Article

Populist ideas on social media: A dictionary-based measurement of populist communication

Johann Gründl [ID]
University of Vienna, Austria

Abstract

This article presents a dictionary-based measurement of populist communication that reaches citizens directly through social media in German. The studied populist messages reflect ideational definitions of populism. Thus, populist messages appeal to the people, dismiss the elites as appalling, or highlight the people's right to unfettered rule. Despite German-speaking countries offering a variety of populist parties, existing automated approaches are rarely applicable to German texts. Furthermore, they are not tailored to social media and often focus only on anti-elitism. I vastly improved existing dictionaries by analyzing populist ideology in its entirety, including people-centrism and the demand for people's sovereignty. The article showed how known populist parties in Austria, Germany, and Switzerland spread these messages to a great extent on social media. Furthermore, I was able to highlight intra-party differences. Finally, the article discusses different aspects of validity and shows that the proposed approach offers high convergent validity and split-half reliability.

Populists are drawn to social media because they provide an ideal opportunity structure for populist communication (see Engesser et al., 2017; Krämer, 2017). On Facebook and Twitter, politicians express populist ideas that directly reach their supporters (Ernst et al.,

Corresponding author:
Johann Gründl, Department of Government, University of Vienna, Kolingasse 14–16, 1010 Vienna, Austria.
Email: johann.gruendl@univie.ac.at

2017a; Groshek and Koc-Michalska, 2017). Following ideational definitions, populism revolves around the ideas that society is "ultimately separated into two homogeneous and antagonistic groups, 'the pure people' versus 'the corrupt elite'" and that "politics should be an expression of the volonté générale (general will) of the people" (Mudde, 2004: 543).

By directly promoting these populist ideas to their supporters on social media, politicians circumvent traditional media outlets and present themselves as close to the people (Jungherr et al., 2019; Krämer, 2017). Social media's tendency toward homophily allows populists and their supporters (i.e. "the people") to strengthen their ties while shielding themselves from divergent opinions (Engesser et al., 2017; Groshek and Koc-Michalska, 2017). Furthermore, social media's algorithms favor controversial content. Such content then often also provokes traditional news media to spread populist messages (Jungherr et al., 2019).

Thus, social media are of utmost relevance when it comes to studying populist communication. However, reliably measuring populist ideas on social media remains challenging. Political actors produce a great wealth of social media content every day. To understand populist social media communication's importance and its potential effects on citizens, we need to analyze large amounts of data; thus, manual coding often becomes unfeasible. Therefore, in this article, I develop and validate an automated dictionary-based content analysis approach to measure populism on social media.

Existing dictionaries provide a starting point but aim at manifestos (Pauwels, 2017; Rooduijn and Pauwels, 2011), party member magazines (Pauwels, 2011), or speeches (Bonikowski and Gidron, 2016a). In contrast to these approaches, I intend to capture populist messages that reach citizens unfiltered and directly on social media. What is more, these attempts to measure populism leave ample opportunity for improvements:

First, available dictionaries lack conceptual scope (see Online Appendix A for terms included in existing dictionaries). They mostly rely on one populist message: *the elites are corrupt and blameworthy*. This message reflects only one side of the coin (Mudde, 2004): The dictionaries neglect the antagonists of the elites, *the good people*, and the populist call for *absolute adherence to the people's will*.

Second, by relying on a few words only, the dictionaries at hand lack reliability. They fail to capture much populist messaging that uses other words. Identifying populism in small amounts of text (e.g. for analyzing short periods or change over time) could become overly dependent on the presence of a few words.

Third, despite the long-standing presence of populist parties in German-speaking countries, most dictionaries are only available in English. If German dictionaries are provided, they lack successful validation (Pauwels, 2017; Rooduijn and Pauwels, 2011). A dictionary tailored to the German-speaking context allows the study of a variety of populist parties. In Austria and Switzerland, established populist right-wing parties exist, the Austrian Freedom Party (*Freiheitliche Partei Österreichs* [FPÖ]) and the Swiss People's Party (*Schweizerische Volkspartei* [SVP]). In Germany, with the Alternative for Germany (AfD), a new party appeals to a populist right-wing electorate (Siri, 2018). The Left (*Die Linke*) is often considered a left-wing populist party (e.g. Rooduijn et al., 2019) or at least a borderline case of populism (Fawzi et al., 2017). However, compared to English, developing dictionaries for languages with more inflections and long compound

words not separated by spaces, such as German, is more complicated. The presented approach should, thus, also provide insights for endeavors to develop dictionaries in similar or more complex languages.

Overall, the article's main contribution lies in developing and validating a proper measure for populism in social media content. In addition, the prevalence of populism in politicians' and parties' communication on Facebook and Twitter in Austria, Germany, and Switzerland is explored. I validated the dictionary against other measures of populism and assessed its split-half reliability. The article shows that the dictionary is a vast improvement over previous measures: It offers more conceptual breadth, better convergent validity, and improved split-half reliability. Descriptive results highlighted that the right-wing populist parties (SVP, FPÖ, and AfD) were also the ones most frequently spreading populist ideas on social media; results for The Left were more ambiguous. In general, populist ideas were more prevalent on Facebook than Twitter.

The dictionary presented in this article is provided as an R package (Gründl, 2020), which allows its easy application to German-language communication by politicians and parties that addresses citizens directly. Functions to assist similar undertakings are also provided as R packages (*multidictR, regexhelpeR*). The dictionary enables researchers to study the role of populist communication on social media and its consequences for citizen's attitudes and political behavior.

Populism in communication

Populist communication was defined in various ways (for detailed discussions, see, for example, Engesser et al., 2017; Reinemann et al., 2017; Wirth et al., 2016). There is work that approaches populism as a set of *ideas* or an *ideology* (Mudde, 2004). Other popular definitions characterize populism as a *style* (Moffitt, 2016) or a political *strategy* (Weyland, 2001). These approaches provide different perspectives on populist communication (Engesser et al., 2017): A focus on strategy contributes to our understanding of *why* politicians employ populist communication; regarding populism as a style focuses research on *how* populists communicate. Research following the ideational approach, as in this article, looks at *what* is said.

In recent years, much empirical research adhered to the ideational approach, which posits that populism is best defined as a set of specific ideas. Mudde's (2004) widely cited definition describes populism as a thin ideology. Others disagree with this conceptualization as an ideology (for a detailed critique, see Aslanidis, 2016) or use different terminology when dealing with populist communication. However—be it a "discourse" (Hawkins and Rovira Kaltwasser, 2018), "discursive frame" (Aslanidis, 2016), the "content of populist communication" (Engesser et al., 2017), or "populist messages" (Reinemann et al., 2017)—most definitions of populist communication rooted in the ideational approach share these three core notions (Mudde, 2004; Wirth et al., 2016: 2):

First, populist communication appeals to the good people (*people centrism*). The people share common wisdom, preferences, and worries. Second, it argues that politics should be an unhindered expression of the people's general will (*people's sovereignty*). Third, it depicts an ill-disposed, corrupted elite in power who betrays, exploits, and patronizes the people. The elite neglects the people's interests and deprives them of their

right to rule (*anti-elitism*). This tension between the decent people and the corrupt elites has a moral or Manichean quality.

In this view, the monist conception of the people does not allow for a plurality of voices. In the populist vision of democracy, a politician's duty is to recognize and fulfill the people's true will. Populists, thus, claim to be the ones genuinely representing the people. By demanding a voice for "the people," populists claim to be the true democrats (Canovan, 1999). Indeed, populism could be a reaction to failings in representative politics because it gives a voice to groups previously not represented by the political elites (Rovira Kaltwasser, 2012). However, it also represents a threat to liberal ideas of democracy because the populist vision of democracy despises compromise and power-sharing mechanisms. Instead, populism favors purely majoritarian democracy: Nothing should limit the unhindered expression of the *volonté générale* (Canovan, 1999; Mudde, 2004).

In this article, communication is considered populist if it conveys the populist ideas mentioned. Since populist communication then is about what is said and not about who said it, populism becomes a feature of the political message rather than the communicator. Broadly, such populist messages are either *advocative* toward the people or *conflictive* toward the elites (Wirth et al., 2016; Wirz, 2018).

Conflictive messages (a) attack the elites and (b) demand to limit their power. These anti-elite messages discredit the elites and detach them from regular citizens. In advocative messages, communicators express (a) people-centric ideas and (b) demand for more popular sovereignty. Such messages stress the people's virtues and achievements, highlight their monolithic character, and announce their collective will. Advocative messages might also demand that the elites should listen to the people or that the majority's power should be increased. Besides, speakers often place themselves close to the people (see Online Appendix B for examples of different populist messages). The distinction between conflictive and advocative populist messages highlights that a focus on conflictive anti-elitism (Rooduijn and Pauwels, 2011) or blaming elites (Hameleers et al., 2016) does not offer the full picture. Anti-elite messages are central to populism, but populism is also about giving a voice to the people (Rovira Kaltwasser, 2012).

Some definitions of populist communication based on the ideational approach go beyond the presented minimal definition and introduce the exclusion of certain groups (e.g. immigrants) from *the people* as an element of *full* or *thick* populism (Jagers and Walgrave, 2007; Reinemann et al., 2017). Every form of populism excludes a part of the population, the elites and their allies. However, an exclusionary stance toward, for example, immigrants is characteristic of a specific type of populist party, the populist radical right (Mudde, 2007). It is inherent to the minimal definition of populism that it usually appears connected with other host or thick ideologies such as socialism or nativism (Mudde, 2004, 2007; Rovira Kaltwasser, 2012). Nevertheless, the set of ideas mentioned above is what unites all populist parties. By tying this article's definition of populist communication to the minimal ideational definition of populism, it captures the populist essence in communication irrespective of the left-wing, right-wing, or other ideological messages it comes along with.

According to definitions of populism as a style, certain stylistic features of political messages might also indicate populist communication. Populist messages, for example, often incorporate colloquial and emotional language or incivility (e.g. Bracciale and

Martella, 2017; Ernst et al., 2019; Moffitt, 2016). However, these features are not necessarily present with each populist actor or in every populist message (Mudde, 2017). Also, communicators might not necessarily use specific communication styles (e.g. emotionalized language) as populist appeals but for other reasons (e.g. to raise attention; Bracciale and Martella, 2017; Wirth et al., 2016).

Relying on the ideational approach provides some advantages, besides being able to identify the ideational populist core (Aslanidis, 2016; Mudde, 2017). The approach provides a framework to study populism as a political actor's ideology or, as in this article, populist communication on social media (e.g. Ernst et al., 2017a). The prevalence of the same populist ideas can also be studied in media reporting (e.g. Müller et al., 2017). Recently, also measures for populist attitudes among the general public were derived from the ideational approach (e.g. Akkerman et al., 2014; Wirth et al., 2016). This versatility allows applying the proposed measurement in integrated research designs such as linkage studies where citizens' attitudes (expressed in surveys) could be related to populist communication exposure. Furthermore, the ideational approach's minimal definition of populism facilitates comparative research across different contexts and cases. The approach's popularity in communication and political science makes the proposed measure connectable to existing research results and questions.

Finally, measuring populism in political communication also provides a way to identify populist actors empirically (Kriesi, 2014). In doing so, populist messages become indicators of the communicator's populist ideas. However, while the ideational approach commonly categorizes political actors as non-populist or populist (i.e. they share all the core ideas), populist appeals in communication are not restricted to populists (Jagers and Walgrave, 2007; Wirth et al., 2016). Populist messages are used across the whole political spectrum (Rooduijn et al., 2014). Nevertheless, populists should rely on populist messages more often and consistently than other actors. Empirically, such a measure of populism becomes a matter of degrees (Mudde, 2017). To identify an actor as populist or not based on the prevalence of populist ideas in its political communication, researchers have to choose cut-off points (i.e. which amount of populist ideas is enough to classify someone as populist).

Measuring populist ideas in text

Jagers and Walgrave (2007) were among the first to use content analysis and manual coding to study populism. Several studies followed in their wake, using different approaches such as holistic grading (i.e. coding whole manifestos or speeches; Hawkins and Castanho Silva, 2019; Hawkins and Rovira Kaltwasser, 2018), coding paragraphs of manifestos (Rooduijn et al., 2014; Rooduijn and Pauwels, 2011), or coding statements or semantic clauses (Aslanidis, 2018; Müller et al., 2017). These approaches might provide reliable insights. However, they require extensively trained coders, are time-consuming, and, thus, often require limiting oneself to a sample of all available communication.

Computerized content analysis has the potential to overcome these limitations. However, few attempts have been made to identify populist ideas in a text through automated coding. Hawkins and Castanho Silva (2019) report mixed results with supervised machine-learning approaches. In a seminal paper, Pauwels (2011) applied a dictionary

approach to party manifestos; Rooduijn and Pauwels (2011) refined this approach further (see also Bonikowski and Gidron, 2016a, 2016b). Applying dictionaries follows a basic bag of words approach (Grimmer and Stewart, 2013). Then, the presence of certain words (e.g. "establishment") indicates populism. This approach's rationale is simple: Across large amounts of text, higher frequencies of *populist words* indicate more populist messages.

As mentioned in the introduction, the existing dictionary approaches for measuring populism are often limited in theoretical and quantitative scope: theoretically, because they mostly rely on words that indicate *anti-elitism* and quantitatively because they rely on only a few, sometimes very infrequent (e.g. "anti-grass-roots" [*anti-basisdemokratisch*]), words. The limited scope poses a few problems: The dictionaries miss out on populist messages not captured by the words included. Furthermore, the results based on such narrow dictionaries are contingent on a few words and, considering the rareness of many included words, very few occurrences. This narrowness produces shaky results. On one hand, populist communication might not be recognized as populist because similar words are used more often than those in the dictionary. On the other hand, communication that uses one of the included words a few times in a non-populist context will quickly be labeled populist. Nevertheless, the existing dictionaries provide a good starting point for developing a broader dictionary.

Although the German-speaking countries host various parties often considered populist, only two German translations of dictionaries exist. Rooduijn and Pauwels (2011) provided a German translation. Pauwels (2017) provides a second German dictionary based on his work with Rooduijn (and also considering Bonikowski and Gidron, 2016a). However, both existing dictionaries seem not fully capable of capturing populism reliably and making valid inferences on how populist text is. Therefore, the authors recommend using the dictionaries only in the first step and then manually checking the identified text passages.

Dictionary development

Dictionary approaches follow a deductive logic; terms are included in the dictionary because they are, supposedly, indicators for a theoretical concept. This logic makes dictionaries an attractive choice for measuring complex social science concepts such as populism (Grimmer and Stewart, 2013; Rauh, 2018). Useful dictionaries provide a very efficient way to reliably code large amounts of text (Grimmer and Stewart, 2013).

Two concepts describe the quality of a dictionary: recall and precision (Grimmer and Stewart, 2013; Stryker et al., 2006). For this dictionary, *recall* describes how good the dictionary is in capturing all populist texts (i.e. *avoiding false negatives*)—it should miss out on as little populist communication as possible. *Precision* describes how good the dictionary is in only capturing texts that are indeed populist (i.e. *avoiding false positives*). The two concepts conflict: If more (or more general) terms are included, the recall will increase, but at the same time, precision would suffer. If only a single word is included that is solely used for populist messages, precision will be high, but the dictionary would, at the same time, miss out on a lot of other populist messages.

I developed and improved my dictionary by (a) expanding the lists of possible terms (i.e. improving recall and theoretical breadth) and (b) selecting or modifying terms to retain high precision. The next sections describe both steps in detail. In developing a dictionary, the two steps are repeated multiple times and often go together. The second step is particularly important because expanding the dictionary involves the risk of sacrificing too much precision. Consequently, the approach would label too many statements populist. Online Appendix C displays the result of these endeavors, a German dictionary for populist communication on social media.

Social media data

To test, develop, and validate the dictionary presented in this article, I relied on Facebook and Twitter data from Austria, Germany, and Switzerland. Facebook is the most widely used social media platform in all three countries. Twitter usage is less common among the general public (Newman et al., 2019); however, its frequent use by politicians and pundits makes it an essential part of the political social media sphere (Ausserhofer and Maireder, 2013).[1] The dataset includes accounts by parties as well as parties' leaders, parliamentary leaders (in Austria and Germany), the highest government officials, and members of the Federal Council in Switzerland. However, in Austria and Germany, party leaders often fulfill multiple roles simultaneously (e.g. leading the party and the parliamentary group). Only content published while political actors were active in leading roles was included (see Online Appendix D for details on used data).

Facebook and Twitter data were accessed through these platforms' application programming interfaces with the self-developed R package *socmedhelpeRs* (Gründl, 2019). Data collection started in June 2017. For Facebook, all available data on public pages at this point were downloaded. Twitter allowed downloading the latest 3200 messages per account. From 2017 on, the data were regularly updated. The analysis in the next section relies on data from 2014 to February 2020.

Broadening the pool of dictionary terms

Building on existing German and English dictionaries, I improved the recall of the dictionary. The aim was to develop a dictionary that is as encompassing as possible (Rauh, 2018). I translated words from available English dictionaries. The semantic network BabelNet (Navigli and Ponzetto, 2012) assisted in identifying related concepts. Furthermore, for all terms, synonyms were looked up automatically through the freely available German-language thesaurus OpenThesaurus.de (https://www.openthesaurus.de).

Moreover, I added terms out of theoretical considerations to also capture advocative messages which express people-centrism and demands for popular sovereignty. Depictions of advocative populist messages and references to "the people" in the literature (e.g. Ernst et al., 2018; Mudde, 2004; Wirz, 2018) served as a starting point. Having advocative populist ideas in mind while developing the dictionary increased the content validity of the measurement (Adcock and Collier, 2001) by fully capturing the concept of populism in texts.

Through looking at the surrounding words where a dictionary term was detected (i.e. looking at the *keywords-in-context*/KWIC), I identified further possibly populist terms. The term "bureaucrat" (*Bürokrat*) might, for example, bring up this sentence: "Schulz is a bureaucrat, far removed from the people—an EU apparatchik." Out of theoretical considerations, I would also add the terms "removed from the people" (*bürgerfern*) and "apparatchik" (*Apparatschik*) as well as their synonyms to the list of possibly populist terms.

Newly added dictionary terms were not confined to single words. It is often useful to include multi-word expressions (e.g. at the "expense of the majority") instead of just one word ("majority" or "expense"). Furthermore, regular expressions were introduced to allow more flexible terms. Like the better known glob-style wildcards (e.g. * or ?), specific patterns act as placeholders for different expressions. For example, "expense (.*) majority" would capture "expense of the poor majority" as well as "expense of a majority." Regular expressions have the advantage of offering more control compared to wildcards; "(crooked|dishonest) politic(s|ian|ians)," for example, captures "crooked politicians" as well as "dishonest politics" (see Online Appendix E for explanations on all used regular expressions).

Using regular expressions allowed me to refrain from stemming the texts. Through stemming, inflected words are reduced to their word stem. Stemming often eases working with texts because one does not have to account for various possible inflections. However, I encountered difficulties regarding some German words. For example, the words *Bürger* (citizen), *bürgen* (to vouch), and *Burg* (castle) would be indistinguishable in their stemmed form *burg*.

Furthermore, keeping the texts in their original form also allows differentiating between plurals and singulars, which might sometimes be useful. For example, *Menschen* (people, humans) is different from *Mensch* (person, human). Therefore, I opted to retain as much control and precision as possible by using regular expressions for capturing different grammatical variations.

Finally, I created a vast amount of new patterns by combining terms that might refer to elites (e.g. elites, politicians) with according adjectives or actions (e.g. dishonest, have betrayed) and terms that might refer to the people (e.g. people, citizens) with the respective descriptions (e.g. ordinary, have had enough). With all these measures, the list of possibly populist terms now includes 12,943 terms. The provided R package also includes this full list of words.

Selecting the relevant and precise terms

In the next step, I kept only relevant and precise terms. Most of the created terms, especially the more complex combinations, rarely or never appear in the available social media corpus. Thus, I could safely exclude them. Next, I strengthened the dictionary's precision by eliminating terms that lack discriminatory power (i.e. are not precise enough).

Terms were validated by looking at the posts and tweets identified by these terms. I excluded words that, more often than not, did not indicate populist messages. For example, the term "the system" (*das System*) did very often not indicate opposition to the

current political system or establishment. Instead, politicians across the ideological spectrum used it to discuss specific policies such as systems for retirement, the health system, or systems for financing public service media. Overall, I kept terms used for populist messages roughly at least 8 out of 10 times.

Multi-word terms are another meaningful way to increase precision. These patterns allowed the inclusion of precise variations (e.g. "so-called experts") of otherwise too general terms (e.g. "experts"). Thus, multi-word dictionary terms help increase recall while still retaining precision (or in regaining precision, without sacrificing much recall).

Further steps helped in evaluating dictionary terms. Terms that are more often used by parties generally not considered populist (e.g. Rooduijn et al., 2019) might not precisely capture populist messages—though they still might. Furthermore, if single terms strongly drive the results of particular parties, this could mean that they are frequently used for specific purposes (e.g. in a policy proposal) and not necessarily indicate populism. Furthermore, based on comparison to expert ratings (see Results section), I computed the correlation between the usage of single dictionary terms by parties with existing measures of how populist these parties supposedly are. These scores also allow an evaluation of individual terms. However, these inductive approaches only highlighted terms to look at more closely; the final decision to include terms was always guided by theoretical considerations.

After selecting the relevant and precise terms, 238 terms remained as indicators for advocative or conflictive populist messages (see Online Appendix C). The full list of terms is still provided with the R package and could serve as a starting point for the application to different contexts (e.g. media reports).

Applying the dictionary

The social media texts from Twitter and Facebook were combined to form a large corpus. I used *R* (version 3.6.3) and *quanteda*[2] (version 1.5.2; Benoit et al., 2018) to apply the dictionary. To facilitate the application of the dictionary, I developed the R package *popdictR* (Gründl, 2020), which is freely available on GitHub. The package allows applying the dictionary to new texts quickly. Apart from functions to work with the dictionary, it also includes the dictionary itself, the full list of possible terms, and information collected during dictionary development such as term frequency or categorizations.

The search patterns included in the dictionary are applied at the sentence level. Otherwise, regular expressions such as ".*" (i.e. any number of any characters) might match a whole paragraph or document. The resulting scores are binary measures that indicate if a sentence contained a populist term.

Aggregated populism scores per party (see Table 1) are percentages of sentences containing a term that indicates populism. For easier comparison and to account for country-level differences in language usage and expert ratings, all measures reported in Figures 1 and 2 were standardized within countries. An alternative operationalization reports the proportion of posts and tweets that include at least one dictionary term. However, this measure does not take differences in the length of posts into account. For example, the German AfD publishes very long texts, making it more likely that a dictionary term (i.e. a populist message) is included in the post.

Table 1. Populist messages by parties on social media (2014–February 2020).

Party	PopuList	Percentage of sentences			Percentage of posts/tweets		
		Facebook	Twitter	Combined	Facebook	Twitter	Combined
Austria							
FPÖ	1	2.31	1.86	2.23	6.42	3.32	5.54
NEOS	0	1.32	0.74	0.93	4.68	1.50	2.12
PILZ	0	0.78	0.90	0.86	2.55	1.99	2.14
SPÖ	0	0.70	0.70	0.70	1.97	1.53	1.68
Grüne	0	0.64	0.52	0.55	1.96	1.10	1.27
ÖVP	0	0.46	0.36	0.40	1.16	0.73	0.86
Germany							
AfD	1	5.05	2.68	4.42	33.98	7.31	17.91
Linke	1	1.92	1.22	1.54	8.68	3.05	4.72
Piraten	0	1.42	0.67	0.77	3.75	1.42	1.64
CSU	0	0.93	0.59	0.75	2.36	1.51	1.91
FDP	0	0.82	0.53	0.65	2.71	1.33	1.82
Grüne	0	0.72	0.46	0.58	2.46	1.11	1.62
SPD	0	0.86	0.41	0.55	3.36	0.98	1.45
CDU	0	0.47	0.27	0.33	1.49	0.65	0.85
Switzerland							
SVP	1	3.88	3.34	3.70	10.01	6.83	8.75
EDU	0	1.75	1.32	1.54	4.62	2.36	3.12
SP	0	1.30	1.05	1.14	3.74	2.09	2.54
EVP	0	1.53	0.78	1.04	2.83	1.36	1.81
BDP	0	0.85	1.09	1.03	1.62	1.75	1.72
Grünliberale	0	1.09	0.99	1.02	3.35	2.21	2.45
CVP	0	1.04	1.00	1.01	2.96	2.09	2.25
Grüne	0	0.60	0.65	0.64	1.55	1.48	1.50
FDP	0	0.57	0.51	0.53	1.57	1.10	1.23

Results and validation

Each of the three included countries is home to a populist radical right party. The SVP has become the largest party in Switzerland. Despite taking governmental responsibilities, it preserved its populist rhetoric (Ernst et al., 2017b). Austria's FPÖ is another "prototypical" populist radical right party (Mudde and Rovira Kaltwasser, 2013: 155). In late-2017, the FPÖ even joined a conservative-led government. In 2019, however, they were ousted from the government in the wake of a scandal (Eberl et al., 2020). In Germany, the AfD was founded in 2013. While less evident in its beginnings, it quickly developed a populist radical right stance (Fawzi et al., 2017; Siri, 2018). All three parties are also regularly classified as populist, for example, in the PopuList project (Rooduijn et al., 2019).

These three parties are also the most populist parties on social media in their countries. Table 1 shows the proportion of populist sentences for all Facebook posts and tweets published by the parties and the respective leading party officials. It also shows
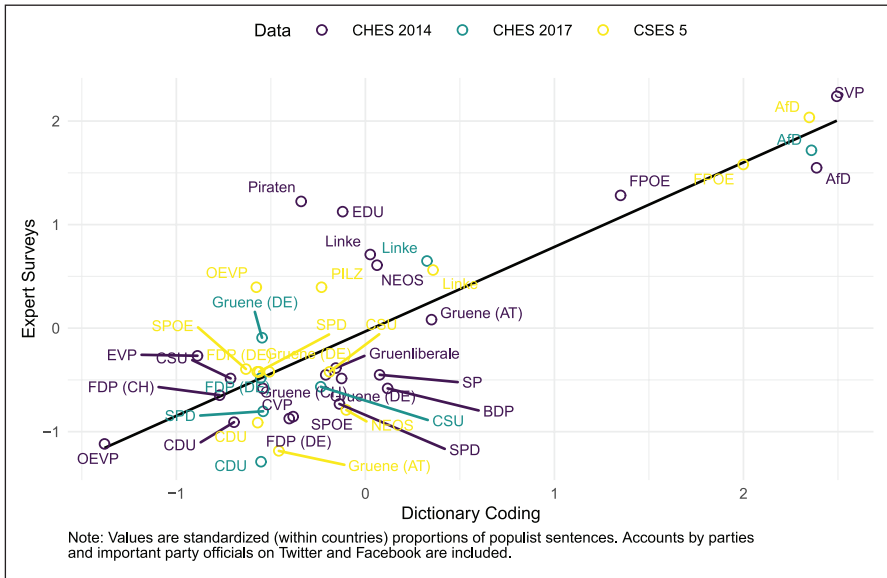
Figure 1. Populism on social media according to the dictionary compared to expert ratings.
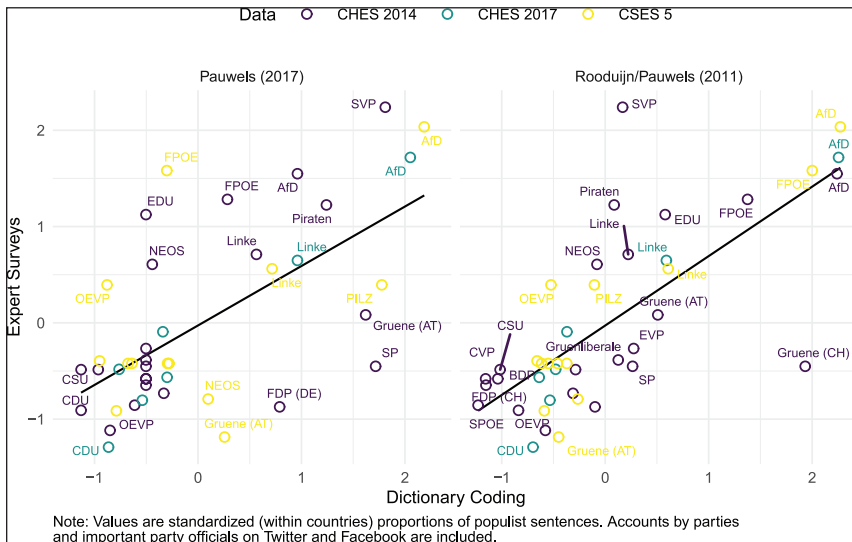

Figure 2. Populism on social media according to Rooduijn & Pauwels (2001) and Pauwels (2017) compared to expert ratings.

the proportion of posts or tweets with at least one populist term. Across the whole corpus spanning from 2014 to February 2020, these parties communicated populist ideas most frequently on Facebook and Twitter.

According to the dictionary, 3.7% of all sentences published by the SVP, the most populist party in Switzerland, included at least one populist term. This result is in line with previous research on populist social media communication by parties that identified the SVP as the Swiss party relying most heavily on populist communication strategies (Ernst et al., 2017b, 2018). In Austria, the FPÖ had the highest proportion of populist sentences (2.2%). For the AfD, the dictionary identified 4.4% of all their sentences as populist.

Results were similar for the proportion of posts or tweets. The right-wing populist parties were again the most populist parties on social media; 10% of the Swiss SVP's posts on Facebook included a populist term; for the Austrian FPÖ, it was 6.4% of all Facebook posts. The proportion of populist posts is exceptionally high for the AfD. Roughly a third of the Facebook posts by the AfD and its officials classified as populist. Their proponents often published lengthy posts on Facebook that included at least one populist term. This higher-than-average length of the AfD's posts made it also more likely to detect a populist message. Weighing the populist sentences by the total number of sentences takes this into account.

In general, populist messages were more prevalent on Facebook than on Twitter across all parties and countries. This result ties nicely with previous findings (Ernst et al., 2017a). Facebook offers unlimited space to make populist arguments, while Twitter limits space and is geared more toward journalists and pundits who use it for exchange and research (Ausserhofer and Maireder, 2013). Thus, politicians in the three countries appeared to prefer Facebook over Twitter to spread populist ideas.

In Germany, the party with the second-highest score was The Left (*Die Linke*; 1.5%). It is often considered a populist party (e.g. Rooduijn et al., 2019). However, others raised doubts about this verdict (Fawzi et al., 2017). Based on the prevalence of populism on their social media accounts, this article's results favor the treatment of The Left as a borderline case rather than a fully populist party. The Left was more populist than other parties; however, they communicated not nearly as much populist messages as the right-wing AfD.

One reason for different perceptions of The Left might lie in intra-party variation (see Online Appendix G for details). Sahra Wagenknecht, one of the leading candidates in 2017 and a prominent proponent of the more left-wing factions within the party, used way more populist messages than all other actors; 4.4% of all sentences on her Facebook page contained populist messages. On the Facebook pages of the party (0.6%) and its other officials (1.3–1.5%), populism was a lot less prevalent. Thus, a very prominent figure such as Sahra Wagenknecht might drive perceptions of The Left as populist. In contrast, the party as a whole was less populist than, for example, the AfD. While there is also variation within the AfD, all studied AfD accounts showed high proportions (at least 3.9%) of populist sentences.

Comparison to expert surveys

The proposed dictionary improved the validity of previous measures by including all relevant aspects of populist communication. Concept validation (Adcock and Collier, 2001) happened through careful examination of texts identified as populist. That way, I also ascertained that indeed different types of populist messages were spotted, and other,

non-populist, types of statements were not. Furthermore, I also looked for congruence with other measures of populism. A dichotomous categorization of parties was already introduced in Table 1. The parties considered populist on the PopuList (Rooduijn et al., 2019) turned out to also use more populist messages on Facebook and Twitter.

The Chapel Hill Expert Survey (CHES) and the Comparative Study of Electoral Systems (CSES) include ratings on parties' degree of populism based on expert surveys. The CHES 2017 includes anti-elitism and a measure on where parties stand on the people/elite divide, but only for German parties. The 2014 CHES includes information on the salience of anti-elitism for all three countries (Polk et al., 2017). The first release of the CSES Module 5 (The CSES, 2019) includes expert ratings on parties' populist stance for the Austrian and German national elections in 2017.

For this analysis, I compared the proportion of populism (at the sentence level) on Facebook and Twitter with how different expert ratings evaluated these parties' populist ideology. Therefore, I analyzed social media data from the periods the corresponding surveys aimed to capture. For the CHES 2014, the analysis considered all Facebook posts and tweets from 2014; for the CHES 2017, it included all texts from 2017. The CSES ratings aim to cover the election campaigns. Accordingly, I included posts and tweets published from the elections' announcement up to the day before the election (May 15–October 14, 2017, in Austria, and January 26–September 24, 2017, in Germany).

Overall, Figure 1 shows that the dictionary picked up on populist ideas in a way that correlated with experts' assessments of how populist or at least anti-elite these parties were. The correlation between all available expert ratings and the dictionary results was $r = .83$. Thus, in addition to face validity and increased concept validity, considerable convergent validity for the proposed approach was established. Results for 2014 varied a little more because less data were available for 2014. Social media played a lesser role back then. Also, data collection started in 2017; therefore, not all content from 2014 was still accessible. However, even given these circumstances, the dictionary approach identified the three populist radical right parties as the most populist party on social media in each country.

Compared to pre-existing dictionaries, my dictionary performed better than the suggestions made by Pauwels (2017) or Rooduijn and Pauwels (2011). According to Pauwels' (2017) dictionary, the FPÖ, for example, was not the most populist party in Austria overall (see Online Appendix F for descriptives from these dictionaries). It also failed to identify the FPÖ as populist in 2014 and during the election campaign in 2017 (see Figure 2). The Swiss SVP had about the same score as the Social Democrats (SP). Rooduijn and Pauwels' (2011) dictionary appeared to do a better job than Pauwels'. However, it also did not undoubtedly identify the Swiss SVP as the most populist party in Switzerland and failed to capture its populism in 2014. At the same time, it attributed the highest score to the Swiss Greens. Altogether, convergent validity regarding the included expert surveys was considerably lower with both pre-existing dictionaries (for Pauwels, $r = .63$; for Rooduijn & Pauwels, $r = .74$; compared to $r = .83$ for the dictionary proposed in this article).

Split-half reliability

Finally, similar to Rooduijn and Pauwels (2011), I assessed my dictionary's split-half reliability. Split-half reliability helps assess the internal coherence of the dictionary (i.e.
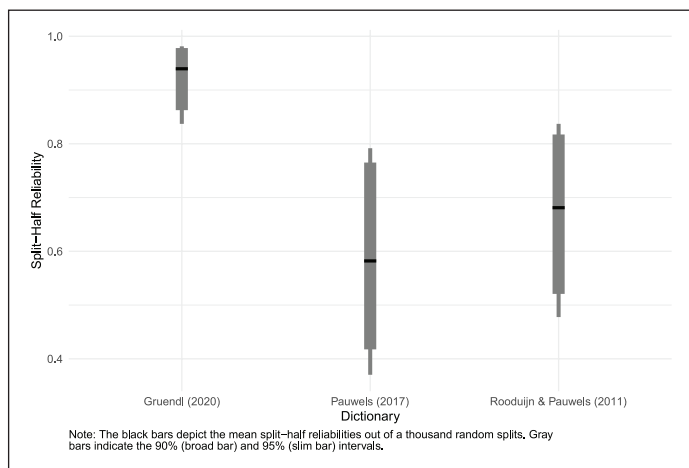
Figure 3. Split-half reliabilities of my dictionary and existing dictionaries by Rooduijn and Pauwels (2011) and Pauwels (2017).

are the included terms indicators for the same theoretical concept). It also provides a means to assert that the dictionary does not overly rely on only a few frequent terms. I split the dictionary into two parts containing a random half of all dictionary terms. Similar to the results in Table 1, a score for each half of the dictionary was computed for every party. Based on the correlation between the party results obtained by the two halves of the dictionary, I computed the split-half reliability using the Spearman–Brown formula.

Moreover, I did not rely on a single random split (cf. Rooduijn and Pauwels, 2011). Instead, I generated a thousand random splits and computed the according split-half reliability. This step allowed me to also make conclusions about the certainty of the split-half reliability estimates. The dictionary achieved an average split-half reliability score of $r = .94$, which is highly satisfactory (above .8). Even the worst random split indicated satisfactory reliability ($r_{min} = .73$).

Figure 3 depicts the split-half reliability and compares it to the results achieved by previously suggested dictionaries on the same sample. It also depicts intervals for 90% and 95% of the obtained reliability scores. Of all thousand random splits, 95% (or 90%) of all values lie within these intervals. Compared to reliability scores reported in Rooduijn and Pauwels (2011) and the results from the two existing German dictionaries on the same data (Pauwels, 2017; Rooduijn and Pauwels, 2011), the split-half test highlights the improvement the proposed dictionary offers over previous suggestions. Rooduijn and Pauwels' (2011) dictionary only had a score of $r = .68$ ($r_{min} = .39$), Pauwels' (2017) dictionary performed worse ($r = .58$; $r_{min} = .13$).

Discussion

In this article, I developed a dictionary-based measurement for the prevalence of populist messages on German-language social media accounts by political actors. Previously,

such a dictionary was not available to researchers. Existing approaches (Bonikowski and Gidron, 2016a, 2016b; Pauwels, 2011, 2017; Rooduijn and Pauwels, 2011) aimed at different types of text or speech, and where German versions exist, they have not been validated successfully. The presented dictionary is released in the R package *popdictR* (Gründl, 2020), which allows applying and adapting it to different texts quickly.

The dictionary was developed with high concept validity in mind. Thus, it includes terms to reflect the theoretical concept of populism fully. It captures conflictive (i.e. against the elites) and advocative (i.e. for the people) populist messages and overcomes severe limitations of previous attempts, which mostly focused on anti-elite messages.

Applied to social media content from Austria, Germany, and Switzerland, the dictionary identified populist messages by different parties in all countries and across the political spectrum. For instance, the dictionary detects populist messages by the left-wing German Linke and the right-wing AfD. These results provide face validity and—considering the dictionary's aim to measure a minimal set of populist ideas rather than one of its variations (e.g. nativist populism)—confidence that the dictionary is agnostic toward thick ideologies.

Results showed that populist radical right parties were using populist messages most frequently on Twitter, and even more so on Facebook. Accounts by the German AfD, the Swiss SVP, and the Austrian FPÖ showcased the highest proportions of populist messages in each country. The AfD was found to spread populist ideas exceptionally frequently on social media.

The findings also highlighted how the left-wing party The Left in Germany should probably be best considered a borderline case of populism (Fawzi et al., 2017). While some consider The Left to be a populist party (Rooduijn et al., 2019), such a binary classification diminishes noticeable differences in the degree of populism between The Left and, for example, the right-wing AfD. The Left relied on populist messages more often than other parties; however, it did so far less than the AfD. What is more, there were remarkable differences within The Left. Sahra Wagenknecht, a prominent proponent of left-wing factions within the party, communicated more populist messages than other party officials or the official party account. For The Left, public perceptions of the party's populism might, thus, depend on whom people were listening to. For the AfD, in comparison, populist messages were very prevalent in all studied accounts.

While the case of The Left emphasized the advantages of being able to assess degrees of populism rather than just classifying it as populist or not, it remains unclear at which level of populist communication on social media parties should be considered populist. For example, a threshold stating that at least 2% of all sentences on Facebook (or a slightly lower number on Twitter) have to encompass populist ideas in order to classify a party as populist would appear to be a sensible choice. It places all parties in the populist camp, which are undisputedly populist. Furthermore, it highlights the Left's status as a borderline case (1.9%) and unambiguously labels Sahra Wagenknecht, its most populist proponent (4.5%), as populist. However, such cut-off points always represent arbitrary decisions by the researcher, which require thorough justification. Nevertheless, the proposed measurement is well-suited to provide assistance and additional validation for such classification tasks.

Convergent validity of the presented dictionary was successfully demonstrated by comparing the results with expert surveys. The dictionary also showed high split-half reliability. Thus, it provides a definite improvement over previous dictionaries available. The presented dictionary, however, also comes with limitations. Unfortunately, it is only available in German. However, the procedures described in the article and the functions provided in the R packages *popdictR* and *multidictR* should allow researchers to develop similar dictionaries for other languages—especially for languages facing similar difficulties (e.g. long compound words and inflections). Aside from that, validation with human-coded material could further increase trust in its results.

Some limitations of the proposed measure are inherent to the dictionary approach. First, dictionaries usually ignore irony; negations might also be challenging (Rauh, 2018). However, in the context of populist messages, irony and negations did not appear to cause significant problems. Second, precise dictionaries will likely underestimate the number of populist messages. I aimed at including as many words as possible to build an encompassing dictionary. However, it is impossible to capture all instances of populist ideas in text through dictionary terms, at least without sacrificing precision. Third, dictionary approaches are limited to textual representations of populist ideas. Much communication on social media also includes visual or audio-visual content. It is difficult to assess how much information is lost by not including visual content because we lack research on the prevalence of visual representations of populist ideas on social media. At least on Twitter and Facebook, visual content is usually accompanied by textual elements. However, the need for more research into visual and audio-visual content becomes apparent when turning toward platforms such as Instagram or Snapchat (Lalancette and Raynauld, 2019).

Future research might also want to explore different populist messages, such as conflictive and advocative messages. Another exciting avenue for further research would be adjusting and applying the dictionary to other types of text, such as news media. Because the dictionary has been developed for texts authored by political actors, it includes terms that capture populist messages if the speaker is a politician. However, in news articles, some of these terms (e.g. "established parties") might not be precise enough (i.e. they might often capture non-populist statements).

Overall, the dictionary developed in this article and the *popdictR* (Gründl, 2020) package for its application to new texts should facilitate studies on populism on social media. It enables researchers to study the role and possible consequences of populist messages directly addressed at citizens.

Funding

ORCID iD

Johann Gründl https://orcid.org/0000-0001-6497-8188

Supplemental material

Supplemental material for this article is available online.

Notes

1. For more details on the adoption of social media by political actors, see, for example, Quinlan et al. (2018) for Germany, Dolezal (2015) for Austria, or Rauchfleisch and Metag (2016) for Switzerland.
2. Unfortunately, *quanteda* does not support regular expressions (or wildcards) that could match multiple words. Thus, I developed the R packages *multidictR* and *regexhelpeR* to overcome these limitations. They are also available on GitHub under https://github.com/jogrue.

References

Adcock R and Collier D (2001) Measurement validity: a shared standard for qualitative and quantitative research. *The American Political Science Review* 95(3): 529–546.

Akkerman A, Mudde C and Zaslove A (2014) How populist are the people? Measuring populist attitudes in voters. *Comparative Political Studies* 47(9): 1324–1353.

Aslanidis P (2016) Is populism an ideology? A refutation and a new perspective. *Political Studies* 64(1S): 88–104.

Aslanidis P (2018) Measuring populist discourse with semantic text analysis: an application on grassroots populist mobilization. *Quality & Quantity* 52(3): 1241–1263.

Ausserhofer J and Maireder A (2013) National politics on Twitter. *Information, Communication & Society* 16(3): 291–314.

Benoit K, Watanabe K, Wang H, et al. (2018) quanteda: an R package for the quantitative analysis of textual data. *Journal of Open Source Software* 3(30): 774.

Bonikowski B and Gidron N (2016a) *Populism in Legislative Discourse: Evidence from the European Parliament, 1999–2004*. Working paper. Available at: https://scholar.harvard.edu/files/bonikowski/files/bonikowski_and_gidron_-_populist_claims-making_in_legislative_discourse.pdf (accessed 23 October 2020).

Bonikowski B and Gidron N (2016b) The populist style in American politics: presidential campaign discourse, 1952–1996. *Social Forces* 94(4): 1593–1621.

Bracciale R and Martella A (2017) Define the populist political communication style: the case of Italian political leaders on Twitter. *Information, Communication & Society* 20(9): 1310–1329.

Canovan M (1999) Trust the people! Populism and the two faces of democracy. *Political Studies* 47(1): 2–16.

Dolezal M (2015) Online campaigning by Austrian political candidates: determinants of using personal websites, Facebook, and Twitter. *Policy & Internet* 7(1): 103–119.

Eberl J-M, Huber LM and Plescia C (2020) A tale of firsts: the 2019 Austrian snap election. *West European Politics* 43(6): 1350–1363.

Engesser S, Fawzi N and Larsson AO (2017) Populist online communication: introduction to the special issue. *Information, Communication & Society* 20(9): 1279–1292.

Ernst N, Blassnig S, Engesser S, et al. (2019) Populists prefer social media over talk shows: an analysis of populist messages and stylistic elements across six countries. *Social Media + Society*. Epub ahead of print 16 January. DOI: 10.1177/2056305118823358.

Ernst N, Engesser S, Büchel F, et al. (2017a) Extreme parties and populism: an analysis of Facebook and Twitter across six countries. *Information, Communication & Society* 20(9): 1347–1364.

Ernst N, Engesser S and Esser F (2017b) Bipolar populism? The use of anti-elitism and people-centrism by Swiss parties on social media. *Swiss Political Science Review* 23(3): 253–261.

Ernst N, Esser F, Blassnig S, et al. (2018) Favorable opportunity structures for populist communication: comparing different types of politicians and issues in social media, television and the press. *The International Journal of Press/Politics* 24(2): 165–188.

Fawzi N, Obermaier M and Reinemann C (2017) Germany: is the populism laggard catching up? In: Aalberg T, Esser F, Reinemann C, et al. (eds) *Populist Political Communication in Europe*. New York: Routledge, pp. 111–126.

Grimmer J and Stewart BM (2013) Text as data: the promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis* 21(3): 267–297.

Groshek J and Koc-Michalska K (2017) Helping populism win? Social media use, filter bubbles, and support for populist presidential candidates in the 2016 US election campaign. *Information, Communication & Society* 20(9): 1389–1407.

Gründl J (2019) socmedhelpeRs. R package. Available at: https://github.com/jogrue/socmedhelpeRs (accessed 12 November 2020).

Gründl J (2020) popdictR. R package. Available at: https://github.com/jogrue/popdictR (accessed 12 November 2020).

Hameleers M, Bos L and de Vreese CH (2016) "They did it": the effects of emotionalized blame attribution in populist communication. *Communication Research* 44(6): 870–900.

Hawkins KA and Castanho Silva B (2019) Textual analysis: big data approaches. In: Hawkins KA, Carlin R, Littvay L, et al. (eds) *The Ideational Approach to Populism: Theory, Method & Analysis*. London: Routledge, pp. 27–48.

Hawkins KA and Rovira Kaltwasser C (2018) Measuring populist discourse in the United States and beyond. *Nature Human Behaviour* 2(4): 241.

Jagers J and Walgrave S (2007) Populism as political communication style: an empirical study of political parties' discourse in Belgium. *European Journal of Political Research* 46(3): 319–345.

Jungherr A, Schroeder R and Stier S (2019) Digital media and the surge of political outsiders: explaining the success of political challengers in the United States, Germany, and China. *Social Media + Society* 5(3): 2056305119875439.

Krämer B (2017) Populist online practices: the function of the Internet in right-wing populism. *Information, Communication & Society* 20(9): 1293–1309.

Kriesi H (2014) The populist challenge. *West European Politics* 37(2): 361–378.

Lalancette M and Raynauld V (2019) The power of political image: Justin Trudeau, Instagram, and celebrity politics. *American Behavioral Scientist* 63(7): 888–924.

Moffitt B (2016) *The Global Rise of Populism: Performance, Political Style, and Representation*. Stanford, CA: Stanford University Press.

Mudde C (2004) The populist zeitgeist. *Government and Opposition* 39(4): 542–563.

Mudde C (2007) *Populist Radical Right Parties in Europe*. Cambridge: Cambridge University Press.

Mudde C (2017) Populism: an ideational approach. In: Rovira Kaltwasser C, Taggart P, Ochoa Espejo P, et al. (eds) *The Oxford Handbook of Populism*. Oxford: Oxford University Press, pp. 27–47.

Mudde C and Rovira Kaltwasser C (2013) Exclusionary vs. inclusionary populism: comparing contemporary Europe and Latin America. *Government and Opposition* 48(2): 147–174.

Müller P, Schemer C, Wettstein M, et al. (2017) The polarizing impact of news coverage on populist attitudes in the public: evidence from a panel study in four European democracies. *Journal of Communication* 67(6): 968–992.

Navigli R and Ponzetto SP (2012) BabelNet: the automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence* 193: 217–250.

Newman N, Fletcher R, Kalogeropoulos A, et al. (2019) *Reuters Institute Digital News Report 2019*. Oxford: Reuters Institute for the Study of Journalism. Available at: https://ora.ox.ac.uk/objects/uuid:18c8f2eb-f616-481a-9dff-2a479b2801d0 (accessed 23 October 2020).

Pauwels T (2011) Measuring populism: a quantitative text analysis of party literature in Belgium. *Journal of Elections, Public Opinion and Parties* 21(1): 97–119.

Pauwels T (2017) Measuring populism: a review of current approaches. In: Heinisch R, Holtz-Bacha C and Mazzoleni O (eds) *Political Populism: A Handbook. International Studies on Populism 3*. Baden-Baden: Nomos, pp. 123–136.

Polk J, Rovny J, Bakker R, et al. (2017) Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 Chapel Hill Expert Survey data. *Research & Politics* 4(1): 2053168016686915.

Quinlan S, Gummer T, Roßmann J, et al. (2018) "Show me the money and the party!" Variation in Facebook and Twitter adoption by politicians. *Information, Communication & Society* 21(8): 1031–1049.

Rauchfleisch A and Metag J (2016) The special case of Switzerland: Swiss politicians on Twitter. *New Media & Society* 18(10): 2413–2431.

Rauh C (2018) Validating a sentiment dictionary for German political language: a workbench note. *Journal of Information Technology & Politics* 15(4): 319–343.

Reinemann C, Aalberg T, Esser F, et al. (2017) Populist political communication: toward a model of its causes, forms, and effects. In: Aalberg T, Esser F, Reinemann C, et al. (eds) *Populist Political Communication in Europe*. New York: Routledge, pp. 12–25.

Rooduijn M and Pauwels T (2011) Measuring populism: comparing two methods of content analysis. *West European Politics* 34(6): 1272–1283.

Rooduijn M, de Lange SL and van der Brug W (2014) A populist zeitgeist? Programmatic contagion by populist parties in Western Europe. *Party Politics* 20(4): 563–575.

Rooduijn M, van Kessel S, Froio C, et al. (2019) The PopuList: an overview of populist, far right, far left and Eurosceptic parties in Europe. Available at: http://www.popu-list.org

Rovira Kaltwasser C (2012) The ambivalence of populism: threat and corrective for democracy. *Democratization* 19(2): 184–208.

Siri J (2018) The Alternative for Germany after the 2017 election. *German Politics* 27(1): 141–145.

Stryker JE, Wray RJ, Hornik RC, et al. (2006) Validation of database search terms for content analysis: the case of cancer news coverage. *Journalism & Mass Communication Quarterly* 83(2): 413–430.

The Comparative Study of Electoral Systems (2019) CSES module 5 first advance release. *Dataset and Documentation*. Available at: https://cses.org/data-download/cses-module-5-2016-2021/

Weyland K (2001) Clarifying a contested concept: populism in the study of Latin American politics. *Comparative Politics* 34(1): 1–22.

Wirth W, Esser F, Wettstein M, et al. (2016) *The Appeal of Populist Ideas, Strategies and Styles: A Theoretical Model and Research Design for Analyzing Populist Political Communication*. 88, Working paper, May. NCCR Democracy. Available at: http://pwinfsdw.uzh.ch/publications/workingpaper/pdf/wp88.pdf (accessed 23 October 2020).

Wirz DS (2018) Persuasion through emotion? An experimental test of the emotion-eliciting nature of populist communication. *International Journal of Communication* 12. Available at: http://ijoc.org/index.php/ijoc/article/view/7846

Author biography

Johann Gründl is currently a predoctoral research associate at the Department of Government at the University of Vienna. His research interests include populism, political communication, emotions and attitudes, and quantitative research methods. He is also part of the Austrian National Election Study (AUTNES).