style	.76	.68
Emotional reporting (1-3)	Negativity and style	.71	.57
Incapability (0-3)	Negativity and style	.66	.54
Negative tonality (0-3)	Negativity and style	.64	.52
Affected groups (0/1)	Policy substance and issue	.83	.66
Decision-making authorities (0/1)	Policy substance and issue	.93	.86
Issue (1-18)	Policy substance and issue	.70	.68
Policy substance (0/1)	Policy substance and issue	.83	.66
Societal actors (0/1)	Policy substance and issue	.84	.69
Media's role (0/1)	Strategy and issue framing	.94	.89
Performance (0/1)	Strategy and issue framing	.71	.43
Public Opinion (0/1)	Strategy and issue framing	.79	.57
Strategic macro-frame (1-2)	Strategy and issue framing	.87	.74
Strategy/Tactics (0/1)	Strategy and issue framing	.83	.66
War/sports language (0/1)	Strategy and issue framing	.70	.40
Wining/losing (0/1)	Strategy and issue framing	.83	.66

Actor level

Gender (0-2)	Formal characteristics	.97	.96
Function (detailed actor list)	Formal characteristics	.94	.94
Favourability (0-3)	Favourability	.69	.58

Note: Based on codings from 14 countries.

Table 3: International inter-coder reliability, per country.

Country	Fretwurst's lotus	Fretwurst's lotus (std.)
Austria	.88	.82
Belgium	.83	.72
Denmark	.78	.60
France	.79	.64
Germany	.86	.76
Greece	.87	.79
Israel	.77	.60
Italy	.73	.52
Netherlands, the	.80	.67
Portugal	.80	.67
Sweden	.83	.71
Switzerland	.72	.52
United Kingdom	.78	.63
United States	.89	.83
Total	.81	.68

Note: Based on the variables mentioned in Table 2.