Creating the EO2griev_hand dataset: Documentation of the human annotation and coding manual

Lea Haiges and Christina Isabel Zuber

1. Development of a novel coding frame and the training of coders p. 1

2. Coding Manual as handed to coders p. 16

1. Development of a Novel Coding Frame and the Training of Coders

Research context

The project studies the conditions under which socio-economic, political, and cultural inequalities between ethnic groups become politically mobilized, focusing on the role **elite discourse** plays in this regard. We start from the well-established argument that collective perceptions of relative deprivation—group grievances—are the key motivating force behind ethnic conflict, We then ask whether and, if so, how ethnic elites can activate such perceptions of being unjustly deprived among group members and thereby contribute to causing a collective response in the form of protest or civil violence. We also aim to find out whether elites can create such perceptions even in the absence of objective inequalities, or whether their rhetoric has to be grounded in objective economic, political, cultural or other inequalities in order to motivate people to take action.

The core concept of relative deprivation can be defined "as actors' perception of discrepancy between their value expectations and their environment's apparent value capabilities" (Gurr 1968: 253). The relevant deprivations driving inter-group conflict are not personal but tied to group identities: individuals perceive the deprivation of the (ethnic) group with which they identify relative to other (ethnic) groups in society, which causes a collective sense of frustration and this frustration can eventually lead to collective aggression.

This document introduces a novel methodology that allows us to systematically identify grievance-activation strategies in the discourse of ethnic elites using qualitative content analysis. We first outline the textual material we use to analyse elite discourse. We then proceed to explain the conceptual logic behind our coding scheme. Finally, we present results from a validation exercise that shows a) that our coding scheme can be used to detect grievance frames in documents issued by ethnic organisations world-wide and b) that pairs of coders achieve very satisfactory levels of inter-coder reliability when coding the same document based on our coding manual and coder training sessions.

The annotated text corpus will be provided to a team of computer linguists (work package 4 of the project) who will take the human identification of grievance frames as a starting point for a systematic investigation into the linguistic cues of grievance frames.

Text Corpus

To capture the messages ethnic elites communicate to group members, we collected:

(1) programmatic documents issued by ethnic organizations, such as:
    a. documents outlining general principles
    b. (electoral) campaign material
    c. declarations on specific issues relevant to group members in the given context
(2) speeches given by ethnic organizations' leaders, such as:
    a. speeches outlining general principles
    b. speeches given during an election campaign
    c. speeches on specific issues relevant to group members in the given context
(3) social media posts on Facebook

To ensure our scheme is broadly applicable, we collected these materials for a variety of different contexts. We first selected 12 countries in six world regions, and for each country (at least) two ethnic organizations. Asia (India, Pakistan), Oceania (Fiji, New Zealand), Europe (Bosnia, Great Britain), Africa (Kenya, Namibia), North & Central America (Canada, Trinidad & Tobago), South America (Bolivia, Peru). The two main selection criteria for choosing these countries were language and geographic location. A focus on former British and former Spanish colonies allowed us to cover a variety of different contexts while dealing with just two languages which is as relevant for human annotation as it is for the potential future linguistic analysis. Furthermore, we aimed to increase diversity by choosing contexts with varying salience of ethnicity, varying occurrence of ethnic violence in the past, and by choosing contexts where ethnic organizations other than ethnic parties claim to represent ethnic groups (for example indigenous non-governmental organizations, as is the case in Peru or Canada).

Defining and Detecting Grievance Frames

We take grievance frames as the observable communicative correlate of a grievance-activation strategy pursued by ethnic elites.

A grievance frame in communication is defined as a set of statements that:

1. portray a given situation in issue area IA as **unjust deprivation of a group** G_1,
2. **blame** another group G_2 for the situation,
3. call on members of G_1 to **act** in order to change the situation.

Conceptually, this definition applies Entman's (1993) understanding of frames to the concrete topic of group grievances. Ethnic elites *define a problem* by portraying a situation as a mismatch between a group's expectations (how things should be) and reality (how things are). They further *diagnose causes* as they attribute blame for the situation to other groups. In portraying the situation as one of *unjust* deprivation, they make *moral judgments*. Finally, in calling on group members to take action, they *suggest remedies*.

A complete grievance would consist of all three elements 1) deploring unjust group deprivation, 2) attributing blame, and 3) calling for action.

Among these, we see the first (deploring unjust group deprivation) as the defining characteristic of a grievance frame, in line with the literature on grievances and conflict.

Portraying a situation as one of unjust deprivation of an explicitly mentioned group is thus *constitutive* of a grievance frame, while blame attribution and calls for action are supplementary features. Unjust group deprivation is thus a necessary condition for coding a section of text as a grievance frame.

Sampling, context and coding units

Following Krippendorf's (2004: 98-109) terminology, the *sampling units* in our project are the documents, respectively the social media posts. As explained above, the documents can be programmatic statements, election manifestos or speeches.

In our project, the documents also serve as *the context units*. This means that coders can draw on each given document to interpret the meaning of each statement in that same document, but should not draw on another document or any additional information for interpreting the meaning of a statement in the document they are currently working with. For example, in a speech given by a Maori leader, he begins by stating that New Zealand's society consists of two classes of citizens, a privileged white class, and an underprivileged brown class. With that initial setting in mind, a reader understands later comparisons between Maori and white New Zealanders differently, than if she had not received this initial information of an unjust situation. [1]

These sampling units have to be unitized into individual *coding units*. Each document has to be broken down into smaller segments of text that are then categorized. These are the chunks of text to which you apply codes in QDA Miner. In our project, we ask coders to search the text for statements that can be interpreted as grievance frames and code these.

That is, we unitize according to a *thematic distinction* (Krippendorf 2004: 107-108). This is a challenging method, because there are no strict a priori rules for what a grievance will look like and where it will appear in the text. But this is precisely one of the goals of our project: figuring out whether there is a common logic to language used to frame a given situation as a group grievance. Since we do not yet know how this logic could be spelled out in general linguistic rules, we need human coders to interpret the material for us and point us towards which part of the text they understand to be portraying a situation as a group grievance, according to our definition. In sum, the text for coders is to both unitize and annotate, as they go though each document.

We allow double coding of our thematic coding units, since the goal is collect information on all grievance frames applied in a given document, not measure the relative salience of some categories over others. If for example one coding unit states that members of G_1 were disadvantaged by fraudulent elections *and* suffer from a bad health infrastructure you should assign both deprivation issue sub-code "political-legal" and deprivation issue sub-code "socio-economic".

Collective Deprivation, Blame Attribution and Calls for Action

Based on two rounds of inductive coding by the project leaders, a coding scheme was developed and subsequently validated on the selection of documents introduced above.

Figure 1 displays the overall coding scheme graphically, while Figure 2 displays the codes for **Deprivation, Blame** and **Action** in more detail.

Following the definition of grievances, we expect ethnic elites to argue that a given situation is unjust for a certain *group G*, and that relative deprivation of the group occurs in a certain *issue area IA*. This is in line with the literature in the field of conflict studies, which focuses on *collective* grievances and also distinguishes different *types* of grievances, for example economic

1 We make one exception to this rule, allowing coders to research if a mentioned party is/was part of the government in order to assign the blame code "government" accordingly (this code is introduced below).

or political (Gurr, 1968; Langer & Stewart, 2014). We therefore ask coders to assign categories for the type of aggrieved group and the issue area concerned.

Our project is primarily driven by interest in the activation of grievance perceptions tied to *ethnic* identity groups. However, following constructivist understandings of social identity, we choose to remain open and see empirically how important ethnic identities really are in the activation of grievance perception in elite language. We therefore list a range of common social identity dimensions in our coding scheme, that aside from ethnicity covers also class, gender, geography and age, as well as a populist cleavage pitting the pure people against the corrupt elite.

Rather than pre-determining the primacy of grievance activation around ethnic identities by design, this approach allows us to assess empirically how important an ethnic framing of grievances really is, compared to other possible group identities to which elites can connect their claims about relative deprivation.

Regarding issue areas, we started deductively with political, socio-economic and cultural, since these are prominent in the literature (Langer & Stewart, 2014). We then added the issue area of security inductively, based on a first round of explorative coding.

We replicated the group identity categories for blame attribution, except for the additional code "government" which was added to the blamed group categories.

Finally, we added categories for types of actions elites could call on group members to engage in, ranging from support (e.g electoral support, voting), via protest to rebellion.

Training and inter-coder reliability test

We trained two coders with an initial workshop and three training rounds. The training material was in English and Spanish, both coders are German native speakers fluent in English and Spanish. During the training phase, we made final edits to the coding manual to clarify any remaining doubts in particular with regard to unitizing and the application of categories. There was no need to add or remove substantive categories, but we had to clarify our rules for unitizing and for how to apply certain categories, in particular the populism categories. After three training rounds, each coder received six documents from six different countries (three manifestos and three speeches) from different contexts (two in Spanish and four in English) to code without any exchange with us supervisors or the other coder. We then implemented inter-coder reliability tests on these documents with the following results:

Document Name	**Country**	**Observed Agreement**	**Krippendorff's Alpha[2]**
BO_MAS_IPSP_S1_2019	Bolivia	100%	
FJ_SODELPA_M1_2018	Fiji	95,8%	0,735
IN_BJP_S1_2020	India	88,6%	0,648
KE_ODM_M1_2017	Kenya	98%	0,806
NA_SWAPO_S1_2019	Namibia	91,7%	0,830

2 All Krippendorff's Alpha values were calculated based on min. 60% overlap between segments marked by the two coders. In the coding of just the social media posts, segmentation did not play a role because one post was to be coded as a whole, considerably easing the coding task.

PE_RUNA_M1_2019	Peru	92,8%	0,501
Total (All 6 Documents)		*98%*	*0,742*

Given that the results were very satisfying in light of the very complex task of thematic unitizing and the rather complex scheme, we decided on the basis of these results that coders were ready to code further material without double coding by the second coder. Note that this task is a lot more involved than the simple classification of whether a given social media post contains or does not contain a grievance we later use to train a classifier.

In order to secure a sufficient inter-coder reliability in the further coding process we regularly checked the inter-coding reliability by providing both coders with respectively two similar document sin a bigger package of documents without them knowing them.

Below you find the results for the two documents in Package #1:

Document Name	**Country**	**Observed Agreement**	**Krippendorff's Alpha**
BO_MDS_S1_2019	Bolivia	100%	
FJ_NFP_M1_2018	Fiji	96,7%	0,832
Total (Both Documents)		*96,8%*	*0,835*

Below you find the results for the two documents in Package #2:

Document Name	**Country**	**Observed Agreement**	**Krippendorff's Alpha**
ES_Jxcat_S2_2020	Spain	71,4%	0,085[3]
ES_VOX_M1_2019	Spain	100%	
Total (Both Documents)		*94,1%*	*0,772*

3 While the observed agreement for this document is acceptable, Krippendorff's Alpha is low. The project leaders evaluated the document and discussed the divergent coding. One mismatch was that Coder#1 identified a sentence as a grievance, which Coder#2 did annotate as Blank. The project leaders evaluated that the sentence is a "can, but does not have to be"-grievance. The second disagreement concerns a paragraph both coders identified as a G2/socio-economic grievance (though coded with different lengths), but the coders identified different additional deprived issues and groups. Both annotations were reasonable. Therefore, value is acceptable because the qualitative differences were not as extensive and reasonable.

Below you find the results for the document in Package #3:

Document Name	Country	Observed Agreement	Krippendorff's Alpha[4]
EC_CONAIE_M3_2020	Ecuador	91,7 %	0,83

In the course of the project, we had to train two new student assistants. Below you find the results for the inter-coder reliability of the new student assistants with programmatic documents of ethnic organizations. During the training process we also coded and compared the annotation of social media posts. Note that due to language abilities the Spanish documents were only coded by one of the new assistants and compared to codings of the previous assistants.

Document Name	Country	Observed Agreement	Krippendorff's Alpha
BO_MAS_IPSP_S1_2019	Bolivia	100%	
FJ_SODELPA_M1_2018	Fiji	96,2%	0,786
IN_BJP_S1_2020	India	93,3%	0,817
KE_ODM_M1_2017	Kenya	98,0%	0,912
NZ_MP_M8_2020	New Zealand	78,8%	0,355[5]
PE_RUNA_M1_2019	Peru	92,6%	0,545
Total English (4 Documents)		*97,8%*	*0,839*
Total Spanish (2 Documents)		*92,8%*	*0,556*

We subsequently let the same coders code the social media messages which turned out to be an overall less complex task because coders did not have to agree on the segmentation, only on the substantive codes.

4 All Krippendorff's Alpha values were calculated based on min. 60% overlap between segments marked by the two coders.

[5] While the observed agreement for this document is acceptable, Krippendorff's Alpha is comparatively low. The project leaders evaluated the document manually. With the exception of one grievance, the coders identified the same paragraphs and annotated them with equal grievance codes. However, the length of the paragraphs differed, which led to a low Krippendorff's Alpha. Therefore, we still went forward since the disagreement concerned only the segmentation, not the substantive classification task.

Bibliography

Entman R.M. (1993). Framing: toward clarification of a fractured paradigm. *Journal of Communication, 43*(4), 51–58.

Gurr, T. (1968). Psychological Factors in Civil Violence. *World Politics*, *20*(2), 245–278. https://doi.org/10.2307/2009798

Krippendorf, K. (2004). Content analysis: An Introduction to Its Methodology (2nd ed.). Thousand Oaks, CA: Sage

Langer, A., & Stewart, F. (2014). Horizontal Inequalities and Violent Conflict: Conceptual and Empirical Linkages. In E. Newman & K. DeRouen (Eds.), *Routledge Handbook of Civil Wars* (pp. 104–118). London [i.a.]: Routledge.

2. Coding manual as handed to coders

Note: This coding manual was handed to student coders alongside training documents and exercises to prepare them for the human annotation of grievance frames used to generate the EO2griev data used in "Reality or Rhetoric? Between-Group Inequality and the Use of Grievance Frames on Social Media". The analysis in that article draws on social media posts only. Since we started the project by annotating also longer programmatic documents of ethnic organisations, the coding manual also includes instructions for coding manifestos or speeches. P. 16 has specific instructions for coding social media posts, the coding categories are identical for all material.

Research context and goals of the coding

The project studies the conditions under which socio-economic, political, and cultural inequalities between ethnic groups become politically mobilized and cause inter-group conflict.

We focus on the role of **elite discourse**. We start from the well-established argument that collective perceptions of relative deprivation—**group grievances**—are the key motivating force behind ethnic conflict. We then ask whether and, if so, how ethnic elites can activate such perceptions among group members and whether elites can thereby stimulate a collective response of the group, in the form of protest or even civil violence.

How do we know whether elites try to activate perceptions of deprivation among group members? We look at how they talk. More concretely, we look at whether they use **grievance frames** in their talk.

> Grievance frame = a statement, or a set of statements, that portray a given situation in an issue area 'IA' as unjust deprivation of a group 'G'

Your job as a coder is to identify grievance frames in political communication.

The following statement illustrates what this will be about:

> "Repeatedly, statistical data informs us of the less than favourable conditions that show great economic disparity between Aboriginal peoples and their Canadian counterparts."

The author is describing the given situation and evaluating it normatively as a problem. It is not a problem for an individual, but a problem of the **group of aboriginal people** so it fits our definition. Aboriginal people are worse off than Canadian people, they are in a situation of relative deprivation. The **issue area** in question is the **socio-economic** one.

This coding manual gives detailed instructions how you can systematically identify and then code such grievance frames.

First Goal: Coding longer documents (manifestos and speeches) that require segmentation:

Getting started

To get started with coding, you need three things:

1. this coding manual, that contains our complete "coding scheme" – the full set of categories you can apply to texts;
2. the software QDA Miner; and
3. the QDA Miner project that contains the documents to be coded.
4. access to a shared spreadsheet to enter some overall information on each document you are coding

This project applies a fairly complex coding scheme. That means that you have to familiarize yourself well with the entire scheme before starting any coding in QDA Miner. Reliable coding requires you to keep all available categories in mind at all times.

At the coding workshop, we will do some exercises in order to help you get to know the coding scheme, but **make sure you read through the entire coding manual several times in advance of the workshop**.

What to code: deciding on segments of text to be discarded for technical reasons, segments of text that are irrelevant, and segments of text that contain grievance frames

In going through each document, you decide to classify the document into three types of segments:

- Segments to be **discarded** for technical reasons (→ apply no code at all)
- Segments you considered, but that do not contain grievances in your judgment (→ apply the "**blank**" code).
- Segments you see as containing **grievances** (→ apply grievance codes introduced below)

We go though each of these in turn in what follows.

Segments to be discarded:

There are some segments of documents you can immediately discard:

Please never code:

- ◊ table of contents
- ◊ headings (this includes any in-text headings and sub-headings that should also not be coded)
- ◊ opening or closing quotes (e.g. quotes at the beginning of a party manifesto, or a poem at the end of a document, etc.)
- ◊ bibliographical information at the end of a document

These parts of the document should not be coded at all. That is, you can simply leave them out of your coding decision about grievances.

Note: Even isolated words (e.g. "Electricity") in a separate line count as a heading.

Irrelevant text ("blank" code)

For the rest of the text that is not to be discarded, you need to decide whether it contains grievance frames or not.

Any part of the text that does *not* contain grievances is to be coded with the "blank" code. This is important for inter-coder reliability calculation, because agreeing on which parts of the text do not contain grievances is part of the challenge the project is mastering.

How to mark "blank" in QDA Miner: Start the "blank" code at the beginning of a document (unless the very first sentence below the heading contains a grievance), or right at the end of a grievance you identified. Pull the "blank" code through until you hit either the next grievance, or a heading (remember that headings are not to be coded) or the end of the text in the document.

Segments that contain grievance frames

To take the decision between grievance segments and "blank" segments, remember:

> Grievance frame = a statement, or a set of statements, that portray a given situation in an issue area IA as unjust deprivation of a group G

You need to search for statements that fit this definition. Note that there may be large chunks of text that you code as "blank" because they do not contain any grievances at all. There may also be entire documents you code as "blank".

Let us now focus on the grievance coding.

The right degree of interpretation

This project asks you to use **qualitative content analysis** to identify "grievance frames" in documents issued by ethnic organizations.

Qualitative content analysis is applied when we want to gain systematic, comparable insights from various documents, but getting at the information we are interested in requires some *interpretation* (so the task cannot (yet) be done by a computer).

While your interpretation and judgment as a coder is necessary to identify a statement as a grievance frame, we want to keep it limited: the golden rule is that another coder who has the same coding manual and training should be able to interpret the same statement in the same way. (We will also test for this; the test is called an "inter-coder reliability test").

How much contextual info can you use to interpret the meaning of a statement?

You can draw on each given document to interpret the meaning of each statement in that same document. You should not draw on another document in your project or on any other additional information for interpreting the meaning of a statement in the document you are currently working with.[6]

> For example, in a speech given by a Maori leader, he begins by stating that New Zealand's society consists of two classes of citizens, a privileged white class, and an underprivileged brown class.
>
> With that initial setting in mind, a reader interprets later comparisons between Maori and white New Zealanders differently, than if she had not received this initial information of an *unjust* overall situation between the groups at the beginning of the same document.

6 We make one exception to this rule: you are allowed to google if a mentioned party is/was part of the government in order to assign the blame code "government" accordingly (this code will be introduced below).

Subtle normative evaluations:

Note that the normative evaluation of a situation as unjust can sometimes be very subtle. Speakers often present statistical information showing that a group is worse off than another group in society. This may be presented in neutral, descriptive language. Then we just have a comparison between groups. But often speakers will present such information using linguistic cues that convey that this situation is not as it should be. Then we have both, the comparison between the own group and other groups / an ideal standard *and* the evaluation of this situation as unjust/undesirable. Consider the following examples to see the subtle differences:

> "The percentage of working-age Aboriginal population having university education is 10%, among the non-Aboriginal population it is 26%."
>
> > [the speaker does state a discrepancy but does not add any connotation of whether this discrepancy is wrong in any way. You may know that it is wrong, based on your own moral judgment, but that is not the question, the question is what the speaker conveys to the receiver here. We do not code this statement as a grievance.]
>
> In contrast:
>
> "The percentage of working-age Aboriginal population having university education has increased from 8% to 10%, but it still lags behind the percentage of non-Aboriginal population is 26.%"
>
> > [by stating that the Aboriginal share still lags behind the non-Aboriginal share, the speaker implies that the shares should be equal and that the current situation is not as it should be. This statement should be coded as grievance (the concrete codes would be: deprived group: ethnicity/G1, deprivation issue: socio-economic)].

The "group plus issue" rule:

Any part of the text that does not connect a group to an issue area in which the group is unjustly deprived can be ignored for the grievance coding since by definition, grievances connect deprived groups to a deprivation issue.

If you find a statement about an unfavourable situation, but this situation is not deplored as disadvantaging for a group, this cannot be a grievance according to our definition. Conversely, If somebody only mentions a group, but not issue on which this group is deprived, it cannot be a grievance frame according to our chosen definition.

→ It needs to be clear from the statement *which societal group is being aggrieved about what*.

> For example:
> "Unemployment in India is too high and there is unfair discrimination on the labour market."
>
> > → this is *not* a grievance frame. Even though the statement is making a moral judgment about a given situation in an issue area that it marks as unjust and diverging from an ideal standard (too high and too much indicate that things should ideally be different), there is no aggrieved group.
>
> In contrast:

"Dalits are particularly hit by unemployment and suffer from unfair discrimination on the labour market."

→ this is a grievance frame – the situation is portrayed as unjust deprivation of a group, namely the Dalit.
(The correct coding would be *deprived group/ethnicity/Dalit* and *deprivation issue/socioeconomic*).

The "grievance frames are at least one sentence long" rule:

As you mark the parts of the text you think relate to grievance frames in QDA Miner, there is one simple syntactic rule for how to mark:

A grievance frame is **at least one sentence long**, but can be longer. Do *not* code partial clauses of sentences as grievance frames. The speaker has to invest at least a full sentence in evoking perceptions of grievance for us to code the segment as a grievance frame.

Such partial grievance clauses are quite common when ethnic organisations make concrete policy proposals, which is most common in the documents in our selection that are election manifestos. The main meaning of these sentence is to propose a policy solution, the grievance is only mentioned to justify the proposed solution. They should therefore not be coded as grievance frames. We want to identify for coherent, complete sentences, or a set of sentences that portray a situation as a collective grievance.

For example:

"We promise to give scholarships to the rural youth, <u>because they are socially disadvantaged</u>".

→ we do not code "because they are socially disadvantaged" as a grievance because it only appears in a clause that justifies a policy proposal and the policy proposal is the main content of the sentence. The grievance does not fill a complete sentence.

In contrast:

"The education of the youth in the countryside is lagging behind the education of the urban youth, every day the future of our rural children is jeopardized by the biased provision of quality education."

→ we identify this as a grievance and code it

(in this case, the correct categories would be *deprived group/age/young* and *deprivation issue/socioeconomic* and *deprived group/geography/rural* and *deprivation issue/socioeconomic*).

Note that for this project, a sentence ends with a proper colon (".") Do not stop your grievance coding at a semi-colon (";").

The issue of time:

We care about elites portraying the **current situation** as a problem, because we are interested in the effect this can have on mobilization.

Therefore, do not code past grievances, unless they are still ongoing / the past injustice is presented as having *ongoing consequences* for the group until the present day. Do also not code hypothetical/future grievances describing that a certain group will be deprived in the future if X or Y happens.

Blaming other groups and calling for action:

We are also interested in

- whether elites **blame** another group for the relative deprivation of their group, and
- whether they **call** on group members **to act** to remedy the situation.

Your complete task is therefore: to search for and code all grievance frames; to check whether elites also blame someone for the grievance; to check whether elites call for action to remedy the injustice.

Blame attribution and calls for action are supplementary, sometimes they occur together with a grievance statement, other times not. (In fact we find that calls for action are extremely rare in the types of documents we code).

Blame and call for action have to be related to and coded *together* with a grievance

IMPORTANT: Do not code blame attribution and calls for action that are not connected to a grievance! We only care about blame and calls for action *in relation to a grievance.* Not per se.

We have an easy rule to ensure that blame and call for action are related to the grievance:

You can only code them if they appear in the text in direct connection with a grievance so that you can mark the grievance statement and the blame / call for action statement together in your coding project.

So when a speaker randomly says "The government is to blame" and it is not clear from a grievance statement that appeared right before in the text that the government is to blame for a collective group grievance, do not apply the blame codes.

In short, to be coded, blame and call for action statements have to be either directly integrated into a grievance segment, or have to appear immediately before or right after the grievance segment. It has to be clear to the reader which grievance the culprit is being made responsible for, respectively which specific grievance is to be remedied in a call for action.

Blame can only be attributed to actors

Blame should only be attributed to individual or collective **actors**, not to abstract objects, events, processes or ideologies (i.e. do not code if "capitalism" or "colonialism" is blamed, but do code if "the capitalists" or "the colonizers" are blamed for the unjust situation that a group finds itself in).

In sum, for the grievance segments you will be coding:

1. elites portraying a given situation in issue area IA as **unjust deprivation of a group** G_1,
2. elites **blaming** another actor/group G_2 for that same unjust situation (in the grievance statement itself, or in the sentence(s) just before or after),
3. elites calling on members of G_1 to **act** in order to change that same situation (in the grievance statement itself, or in the sentence just before or after).

Double coding in the grievance segments

We **allow double-coding** of segments for the grievance parts of the text. A grievance statement can be coded with two or more sub-codes if applicable.

If for example one and the same sentence states that members of G_1 were disadvantaged by fraudulent elections *and* suffer from a bad health infrastructure you should assign both deprivation issue "political-legal" and deprivation issue "socio-economic". (Codes will be introduced below).

Grievances on behalf of ethnic and social groups in other countries

In this project, we are only interested in ethnic and social groups within the respective country. If a document refers to an ethnic group in another country (e.g. a Spanish party talks about the situation of Roma people in Italy), do not code this as a grievance. While the Roma people are also a relevant ethnic group in Spain, we are only interested in grievances made on behalf of the Spanish Roma by Spanish organizations. The same holds for societal groups (e.g. a Spanish party posting about farmers in Tunisia should not be coded).

Illustration

The following example serves as an illustration of the overall logic of identifying grievances, blame and calls for action (sentences are numbered for convenience, but remember: grievances can be several sentences long!):

"(1)We clearly saw electoral fraud occurring in the last election, our fellow group G_1 did not receive as many votes in the official vote count as they should have.

(2)I am therefore calling on all members of G_1 to demonstrate against the fraudulent elections, which were rigged by the corrupt G_3 politicians!"

[….. speech goes on….]

[….. speech goes on….]

"(30)And have you noticed that all natural resources in the region are exploited by G_2, while our G_1 and the rural population living in the area do not have any access to the mines?

(31) Also, our fellow G1s are excluded from all political positions in the region!
(32)This cannot go on this way!.

→ The first sentence frames a given situation as a grievance according to our definition. The speaker argues that deprivation of a group concerning a certain issue is present.

The deprived group is G1 and the issue is a political-legal one. We code these two sentences as *deprived group/ethnicity/ G1* and *deprivation issue/political-legal.*

The second sentence relates to the same grievance and **blames** another group G_3 for the election fraud and calls for **action.** Therefore, we assign *blame/ethnicity/#G3* and *action/protest*. We mark both sentences together and assign the grievance and the blame codes to the set of sentences 1 and 2.

Somewhere later in the same speech, our coder finds another interesting statement in sentence 30 ("and have you noticed…"). Now the party leader talks about another grievance of the same deprived group (*ethnicity/G_1*), but the deprivation issue is different (*socio-economic*). We apply the codes: *ethnicity/G_1* and *socio-economic.*

Furthermore, in the same statement, the speaker also states that the rural population is being treated unfairly in the socio-economic realm. This illustrates also that there can be more than one grievance in a sentence. The correct additional codes for sentence 30 are *deprived group/geography/rural* and *deprivation issue/socio-economic).*

In sentence 31 and 32, the speaker deplores again a political-legal issue as unacceptable. The deprived group is again G1. As before, we assign (*ethnicity/G_1*) and *deprivation issue/political-legal.* This time there is no blame or call for action connected to the grievance.

Second Goal: Coding Social Media Posts (Facebook) that require no Segmentation

Grievance Coding Rules-Adaptions for Social Media Posts

1. No segmentation- If you find a grievance in a post then code the whole post (this also includes urls, dates, headlines etc.). If a post does not contain a grievance code whole post as #blank. There should not be any post in which you assigned blank code AND a grievance code.

2. Context Unit- All posts by one ethnic group serve as your context information. This means that if the third post by one ethnic group contains information that help you interpret the first post, you can use this information to make your coding decision.

3. Google/additional information- You are NOT allowed to search for further context information outside of the posts and google. The only exception are ethnic group names and if you for example are unsure whether the ethnic group Guarani belongs to the lowland indigenous people or highland indigenous ones.

4. Quotes- In social media like twitter and co. there can be high usage of quotes. Include them and treat them as normal text that also needs to be coded.

5. Symbolic Individuals- If a post only talks about an individual person and not a group, but the person and what happened to her serves as a symbol for a societal group (eg. Roma girl) code this as a grievance.

6. Issue Unclear-Furthermore, it can be more difficult to classify the issue because the post is too short. In such cases just use the "Deprived Issue/Other" Issue code.

7. Hashtags – If a post contains a grievance and a call to action is made in a hashtag as part of this post, assign a call to action code.

The Coding Scheme

The remainder of this document will now present you with the coding scheme including all categories in detail.

Figure 1 displays the overall coding scheme, including the overall classification of the document we ask you to do in the xls spreadsheet.

Figure 2 displays the various Grievance codes and the "Blank" codes to be applied in QDA Miner. Note that the grievance elements and the blank code are mutually exclusive: you can either mark a segment as "blank" or annotate it with the grievance elements, but not both. Once you identify a segment as containing grievances, you can apply multiple group, issue, blame or action codes as you see fit.

The figures are followed by tables that describe each code in detail.

The overall coding of the document in a spread sheet is explained at the very last, at the end of this document.

Figure 1. Overview of all codes

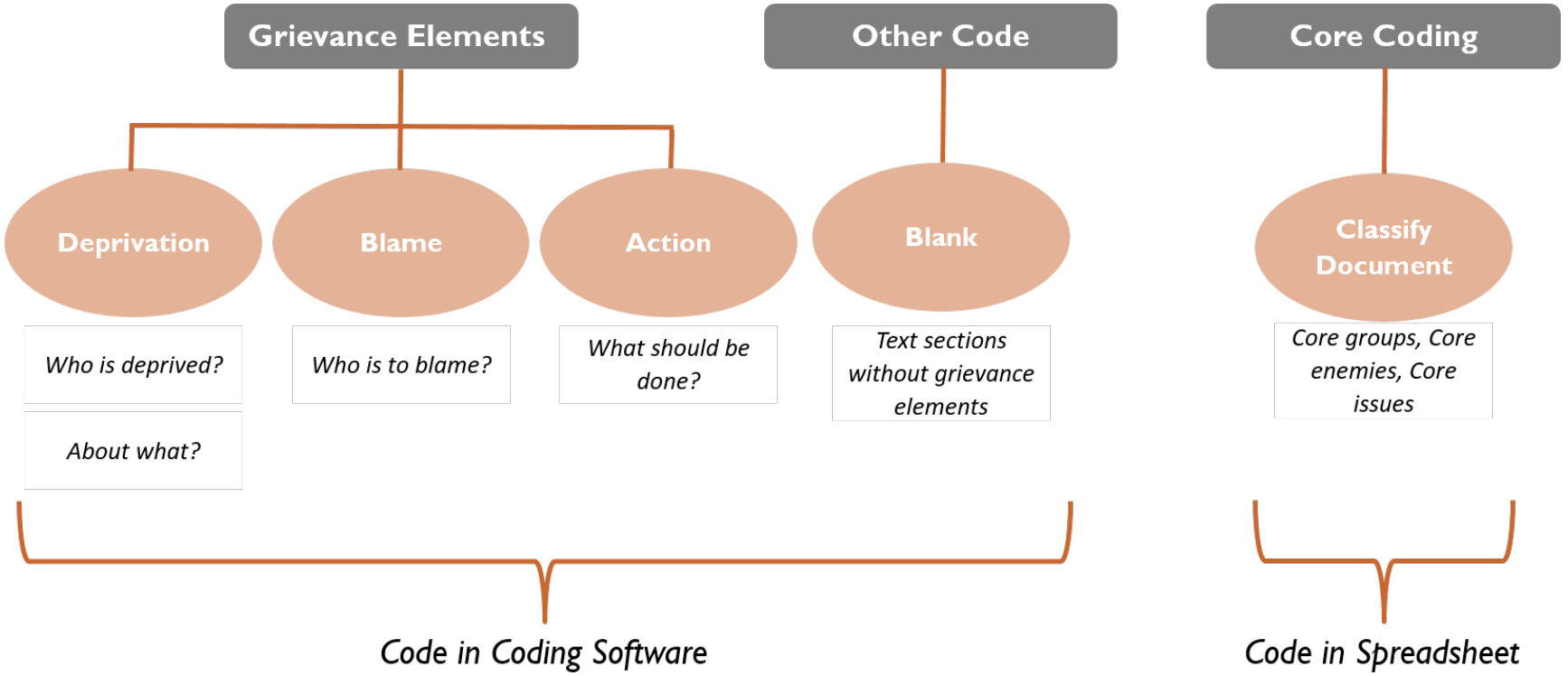

Figure 2. Coding scheme for coding grievances in coding software

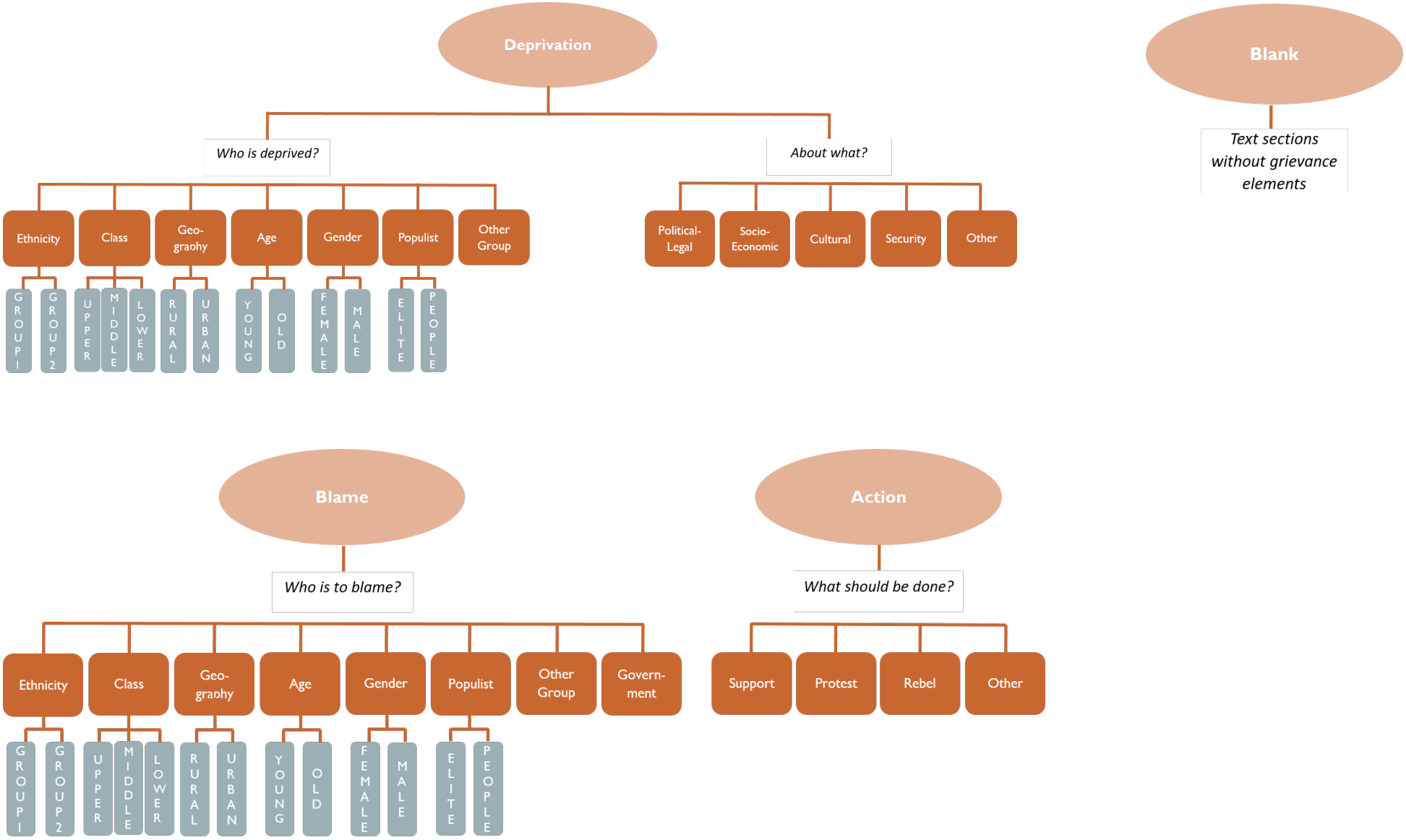

Table 1. Coding instructions for coding grievances in the coding software

Grievance element	Definition	Coding rule	Example
Deprivation (basis for identifying the grievance frame)	Statements that characterize a situation in an *issue* area as unjustly diverging from a (real or imagined, explicit or implicit) desirable standard for a certain *group*	Choose type of group (deprivation for whom?) Choose type of issue (what kind of deprivation?) from sub-categories of "deprivation" explained below. Do not code as grievance if a situation is described as unjust, but no aggrieved group is mentioned. Grievances are at least one sentence long, but one sentence can contain multiple grievances. Do not assign code if statement describes how a future situation should be, but does not deplore the current situation as unjust. Do not assign code if statement is a "hypothetical grievance" (e.g. if the government had pushed through with this policy, our pensioners would have been disadvantaged).	"Every statement I intend to make is evidence of the existence of two classes of citizen – first class white second class brown. The facts cannot be denied, they cannot be rejected and that makes their existence totally reprehensible, totally unacceptable."
Blame	Statements that blame another group for the problem	Choose type of group from the sub-categories of "blame". Do not assign code if it is a statement in passive voice and no specific group is blamed (e.g. the Indians were discriminated against in the past). Blame has to be connected by the speaker directly to the grievance and has to appear in the immediate vicinity of the grievance in the text.	"The day though, that all 13 of Labour's Māori MPs have been silenced […] – by a Māori minister with the support of his Māori caucus and being endorsed 100% in these activities by the Prime Minister"
Action	Statements that call on group members to act against the problem	Choose type of action from the sub-categories of "action". Note: mobilizing statements are different from policy proposals. Only use "action" if the statement calls on group members to act, not if another remedy such as a policy reform is suggested. Call for action has to be connected by the speaker directly to the grievance and has to appear in the immediate vicinity of the grievance in the text.	"The dominant voice in the Tory party has been of right-wing English nationalism. They have disregarded the voices of citizens in Scotland and the north of Ireland to remain. […] Your vote for remain candidates in this election is again about sending a message […]"

Table 2. Coding instructions for type of deprivation issue

Note: When assigning issue areas, respect the following rule: the **issue area is defined by the consequences**, not by the cause of the problem. If for example a document states that ethnic group G_1 does not have access to natural resources because of laws adopted by the elite, the issue area "socio-economic" should be assigned. While the cause of the problem is "political-legal", the consequences defining the currently suffered grievance are "socio-economic".

Issue area	Definition	Coding rule	Example
Political-legal	Statements that deplore a deprivation in terms of political, civic or human rights	Code here grievances about corruption, if it is unclear whether corruption occurs in politics or in economy. Code contains grievances about justice and discrimination, unless grievance is explicitly linked to another issue (e.g. lack of income justice should be coded as socio-economic grievance).	"It is a sad day indeed for the i Taukei in Fiji when their indigenous rights under ILO 169 and United Nations Declaration on the Rights of Indigenous People (UNDRIP) had been taken away by the false Constitution and Decrees made beforehand and locked by Sections 173 and 174 of the draconian Constitution of 2013. These Decrees do take away the rights and authority of the indigenous i Taukei people here in Fiji."
Socio-economic	Statements that deplore a deprivation in terms of access to socio-economic resources or services	Code contains education grievances, except if education deprivation has cultural consequences (e.g. language use in school should be coded as cultural grievance). Code contains migration grievances, unless grievance is explicitly linked to another issue (e.g. migrants ruin the cultural values of Pakistan people should be coded as cultural grievance). Code contains climate change grievances, unless grievance is explicitly linked to another issue.	"The Congress of Aboriginal Peoples (CAP) represents the status and non-status First Nations off-reserve, Métis and southern Inuit, who are the most socio - economically disadvantaged group in Canada. Repeatedly, statistical data informs us of the less than favourable conditions that show great disparity between Aboriginal peoples and their Canadian counterparts."
Cultural	Statements that deplore a deprivation in terms of culture, language,	Code here claims about education if the primary consequence of the deprivation is a disadvantage related to a group's language or culture.	"However, despite there being no evidence to support the position, the Court decided that the significance of the land to Maori could be accommodated within urban development. In essence the

	identity		Court set aside specific provisions reflecting Maori values and concepts, that were recognised by local authorities, in order for a landowner to realise an increased economic return."
Security	Statements that deplore a deprivation in terms of security	Code contains cyber security and safety of personal information in the internet.	"Honourable Members, we have all been witness to the increase in the atrocities on minorities in Pakistan over time. We have all seen what happened in Nankana Sahib recently. It is the responsibility of all of us to bring the atrocities being committed in Pakistan to the notice of the global community. While condemning the atrocities on the minorities in Pakistan, I urge the world community to take cognizance of it and take necessary steps in this regard."
Other	Residual category	Use for issues that do not fit under any of the other issue areas. Only use as last resort if statement fits neither of the other issue areas.	"The root cause of the appalling conditions of Aboriginal peoples can be traced to the ongoing negative effects brought on by colonialism, which assumed total control and management of their lands and resources and which continues to negatively impact their lives to this day."

Table 3. Coding instructions for deprived group

Note that in order to count as any group at all, the identity category in question must be distinguishable from all people in the respective country. That is, a group is defined against what it is not within the same societal context. All of Myanmar's people is not a distinguishable sub-set of the whole population, Mons or Zomis are a sub-set of Myanmar society and should be coded.

Deprived group	Definition	Coding rule	Example
Class: upper/middle/ lower	Statements that deplore a deprivation along class identities:	Class can be defined in terms of income or in terms of profession	"The latest trade agreements negotiated and signed by Ottawa have done nothing but create uncertainty in Quebec's towns and regions, in particular among farm owners, who are generally the ones who stimulate economic growth in their communities." (code: deprived group/middle class)
	Rich/upper class	Includes big entrepreneurs, big business owners, bankers and large landowners, capitalists	
	Middle class	Includes small entrepreneurs, small businessmen, shop owners, farm owners, craftsmen with own shop	
	Poor/lower class	Includes workers, agricultural and migrant workers/Employees, farmers, unless their farm owning states is emphasized (then code middle class), traders if clear from context that are small scale street vendors	
Populist: the elite/the people	Statements that deplore a deprivation along the lines of "the people" imagined as sovereign and pure vs. the "corrupt elite", imagined as taking away power from the people.	Choose between "the people" and "the elite" subcategories. Do not use this code if "the elite" or "the people" are clearly defined by income or profession or other identities, in the latter case use relevant categories (e.g. the rich elite should be coded as "class/rich").	TBA
Ethnicity/G1, G2…	Statements that deplore a deprivation along the lines of ethnicity	Assign group number based on the list of groups in each context in the google spreadsheet. If statement refers to the group with a label derived from its homeland region, code also. Do not code mere references to the region itself. Do code it, if the people in the region are referred to (e.g. references to the "Quebecois" or "the people of Quebec" or "Quebeckers" would be coded, references to just "Quebec" not).	"Aboriginal people are the poorest, least healthy, less educated, less employed and most incarcerated group in Canada. They are subject to racism and discrimination that are reflected in the policies that manage their lives."

		If statement refers to a group which name is equal to the country name (e.g. the ethnic group of Fijians in Fiji or Spanish in Spain), the coder has to decide based on the context whether Fijian/Spanish is used to refer to ethnic group or to the whole nation. Assign this code also if ethnic group is not mentioned by name but with "our people" "our community" if the context and previous statements make clear that "our" refers to an ethnic group. Do not assign code to mere pronouns (e.g. do not code "our families" as G1, even if it is clear from context that "our" indicates G1 because statement refers primarily to families. In short, code groups that appear as nouns.	
Gender: male/female	Statements that deplore a deprivation along the lines of gender identities	Choose between male and female subcategories.	"Specifically, we have highlighted among others, the high unemployment rate among the youth; unacceptable incidents of crime and gender-based violence targeting women and girls."
Age: old/young	Statements that deplore a deprivation along age groups	Choose between old and young subcategories.	"The increasingly youthful population is faced with unemployment and diminishing opportunities for self-actualisation. While inter-generational conflict for leadership continues to intensify, the government has, since Independence, not been able to invest in infrastructure that is capable of assuring the youth of a better today and the best of tomorrow."
Geography: urban/rural	Statements that deplore a deprivation along the urban rural cleavage	Choose between rural and urban subcategories.	"Rural people are treated like second-class citizens. They do not have access to services and technology in urban centres. They have to travel several times to towns and cities to meet simple needs, wasting time and money."
Other group	Statements that link deprivation to a group not listed	Use for statements that clearly link deprivation to a group but the group does not fit any of our predefined social identities.	"Madam Speaker, the later coups, the year 2010 was held under the management and control of the current Prime Minister as the Military Commander at that time. This coup with its

		Assign this code if a party is mentioned, but was not part of the government at the respective time period.	following associated mutiny, some of our soldiers were killed or murdered, and the law has not been allowed to take its course to-date. How can one keep harping about peace and prosperity when the aggrieved families are still crying today? How foolish one can be?"

Table 4. Coding instructions for sub-categories of blame

Blamed group	Definition	Coding rule	Example
Same sub-categories as aggrieved group with the addition of the code "Government"	Statements that blame members or elites of another group or the group itself for the situation of unjust deprivation	Do assign this code if an individual or a collective **actor** is blamed (e.g. the colonial government). Do not assign this code if a policy, abstract object, ideology, or historic event is blamed that is not an actor (e.g. colonialism, capitalism).	"MQM is the only political party in Pakistan which represents and comprises of ninety eight percent working, middle class and poor masses of the country who are presently down trodden, disadvantaged and exploited by the two percent ruling elite." (code: blame/populist/elite)
Government	Statements that blame the government for the situation of unjust deprivation	Use for statements that clearly blame the government for the situation of unjust deprivation. This includes actions directly by the government and governing parties / actors (e.g. president) but also actions by state officials, such as administrators, the police or the military. If ethnic group G_x is mentioned later as party in government, then code both "government" and G_x if applicable. Do not assign code if statement blames a policy for the situation without mentioning a government that adopted the policy. Do not assign code if it is a passive statement and no specific group is blamed (e.g. the Indians were discriminated against in the past). Do assign this code if statement uses capital name instead of	"Over the last 5 years of the BJP government, India has suffered under the 'Modi model,' which has devastated our economy, impoverished our farmers, destroyed jobs, divided communities, stoked hatred and spread fear. It is a model that has stripped the average Indian of dignity, voice and trust. This is not the model India deserves. This is a model that all Indians must reject."

		"government". Do assign code if statement mentions party name and this party is/was part of the government for the respective time period (in this case you are allowed to google). If the mentioned party is/was not part of the government code party as "other group". Use only for country's government. External actors should be coded as "other group"	

Table 5. Coding instructions for sub-categories of call for action

Call for action	Definition	Coding rule	Example
support	Statements that call on group members to support the activities of the organization, for example by voting for them	Only assign this code if statement explicitly calls for action. Do not assign code based on an interpretation of a statement as call for support.	"The dominant voice in the Tory party has been of right-wing English nationalism. They have disregarded the voices of citizens in Scotland and the north of Ireland to remain. But the game is far from over. Your vote for remain candidates in this election is again about sending a message about which union you want to belong to; Ireland in Europe or Britain."
protest	Statements that call on group members to demonstrate/protest against the situation of deprivation		
rebel	Statements that call on group members to turn to violent means to change the situation		
Other	Residual category for calls for action that do not fit under either of the other sub-categories of calls for action		"The wheels of the sunset clause under the cultural autonomy is now rolling and it will be executed faster in the next four years. Brace yourselves!" (Here, "Brace Yourself" is the central part determining the Call for Action code assignment)

Coding the entire document in the spreadsheet:

Core Groups, Core Enemies, Core Issues

In addition to coding grievance frames within the document, we ask coders to classify each document as a whole according to its core groups, core enemies and core issues. For each document up to two groups, two enemies, and two issues can be identified. This is so we have quickly accessible data on what each document is primarily about, which group is primarily appealed to and against whom.

For this overall coding of the entire document, we use a spreadsheet available on *googledocs*, allowing all coders to enter information simultaneously online.

Coders read through the *entire* document with three questions in mind:

1. Are there one or more *issues* that appear as central to the organization's/speaker's programmatic statement, that is, are there issues the organization/speaker says is more important than others/crucial to the organization's agenda (e.g. the discrimination of Fijians in parliament)? Note that these do not have to be presented as grievances.
2. Are there a *group* or groups that the party primarily appeals to, such as a group or groups of which the organization/speaker claims to be the true defender/representative (e.g. Indians, workers, the poor etc.)?
3. Are a key *enemy* or enemies singled out by the party, another social group or groups that the party sees as especially dangerous or particularly responsible/blame-worthy (the rich, the Bosniaks etc.)?

In the exceptional case that the answer to either of these questions is no, we insert "none" into the google spreadsheet but the goal should be to identify up to two overall core issues, group and enemies for each document.

For core group and core enemy, first enter the same standardized categories as in "deprivation for whom" or "blame". In addition, enter the names of core groups and core enemies as labeled in the actual document. Analogously, for core issues, assign the standardized codes for issue areas used in "what kind of deprivation" and enter the dominant issue as labeled in the actual document.

If one of the two groups/enemies/issues is more important than the other, note the more important one as group1/enemy1/issue1. If both appear as similarly important enter them in whichever order you prefer. The entries for core issue 1 and 2 can be the same (e.g. socio-economic for both cases), as long as the issue labels are different.

Table 2. Template and example for coding core groups, core enemies and core issues

document	core group 1	group label 1	core group 2	group label 2	core enemy 1	enemy label 1	core enemy 2	enemy label 2	core issue 1	issue label 1	core issue 2	issue label 2
FJ_SODEL-PA_M1_2018	ethnic group 1	fijians	poor	workers	Ethnic, group 2	indians			political-legal	Discrimination of Fijians in parliament		