

*This version of the article has been accepted for publication, after peer review and is subject to Springer Nature's [AM terms of use](#), but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The version of Record is available online at: <http://dx.doi.org/doi:10.1007/s11135-017-0650-0>.*

Ce manuscrit est la version acceptée pour publication, après évaluation par les pairs, de l'article suivant : Daigneault, Pierre-Marc, Dominic Duval et Louis M. Imbeau « Supervised scaling of semi-structured interview transcripts to characterize the ideology of a social policy reform », publiée sous forme définitive dans *Quality & Quantity* à l'url suivante : <http://dx.doi.org/doi:10.1007/s11135-017-0650-0>. Cet article peut être utilisé à des fins non-commerciales en conformité avec les modalités et conditions de Springer relativement à l'utilisation des versions auto-archivées.

## Supervised scaling of semi-structured interview transcripts to characterize the ideology of a social policy reform\*

Pierre-Marc Daigneault, Dominic Duval et Louis M. Imbeau

Université Laval

**Abstract:** Automated content analysis methods treat “text as data” and can therefore analyze efficiently large qualitative databases. Yet, despite their potential, these methods are rarely used to supplement qualitative analysis in small-N designs. We address this gap by replicating the qualitative findings of a case study of a social policy reform using automated content analysis. To characterize the ideology of this reform, we reanalyze the same interview data with Wordscores, using academic publications as reference texts. As expected, the reform’s ideology is center/center-right, a result that we validate using content, convergent and discriminant strategies. The validation evidence suggests not only that the ideological positioning of the policy reform is credible, but also that Wordscores’ scope of application is greater than expected.

**Keywords:** Automated content analysis; Supervised scaling; Quantitative text analysis; Social policy; Mixed methods; Case study

\*The online version of this article (<https://doi.org/10.1007/s11135-017-0650-0>) contains supplementary material, which is available to authorized users.

### 1. Introduction

Automated content analysis methods treat text as data and, as a result, allow researchers to analyze efficiently a large volume of data (Cousins and McIntosh 2005; Grimmer and Stewart 2013; Krouwel and van Elfrinkof 2013). Wordscores is a supervised scaling method that estimates the position of “virgin texts” on an *a priori* dimension operationalized by “reference texts” with known or assumed position selected by the researcher (Laver and Benoit 2002; Laver, Benoit and Garry 2003). Researchers have used Wordscores with considerable success to estimate the position of

political actors using data derived from various sources such as political manifestos (e.g., Laver, Benoit and Garry 2003; Collette and Pétry 2014; Gauvin, Chhim and Medeiros 2016), political speeches (e.g., Imbeau 2009; Baturu and Mikhaylov 2013; Herzog and Benoit 2015), official policy and administrative documents (e.g., Charbonneau, 2009; Klüver, 2009; Benoit, Cantin and Duchesne 2013), news articles (e.g., Marzagão 2015; Montpetit 2016), and even Twitter feeds (e.g., Toff and Kim 2013). Yet, researchers have not used Wordscores to analyze interviews or other sources of qualitative data derived within qualitative case studies. Because qualitative researchers assemble a rich corpus of archives, official documents, and interview transcripts, automated methods can supplement qualitative methods such as thematic analysis in small-N research. We argue that Wordscores is particularly promising in that regard because its deductive logic fits well with a significant proportion of qualitative research in political science in which theory is central, variables are measured on an ordinal scale and cases are ranked accordingly.

We address this gap by attempting to replicate with Wordscores the qualitative findings of a published case study on a social policy reform (Daigneault 2015). We quantitatively reanalyze the same data, that is, semi-structured interviews, using excerpts from academic publications as reference texts. In addition, we rely on three strategies to assess the validity of the estimates generated by Wordscores, namely content, convergent and discriminant validation (Adcock and Collier 2001).

Validating the results of automated content analysis methods such as Wordscores is always a good practice because these methods are premised on incorrect models of language in which words are taken out of context (Grimmer and Stewart 2013; Krouwel and van Elfrinkhof 2013). In the context of this study, however, four more reasons militate in favor of validating the results. First, we estimate the position of a single case (virgin text), which reduce our ability to assess whether the placement of this case by Wordscores is meaningful and valid because we cannot compare it to other cases (virgin texts). Second, our reference texts are short (less than 2,000 words), which reduces the proportion of scored words and increases the uncertainty of the estimates produced by Wordscores (Laver, Benoit and Garry 2003; Grimmer and Stewart 2013). Third, “the technique is perhaps best applied to contexts where language use is relatively normalized (e.g., legal arguments)” (Cousins and McIntosh 2005, p. 590), which is clearly not the case in semi-structured interviews. Indeed, interviews are likely to contain more “noise” (i.e., uninformative words which distracts researchers from the “signal” or informative words regarding the dimension of interest) than other types of text. They generally contain the questions and comments of the researcher, small talk used to build rapport between the researcher and the participant, anecdotes and referrals. In addition, researchers frequently address various topics within the same interview (e.g. the ideology of a reform, its origins and its impact).<sup>1</sup> The fourth and more fundamental reason to validate the results is that our reference and virgin texts, respectively academic publications and interview transcripts, are significantly different. Whereas academic publications are formal, well-structured texts intended for an academic audience that contain jargon and scientific terms, semi-structured interview transcripts contain spoken text and are relatively informal and unstructured. Yet, because a different lexicon will translate into a low proportion of scored words and will increase the uncertainty of the estimates, Wordscores requires reference texts that use the same lexicon in the same context as virgin texts. For that reason, Laver, Benoit and Garry (2003) do not recommend using party manifestos to estimate the position of

---

<sup>1</sup> Although we can edit the interview transcripts to improve the signal/noise ratio, this requires time and resources, thereby decreasing the comparative advantage of Wordscores for analyzing interview data.

legislative speeches. Yet, we argue that the compatibility of reference and virgin texts of a different nature should not be ruled out *a priori* in Wordscores analyses. Indeed, researchers have obtained convincing results with Wordscores using formal written texts to analyze informal written texts (Toff and Kim 2013) or formal spoken texts (Weinberg 2010; Warwick 2015). Whether the results generated by Wordscores using academic publications and interview transcripts are valid is therefore an empirical question.

In this article, we first outline the case and Daigneault's findings regarding the ideological orientation of the reform. Second, we describe the research design and methods for this study. We present the empirical results regarding the position of the reform and validation results in the third section. In the last section, we conclude by discussing the validity of the inferences generated by Wordscores with respect to the ideological orientation of our case in the less-than-ideal conditions of this study.

## 2. The case and its ideological orientation

The case examined in this study is, *Building Independence* (hereafter "the reform"), an important social policy initiative launched in a Canadian province in the late 1990s. The reform, which was primarily aimed at supporting low-income working parents, comprised two income supplements, supplementary health benefits and an allowance for welfare clients enrolled in basic education and related courses.

Using qualitative methods, Daigneault (2015) characterized the ideology of this social policy reform. His study was structured around a typology of social assistance (welfare) paradigms (Daigneault 2014) which, in turn, was inspired from the work of Levitas (2005) on social exclusion. The typology contains three paradigms that Daigneault (2014) characterized in terms of ideology: *entitlement* (left), *activation* (center)<sup>2</sup> and *workfare* (right). The entitlement paradigm is premised on a social democratic ideology. A decent level of income is guaranteed as a matter of rights for citizens in order to fight poverty. The workfare paradigm is premised on neoconservatism and neoliberalism. Social assistance is considered a privilege, not a right. To fight the so-called 'culture of dependency', welfare benefits should be meager, conditional, and subject to various control measures, in particular for "undeserving" clients such as single employable persons. The activation paradigm is based on a centrist, third way ideology that emphasizes reciprocity, productivity, and equality of opportunity. Social assistance is a contract between the state and the individual aimed at reintegrating the latter into employment. To that end, the basic welfare benefit should be low but complemented by generous conditional income supplements aimed at 'making work pay', and various 'work enabling' measures (e.g., job-search, training and employability).

Daigneault's (2015) data are derived from twelve face-to-face semi-structured interviews conducted in 2012 with a purposive sample of policy actors (politicians, civil servants and activists). Respondents were selected for their intimate knowledge of the reform and for their diversity in terms of ideological and organizational perspectives. The interview script included questions on various aspects of the reform (its origins, its philosophy, the policy problem it sought to address, its objectives, the choice of policy instruments, etc.). Interviews, which lasted an hour on average,

---

<sup>2</sup> Although the ideological underpinnings of activation is to some extent a matter of debate, we argue that it epitomizes a predominantly centrist, third way ideology (see Driver and Martell 2000; Van Berkel and Møller 2002; White 2004; Huo, 2009).

were transcribed and respondents had an opportunity to check their transcripts for accuracy and/or to provide additional comments (facts, interpretations, etc.). Daigneault (2015) used qualitative thematic analysis (see Braun and Clark 2006) to analyze the interviews and generated nine crosscutting themes related to the ideological underpinnings of the reform. He then qualitatively assessed, using a three-point scale, the extent to which each of these themes is congruent with entitlement, workfare and activation. He found that the reform aligned primarily with activation, but had secondary alignments with the two other paradigms, in particular workfare. Assuming the validity of this result, we expect that the reform's position according to Wordscores to be slightly to the right of the scale (center/center-right).

### 3. Research design, methods and data

We pursue two objectives with this study. First, we attempt to replicate Daigneault's (2015) qualitative finding regarding the ideology of the reform using supervised scaling, based on the same data and conceptual framework. For that purpose, we arbitrarily define an ideological dimension ranging from  $-1$  (left) to  $+1$  (right). We assume that entitlement and workfare are respectively located at each of these extremes.<sup>3</sup> Second, we assess the validity of the estimates generated by Wordscores for this case.

#### 3.1 Reference texts

We use only two reference texts in this study because a middle reference document anchored on a single dimension can be only a mixture of both extremes that will cause a drag toward the middle of the dimension without adding anything much to the analysis (Kenneth Benoit, personal communication, March 27 2015). The two reference texts, namely *entitlement* (1,318 words<sup>4</sup>,  $-1.00$ ) and *workfare* (1,667 words,  $+1.00$ ), are concatenated documents composed of academic publications drawn from two sources. We first used the text relevant to entitlement and workfare in Daigneault's (2014), to which we respectively added Levitas' (2005, p. 9-21) characterization of the redistributive (RED or left) and moral underclass (MUD or right) discourses. Reference texts are available as supplementary material to this article (we randomized word order to protect copyrights).

#### 3.2 Virgin texts

We used two types of virgin texts in this study, one to estimate the position of the reform, the other to validate the results. Because the interviews aimed at characterizing the reform rather than individual respondents, we merged the twelve individual transcripts (from 1,224 to 4,287 words) into a single concatenated text titled *Reform* (32,942 words). We added the whole content of the interviews into this concatenated document, not just the sections that addressed the ideology of the reform. The *Reform* text is available as supplementary material to this article (we randomized word order to protect the anonymity of our respondents).

The second type of virgin text are academic publications (four encyclopedia entries, four articles and one book chapter) from which we created three concatenated documents to reduce the variability due to the idiosyncrasies of each text and increase the robustness of our estimates.

---

<sup>3</sup> The policy paradigms are 'types' and are therefore qualitatively distinct. However, their ideological characterization by Daigneault's (2014, 2015) allows us to position them on the same left-right dimension.

<sup>4</sup> The word counts for reference and virgin texts are those after preprocessing.

We selected these publications because, at face value, their content is clearly and univocal representative of the ideology of each social assistance paradigm: *left* (Morel 2006; Dean 2008; Stephens 2010); *center* (Walters 1997; Jenson and Saint-Martin, 2004; Williams 2006) and *right* (Ralph and Stobbe 1997; O'Connor 2001; Welshman 2006: 6,331 words). To increase our confidence into our sample, we asked an internationally renowned social policy scholar, Daniel Béland, to classify each text, working independently, in one of the three paradigms. The agreement rate was 9/9 or 100%. The *left*, *center* and *right* texts are available as supplementary material to this article (we randomized word order to protect copyrights).

### 3.3 Procedures

We used the R (package *quanteda*) implementation to perform preprocessing and word scoring (Benoit and Nulty 2017). We preprocessed the corpus in a typical fashion, discarding punctuation, capitalization, numbers, and stop words. We then removed the affixes of the word keeping only the stem and discarded words that occurred only once in our corpus, as they could not be scaled, which removed most misspellings that remained in the interview transcripts. Wordscores benefits from removing words that have no meaning related to the studied dimension (e.g., words serving grammatical functions), as they tend to drag scores toward the middle of the scale. Whereas these words are assumed to be akin to random noise and are in equal proportion in each document, it is safer to remove them and doing so increases Wordscores' performance (Ruedin 2013; Collette and Pétry 2014). We used the Martin and Vanberg (2008) – hereafter MV – rescaling procedure to transform the raw scores into scores that are more easily interpretable and because it corrects the drag toward the bigger reference text occurring with Wordscores when reference texts are not of the same length. Moreover, the MV rescaling is recommended when we compare a small number of virgin texts directly to two reference texts, as is the case here (Benoit and Laver 2008).

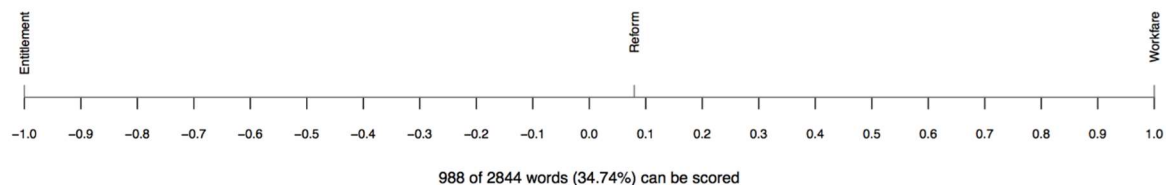
## 4. Results

Despite the low proportion of scored words (34.74%), the empirical results regarding the position of the reform are very positive (Table 1 and Figure 1). Indeed, Wordscores estimates the position of the reform at +0.08 (MV) on our ideological dimension, which is clearly in line with our expectations that the reform is located slightly to the right of the dimension's midpoint (center/center-right). At this point, however, we are unsure about the scored words' meaningfulness because the centrist position of the reform could result from uninformative words (e.g., those playing linguistic functions) contained in the text rather than truly informative, centrist words (Lowe 2008; see also Brier and Hopp 2011). To rule out this possibility, we followed two other types of validation strategies.

**Table 1 Rescaled and original results – Reform.**

Texts	Rescaled Text Scores (MV)	Original Text Scores	Original C.I. Low	Original C.I. High	Number of Words
Entitlement	-1	-0.48	-0.50	-0.45	1,318
Workfare	1	0.48	0.45	0.50	1,667
Reform	0.08	0.04	0.03	0.05	32,942

Source: Authors



**Fig. 1** MV Text Score – Reform

First, we used content validation to check whether the specific words that are scored (matched) their frequencies and their ‘word score’ (i.e., position on the dimension of interest) make sense from a substantive perspective (Slapin and Proksch 2008; Weinberg 2010; see also Adcock and Collier 2001). We browsed through the list of scored words used by Wordscores to estimate the position of the reform. Based on our intimate knowledge of the theory and of the case, we selected informative words clearly aligned with either a left- or right-wing ideology. The results are moderately positive. On the one hand, only approximately 15% of the words have clear and unambiguous informative content. The 15 most influential words in estimating the position of the reform<sup>5</sup> are ambiguous and therefore cannot be easily interpreted in the context of this study (Table 2). For instance, the word *right* can refer to an ideological orientation (e.g., right-wing, +1), social or human rights (-1), an adverb (e.g., to do the right thing) or, more probably in the context of this study, to a common interjection in an interview (e.g., “right?”). We had similar problems with *just* – a noun or adjective related to social justice (-1) or a common adverb (e.g., “we just assumed that...”) – and *social* which could be associated with the left (social-democratic, social rights, etc.) or could simply be used as neutral, generic terms (social assistance). Finally, although *independ\** is associated with workfare in theory, its meaning is ambiguous here because of the name of the reform (Building Independence). Although word combinations (e.g., *social justice*) would partly offset the ambiguous meaning of these scored words, using them would violate Wordscores’ “bag of words” assumption to the effect that individual words are the unit of analysis and that their order is irrelevant (Laver, Benoit and Garry 2003; Grimmer and Stewart 2013). Furthermore, at face value, the ambiguous words that we identified do not display a clear pattern (bias) in either direction, which suggests that uninformative words cancel each other out.

---

<sup>5</sup> The word frequencies displayed in Tables 2 and 3 and discussed in the text include all the texts. For example, the word frequency for “think” is 647, which means that this word was scored 647 times in the virgin and reference texts.

Table 2 Most influential uninformative words – Reform.

Words	Frequency	Word Score	Influence
<i>think</i>	647	-1	-647
<i>right</i>	630	-0.59	-371.7
<i>well</i>	428	-0.73	313.39
<i>just</i>	292	-1	-292
<i>social</i>	467	-0.48	-224.17
<i>realli</i>	217	1	217
<i>get</i>	271	1	271
<i>independ</i>	199	1	199
<i>like</i>	431	0.40	172.4
<i>kind</i>	164	1	164
<i>now</i>	152	1	152
<i>actual</i>	150	1	150
<i>assist</i>	285	0.51	145.59
<i>feder</i>	140	1	140
<i>back</i>	136	1	136

Source: Authors

On the other hand, the 150 informative words or so that we have identified make much sense substantively (Table 3 presents the 25 most influential words on the left and right). Influential words on the left refer mainly to the structural challenges, social problems and barriers related to living on social assistance (*poverty, child, health, hous\*, food, lack, etc.*) and, to a lesser extent, to a vision of a society that should care more for worst-off people (*democrat\*, citizen, progres\*, etc.*). Influential words on the right, by contrast, underline the importance of work (*work, employ\*, job, etc.*), the perception that social assistance claimant are passive (*depend\*, recipi\*, client, etc.*), morally responsible for their fate (*poor, blame, cultur\*, attitud\*, ethic, underclass, etc.*) and must be pushed off welfare (*incent\*, workfare, require\*, punit\*, etc.*). Moreover, we noted relatively few “anomalies” (i.e., misplacements).

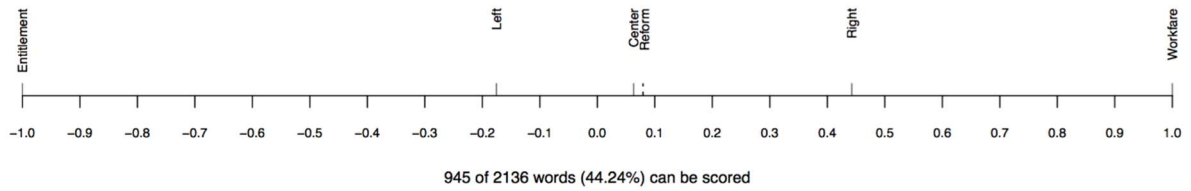
**Table 3 Most influential informative words – Reform.**

Left				Right			
Words	Frequency	Word Score	Influence	Words	Frequency	Word Score	Influence
<i>child</i>	135	-0.68	-91.8	<i>work</i>	359	0.26	93.34
<i>health</i>	86	-1	-86	<i>incent</i>	67	1	67
<i>poverti</i>	149	-0.46	-68.54	<i>poor</i>	60	0.88	52.8
<i>hous</i>	59	-1	-59	<i>cultur</i>	56	0.82	45.92
<i>public</i>	109	-0.44	-47.96	<i>job</i>	83	0.52	43.16
<i>food</i>	39	-1	-39	<i>workfar</i>	34	1	34
<i>resourc</i>	37	-1	-37	<i>underclass</i>	33	1	33
<i>care</i>	31	-1	-31	<i>depend</i>	45	0.53	23.85
<i>live</i>	43	-0.59	-25.37	<i>young</i>	22	1	22
<i>lack</i>	37	-0.68	-25.16	<i>ethic</i>	21	1	21
<i>earn</i>	25	-1	-25	<i>recipi</i>	18	1	18
<i>exclus</i>	44	-0.50	-22	<i>blame</i>	16	1	16
<i>financi</i>	22	-1	-22	<i>moral</i>	15	1	15
<i>tax</i>	48	-0.44	-21.12	<i>stigma</i>	14	1	14
<i>labour</i>	34	-0.59	-20.06	<i>requir</i>	20	0.69	13.8
<i>secur</i>	26	-0.77	-20.02	<i>attitud</i>	26	0.51	13.26
<i>barrier</i>	20	-1	-20	<i>rule</i>	13	1	13
<i>home</i>	20	-1	-20	<i>target</i>	13	1	13
<i>inequ</i>	23	-0.76	-17.48	<i>employ</i>	176	0.07	12.32
<i>democrat</i>	17	-1	-17	<i>communiti</i>	60	0.21	12.6
<i>citizenship</i>	29	-0.53	-15.37	<i>singl</i>	40	0.32	12.8
<i>citizen</i>	18	-0.77	-13.86	<i>client</i>	30	0.4	12
<i>equal</i>	12	-1	-12	<i>workforc</i>	10	1	10
<i>insur</i>	12	-1	-12	<i>punit</i>	9	1	9
<i>progress</i>	12	-1	-12	<i>mother</i>	13	0.69	8.97

Source: Authors

To ensure that our results are valid, we further tested the capacity of Wordscores, using the same reference texts, to estimate correctly the position of three concatenated academic texts whose content’s ideological position is clearly left, center or right. The results are very convincing (see Table 4 and Figure 2). First, Wordscores correctly ranks, from left to right, the academic texts on the ideological dimension and, despite a low proportion of scored words (44.59%), we observe significant differences between their estimated positions (discriminant validation). Second, whether we look at the original or rescaled results, the reform’s position (cf. Tables 1 and 4) is very close to the position of the center text. In fact, the confidence intervals of the original scores for the reform and center texts do overlap, which suggests that we cannot reject the hypothesis that they share an identical position (convergent validation). Moreover, as expected, the reform’s estimated position is significantly different from the position of the left and right academic texts (discriminant validation).





**Fig. 2 MV text scores – Academic texts**

Note: The reform’s position is shown for comparison purpose only.

**Table 4 Rescaled and original results – Academic texts.**

Texts	Rescaled Scores (MV)	Text Scores	Original Text Scores	Original C.I. Low	Original C.I. High	Number of words
Entitlement	-1	-0.48	-0.50	-0.45	-0.45	1,318
Workfare	1	0.48	0.45	0.50	0.50	1,667
Left	-0.17	-0.08	-0.10	-0.06	-0.06	6,435
Center	0.06	0.03	0.01	0.04	0.04	9,280
Right	0.43	0.21	0.19	0.23	0.23	6,331

Source: Authors

## 5. Discussion and conclusion

In this study, we replicated Daigneault’s (2015) qualitative finding regarding the ideology of a social policy reform, using the same data and conceptual framework using Wordscores, an automated content analysis method. Second, we used content, convergent and discriminant validation (Adcock and Collier 2001; Slapin and Proksch 2008; Grimmer and Stewart 2013) procedures to demonstrate the validity of our results. This step is particularly important since we used short, formal, written reference texts (academic publications) to analyze long, informal and spoken virgin texts (interview transcripts). As expected, Wordscores estimated the position of the reform slightly to the right of the midpoint of the scale, which suggests a predominantly center/center-right ideology. Furthermore, the validation evidence suggests that this result is highly credible (Table 5).

**Table 5 Summary of validation evidence.**

Validation Strategy	Validation Question(s)	Summary of Findings	Evidence Assessment
Convergent I	Does the ideological position of the reform as estimated by Wordscores converge with the findings of a previous qualitative study?	Yes, there is strong convergence between the position of the reform as estimated by Wordscores (MV score = 0.08) and our qualitative expectation (a position slightly to the right of the midpoint).	Strongly positive
Content	Do the matched words and their scores make sense substantively (e.g., words with a negative score should be associated with a left-wing ideology)?	Yes, in part. The proportion of words that have unambiguous informative content is low (around 15%). However, these words seem to capture adequately a left- or right-wing ideology and there are few errors.	Moderately positive
Discriminant I	Are the estimates for three sets of academic texts representing a left (-1), center (0) and right (+1) ideology congruent with theoretical expectations?	Yes, the MV scores for the left (-0.17), center (+0.06) and right (+0.43) texts display significant differences and are positioned according to expectations.	Strongly positive
Discriminant II	Does Wordscores estimate the position of the reform to be significantly different from the positions of the left and center academic texts?	Yes, the position of the reform (MV score = 0.08) differs significantly from the position of the left (MV score = -0.17) and right (MV score = +0.43) texts.	Strongly positive
Convergent II	Does the estimate of the ideological position of the reform converge with the estimate of the center academic text?	Yes, there is almost perfect convergence (MV score for reform text = 0.08; MV score for center text = 0.06).	Strongly positive

Source: Authors

The main limitation of this study is the low proportion of words used by Wordscores to estimate the position of the reform. Indeed, only 34.74% of the words were scored in interview transcripts, a proportion which improved (44.24%) when we used academic publications. Researchers are rarely transparent regarding the proportion of scored words but, for those who are, figures are in the 85–95% range (e.g., Laver, Benoit and Garry 2003; Galli, Grembi and Padovano 2009; Weinberg 2010). To our knowledge, the lowest proportion of scored words – in the 50–80% range – was reported in an unpublished study that used political manifestos to analyze Twitter feeds (Toff and Kim 2013). We believe that the high proportion of scored words reported in the literature could be explained by more aggressive preprocessing procedures than the one we applied in this study.

Based on the empirical evidence presented in this study, we suggest that Wordscores' scope of application is greater than what Laver, Benoit and Garry (2003) envisioned. Indeed, Wordscores seems to perform well even when significantly different reference and virgin texts are used or when a single text (case) is analyzed, thereby opening the door to a host of novel and interesting applications beyond the traditional use of the method. In particular, we believe that Wordscores holds much potential to supplement qualitative analysis in small-N research within mixed- or multi-methods research (Collier, Brady and Seawright 2010; Johnson, Onwuegbuzie and Turner 2007). If we assess automated content analysis methods according to their utility with respect to the task they perform for political scientists (Grimmer and Stewart 2013), then Wordscores is of great value. Using this method, we were able to analyze efficiently a relatively large corpus made of interview transcripts and academic publications *and* to produce valid inferences on policy position.

We conclude by suggesting the following avenue for future research. We need to assess whether the gain in Wordscores' performance that we would obtain by 'cleaning' texts better, that is, by removing more uninformative words and by working exclusively with the most relevant extracts of interviews, is worth the increased investment in time and resources.

## References

- Adcock, R., Collier, D.: Measurement Validity: A Shared Standard for Qualitative and Quantitative Research. *American Political Science Review* **95**(03), 529-546 (2001)
- Baturo, A., Mikhaylov, S.: Life of Brian Revisited: Assessing Informational and Non-Informational Leadership Tools. *Political Science Research and Methods* **1**(01), 139-157 (2013)
- Benoit, K., Laver, M.: Compared to What? A Comment on "A Robust Transformation Procedure for Interpreting Political Text" by Martin and Vanberg. *Political Analysis* **16**(1), 101-111 (2008)
- Benoit K., Nulty P. *Quanteda: Quantitative Analysis of Textual Data*. Version 0.9.9-24. (2017)
- Benoit, M., Cantin, M.-H., Duchesne, É.: Une nouvelle culture émergente ? L'évaluation du positionnement des groupes d'intérêt anglophones face au Cadre stratégique agricole canadien de 2007. *Canadian Journal of Political Science/Revue canadienne de science politique* **46**(04), 921-950 (2013)
- Braun, V., Clarke, V.: Using thematic analysis in psychology. *Qualitative Research in Psychology* **3**(2), 77-101 (2006)
- Brier, A., Hopp, B.: Computer assisted text analysis in the social sciences. *Quality & Quantity* **45**(1), 103-128 (2011)
- Charbonneau, É.: Talking like a tax collector or a social guardian? The use of administrative discourse by U.S. State lottery agencies. In: Imbeau, L.M. (ed.) *Do they walk like they talk? Speech and action in policy processes*. pp. 223-240. Springer, New York (2009)
- Collette, B., Pétry, F.: Comparing the position of Canadian political parties using French and English manifestos as textual data. In: Kaal, B., Maks, I., Elfrinkhof, A.v. (eds.) *From text to political positions: text analysis across disciplines*. pp. 27-44. John Benjamins, Amsterdam (2014)
- Collier, D., Brady, H.E., Seawright, J.: Outdated Views of Qualitative Methods: Time to Move On. *Political Analysis* **18**(4), 506-513 (2010)
- Cousins, K., McIntosh, W.: More than Typewriters, More than Adding Machines: Integrating Information Technology into Political Research. *Quality and Quantity* **39**(5), 581-614 (2005)
- Daigneault, P.-M.: Three paradigms of social assistance. *Sage Open* (2014).  
<http://dx.doi.org/doi:10.1007/s11135-017-0650-0>.

- Daigneault, P.-M.: Ideas and Welfare Reform in Saskatchewan: Entitlement, Workfare or Activation? *Canadian Journal of Political Science/Revue canadienne de science politique* **48**(1), 147-171 (2015)
- Daigneault, P.-M., Duval, D., Imbeau, L.M.: L'analyse de contenu automatisée et les entretiens font-ils bon ménage ? Caractériser l'idéologie d'une réforme de politique sociale à l'aide de Wordscores. In: Daigneault, P.-M., Pétry, F. (eds.) *L'analyse textuelle des idées, du discours et des pratiques politiques*. pp. 267-292. Presses de l'Université Laval, Québec (2017)
- Dean, H.: Social Policy and Human Rights: Re-thinking the Engagement. *Social Policy and Society* **7**(01), 1-12 (2008)
- Driver, S., Martell, L.: Left, Right and the third way. *Policy & Politics* **28**(2), 147-161 (2000)
- Galli, E., Grembi, V., Padovano, F.: Would you trust an Italian politician? Evidence from Italian regional politics. In: Imbeau, L.M. (ed.) *Do they walk like they talk? Speech and action in policy processes*. pp. 109-129. Springer, New York (2009)
- Gauvin, J.-P., Chhim, C., Medeiros, M.: Did They Mind the Gap? Voter/Party Ideological Proximity between the BQ, the NDP and Quebec Voters, 2006–2011. *Canadian Journal of Political Science/Revue canadienne de science politique* **49**(2), 289-310 (2016)
- Grimmer, J., Stewart, B.M.: Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis* **21**(3), 267-297 (2013)
- Herzog, A., Benoit, K.: The Most Unkindest Cuts: Speaker Selection and Expressed Government Dissent during Economic Crisis. *The Journal of Politics* **77**(4), 1157-1175 (2015)
- Huo, J.: *Third way reforms: Social democracy after the golden age*. Cambridge University Press, New York; Cambridge (2009)
- Imbeau, L.M.: Dissonance in fiscal policy: a power approach. In: Imbeau, L.M. (ed.) *Do they walk like they talk? Speech and action in policy processes*. pp. 167-184. Springer, New York (2009)
- Jenson, J., Saint-Martin, D.: Building blocks for a new social architecture: the LEGO paradigm of an active society. *Policy & Politics* **34**(3), 429-451 (2006)
- Johnson, R.B., Onwuegbuzie, A.J., Turner, L.A.: Toward a definition of mixed methods research. *Journal of Mixed Methods Research* **1**(2), 112-133 (2007)
- Klüver, H.: Measuring Interest Group Influence Using Quantitative Text Analysis. *European Union Politics* **10**(4), 535-549 (2009)
- Laver, M., Benoit, K.: Locating TDs in Policy Spaces: The Computational Text Analysis of Dáil Speeches. *Irish Political Studies* **17**(1), 59-73 (2002)
- Laver, M., Benoit, K., Garry, J.: Extracting Policy Positions from Political Texts Using Words as Data. *American Political Science Review* **97**(02), 311-331 (2003)
- Levitas, R.: *The inclusive society? Social exclusion and New Labour*. Palgrave Macmillan, New York (2005)
- Martin, L.W., Vanberg, G.: A Robust Transformation Procedure for Interpreting Political Text. *Political Analysis* **16**(1), 93-100 (2008)
- Marzagão, T.: *Using NLP to measure democracy*. (2015)
- Montpetit, É.: In defense of pluralism : policy disagreement and its media coverage. *Cambridge studies in comparative public policy*. Cambridge University Press, Cambridge, United Kingdom (2016)
- Morel, N.: Social rights. In: Fitzpatrick, T., Kwon, H.-j., Manning, N., Midgley, J., Pascall, G. (eds.) *International encyclopedia of social policy*. pp. 1291-1294. Routledge, Abingdon, Oxon ; New York, NY (2006)
- O'Connor, B.: The intellectual origins of 'welfare dependency'. *Australian Journal of Social Issues* **36**(3), 221-212 (2001)

- Ralph, D., Stobbe, M.: Welfare reform as moral reform in Saskatchewan. In: Biggs, L., Stobbe, M. (eds.) *Devine rule in Saskatchewan: A decade of hope and hardship*. pp. 267-281. Fifth House Publishers, Saskatoon, SK (1991)
- Ruedin, D.: The Role of Language in the Automatic Coding of Political Texts. *Swiss Political Science Review* **19**(4), 539-545 (2013)
- Slapin, J.B., Proksch, S.-O.: A Scaling Model for Estimating Time-Series Party Positions from Texts. *American Journal of Political Science* **52**(3), 705-722 (2008)
- Stephens, J.D.: The social rights of citizenship. In: Castles, F.G., Leibfried, S., Lewis, J., Obinger, H., Pierson, C. (eds.) *The Oxford Handbook of the Welfare State*. pp. 511-525. Oxford University Press, Oxford (2010)
- Toff, B.J., Kim, Y.M.: Words that matter: Twitter and partisan polarization. Paper presented at the UW-Madison's Political Behavior Research Group meeting, Madison, Wisconsin.
- van Berkel, R., Møller, I.H.: The concept of activation. In: van Berkel, R., Møller, I.H. (eds.) *Active social policies in the EU: Inclusion through participation?* pp. 45-71. The Policy Press, Bristol, UK (2002)
- Walters, W.: The 'Active Society': new designs for social policy. *Policy & Politics* **25**(3), 221-234 (1997)
- Warwick, P.V.: Public opinion and government policy in Britain: A case of congruence, amplification or dampening? *European Journal of Political Research* **54**(1), 61-80 (2015)
- Weinberg, M.: Measuring Governors' Political Orientations Using Words as Data. *State Politics & Policy Quarterly* **10**(1), 96-109 (2010)
- Welshman, J.: Workfare. In: Fitzpatrick, T., Kwon, H.-j., Manning, N., Midgley, J., Pascall, G. (eds.) *International encyclopedia of social policy*. pp. 1559-1561. Routledge, Abingdon, Oxon ; New York, NY (2006)
- White, S.: Welfare Philosophy and the Third Way. In: Lewis, J., Surender, R. (eds.) *Welfare state change: Towards a Third Way?* pp. 25-46. Oxford University Press, Oxford; New York (2004)
- Williams, C.C.: Active welfare. In: Fitzpatrick, T., Kwon, H.-j., Manning, N., Midgley, J., Pascall, G. (eds.) *International encyclopedia of social policy*. pp. 7-8. Routledge, Abingdon, Oxon ; New York, NY (2006)