COGITATIO

Article

A New Approach to Map and Quantify Representative Claims and Measure Their Validation: A Case Study Analysis

Viola Joschko [1],* and Luis Glaser [2]

[1] Department of Political Science, Goethe University, 60323 Frankfurt, Germany; E-Mail: mail@viola-joschko.de
[2] Department of Computer Science, Goethe University, 60323 Frankfurt, Germany; E-Mail: lglaser@cs.uni-frankfurt.de

* Corresponding author

Abstract
Relying on the theory of Saward (2010) and Disch (2015), we study political representation through the lens of representative claim-making. We identify a gap between the theoretical concept of claim-making and the empirical (quantitative) assessment of representative claims made in the real world's representative contexts. Therefore, we develop a new approach to map and quantify representative claims in order to subsequently measure the reception and validation of the claims by the audience. To test our method, we analyse all the debates of the German parliament concerned with the introduction of the gender quota in German supervisory boards from 2013 to 2017 in a two-step process. At first, we assess which constituencies the MPs claim to represent and how they justify their stance. Drawing on multiple correspondence analysis, we identify different claim patterns. Second, making use of natural language processing techniques and logistic regression on social media data, we measure if and how the asserted claims in the parliamentary debates are received and validated by the respective audience. We come to the conclusion that the constituency as ultimate judge of legitimacy has not been comprehensively conceptualized yet.

1. Introduction

In the political will-formation processes, increasingly complex constellations of actors exist. Traditionally, representatives are authorized by and held accountable through an election by territorially defined constituencies. Due to the increasing passiveness of voters, reflected by decreasing voter turnout, the loss of faith in political institutions, a decline of party loyalty and growing political power of non-governmental organizations, tensions concerning who has the democratic right to represent the people emerged (Näsström, 2011; van Biezen, Mair, & Poguntke, 2012). Nowadays, however, not only the elected political elite but also self-appointed actors feel entitled to make representative claims and voice political interests. The inclusion of these new voices might provide novel avenues in the production of legitimacy within democratic systems (Mair, 2009; van Biezen, 2014; Young, 2002).

It has become apparent that the traditional models of electoral accountability and constitutional representation might not be sufficient anymore. How democratic accountability and legitimacy are operationalised in these decentralized arrangements, made up of a broad variety of representative claims and claim-makers, becomes crucial in order to assess and attain democratic legiti-

macy (Guasti & Geissel, 2019). These societal developments revived the scientific debate about political representation, oftentimes referred to as "representative turn" (Disch, 2011; Näsström, 2011, 2015; Urbinati & Warren, 2008). Representation is no longer bound to a static principal-agent relationship, but is seen as a dynamic process in which elections are no guarantee that representation will be democratic (Disch, 2015, p. 487). Representative claims—i.e., claims of a subject wherein she or he vouches for a certain proposal (interest) or constituency (with that particular interest)—are mechanisms of the dynamic processes of representation.

Especially Saward's framework of representative claim-making became very influential. He sees claim-making as the core of representation: "'Representation' can be said from this perspective not to exist; what exists are claims and their receptions." According to him, "Representation is produced by processes of claim-making and consequent acceptance or rejection by audiences or parts of audiences" (Saward, 2006, p. 306).

In this claim-making process, he sets the focus on the claimed constituency and the audience respectively. He explicitly emphasizes the importance of the claimed constituency, in particular as he criticizes Pitkin (1967) for focusing solely on the representatives, while "the represented is taken as unproblematically given" (Saward, 2006, p. 300). With this, he challenges the traditional assumption of an objectively pre-existing and manifest constituency with apparent needs and interests. He puts forward the idea that in their claims, political representatives verbally construct their constituencies, e.g., as "hard working, highly educated women" and imply that they are their best representatives. However, every claim of representation leaves space for contestation or rejection by the ostensible constituency (or other political actors) (Saward, 2006, p. 302).

By emphasizing the role of the constituency, Saward subsequently derives democratic legitimacy from the reception of the claim by the audience. He considers this as a valid form of authorization:

> If we want to assess the democratic legitimacy of representative claims in a democratic 'way' (and, as democratic theorists, we should), then we must leave it up to the 'would-be constituents of claims' to decide whether or not to accept them." (Saward, 2010, pp. 144–147)

Therefore, claims are legitimate when they are perceived as such "by appropriate constituencies under reasonable conditions of judgement" (Saward, 2010, pp. 144–147).

Due to this broader perspective, which neither conceives representation as a given outcome of an electoral process any longer, nor accepts constituencies as undeniable, a necessity to examine the reception and therefore legitimacy of representative claims arises.

To allow for a systematic empirical analysis in this regard, Disch (2015) vows for a separation of three stages of claims making, whereas the first is the making of the claim and the second is the reception of the claim. The reception stage itself consists of two steps, first the audience must recognize the claim in order to then validate it, resulting in acceptance or rejection of the claim.

The most prominent systematic approach analysing representative claims can be accredited to De Wilde (2013). With his "Representative Claim Analysis" (RCA) he combines Saward's theoretical terms (claim-maker, subject, object) with the method of political claims analysis (Koopmans & Statham, 1999). In the tradition of political claims analysis, he sticks with qualitatively analysing newspaper articles, forfeiting the assessment of unaltered claims in direct speech. Here, the reception of the claim (the second stage in the process of claim-making; Disch, 2015) remains left aside.

Although a lot of research has applied some sort of representative claims analysis, it seems that no one has ever assessed representative claims from the citizens' standpoint, e.g., with a focus on the reception by a claimed constituency. For example, Kuyper (2016) proposed to derive the validity of claims from a normative evaluation by experts, instead of considering the claimed constituency as authority. In contrast, we aim at contributing to the gap in the literature by clarifying empirically who claim-makers claim to represent, i.e., which constituencies they are creating, to eventually measure the reception of the claim by the claimed constituency (measured as audience).

We do this with the aid of a case study: the introduction of the women's quota in German supervisory boards. The reason for this choice is that it was a highly polarized debate with a clear outcome, where a multitude of opposing claims were made. Also, we can test the semi-automated structuring of the parliamentary claims by comparing our results with what is commonly known about the position of the involved parties towards the quota and thus validate the application of the MCA as a basis for the assessment of the validation. Plus, if the semi-automated structuring of the claims replicates commonly known results, an application to other novel and more complex cases involving non-elected political claim-makers may be considered.

We conduct three analyses. First, claims made in the German parliament are systematically analysed within the representative claims framework. Then, we assess the reception of the claims on Twitter and determine its validity for the purpose of, finally, measuring the validation of the claims made in the parliament. This work will cover the first and second step of the claim-making process (Disch, 2015): the analysis of representative claims in the parliament and the assessment of their reception on social media. With this we provide an empirical contribution to a theoretically elaborated scientific debate about representative claim-making.

As Näsström (2011, p. 502) points out, it is important to differentiate between "diagnostic or normative work". We also acknowledge that prior to normative de-

bates about the validation of representative claims, the empirical status quo needs to be captured. Our approach is purely diagnostic, as we aim to provide a method to systematically assess representative claims and operationalise validation. However, a theoretical reflection on which stipulations need to be met in order to define claims as accepted or declined lies beyond the scope of this work and needs to be addressed in the future.

2. Conceptual Framework

2.1. Making of Representative Claims

According to Saward, "a representative claim is a claim to represent or to know what represents the interests of someone or something" (Saward, 2006, p. 305). As he stresses, representation does not rely on a real constituency, however, creates an idea about it: "Representing is depicting...a constituency as this or that, as requiring this or that, as having this or that set of interests.". This stance hints at the risk of taking the constituencies' interest at face value, e.g., taking a feminist mindset for granted when referring to women as a constituency. To put it differently, if an MP claims to represent women and therefore affirms the introduction of the women's quota, it is tempting to evaluate this representative claim as valid because, at face value, women benefit from the quota and one would therefore expect them to be proponents of it. Keeping this in mind, we emphasize the importance of a systematic analysis of the constituencies' reception of the claims, all preconceptions of the researcher aside.

For us, a representative claim consists of at least three defining elements: an "issue", which in our case is always the introduction of the women's quota, a "stance", which is a position indicating a preference regarding the issue of relevance (pro, contra or neutral) and thirdly, a "claim-maker" who voices the opinion. Further elements can be a "claimed constituency", either humans (e.g., all women in Germany) or normative schemes (e.g., justice or equality) claimed to be represented, and there also might or might not be a verbalized "linkage", which refers to the relationship between claim-maker and constituency (see Figure 1). Guasti and Geissel (2019) take into account whether and how the claim-maker claims a linkage between him/herself and the represented, resulting in three types of claims: "claim of representation", "claim of values and interests" and "proclamation".

To this we add the element of "justification", conceptually referring to Saward (2010) where he introduces the "cultural code". De Wilde subsequently picks up this notion and renames it to "frame". He defines it as follows:

> In this sense, the 'framing' variable of the claims analysis method clearly resonates with what Saward calls a 'cultural code' in which a representative claim is made. The 'frame' in claims analysis is of internal character in the sense of a justification or meaning provided by the claimant. (De Wilde, 2013, p. 287)

We therefore adopt the concept of cultural code/frame, but name it "justification". We define it as the explana-
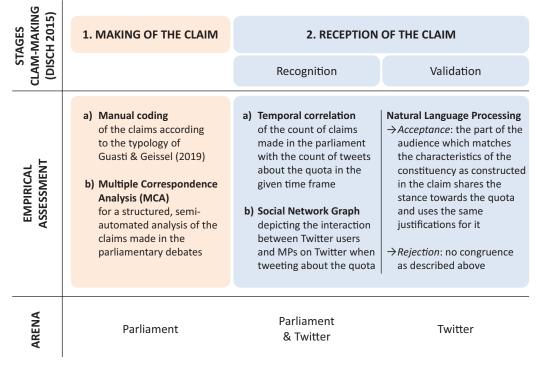
STAGES CLAM-MAKING (DISCH 2015)	1. MAKING OF THE CLAIM	2. RECEPTION OF THE CLAIM	
		Recognition	Validation
EMPIRICAL ASSESSMENT	a) **Manual coding** of the claims according to the typology of Guasti & Geissel (2019) b) **Multiple Correspondence Analysis (MCA)** for a structured, semi-automated analysis of the claims made in the parliamentary debates	a) **Temporal correlation** of the count of claims made in the parliament with the count of tweets about the quota in the given time frame b) **Social Network Graph** depicting the interaction between Twitter users and MPs on Twitter when tweeting about the quota	**Natural Language Processing** →*Acceptance*: the part of the audience which matches the characteristics of the constituency as constructed in the claim shares the stance towards the quota and uses the same justifications for it →*Rejection*: no congruence as described above
ARENA	Parliament	Parliament & Twitter	Twitter

Figure 1. Overview of the conceptual framework.

tion for the stance of the claim-maker towards the issue, e.g., the explanation why the issue (introduction of the quota) is supposed to be beneficial or detrimental to the constituency according to the claim-maker.

In textual terms, a claim can be comprised of a single sentence or several connected sentences. With the sole analysis of newspaper articles, the original claims are already preprocessed and filtered through the lens of journalists. Therefore, we advocate for the assessment of direct speech of the claim-makers. In our case, we analyse parliamentary speeches, as they serve as the most accurate reflections of the MP's self-representation.

2.2. Reception of Representative Claims

Representation through claim-making only works for Saward, if claims are received by an audience. De Wilde identifies some problems for empirical inquiry in Saward's analytical framework: the operationalisation of the difference between constituency and audience. He states that the problem lies in the difficulty of identifying a passive and often not specifically mentioned audience. Therefore, he rejects the distinction as "needlessly complicated" and excludes the audience from his claims analysis (De Wilde, 2013, p. 284). Disagreeing with De Wilde, we argue that the difference between audience and constituency becomes necessary for the empirical

assessment of the validation of claims. In our view, the constituency is an abstract entity, verbally constructed by the claim-maker in the claim. The audience is defined as the people who match the defined characteristics of the respective constituency as verbalized in the claim and actively accept or reject the claims. If a claim-maker states that "I am representing the men in Germany and oppose the quota, because it discriminates against men", then men in Germany are the constituency and the German men who voice that they support/oppose the quota are defined as the audience. If a claim-maker claims to represent justice (a normative scheme), then a corresponding audience does not exist, as the validation of claims necessarily entails the active expression of acceptance or rejection and thus, only human beings can be considered to be an "audience".

We define the reception of the claim as being twofold: first, the claim needs to be directly or indirectly recognised and second, it needs to be validated, resulting in acceptance or rejection of the claim by the respective audience. A claim is considered accepted if the part of the audience matching the depiction of the claim-maker's constituency shares the respective stance and justification (see Figure 2).

Before the validation of the claims can be scrutinized, we have to make sure that we can assume an audience on Twitter. Do people on Twitter actually care and talk
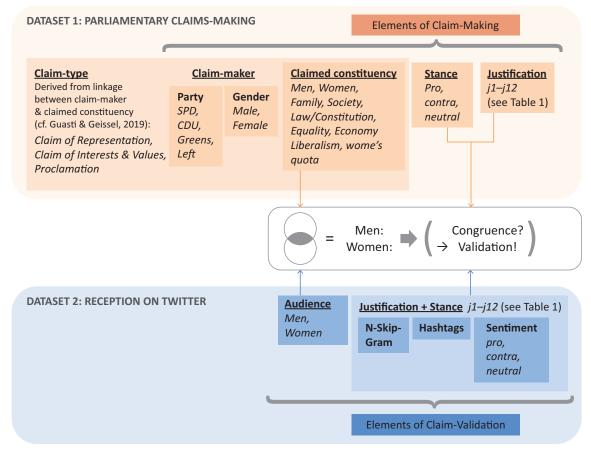

Figure 2. Overview of data structure and assessment of validation.

about the introduction of the quota? If not, it would not be reasonable to use Twitter data for validation purposes. To find out whether the Twitter debate about the women's quota is interconnected with the parliamentary debate, we first examine if MPs and Twitter users interact on Twitter by talking about the quota with each other. Second, we check the assumption of a temporal interdependence: if the parliament debates about the quota (i.e., many claims are made), we assume an increase of tweets about the quota as well. We do not think it is necessary that the audience has been directly confronted with the claims by watching parliamentary debates. We assume that the stance towards the quota and the respective justifications used for or against it in the claims could also be communicated via other tweets on the topic, newspaper articles covering the parliamentary debate, or other secondary sources.

Considering the restrictions the analysis of social media data bears, it still appeared as the most suitable source for our purpose, as people are able to express their opinion and sentiment towards an issue in an uncensored manner, directly and in an open text format. For these reasons, micro-blogging services are widely studied as a source of data on public opinion (Hassanpour, 2013; Jungherr, Schoen, & Jürgens, 2016; O'Connor, 2010; Tumasjan, Sprenger, Sandner, & Welpe, 2010). In contrast, standardized forms of public opinion surveys usually solely depict (dis)agreement on Likert scales, which are easy to interpret but poor regarding wealth of information. Plus, they are sometimes affected by social desirability bias, one of the most common biases affecting survey research (Barberá & Rivero, 2015; King & Bruner, 2000).

In sum, we operationalise the recognition of claims twofold. First, as a correlation of the count of claims made in the parliament with the count of tweets about the quota in the given time frame. Second, we check whether Twitter users and MPs interact with each other on Twitter. We operationalise interaction as a visibly interconnected social network on Twitter where the MP nodes are centrally located. This suggests that the MP is a relevant actor in the conversation. If both of these assumptions are met, we can assume the existence of an audience which validates (accepts or rejects) the parliamentary claims on Twitter.

We define validation as the congruence between claimed constituencies (as constructed in the claims) and audience (people validating the claims) in terms of stance and justification. We do not presume any kind of descriptive representation, meaning that an MP does not need to fit the descriptive characteristics of the claimed constituency. In this case this means that the gender of the MP does not play any role in the validation. If a male claim-maker claims to speak for women, then women can validate this claim.

We will not verify the factual correctness of the claims, e.g., if the MP makes claims in favour of the quota and uses the justification that the introduction of the quota is good for economy, we will not fact-check that. We will only examine whether the according audience (on Twitter) also believes so and favours the quota (stance) for this reason (justification). Thus, to consider a claim to be accepted, two conditions have to be fulfilled. The stance has to be congruent (measured as sentiment) and the justifications used in the parliament have to resonate with the audience.

We explicitly do not operationalise any threshold in terms of how congruent the claim and validation need to be in order to consider the claim accepted, because this definition of a benchmark should be set by a normative instead of an empirical approach. The implications of the findings will be discussed in the end. However, we restrict this work to the nature of a methodological trial to assess reception, not a normative evaluation thereof.

3. Making of the Claim: Empirical Analysis 1

The aim of this first empirical analysis is to map and to quantify claims as a prerequisite to measuring the reception of the claims.

3.1. Data Collection and Variables

To identify claims about the women's quota in the parliament, we analyse all the speeches of the 18th legislative period of the German Bundestag (22 October 2013–24 October 2017). To create the data set, we kept only those passages which were explicitly related to the introduction of the women's quota in German supervisory boards. Using this data, we coded the claim-maker (party and gender of MP), claimed constituency, stance (pro, contra, neutral) justification (open coding of claims) and claim type according to the claims typology of Guasti and Geissel (2019).

3.2. Method

To provide claim clusters, we examine which are the most similar and most dissimilar claims. Before the clustering, we need to quantify the claims. For this purpose, we apply multiple correspondence analysis (MCA). MCA is a dimension reduction technique for categorical and mixed data and performs data reduction by representing multidimensional data as points in a low-dimensional Euclidean vector space. The reasoning is similar to factor analysis. The goal is to measure the similarity (as distance) of single claims to cluster similar ones together. This helps to structure the claims according to the most relevant (non-)shared variable categories, representing latent structures (here: claim clusters) in the data (Rencher & Christensen, 2012). The clustering of claims facilitates the interpretability of the variety of claims made within the debate, eliminating the need to assess the reception of each single claim. Instead we assess the reception of the resulting claim clusters, which depict the relevant positions in the debate represented

Table 1. List of justifications in parliament and social media.

id	justification
1	"Quota is not an important topic, others are more pressing"
2	"Quota discriminates against men"
3	"Quota conflicts with constitution"
4	"Gender instead of qualification as the hiring criteria"
5	"Quota damages the economy"
6	"Quota benefits the economy"
7	"Quota leads to more equality, fairness"
8	"Quota is overdue and a step in the right direction"
9	"Quota is not sufficiently far-reaching"
10	"Quota leads to more qualification in board rooms"
11	"Voluntary quota is not enough, the state needs to intervene"
12	"Quota must not conflict with care work"

Note: Throughout this article we will reference these justifications as j_{id}, e.g., we refer to the first one as j1.

by the MPs in the parliament. Therefore, this first analysis provides the basis for the ensuing measurement of the reception of the claims.

We analyse a data frame of 320 claims × 31 variable categories, with claims as rows and all of the categories for every variable j in the columns. This table therefore has $I \times K$ dimensions and is composed of 1s and 0s. An element x_{ik} of this table has a value of 1 if the claim i carries category k, and a value of 0 if it does not (Husson, Lê, & Pagès, 2017, p. 134).

The distance between claims is calculated by summing up the squared differences between categories $(x_{ik} - x_{i'k})^2$ and dividing (and thereby balancing) by I_k, the number of individuals carrying category k. This distance can be expressed as:

$$d^2_{i,i'} = \frac{I}{J} \sum_{k=1}^{K} \frac{1}{I_k} \frac{(x_{ik} - x_{i'k})^2}{I_k}$$

where J is the number of variables, each having K categories.

The distance between the two categories k and k' is calculated by counting the individuals which carry either category k or k' and divide it by I_k and $I_{k'}$ respectively, the number of individuals carrying category k or k' (Husson et al., 2017, p. 135). The squared distance between two categories can be expressed as:

$$d^2_{k,k'} = I \sum_{i=1}^{I} \left(\frac{x_{ik}}{I_k} - \frac{x_{ik'}}{I'_k} \right)$$

where I is the number of individuals.

We further demonstrate that it is possible to use MCA as a preprocessing technique to eventually perform a clustering method on the principal components (Husson et al., 2017), which defines homogeneous claim clusters. Due to the optimal allocation of quantitative values to categorical variables, further multivariate analyses are also facilitated (Di Franco, 2016).

3.3. Descriptive Results

In total, we found N = 310 claims in the parliament, with 70% of them made by female MPs. A total of 33 unique constituencies were referred to by the claim-makers in the parliament. For further analysis, we grouped them together, resulting in 9 different constituencies: men, women, family, society, law/constitution, equality, economy, liberalism and interestingly, after a certain point, "women's quota" itself became a constituency (normative scheme and therefore value worth representing; c.f., Guasti & Geissel, 2019). The justifications we found due to open coding in the claims are listed in Table 1.

3.4. Mapping of Claims: Analysis 1

The results of the MCA suggest to restrict the analysis to the description of the first three dimensions, which express 71.2% of the explained variance (JCA corrected principal inertia/"R^2"; c.f., Greenacre, 2017, p. 146). As in factor analysis, the most important (in terms of explained variance) dimension is the first one. The overall categories with the highest squared cosine (thus which appear most suitable to structure the whole parliamentary debate) are pro, contra and neutral and belong to the variable stance (Figure 3).

The relationship between the variable categories is shown in Figure 4. Categories with a similar profile group together, while negatively correlated variable categories are positioned on opposed quadrants. The resulting three claim clusters (Figure 3) become visibly distinguishable: the neutral cluster is depicted in the upper centre of the graph, the positive cluster in the lower left and the negative cluster in the lower right.

Figure 4 depicts the representation of the variable categories in a two-dimensional chart. Those combinations of variable categories, which are typically shared by claims belonging to the same cluster, group together in the graph. For example, in claims where the claimed con-
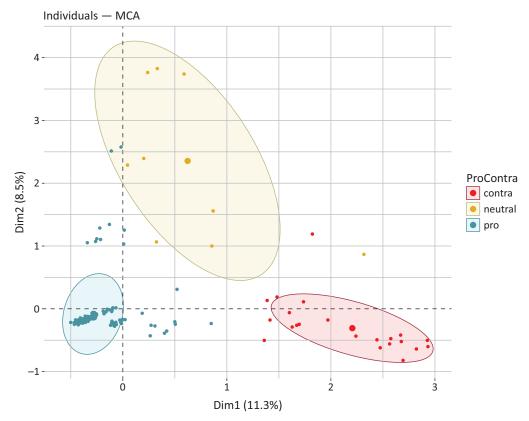
Figure 3. As a result of the MCA, we see every claim mapped as a dot on the first two dimensions, coloured according to its stance regarding the quota. The variable stance (pro, contra, neutral) appears most suitable in order to structure the claims made within the parliamentary debate about the women's quota, as they noticeably cluster together.

stituency is "men", usually the stance is contra the quota and common justifications are j1, j2, j3, j4 (lower right of the graph). These three resulting claim-clusters (see Table 2) will be objected to the validation in the end.

3.4.1. Description of the Dimensions 1–3

Dimension 1: Claims characterized by a positive coordinate on the x-axis are mostly claims that speak for

Table 2. The first analysis reveals 3 claim clusters, which are structured along the variable stance. Only significantly correlated variable categories are listed here.

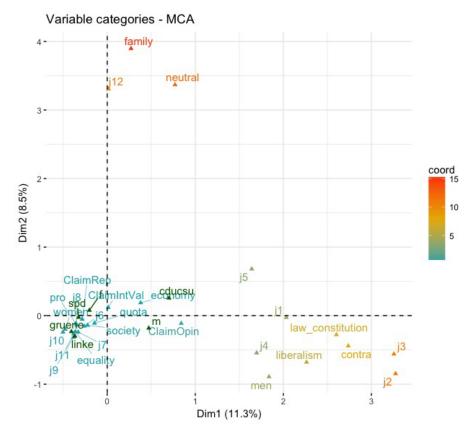| Claim Clusters		
CLUSTER 1: NEGATIVE STANCE	CLUSTER 2: NEUTRAL STANCE	CLUSTER 3: POSITIVE STANCE
Claim-maker • Party: Conservatives • Gender: Male	**Claim-maker** • Party: Conservatives • Gender: Female	**Claim-maker** • Party: Greens, Left, Social Democrats • Gender: Female
Claimed Constituency: Men, Law/Constitution	**Claimed Constituency**: Family	**Claimed Constituency**: Women, Equality
Justification: j1 (Quota is not an important topic, others are more pressing), j2 (Quota discriminates against men), j3 (Quota conflicts with constitution) and j4 (Gender instead of qualification as the hiring criteria)	**Justification**: j12 (Quota must not conflict with care work)	**Justification**: j7 (quota leads to more equality, fairness), j8 (quota is overdue), j9 (quota is not far-reaching enough), j10 (quota leads to more qualifications in board rooms) and j11 (voluntary quota is not enough, the state needs to intervene
Claim-type: proclamation	**Claim-type**: no claim-type sign. associated with this claims cluster	**Claim-type** Claim of representation

Figure 4. Relationships between the variable categories. Notes: The variable categories with similar profiles group together. The more they turn red the more distinct these categories represent the according dimension. Categories in opposed quadrants are negatively correlated. For the categorical variables, an ANOVA-model with one factor is done for each dimension. An F-test is derived to see whether the variable has an influence on the dimension and t-tests are done category by category. We can see if the coordinates of the individuals of the sub-population defined by one category are significant.

men (t = 0.27, p ≤ .001) and in the name of the law/constitution (t = 1.55, p ≤ .001). The stance is negative (t = 1.36, p ≤ .001) and the justification for this is mostly j1 (Quota is not an important topic, others are more pressing, t = 0.99, p ≤ .001), j2 (Quota discriminates against men, t = 1.99, p ≤ .001), j3 (Quota conflicts with constitution, t = 1.97, p ≤ .001) and j4 (Gender instead of qualification as the hiring criteria, t = 0.71, p ≤ .001). The type of claim closest to this group is the *proclamation* (t = 0.52, p ≤ .001). Supplementary variables associated with this dimension are *CDU* (Christian Democrats) (t = 0.65, p ≤ .001) and *male* (t = 0.27, p ≤ .001), which means these sort of claims can be attributed to male MPs from the CDU party.

In the space of negative *x*-coordinates, we find claims with a positive stance (pro introduction of the quota, t = −1.13, p ≤ .001) and women as constituency (t = −0.93, p ≤ .001), plus MPs who claim to speak in the name of equality (t = −0.84, p ≤ .001). Typically used justifications are j7 (Quota leads to more equality, fairness, t = −0.96, p = .003), j8 (Quota is overdue and a step in the right direction, t = −0.93, p ≤ .001), j9 (Quota is not sufficiently far-reaching, t = −0.96, p = .006), j10 (j10 = Quota leads to more qualification in board rooms, t = −1.05, p = .03) and j11 (Voluntary quota is not enough, the state

needs to intervene, t = −0.96, p = .02). The type of claim most common in this group is the *claim of representation* (t = −0.39, p ≤ .001). Supplementary variables associated with this dimension are *Greens* (t = −0.25, p ≤ .001), *the Left* (t = −0.22, p = .016), *SPD* (t = −0.18, p ≤ .001) and *female* (t = 0.27, p ≤ .001). The first dimension is predominantly defined by the *stance* variable (R² = 0.88, p ≤ .001), opposing pro-claims in the name of women with contra-claims in the name of men.

Dimension 2: Here, claims with high and positive coordinates often refer to families as constituencies (t = 2.6, p ≤ .001), claims have a neutral stance (t = 1.7, p ≤ .001) and use justification j12 (Quota must not conflict with care work) (t = 2.28, p ≤ .001). Supplementary variables associated with this dimension are *CDU* (t = 0.23, p ≤ .001) and *female* (t = 0.09, p = .036), which means that female MPs from CDU are those making these kinds of claims. The second dimension, therefore, is predominantly defined by the variables *constituency* (R² = 0.69, p ≤ .001) and *justification* (R² = 0.75, p ≤ .001) and distinguishes neutral claims in the name of families from all the other categories.

Dimension 3: Variables with an over-average contribution to dimension 3 refer to the economy as constituency (t = 1.6, p ≤ .001), use justification j5 (Quota

damages the economy) (t = 2.1, p ≤ .001) and j6 (Quota benefits the economy) (t = 0.77, p ≤ .001) and *claims of interests and values* (t = 0.41, p ≤ .001). Claims with positive coordinates on this dimension therefore tend to share the categories mentioned above. There are no supplementary variables significantly associated with this dimension, meaning that there is no difference regarding party and gender of MPs making these kinds of claims. The third dimension therefore is predominantly defined by the variables *constituency* ($R^2 = 0.73$, p ≤ .001) and *justification* ($R^2 = 0.68$, p ≤ .001) and distinguishes claims made in the name of economy from the rest.

3.4.2. Grouping Claims by Supplementary Variables

The supplementary variables gender and party of MP do not contribute to the plane construction, meaning they have not served to define the distance between the individual claims, but the categories of supplementary variables are eventually superimposed on the map. This makes it possible to investigate the influence of the gender and party of the claim-maker on their stance and claimed constituency. However, the two complementary variables are not considered for the validation analyses.

On all extracted dimensions with p ≤ .001, *speaker party* is the supplementary variable which, in comparison with *gender*, best explains the differences between the claims with respect to their stance (pro, contra, neutral). It is also possible to explore the meaning of the di-

mensions by examining the content of the claims with extreme values on the respective dimensions, but this is out of scope for our work.

3.4.3. Hierarchical Clustering on Principal Components

A hierarchical clustering on principal components (HCPA) is performed using the Ward's criterion on the results of the MCA in order to identify clusters of similar claims and validate the structure found beforehand (Figure 5). The classification made on individuals due to the HCPA analysis reveals 3 clusters, corresponding to the stance variable categories pro, contra and neutral. Therefore, the HCPA results confirm the insights gained through the MCA.

3.5. Conclusion

The results of the MCA and the HCPA suggest three claim clusters: pro claims, contra claims and neutral claims. Thus, the representative claims made in the parliament can be mainly structured along those lines: claims made in favour of the quota claim to represent women and equality, use justifications j7 (quota leads to more equality, fairness), j8 (quota is overdue), j9 (quota is not far-reaching enough), j10 (quota leads to more qualifications in board rooms) and j11 (voluntary quota is not enough, the state needs to intervene) and are constructed with the claim-type *claim of representation*. Claim-makers tend to be female and either Greens, Lefts or SPD. Claims
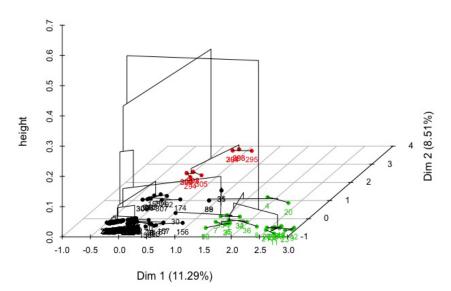

Figure 5. The dendrogram resulting from the HCPA analysis suggests a 3 clusters solution; claims are coloured according to the clusters they belong to: pro (black), contra (green) and neutral (red).

with a neutral stance regarding the quota claim to represent families, usually stressing that a quota needs to be compatible with family politics and care work. Claim-makers are usually female CDU MPs. Claims opposing the introduction of the quota are claiming to represent men and the law/constitution, use justifications j1 (quota is not an important topic, others are more pressing), j2 (quota discriminates against men), j3 (quota conflicts with constitution) and j4 (gender instead of qualification as the hiring criteria) and are constructed as *proclamation*. Claim-makers are mostly male and from CDU.

Considering that the main aspect for the differentiation between the claim-types is the claimed linkage between claim-maker and constituency (c.f., Guasti & Geissel, 2019), it appears interesting that for claims opposing the quota, oftentimes a linkage is not constructed. When claim-makers endorse the quota, they usually construct a linkage. Further research on this phenomenon is necessary and could provide new insights regarding the relationship between the represented and the representative.

We also found that party-alignment plays a more important role than gender when it comes to the stance of the MP towards the quota, which is in line with recent research (Espírito-Santo, Freire, & Serra-Silva, 2018).

MCA proved to be a suitable tool for structuring and quantifying the extracted claims. The underlying estimated structure of the debate resulting from the MCA aligns with common knowledge regarding the polarization of the debate, the different arguments and stances of the respective parties. Therefore, it seems to provide methodological face validity. Consequently, an application of this method to a case which has been less investigated seems promising.

4. Reception of the Claim: Empirical Analysis 2

4.1. Data Collection and Variables

We collected all tweets between 2013 and 2017 containing at least one of the German terms for "women's quota", "gender quota" or "Flexi-quota". We also collected the user profile information of people who posted the tweets. These profiles contain metadata including their username and occasionally their real name and/or location. We extrapolated the gender of the users, using the genderize.io API. Furthermore, we extracted both hashtags and mentions to infer topics of a tweet.

The variable *stance* was computed with the aid of a sentiment analysis (Hutto & Gilbert, 2014; Rill et al., 2012). Sentiment analysis is a way to automatically measure opinion in texts on the basis of single or multiple words which express a judgement (Taboada, Brooke, Tofiloski, Voll, & Stede, 2011). The terms are assigned a value between −1 (highly negative) and 1 (highly positive), e.g., "love" has a value of 0.82, while "not good" has a value of −0.63 (Rill et al., 2012). This results in an average sentiment per tweet, ranging from −1 to 1.

For the variable *justification*, we randomly picked 500 tweets and manually coded the justifications used. Of those, we selected meaningful 2-skip-2-grams. 2-skip-2-grams are combinations of each single word of a sentence paired with another single word within a 3-word vicinity of the initial word (e.g., "Men are being discriminated against" has the 2-skip-2-gram "men, discriminated"). Then we extracted the hashtags from the tweets. The combination of 2-skip-2-grams or hashtags with the sentiment led to the annotation of the justification. If, for example, the hashtag "*#economy*" was combined with a positive sentiment, then this led to the annotation of justification 6. If the sentiment was negative, this led to justification 5.

4.2. Descriptives

Our data collection comprised 53,807 tweets. For 18,731 Twitter users in our dataset, we were able to annotate the gender based on the first names. For the remaining part, an assumption was not possible as they used a pseudonym or did not declare their first name. For further analyses, we kept only the users for which we could estimate the gender. These 12,287 (65.6%) users accounted for 20,079 tweets.

We also checked for a regional bias. We were able to extract the location of 6,578 (35.12%) users. We found that the three most frequent states people tweeted from are the three German city states. These also have the youngest population (Statistisches Bundesamt, 2017), suggesting our data is slightly biased towards a younger audience. Manual coding of the 500 tweets resulted in 1549 2-skip-2-grams, which yielded 195 relevant bigrams we used for further analyses.

4.3. Recognition of Parliamentary Claims on Twitter

4.3.1. Recognition I: Temporal Trend

A Pearson product-moment correlation coefficient was computed to assess the relationship between the amount of tweets about the quota and the amount of claims in the parliament, showing that there is a strong, positive correlation between the two variables, ($r = 0.754$, CI[0.55, 0.87], $p < 0.001$). The line-chart in Figure 6 summarizes the results, showing that as the debate in the parliament became more vivid, people also tweeted more about the topic and the other way round.

4.3.2. Recognition II: Interaction on Twitter

Figure 7 shows the interaction between people talking about the women's quota on Twitter. Women are displayed as green nodes, men as yellow nodes and MPs as red nodes. If a user A mentions user B, we draw a directed edge from A to B. Strongly connected clusters containing German MPs suggest that MPs are directly engaged with the Twitter users that were included in our analysis.
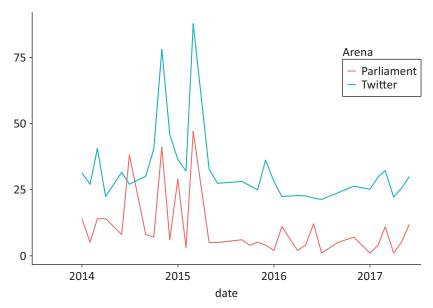
Figure 6. Recognition as count of claims in parliament and tweets about the quota.

We find that users are clustered in two large groups, ignoring small sub groups that are not connected to the largest graph. The number of male and female users is balanced. Figure 8 shows the larger group that includes several MPs with Manuela Schwesig (SPD), Heiko Maas (SPD) and Katrin Göring-Eckardt (Greens) as central points of gravity, who were heavily involved in public discussion about the women's quota.

Overall, the time correlation as well as the Twitter network-graph suggest that the parliamentary debate and the Twitter debate about the introduction of the women's quota are clearly intertwined. Having members of parliament as central cluster nodes reinforces this impression. In consequence, it seems adequate to consider Twitter as an appropriate social media platform to assess the validation of the claims by the claimed constituency.
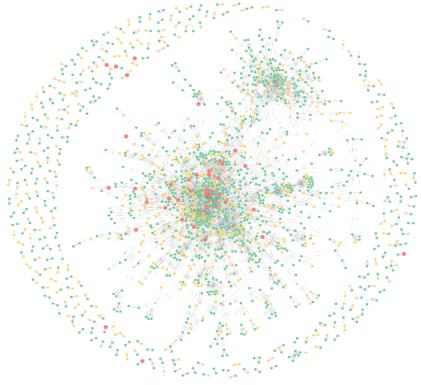

Figure 7. The dots represent Twitter users (female in yellow, male in green) and MPs (red) tweeting about the women's quota. Lines between the dots represent communication in the form of messages directed from one dot to another. The MPs appear to be central in the Twitter debate about the quota and interact with the audience.
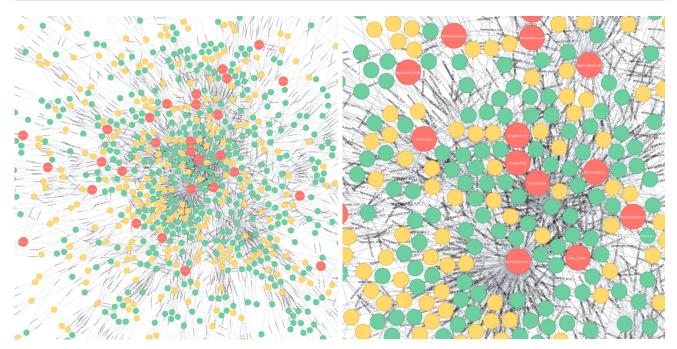
Figure 8. Central cluster of users tweeting about women's quota: MPs in red, female Twitter users in yellow, male Twitter users in green.

4.4. Validation of Parliamentary Claims on Twitter

About 70% of tweets held a neutral sentiment. Extremely positive or negative tweets were scarce. The difference between tweets created by male or female users was insignificant with Cramér's V of $\phi = 0.034$ and $\chi^2 < .001$ for tweets with a negative sentiment, $\chi^2 < .001$ for neutral sentiment tweets and $\chi^2 < .001$ for tweets with a positive sentiment. This shows that considering only the sentiment of a tweet, irrespective of the justification, is not sufficient to say that men oppose and women support the quota.

We then conducted a binary logistic regression with the Twitter users' gender as the dichotomous criterion variable (1 = female, 0 = male) and the justification the Twitter users mentioned for their stance towards the quota as predictor variable. The contrast specified was a sum contrast, which means that the intercept of the model is the Grand Mean. For the 12 justifications, 0 indicates the absence and 1 the presence of the justification in a tweet. The betas represent the change in log odds of that justification being used by a woman instead of a man (Table 3).

The logistic regression model was significant with a χ^2 of 14.8 (p = .002). The model explained 19% of variance (Nagelkerke's R^2) in the gender of Twitter users.

Significant (alpha 5%) are j3 (Quota conflicts with constitution, b = −3.26, p < 0.001) and j4 (Gender in-

Table 3. Binomial logistic regression results using gender as the criterion.

Predictor	beta	SE	Wald	p value	OR (exp)	Fit
Justification 1	−0.34	0.293	−1.149	0.250	0.714	
Justification 2	−0.65.	0.356	−1.827	0.067	0.521	
Justification 3	−3.26*	1.018	−3.199	0.001	0.038	
Justification 4	−1.05**	0.310	−3.381	0.000	0.350	
Justification 5	−1.95.	1.069	−1.820	0.068	0.142	
Justification 6	−0.13	0.517	−0.258	0.796	0.875	
Justification 7	−0.61.	0.344	−1.780	0.075	0.541	
Justification 8	−0.03	0.244	−0.122	0.902	0.970	
Justification 9	−0.33	0.363	−0.894	0.371	0.722	
Justification 10	0.22	0.223	0.333	0.739	1.250	
Justification 11	−0.45	−0.451	−0.935	0.349	0.636	
Justification 12	−0.41	−0.405	−1.088	0.276	0.666	
						Nagelkerke R^2 = 0.19

Note: . indicates p < .10; * indicates p < .05; ** indicates p < .01.

stead of qualification, b = −1.05, p < .001) and on a 10% alpha level j2 (Quota discriminates against men, b = −0.65, p = 0.067) and j7 (Quota leads to more equality, b = −0.61, p = 0.075). These justifications, therefore, are more likely to be used by men.

The results indicate that gender is not decisive in determining the general stance towards the quota. Only sentiment in combination with the justification allows for differentiation. The negative stance towards the quota in combination with the justifications 2, 3 and 4 is in line with the respective claim-cluster, which therefore can be considered accepted. Justification 7, which is a positive justification, in combination with negative sentiment, is also significantly more likely to be used by men. This might seem counterintuitive, however, could be interpreted as sarcasm. This is very common in social media but difficult to assess. However, is should be taken into account that this analysis is just illustrative, aiming at outlining how such an analysis could be conducted.

5. Summary and Conclusion

Relying on the theory of Saward (2010) and Disch (2015), we studied political representation through the lens of representative claim-making. Our aim was to contribute to the literature by first clarifying empirically which constituencies were being constructed in the parliamentary debates on the introduction of the women's quota in German supervisory boards (2013–2017) and which justifications were used to explain the respective stance towards the quota. By those means, we found three distinct claim-clusters. Then, we assessed if and how the claims made in the parliament were received and validated (accepted or rejected) by the audience on Twitter.

We ensured that it is adequate to consider Twitter users as audience for the claims made in the parliament by analysing the temporal course of the debate between October 2013–October 2017 in both arenas, parliament and Twitter, as well as the interconnection of members of the parliament and audience on Twitter with a social network analysis. We found that when the quota was discussed in the parliament, the number of tweets about the women's quota increased and vice versa. Furthermore, the German members of parliament and Twitter users of the general public frequently interacted with each other on Twitter when talking about the women's quota. On these grounds we could assume an audience on Twitter which received the claims.

We then analysed all of the German tweets on the quota. We conducted a sentiment analysis to calculate the stance towards the quota and deduced the respective justifications. Finally, the relationship between these justifications, the user's gender and the sentiment expressed towards the women's quota were examined using logistic regression analysis.

The results show that gender alone is not decisive in determining the general stance towards the quota. When it comes to opposing the quota, men are signif-

icantly more likely to use the justifications "Quota discriminates against men", "Quota conflicts with constitution" and "Gender instead of qualification as the hiring criteria" than women. This corresponds to one of the three claim-clusters, characterized by a negative stance towards the introduction of the women's quota, referring to men as their constituency and making use of the same justifications.

At first sight, this conveys the impression that the respective claims are accepted. Nonetheless, we found that in plain numbers, men do support the introduction of the quota more often than they oppose it. However, 37% of all men express negative sentiments towards the quota, compared to 23% of women. In the end, which sort of threshold ought to be applied here in order to decide whether these claims are now considered to be accepted or rejected? Is it just to say that the majority of men supports the quota, therefore the contra-claims voiced by men of the CDU are illegitimate?

The representative turn (Disch, 2015; Näsström, 2011, 2015; Urbinati & Warren, 2008) and most prominently Saward (2010) refrained from bounding democratic legitimacy to the traditional models of electoral accountability. Instead they put forward the idea of a non-institutionalized, more flexible view on representation by looking at representative claims, basing legitimacy on their validation. Saward states that the constituency is the ultimate judge of legitimacy and without even refuting this position, we empirically reached the point where we have to admit that this is not a sufficient directive. It becomes evident that this proposal has not been comprehensively conceptualized. A crucial gap in the concept remains to be resolved: claims cannot be validated without institutionalized agreements based on a societal consensus. We need to define a threshold for acceptance and by that ask: how much acceptance is enough? Is validation to be treated equivalently to a majority voting and therefore the CDU claims in this debate must be considered rejected? Or are CDU men legitimately speaking for those 40% of men (which equals about 20% of the population) who oppose the quota for the same reasons as CDU claims and therefore they, as a minority, are empowered to validate the CDU to represent them? It becomes clear that even if we accept the constituency (audience) as normative authority, common evaluation criteria we are building our framework on need to be negotiated. This in turn requires a normative assessment beyond that of the affected constituency and therefore beyond Saward's conceptualization of reception by constituency.

Conflict of Interests

The authors declare no conflict of interests.

References

Barberá, P., & Rivero, G. (2015). Understanding the political representativeness of Twitter users. *Social Science Computer Review*, *33*(6), 712–729.

De Wilde, P. (2013). Representative claims analysis: Theory meets method. *Journal of European Public Policy*, *20*(2), 278–294.

Di Franco, G. (2016). Multiple correspondence analysis: One only or several techniques? *Quality & Quantity*, *50*(3), 1299–1315.

Disch, L. (2011). Toward a mobilization conception of democratic representation. *American Political Science Review*, *105*(1), 100–114.

Disch, L. (2015). The "constructivist turn" in democratic representation: A normative dead-end? *Constellations*, *22*(4), 487–499.

Espírito-Santo, A., Freire, A., & Serra-Silva, S. (2018). Does women's descriptive representation matter for policy preferences? The role of political parties. *Party Politics*. https://doi.org/10.1177/1354068818764011

Greenacre, M. J. (2017). *Correspondence analysis in practice* (3rd ed.). Boca Raton, FL: CRC Press.

Guasti, P., & Geissel, B. (2019). Saward's concept of the representative claim revisited: An empirical perspective. *Politics and Governance*, *7*(3), 98–==101==.

Hassanpour, N. (2013). Tracking the semantics of politics: A case for online data research in political science. *Political Science & Politics*, *46*(2), 299–306.

Husson, F., Lê, S., & Pagès, J. (2017). *Exploratory multivariate analysis by example using R* (2nd ed.). Boca Raton, FL: CRC Press.

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the eighth international conference on weblogs and social media* (pp. 216–225). Ann Arbor, MI: ICWSM.

Jungherr, A., Schoen, H., & Jürgens, P. (2016). The mediation of politics through Twitter: An analysis of messages posted during the campaign for the German federal election 2013. *Journal of Computer-Mediated Communication*, *21*(1), 50–68.

King, M. F., & Bruner, G. C. (2000). Social desirability bias: A neglected aspect of validity testing. *Psychology and Marketing*, *17*(2), 79–103.

Koopmans, R., & Statham, P. (1999). Political claims analysis: Integrating protest event and political discourse approaches. *Mobilization. International Journal of Research and Theory about Social Movements, Protest and Collective Behavior*, *4*(2), 203–221.

Kuyper, J. W. (2016). Systemic representation: Democracy, deliberation, and nonelectoral representatives. *American Political Science Review*, *110*(2), 308–324.

Mair, P. (2009). *Representative versus responsible government* (Working Paper No. 8). Cologne: Max Planck Institute for the Study of Societies.

Näsström, S. (2011). Where is the representative turn going? *European Journal of Political Theory*, *10*(4), 501–510.

Näsström, S. (2015). Democratic representation beyond election. *Constellations*, *22*(1), 1–12.

O'Connor, B. (2010). From tweets to polls: Linking text sentiment to public opinion time series. In *Proceedings of the fourth conference on weblogs and social media* (pp. 122–129). Washington, DC: The AAAI Press.

Pitkin, H. F. (1967). *The concept of representation* (1st ed.). Berkeley, CA: University of California Press.

Rencher, A. C., & Christensen, W. F. (2012). *Methods of multivariate analysis* (3rd ed.). Hoboken, NJ: Wiley.

Rill, S., Adolph, S., Drescher, J., Reinel, D., Scheidt, J., Schütz, O., … Korfiatis, N. (2012). A phrase-based opinion list for the German language. In J. Jancsary (Ed.), *Proceedings of KONVENS 2012* (pp. 305–313). Vienna: ÖGAI.

Saward, M. (2006). The representative claim. *Contemporary Political Theory*, *5*(3), 297–318.

Saward, M. (2010). *The representative claim* (1st ed.). Oxford: Oxford University Press.

Statistisches Bundesamt. (2017). *Statistisches Jahrbuch Deutschland 2017* [Annual Abstract of Statistics]. Wiesbaden: Statistisches Bundesamt.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics*, *37*(2), 267–307.

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with Twitter: What 140 characters reveal about political sentiment. In *Proceedings of the fourth conference on weblogs and social media* (pp. 178–185). Washington, DC: The AAAI Press.

Urbinati, N., & Warren, M. E. (2008). The concept of representation in contemporary democratic theory. *Annual Review of Political Science*, *11*(1), 387–412.

Van Biezen, I. (2014). The end of party democracy as we know it? A tribute to Peter Mair. *Irish Political Studies*, *29*(2), 177–193.

Van Biezen, I., Mair, P., & Poguntke, T. (2012). Going, going,…gone? The decline of party membership in contemporary Europe. *European Journal of Political Research*, *51*(1), 24–56.

Young, I. M. (2002). *Inclusion and democracy* (1st ed.). Oxford: Oxford University Press.

About the Authors

Viola Joschko studied Psychology and Political Science at the University of Frankfurt. Her research interests involve applied statistics and the use of novel methodologies in social sciences. She was a research assistant in the DFG/ANR-funded project "(New) Political Representative Claims: A Global View (France, Germany, Brazil, China, India)", aiming to investigate and understand the variety of representative claims emerging around the world.

Luis Glaser is a Graduate Student in Cognitive Systems at Potsdam University and studied Computer Science and Political Science at the Goethe University Frankfurt. He is a Student Assistant at the Applied Computational Linguistics Lab at the Goethe University Frankfurt. His current main interests are NLP, reinforcement learning and discourse relations.