Text-Mining Obituaries Between 1953 and 2019 Revealed that Women Leaders are Described Increasingly Like Men Leaders, But Yet Evaluated Differently

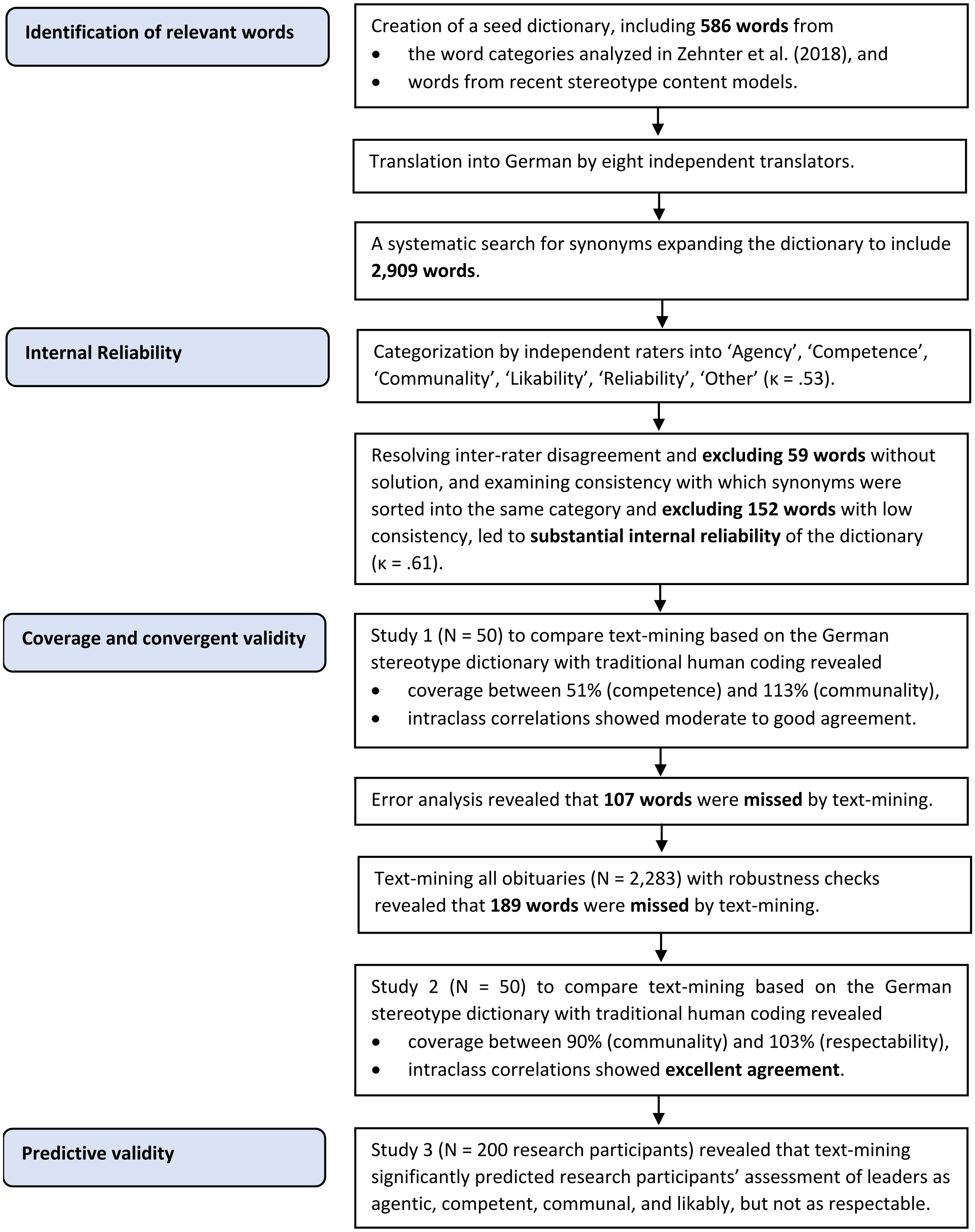
Supplementary Material

Miriam K. Zehnter, Christoph U. Wolfmayr, Leona A. Andriopoulos, Erich Kirchler, Martin Voracek, and Michelle K. Ryan

Supplement 1.

Development and Quality Assessment of the Text-Mining Dictionary

As a basis for the text-mining of leader obituaries, we created a dictionary with relevant descriptions and evaluations. Following Nicolas and colleagues' (2021) recommendations (who developed the English comprehensive stereotype content dictionaries), we proceeded in several steps: We (1) identified relevant words for the dictionary, (2) examined its internal reliability, and (3) tested its convergent validity. In an additional step, we tested the predictive validity of the text-mining dictionary. Specifically, we examined whether the word counts obtained through text-mining predicted how research participants assessed leaders based on their obituaries. Notably, the creation and validation of the dictionary was an iterative process including several rounds of assessment and improvement. Figure 1 visualizes this process and key results.



Supplement 1: Figure 1.

Flowchart of the Development and Quality Assessment of the Text-mining Dictionary

Identification of Relevant Words

First, we created a seed dictionary with relevant words drawing on Zehnter and colleagues' work (2018) and other recent stereotype content models and empirical studies (Abele et al., 2016; Hentschel et al., 2019; Pietraszkiewicz et al., 2019; Prentice & Carranza, 2002). This resulted in a list of 586 words. Eight individuals (undergraduate psychology students, seven women, one man) translated these words to German. As all translators had German as first language and professional proficiency in English, we did not use forward-backward translation. Instead, the translators convened to discuss their translations and specific word choice and tone. Subsequently, we conducted a systematic search for synonyms using the online Duden as the most comprehensive German language dictionary (Dudenredaktion, 2019). This resulted in a list of 2,909 words with 1,686 unique roots.

Internal Reliability

Unlike most previous dictionaries of gender stereotypes (Nicolas et al., 2021; Pietraszkiewicz et al., 2019), our text-mining dictionary distinguishes between agency and competence as they were independent factors in previous research (Rogers et al., 2013) and their gender-stereotypically appears to be changing over time (agency continued to be associated with men, while competence was increasingly associated with women) (Eagly et al., 2020). In addition, our dictionary includes two categories of evaluative words: likability and respectability. Thereto, we asked eight raters (undergraduate psychology students, four women, four men) to categorize all words into the categories 'Agency', 'Competence', 'Communality', 'Likability', 'Respectability', and 'Other'. To reduce fatigue, each rater categorized only one-quarter of the dictionary.

The initial inter-rater agreement was moderate, Cohen's Kappa = .53, which was below our aim of substantial internal reliability. A third rater (a woman undergraduate psychology student) resolved interrater disagreement. Fifty-nine words (3.55%) for which interrater disagreements could not be resolved were excluded from the dictionary, as they may not be reliable indicators for their categories.

As an additional indicator for internal reliability, we used the synonyms of each word to examine how consistently they were sorted into the same category. Thereto, we created an index ranging from 0 to

1, where 1 indicates that all synonyms were sorted into one category. This index was the number of a words' synonyms in its most dominant category minus the number of synonyms in its second most dominant category divided by the total number of synonyms. For example, if word A had four synonyms and all synonyms were sorted into the same category, the index would be (4-0)/4 = 1. In another example, Word B also has four synonyms of which 2 were sorted into Category 1, one into Category 2, and one into Category 3. In this case the consistency index would be (2-1)/4 = 0.25.

As a cut-off value, we decided that the consistency value should be at least 0.25, as around this value one category became clearly dominant among a word's synonyms, whereas in values below, a word's synonyms tended to be spread out in different categories. One-hundred fifty-two words (9.15%) with consistency values under the cut-off criterion were excluded from the dictionary.

After excluding words for which inter-rater agreement could not be reached or the synonym consistency was below our cut-off criterion, the dictionary's internal reliability was substantial, Cohen's Kappa = .61 [.58, .64].

Study 1. Coverage and Convergent Validity of the Initial Dictionary

Coverage and convergent validity are important indicators for the quality of gender stereotype dictionaries (Nicolas et al., 2021). Coverage refers to the overall proportion of words that text-mining identifies in comparison to human coders, whereas convergent validity means that there is agreement between the specific words identified by text-mining vs human coders. To examine our dictionary's coverage and convergent validity, we examined whether text-mining obituaries using our newly created dictionary would yield similar results than using traditional human coding. We selected a random subset of 50 obituaries using a random number generator. Then, following the recommended process for content analysis (Neuendorf, 2002), two coders (undergraduate psychology students, one woman, one man) read through the obituaries, identified relevant words and categorized then into the five categories 'Agency', 'Competence', 'Communality', 'Likability', and 'Respectability'.

Coverage. To meet the criterion of coverage, text-mining obituaries with our dictionary should yield an outcome that was equally or more reliable than the traditional human coding approach.

Specifically, the dictionary should identify as many relevant words as possible (i.e., 90 percent of those found by human coders), while not counting context-irrelevant words (i.e., less than 10 percent irrelevant counts). To reduce the risk of context irrelevant words a priori, we only analyzed the main text of the obituaries containing the descriptions of the deceased leaders. Additional text, such as details about the funeral, names of companies, co-workers, and family members, as well as Bible quotes and poetry were not analyzed.

Text-mining the subset of 50 obituaries based on our dictionary found 79 percent of the agency words identified by the human coders, only 51 percent of competence words, 113 percent of communality words, 65 percent of likability words, and 58 percent of respectability words. The value exceeding 100 percent means that the text-mining approach identified greater frequencies of communality words than the human coders, indicating that context irrelevant words have been counted. Example 1 in the below-presented qualitative demonstration of our text-mining approach (see Supplement 2) includes a so-called over-count. Concretely, the obituary concludes with the sentence ‘We will always remember him *faithfully*’. Here, the word faithfully was incorrectly counted as a communal description of the deceased leader, when in fact it expressed the organizations commitment to commemorate their leader.

Convergent validity. To meet the criterion of convergent validity, there should be good agreement, that is, intra-class-correlations (ICC) > .75, between the words identified through text-mining vs the human coders. The ICC indicated good reliability for ‘Agency’ (ICC = .79), moderate to good reliability for ‘Competence’ (ICC = .74), ‘Communality’ (ICC = .73), and ‘Likability’ (ICC = .74), and moderate reliability for ‘Respectability’ (ICC = .56).

Error analysis. The coverage and convergent validity of our initial dictionary were below our criteria for good reliability. Thus, in a next step, we examined the outputs of both the text-mining and the human coding for potential errors. The most prevalent error in the text-mining approach was an insufficient coverage of relevant words. In total, we found 107 relevant words that had not been included in our dictionary. Mostly, these words were synonyms of included words. The most prevalent error of human coders was the counting of words that we did not classify as indicative for our categories. For example, human coders counted academic titles of deceased leaders as respectability words.

Given the relatively low coverage of our initial dictionary, we also text-mined all obituaries (N = 2,283) using additional robustness checks. Specifically, we generated lists of all words that were counted into the five categories ‘Agency’, ‘Competence’, ‘Communality’, ‘Likability’, and ‘Respectability’. In addition, we generated a list of all uncounted words. We screened the former list to identify incorrect counts and the latter to identify relevant words (e.g., synonyms, word flexions) that were missed. This led to the discovery of 44 additional agency words with seven unique roots, 33 competence words with 18 unique roots, 47 communality words with 20 unique roots, twenty likability words with five unique roots, and 44 respectability words with 15 unique roots. These words were added to the dictionary.

Study 2. Coverage and Convergent Validity of the Final Dictionary

We selected another random subset of 50 obituaries, using a random number generator, and examined whether text-mining this subset of obituaries with our final dictionary would yield (1) substantial coverage (the dictionary should identify a minimum of 90 percent words of those found by human coders, but should identify less than 10 percent irrelevant words) and (2) good agreement (ICC > .75) between the words identified through text-mining vs a human coder. This time, the first author read through the obituaries and identified relevant words and categorized them into our five categories ‘Agency’, ‘Competence’, ‘Communality’, ‘Likability’, and ‘Respectability’.

Coverage. Text-mining the obituaries based on the final dictionary identified very high percentages of the words identified through human coding (94% agency, 96% competence, 90% communality, 100% likability, 103% respectability). Again, the last value (exceeding 100 percent) means that the text-mining approach identified greater frequencies of respectability words than the human coder, which indicates that context irrelevant words were counted.

Convergent validity. The intra-class correlations (ICC) between the words identified through text-mining vs the human coder indicated excellent agreements for ‘Agency’ (ICC = .96), ‘Competence’ (ICC = .91), ‘Communality’ (ICC = .90), ‘Likability’ (ICC = .93), and ‘Respectability’ (ICC = .96).

Study 3. Predictive Validity

To examine the predictive validity of our dictionary, we tested whether the words text-mined from the leader obituaries predict how research participants would describe and evaluate the leaders based on their obituaries. Using stratified sampling (Neuendorf, 2002), we randomly selected four obituaries for women leaders and four obituaries for men leaders. The four obituaries for women leaders were published in 1962, 2001, 2016, and 2019: the four obituaries for men leaders in 1956, 1995, 1998, and 2016. The selected obituaries varied in the frequencies of words signifying agency (0-7), competence (0-6), communality (0-9), likability (0-5), and respectability (0-7). To mitigate the risk that research participants assess the deceased leaders based on their own gender stereotypes, we revised the obituaries to conceal the leaders' gender.

Sample. Two hundred research participants assessed the deceased leaders based on their obituaries. To avoid fatigue, each participant read one randomly assigned obituary. Research participants were recruited via Prolific and paid between 1.00 and 2.50 Euros for the time spent on the survey. All participants resided in Germany and were fluent in German, 72 (36%) were women and 128 (64%) men. They were between 18 and 68 years of age (M = 31.48, SD = 9.82), and 42 (21%) held leadership positions themselves.

Procedure. After reading one obituary, the participants assessed the deceased leaders based on the five dimensions of our text-mining dictionary, each of which was operationalized with two items: agency (influential, assertive), competence (competent, experienced), communality (humane, warm), likability (popular, like a friend), and respectability (esteemed, respected). To get an overall impression of how gender stereotypical the selected obituaries were, we then asked participants to identify the leader's gender based on their obituary. Correct responses were rewarded with a bonus of 20 Cents.

Results and Discussion. Because of potential changes in verbal expressions over time, there may be greater alignment between the text-mining of more recent obituaries and research participants' assessment of leaders. To test, whether this was the case, we performed a hierarchical linear regression analysis, in which we regressed research participants' assessment of the deceased leaders on the text-mining counts, the decade in which an obituary was published, and the interaction Text-mining counts × Decade. To control for individual differences between participants' assessments, we included a random intercept for research participants. To control for differences between the selected obituaries (other than their text-mining counts), we also included a random intercept for obituaries. If the text-mining of more recent obituaries aligns better with research participants' assessments (suggesting better predictive validity of the text-mining approach for more recent obituaries), the interaction Text-mining counts × Decade should be statistically significant.

Results (summarized in Table 1) revealed large variance between research participants, indicating that (unknown) individual differences between research participants predicted their assessment of the deceased leaders. The variance between obituaries was relatively small, indicating that unknown differences between the selected obituaries played a small role in the prediction of participants' assessment of the deceased leaders. Text-mining counts significantly predicted research participants' assessments of the deceased leaders, supporting the overall predictive validity of our text-mining dictionary. Notably, neither Decade, nor the interaction Text-mining counts × Decade were significant predictors of research participants' assessments. These results suggest that the predictive power of our text-mining approach is 'timeless', that is, it does not depend on the decade in which an obituary was published.

Supplement 1: Table 1.

Hierarchical Regression Predicting Participants' Assessment of Leaders by Text-mining Counts × Decade

Random Effects (Intercepts)	σ^2	SD		
Research participants	0.60	0.77		
Obituaries	0.06	0.25		
Fixed Effects	B	SE	z	p
Text-mining counts	0.12	0.04	3.37	.001
Decade	0.09	0.08	1.19	.273
Text-mining counts × Decade	-0.01	0.01	-1.36	.173

When identifying leaders' gender, success rates were highest – 92 and 80 percent – for the two men leaders whose obituaries included high counts of agency words (in line with gender stereotypes), and lowest – 27 percent – for a woman leader whose obituary included zero communality words (contrasting gender stereotypes). In general, success rates of correctly identifying a leaders' gender were higher for men leaders (50-92%) than for women leaders (27-58%), likely because woman leaders occupy a counter-stereotypical role. Table 2 shows the eight obituaries presented to the research participants alongside the word counts obtained through text-mining and participants' assessments of the leaders. The obituaries were translated from German to English by the first author.

Supplement 1: Table 2.

Word Counts Obtained Through Text-Mining vs Ratings from Research Participants

Obituary 1: Woman business leader, published 1962 in conservative newspaper

Today, after short, severe sickness, the co-founder of our company EP passed away. During EPs life, advice, concern, and care went to the employees and the factory, in which EP lived amongst us even after the death of their spouse. Through kindheartedness and human sympathy, EP will remain unforgotten.

	Agency	Competence	Communality	Likability	Respectability
Text-mining counts (N)	1	1	5	0	0
Mean ratings M (SD) by research participants (N = 24)	4.65 (1.09)	5.40 (0.92)	6.02 (0.81)	5.44 (1.12)	5.83 (0.75)
Identified as woman by	58%				

Obituary 2: Woman business leader, published 2001 in liberal newspaper

We are grieving our shareholder, certified political economist LO, bearer of the Federal Cross of Merit, bearer of the Bavarian state medal for social services / ‘Construction South’. Over five decades, LO tirelessly championed the interests of the company group and contributed with unerring decisions to the success of the company. As editor, publisher, and designer of our magazine ‘Construction South informs’, leader of the entire advertisement and PR department, LO significantly shaped the image of the company. LOs great personal dedication to support the medical treatment of sick children from Ukraine and Romania and LOs championship for social housing were honored with the Federal Cross of Merit in 1999, which was personally handed out by the Federal Minister Roman Herzog. All employees appreciated LOs extremely generous and kind support and the ever-present certainty of being able to reach out to LO with professional and personal concerns and to receive competent advice. With LO, we lose a personality who is irreplaceable. With the utmost respect for LOs life achievements, we will keep LO an honorable memory. Deeply grateful, we say farewell.

	Agency	Competence	Communality	Likability	Respectability
Text-mining counts	7	4	5	0	7
Mean ratings M (SD) by research participants (N = 24)	5.25 (1.22)	5.83 (1.50)	5.77 (1.37)	5.42 (1.21)	5.88 (1.56)
Identified as woman	46%				

Obituary 3: Woman business leader, published 2016 in liberal newspaper

We are shocked and deeply grieving the death of our highly esteemed colleague Dr SK, leader of the department ‘Group Accounting and Reporting’ in the executive board for finance, controlling, and risk at Allianz SE. Dr K died after severe sickness at the age of only 56 years. In 23 years at Allianz SE, SK has inspired many colleagues with professional knowledge, devotedness, and reliability. SK was an everywhere respected member of the Allianz family, who radiated competence, openness, and collegiality inside and outside the company. We will always remember SKs friendliness and an honest interest in co-workers and colleagues. We are grateful for our time together and for the great accomplishments that Dr K has achieved for the Allianz. Our sympathy goes to SKs family and friends, who accompanied SK devotedly in the last couple of months.

	Agency	Competence	Communality	Likability	Respectability
Text-mining counts	3	3	9	4	2
Assessment from research participants (N = 25)	4.76 (0.89)	5.80 (0.68)	5.68 (0.81)	5.10 (0.91)	5.76 (0.82)
Identified as woman	44%				

Obituary 4: Woman academic leader, published 2019 in liberal newspaper

The department for law of the Goethe University is mourning Prof Dr BH, LL.M. (Univ. of Chicago). With BH, the department loses an outstanding representative of international economic law and an internationally recognized colleague who was equally committed to research, teaching, the promotion of junior researchers, and self-management. BH was associated with the department since 2004 and held the professorship for civil law, German, European, and international economic law and finance and comparative law. The interdisciplinary connections of the law to the economy in particular and the comparison with other European law systems and the US-American law were in the focus of BHs scientific work. The department owes BH important impetus, in particular for the internationalization of teaching and will keep this esteemed colleague a worthy memory.

	Agency	Competence	Communality	Likability	Respectability
Text-mining counts	2	2	0	1	3
Assessment from research participants (N = 26)	5.25 (0.80)	6.17 (0.66)	4.63 (1.15)	4.27 (1.17)	5.94 (0.77)
Identified as woman	27%				

Supplement 1: Table 2.

Word Counts Obtained Through Text-Mining vs Ratings from Research Participants

Obituary 5: Man business leader, published 1956 in conservative newspaper
Completely unexpected, our executive board member, district master craftsman/craftswoman ER passed away at the age of 58. The deceased advised us for years in a selfless manner and significantly promoted our social duties. With the deceased, we lose a person, who enjoyed great appreciations thanks to an honest character and human sympathy. We will keep ER an honorable memory.

	Agency	Competence	Communality	Likability	Respectability
Text-mining counts	0	0	5	0	2
Assessment from research participants (N = 24)	4.42 (1.38)	5.38 (1.39)	5.16 (1.49)	4.80 (1.40)	5.28 (1.46)
Identified as man	68%				

Obituary 6: Man business leader, published 1995 in conservative newspaper
Mourning we say farewell to PEB who passed away suddenly and unexpectedly at the age of 53. More than 22 years, PEB was associated with our company. Facing challenging tasks, among other in Great Brittain, Italy, and Brazil, PEB championed for our corporation with great expertise and ever-present commitment. Since 1994, PEB was successfully acting as the manager of our holding company in Bangkok/ Thailand. During a successful professional career, PEB enjoyed recognition and appreciation as leader and as colleague. PEBs early death affects us deeply. Our appreciation goes to a life-affirming, dutiful person and a great personality. We will keep PEBs work in good memory.

	Agency	Competence	Communality	Likability	Respectability
Text-mining counts	7	1	1	1	4
Assessment from research participants (N = 24)	5.28 (0.87)	6.18 (0.66)	4.76 (1.01)	4.68 (1.20)	6.00 (0.84)
Identified as man	92%				

Obituary 7: Man business leader, published 1998 in conservative newspaper
The former member of our executive board, Dr KH passed away at the age of 65. Dr H started a professional career at the former Cassella Farbwerke Mainkur AG in 1959. In the year 1972, KH took over the leadership of the production group for dyes in our company and was, before becoming member of the executive board in 1986, leader of the department for dyes, pigments, and precursors. Until the retirement in 1994, Dr H was in the executive board of our company and responsible for the department of fine chemicals and dyes, as well as for the department of environmental protection and security. With great professional knowledge and a distinct talent to lead people, Dr H filled this area of responsibility confidently and pragmatically. A systematic way of working, great commitment, and an exemplary social competence ensured KH the respect and appreciation of the colleagues in the executive board and the affection of the employees. The problems of the people in the factory were especially close to KH's heart. In difficult times, KH's extraordinary commitment went to environmental protection and the security of the company. In all departments of the company, KH championed for the expansion of integrating environmental protection of the production through developing procedures and products, which yielded progress for the environmental protection and security. KH always felt responsible for the welfare of the company and employees. With gratitude and mourning, we say farewell to an amiable person, who enjoyed great appreciation. KHs personality and actions in our company, which were characterized by warmhearted humanity, will remain unforgotten.

	Agency	Competence	Communality	Likability	Respectability
Text-mining counts	7	6	4	2	7
Assessment from research participants (N = 24)	5.72 (0.75)	6.64 (0.45)	5.86 (1.01)	5.36 (0.92)	6.42 (0.59)
Identified as man	80%				

Obituary 8: Man academic leader, published 2016 in liberal newspaper
In great grief and disbelief, we say farewell to our dear former colleague Prof Dr med SM. SM left us suddenly and unexpectedly. During SMs time as head physician in the department of internal medicine of the university hospital, SMs particular care was for the patients and SMs commitment for research and teaching. As former president of the German Society for Diabetes, president of numerous scientific congresses and as physician, SM shaped the German diabetology like few others. All actions were permeated with a deeply ethical conduct and deep compassion for the sick and the weak. Much too early, we lost an extraordinary person, a dear colleague, friend, and advisor. We will miss SMs humor and human warmth. Our deep condolences go to SMs family.

	Agency	Competence	Communality	Likability	Respectability
Text-mining counts	3	2	5	5	2
Assessment from research participants (N = 24)	5.52 (0.89)	6.42 (0.64)	6.12 (0.77)	5.87 (0.94)	6.38 (0.84)
Identified as man	50%				

For a test of the predictive validity of our text-mining approach by category, we examined whether the word counts obtained from text-mining would predict research participants' assessment of the deceased leaders, using general linear regressions analyses. In each model, we regressed the outcome variable – participants' assessment of the leader (as agentic, competent, communal, likeable, or respectable) – on the respective text-mining counts and the two predictors 'leaders' true gender', and 'perceived leader gender'. Thus, these models provided not only insight into the predictive power of our text-mining approach, but a significant predictor of true gender indicated that the obituaries communicate gender stereotypes in additional ways that are not captured by our text-mining approach. Vice versa, a significant predictor of perceived gender indicated that participants assessed leaders based on their own gender stereotypes.

Results from the regression analyses, shown in Table 2, revealed that text-mining counts were significant predictors of participants' assessment of the deceased leaders as agentic, competent, communal, and likable. In addition, leaders' true gender significantly predicted participants' competence assessments such that men leaders were rated as more competent than women leaders. Thus, obituaries may include gendered competence signals (e.g., job titles) that are not captured by our text-mining approach. Moreover, participants' perceptions of leaders' gender significantly predicted their ratings of leaders as communal and likable, such that presumed women leaders were rated as more communal and more likable. These results suggest an association between participants' gender stereotypes and their assessments of the deceased leaders. In the masculine context of leadership, research participants may have relied on words signaling communality and likably to identify women leaders. Alternatively, participants' preconceptions of the deceased leaders' gender may have affected their assessments.

Neither text-mining counts, nor leaders' true gender, nor perceived leaders' gender significantly predicted participants' assessment of leaders as respectable. Notably, respectability ratings were high across all obituaries. Possibly, just receiving an obituary signaled respectability to participants. This would be in line with results from the above-presented validation studies, which compared word counts extracted through text-mining vs human coders. Here, human coders also counted information other than respectability words (e.g., academic titles) as signals for respectability.

Supplement 1: Table 3.

General Linear Regression Analyses Predicting Research Participants' Assessment of Leaders

	Agency		Competence		Communality		Likability		Respectability	
	z	p	z	p	z	p	z	p	z	p
Text-mining counts	3.79	**<.001**	3.98	**<.001**	4.28	**<.001**	2.12	**.036**	1.70	.091
True leader gender	-0.74	.463	-2.67	**.008**	-0.94	.348	-0.90	.368	-1.00	.319
Perceived leader gender	-0.75	.452	-0.87	.383	2.12	**.035**	3.14	**.002**	0.81	.420

Note. This table summarizes the results from general linear regression analyses on participants assessment of leaders as agentic, competent, communal, likable, and respectable. In in each model, we predicted participants' assessment on the text-mining counts (the frequencies of words from this category identified through text-mining), true leader gender (which was blinded), and perceived leader gender (participants were asked to guess leaders' gender).

Conclusion. In sum, the results from Study 3 support the predictive validity of our text-mining dictionary. At large, the word counts text-mined from leader obituaries were coherent with the ways research participants assessed deceased leaders based on their obituaries. Thus, our text-mining approach captures meaningful signals of the descriptions and evaluations entailed in the obituaries of leaders.

General Conclusion

The final text-mining dictionary included 2,622 nouns, adjectives, and verbs with 1,468 unique roots. Nine hundred ninety-four words with 522 (35.56%) unique roots signified agency, 475 words with 284 (19.35%) unique roots competence, 825 with 480 (32.70%) unique roots communality, 128 words with 70 (4.77%) unique roots likability, and 219 words with 112 (7.23%) unique roots respectability.

This dictionary has substantial internal reliability, excellent coverage and convergent validity (Study 2) and predictive validity (Study 3).

Limitations and User Suggestions for the Text-Mining Dictionary

The text-mining dictionary is available on the Open Science Framework (https://osf.io/7p29f/?view_only=07061ecd9ddf4f6e814b55963dccde36). However, in its current form, the dictionary may not be suitable for study contexts other than the analysis of obituaries. Specifically, its usefulness in contexts that include blatant negative descriptions and evaluations (e.g., comments on social media, newspaper articles) may be limited. Thus, before using our dictionary in other contexts, we suggest that researchers systematically expand it with relevant antonyms that capture negative descriptions and evaluations usually not present in obituaries. Upon text-mining, we also strongly advise researchers to employ the robustness checks provided in our text-mining R-script (lines 195-277). These robustness checks allow researchers to identify incorrect and missed counts.

References

Abele, A., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the Fundamental Content Dimensions: Agency with Competence and Assertiveness – Communion with Warmth and Morality. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01810

Bundesagentur für Arbeit. (2022). *Die Arbeitsmarktsituation von Frauen und Männern 2021 [Labourmarket conditions for women and men in 2021]*.

Dudenredaktion. (2019). *Duden - Das Synonymwörterbuch: Treffend formulieren mit 300000 sinnverwandten Wörtern[Duden - The glossary of synonyms: Fitting expressions with 300,000 synomynous words]*. Duden.

Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *American Psychologist*, *75*(3), 301–315. https://doi.org/10.1037/amp0000494

Hentschel, T., Heilman, M. E., & Peus, C. V. (2019). The multiple dimensions of gender stereotypes: A current look at men’s and women’s characterizations of others and themselves. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00011

Nicolas, G., Bai, X., & Fiske, S. T. (2021). Comprehensive stereotype content dictionaries using a semi-automated method. *European Journal of Social Psychology*, *51*(1), 178–196. https://doi.org/10.1002/ejsp.2724

Pietraszkiewicz, A., Formanowicz, M., Gustafsson Sendén, M., Boyd, R. L., Sikström, S., & Sczesny, S. (2019). The big two dictionaries: Capturing agency and communion in natural language. *European Journal of Social Psychology*, *49*(5), 871–887. https://doi.org/10.1002/ejsp.2561

Prentice, D. A., & Carranza, E. (2002). What Women and Men Should Be, Shouldn’t be, are Allowed to be, and don’t Have to Be: The Contents of Prescriptive Gender Stereotypes. *Psychology of Women Quarterly*, *26*(4), 269–281. https://doi.org/10.1111/1471-6402.t01-1-00066

Rogers, K. B., Schröder, T., & Scholl, W. (2013). The affective structure of stereotype content. *Social Psychology Quarterly*, *76*(2), 125–150. https://doi.org/10.1177/0190272513480191

Zehnter, M. K., Olsen, J., & Kirchler, E. (2018). Obituaries of female and male leaders from 1974 to 2016 suggest change in descriptive but stability of prescriptive gender stereotypes. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.02286

Supplement 2.

Qualitative Demonstration of Text-mining Leader Obituaries

In the following section, we show the outcome from text-mining eight leader obituaries. The aim of these examples is to illustrate the process of text-mining to readers unfamiliar with this method. Aiming for greater methodological transparency, we also indicate text-mining errors, which despite all diligence can never be fully avoided. The presented obituaries were translated from German by the first author. Please note that in some cases the English translation may result in two or more words for one German words. In other cases, it may not fully capture the cultural connotations associated with a specific word.

Example 1.

Business leader, Man, 1962, published in liberal newspaper

Our **revered**, everywhere **popular** boss Mr. *Name*, managing partner of *Name of Company* has passed away on *date of death*. With him, we lose more than a supervisor, because he **led** with his **heart**, **prudence**, and **outstanding professional knowledge** in an **exemplary** way, and he won our **particular respect** and **reverence** through his **humane** and **companionable** conduct. We will always remember him **faithfully**.

Category	*Count*	*Words*
Agency	1	led
Competence	2	outstanding, professional knowledge
Communality	5	heart, prudence, humane, companionable, faithfully[1]
Likability	1	popular
Respectability	5	revered, exemplary, particular, respect, reverence

[1]Note: Example of an over-count.

Example 2.

Business leader, Man, 1953, published in liberal newspaper

After almost 50 years of **devoted** work on his **oeuvre**, the founder and senior manager of our publisher *name of organization* passed away after long suffering. His employees mourn an always **helpful** boss and **benevolent** man.

Category	*Count*	*Words*
Agency	0	
Competence	1	oeuvre
Communality	3	devoted, helpful, benevolent
Likability	0	
Respectability	0	

Example 3.

Business leader, Woman, 1980, published in conservative newspaper

With **appreciation**, we say farewell to our longstanding **colleague** and shareholder *Name*. Her sudden death leaves us deeply mourning. With her **exemplary performance of duties** and **loyalty**, she was **devoted** to our company. Due to her **cheerful**, **life-affirming** nature, she leaves many **friends**.

Category	*Count*	*Words*
Agency	0	
Competence	0	
Communality	5	performance of duties, loyalty, devoted, cheerful, life-affirming
Likability	2	colleague, friends
Respectability	1	appreciation, exemplary

Example 4.

Business leader, Man, 1989, published in conservative newspaper

On *date*, our retired, former managing director, Mr Prof Dr Dr *Name* died unexpectedly at the age of 77 years. Until his retirement in 1977 the deceased was associated with our house for almost 40 years. During this time, he crucially **shaped** the development of our company as chief manager of the former *name of company* and member of the executive board of *name of company*. Equipped with always **new** **ideas**, which he **knew** to implement in a **scientifically** **profound** way, he has developed many procedures that are used around the world. With his **entrepreneurial** **farsightedness** and his great personal **commitment**, he has crucially **contributed** to the international **recognition** of our house. His **exemplary** **care** was directed towards his employees. We mourn the loss of a great **entrepreneurial personality,** a **renowned** scientist and **brilliant** man, whose life and **actions** were characterized by **persuasiveness**, **simplicity**, and dignity. Until the end, the deceased was a **faithful** **friend** and **valuable** **advisor.** His death leaves us mourning. We will keep an **honorable** memory of him.

Category	*Count*	*Words*
Agency	7	shaped, new, entrepreneurial, commitment, entrepreneurial personality, actions, persuasiveness
Competence	6	knew, scientifically, profound, farsightedness, brilliant, advisor
Communality	4	contributed, care, simplicity, faithful
Likability	1	friend
Respectability	5	recognition, exemplary, renowned, valuable, honorable

Example 5.

Business leader, Woman, 2016, published in the liberal newspaper

Deeply mourning, we say farewell to the co-founder of our corporate group *name*. An **extraordinarily** fulfilled and **happy** life comes to its end. With *name*, we lose a **brilliant** **entrepreneur** and an **affectionate** person. With **passion** and **vigor**, she built with her husband *name* our corporate group. Until the end, the employees were her priority. She was **esteemed** and **revered** among our employees for her personal **commitment**. For many she was a **mentor** and **inspiration** *Name* was born *year of birth* in *place of birth*. 1962, she married *name of husband*. **Together** *name of deceased* and *name of husband* founded our corporate group. After the early death of her husband in 1987, she took over the **leadership** of her **family** as a **guiding** and driving **force** and **supported** her children in the **expansion** of the corporate group. We **appreciate** *name* and **bow down** to a **brilliant** personality. We will keep her in **honorable** memory. Our sympathy goes to her **family** who we send our deepest condolences.

Category	*Count*	*Words*
Agency	8	entrepreneur, passion, vigor, commitment, leadership, guiding, force, expansion
Competence	2	brilliant, brilliant
Communality	6	happy, affectionate, together, family, supported, family
Likability	0	
Respectability	8	extraordinarily, esteemed, revered, mentor, inspiration, appreciate, bow down, honorable

Example 6.

Academic leader, Woman, 2004, liberal newspaper

The medical profession is mourning Dr. *Name*. Since 1996 she was president of the Medical Association of *City in Germany* and since 1999 Vice-president of the *Medical Association in Germany*. With *Name*, the medical profession has not only lost one of their **outstanding** personalities, but also an exceptional **lovable** and **warmhearted** **colleague**, who enjoyed great **sympathy** far beyond our profession. For her enduring **merits** for the healthcare system and the medical profession in Germany, *Name* was **honored** the *Name of Award* from the *Name of Awarding Institution*.

Category	*Count*	*Words*
Agency	0	
Competence	2	outstanding, merits
Communality	1	warmhearted
Likability	3	loveable, colleague, sympathy
Respectability	1	honored

Example 7.

Academic leader, Man, 2004, liberal newspaper

The German Society for *Blinded* Medicine is mourning Prof. Dr. Dr. *Name*. With *Name*, the German Society for *Blinded* Medicine loses their nestor and former chairman. Professor *Name* was a **highly esteemed** university teacher, scientist, and **committed** doctor, who for decades has earned great **merits** in internal medicine nationally and internationally. With his **understanding** of science, teaching, and patient care, he has **shaped** a whole generation of physicians. Until the end, he participated in the annual congresses of the society, in the **scientific progress** of internal medicine, and the development of the health care system with an alert mind. For his **exemplary** medical demeanor, outstanding **commitment**, **balanced** judgement, and his **commitment** to internal medicine as well as his patients, we owe him extraordinary gratitude.

Unforgotten will *Name* go down in the history of our society.

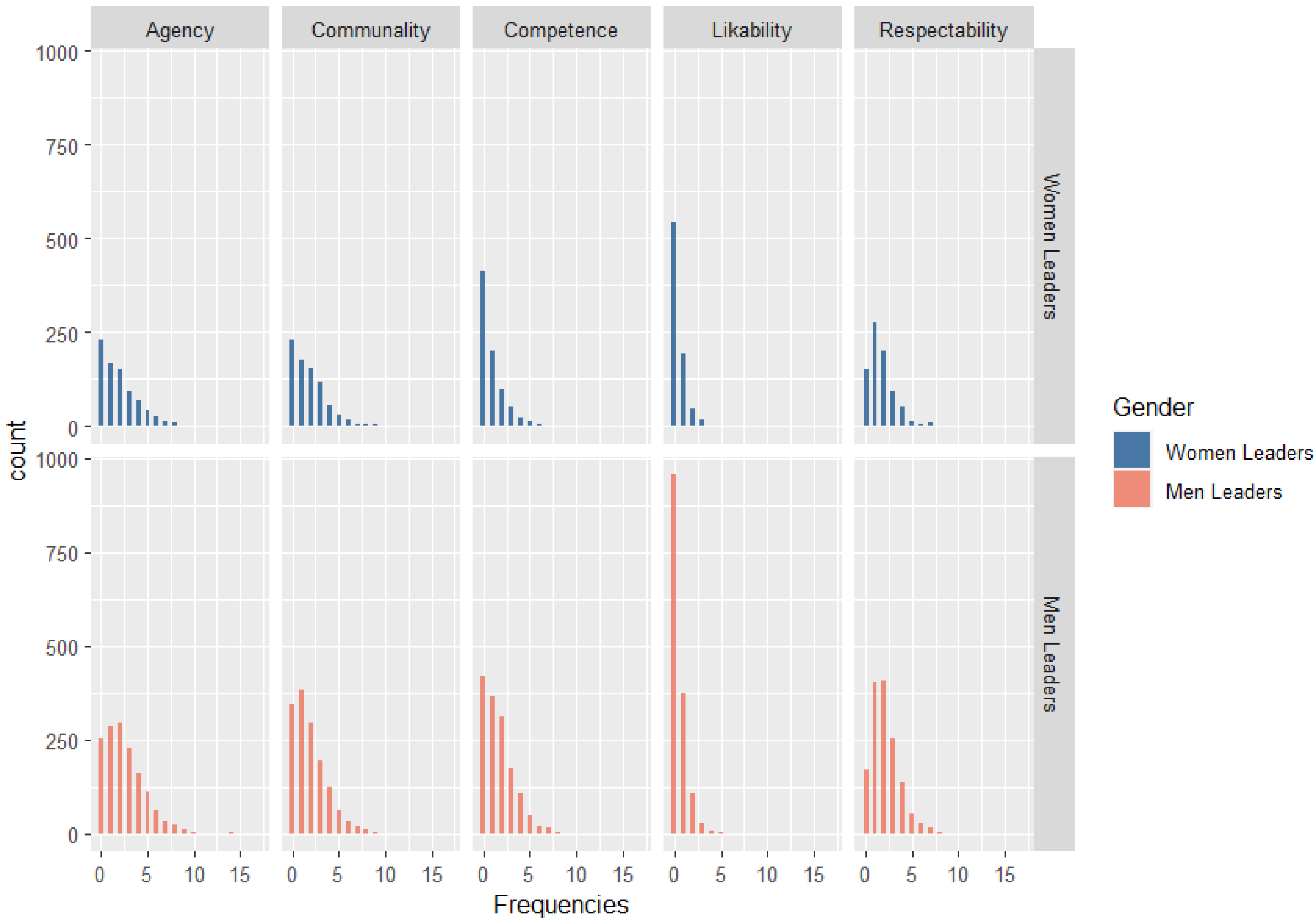
Category	*Count*	*Words*
Agency	4	committed, shaped, commitment, commitment
Competence	2	merits, understanding of science
Communality	2	understanding, balanced
Likability	0	
Respectability	2	esteemed, exemplary

Example 8.

Charity leader, Women, 2019, conservative newspaper

Sad, but endlessly **appreciative**, we say farewell to *Name of deceased*. Our **appreciated honorable president** and wearer of the ring of honor of the city of *City* passed away shortly before her 89th birthday in her hometown *City*. This news hit us unexpectedly. We are deeply mourning, because *Name of the deceased* was a **role model** to us. She was a **smart** and **canny** **pioneer** for the preservation of nature in Germany. We will painfully miss her **courage**, her **prudence**, and her **farsightedness**.

Category	*Count*	*Words*
Agency	2	pioneer, courage
Competence	3	smart, canny, farsightedness
Communality	1	prudence
Likability	0	
Respectability	4	appreciative, appreciated, honorable, role model



Supplement 3.

Count Distribution of Outcome Variables

Note. Agency: Max = 17, 75th percentile = 4; Competence: Max = 10, 75th percentile = 2; Communality: Max = 12; 75th percentile = 3;

Likability: Max = 6, 75th percentile = 1; Respectability = 10; 75th percentile = 3.

Supplement 4.

Robustness Check 1: Re-testing Hypotheses 1 and 2 with Negative Binomial Hurdle Regressions

As a robustness check, we present negative binomial hurdle regression models, which unlike the negative binomial models presented in the main manuscript separate zero counts from positive counts. Thus, hurdle regressions combine a binomial zero-inflated model predicting the odds of zero vs at least one count with a model predicting counts above zero (Blevins et al., 2015; Feng, 2021). Re-testing our hypotheses, hurdle regressions allowed us to account for the zero-inflation present in our data, while including the same predictors and control variables in both the zero-hurdle and the count model.

At large, the results from these analyses were consistent with the results presented in the main manuscript, except for agency. In the initial hurdle model on agency, there were neither significant gender differences in the likelihood to receive zero agency words, nor significant change thereof. Similaritly, in the count model, there were neither significant gender differences in the frequencies of agency words, nor significant change thereof. Upon removing the non-significant interaction Gender × Decade from the regression equation, in the zero-inflated model, the main effect of gender emerged as statistically significant with women leaders being more likely described with zero agency words than men leaders. In the count model, there were no significant gender differences in the frequencies of agency words. Hence, unlike the results from linear and non-linear negative binomial regressions presented in the main manuscript, results from negative binomial hurdle regression do not support Hypothesis 1a. Women leaders were more likely described with zero agency words and this gender gap did not decrease over time.

Supplement 4: Table 1.

Negative Binomial Hurdle Regression Predicting the Frequencies of Agency, Competence, and Communality Words

	Agency			Competence			Communality[2]		
Zero-hurdle Model	b (SE)	z	p	b (SE)	z	p	b (SE)	z	p
Obituary Length	0.05 (0.00)	16.55	**<.001**	0.03 (0.00)	16.02	**<.001**	0.04 (0.00)	16.01	**<.001**
Newspaper (conservative)	0.04 (0.13)	0.34	.733	0.48 (0.10)	4.56	**<.001**	-0.04 (0.11)	-0.32	.747
Academic leader	1.11 (0.62)	1.79	.074	0.80 (0.56)	1.43	.153	-1.73 (0.82)	-2.11	**.035**
Business leader	1.20 (0.59)	2.03	**.043**	0.52 (0.54)	0.97	.330	-1.08 (0.81)	-1.34	.180
Charity leader	1.20 (0.63)	1.91	.056	0.30 (0.56)	0.54	.593	-0.83 (0.83)	-1.00	.319
Political leader	0.60 (0.64)	0.94	.347	0.06 (0.57)	0.10	.916	-1.75 (0.83)	-2.10	**.036**
Leadership level (high)	-0.22 (0.14)	-1.59	.113	0.33 (0.17)	2.00	**.046**	-0.06 (0.12)	-0.52	.602
Founder	0.67 (0.22)	3.08	**.002**	-0.23 (0.12)	-2.00	**.046**	0.09 (0.18)	0.53	.594
Decade	0.20 (0.03)	5.65	**<.001**	0.00 (0.04)	0.09	.929	-0.08 (0.03)	-2.45	**.014**
Gender (women leader)	-0.30 (0.13)	-2.38	**.018**	-0.87 (0.26)	-3.40	**.001**	0.20 (0.12)	1.78	.076
Gender × Decade	-	-	ns[1]	0.05 (0.05)	0.88	.380	-	-	ns[2]
Count Model									
Obituary Length	0.01 (0.00)	21.79	**<.001**	0.01 (0.00)	15.44	**<.001**	0.01 (0.00)	16.28	**<.001**
Newspaper (conservative)	0.14 (0.04)	3.90	**<.001**	0.00 (0.05)	0.05	.963	0.06 (0.05)	1.24	.216
Academic leader	-0.14 (0.19)	-0.75	.456	0.66 (0.33)	2.00	**.046**	0.13 (0.25)	0.53	.598
Business leader	0.16 (0.19)	0.86	.390	0.68 (0.32)	2.10	**.035**	0.42 (0.24)	1.73	.084
Charity leader	-0.05 (0.19)	-0.26	.796	0.37 (0.34)	1.09	.278	0.46 (0.25)	1.85	.064
Political leader	0.05 (0.20)	0.26	.797	0.41 (0.09)	1.20	.229	0.25 (26)	1.00	.320
Leadership level (high)	-0.00 (0.04)	-0.05	.964	-0.05 (0.09)	-0.52	.605	-0.05 (0.05)	-0.97	.333
Founder	0.25 (0.05)	4.74	**<.001**	-0.13 (0.06)	-2.25	**.024**	0.05 (0.07)	0.70	.482
Decade	0.05 (0.01)	5.14	**<.001**	0.04 (0.03)	1.52	.130	-0.03 (0.01)	-2.78	**.005**
Gender (women leader)	-0.03 (0.04)	-0.72	.474	-0.63 (0.17)	-3.74	**<.001**	0.17 (0.05)	3.59	**<.001**
Gender × Decade	-	-	ns	0.08 (0.03)	2.66	**.008**	-	-	ns
Pseudo R^2	.434			.348			.291		
Unique R^2 Gender	.002			.016			.005		
Unique R^2 Gender × Decade	-			.003			-		
Hypothesis testing	**H1a: Contradicted**			**H1b: Supported**			**H1c: Supported**		

[1]*Note*. In the initial zero-hurdle model on **agency**, gender (b = 0.08, z = 0.26, p = .793) and the interaction Gender × Decade (b = -0.10, z = -1.43, p = .152) were not significant. Similarly, in the initial count model on agency, gender (b = -0.21, z = -1.82, p = .052) and the interaction Gender × Decade (b = 0.04, z = 1.82, p = .069) were not significant.
[2]*Note*. In the initial zero-hurdle model on **communality**, gender (b = 0.29, z = 1.02, p = .307) and the interaction Gender × Decade (b = -0.02, z = -0.32, p = .746) were not significant. Similarly, in the initial count model on communality, gender (b = 0.09, z = 0.77, p = .441) and the interaction Gender × Decade (b = 0.02, z = 0.85, p = .393) were not significant.

Supplement 4: Table 2.

Negative Binomial Hurdle Regression Predicting the Frequencies of Likability and Respectability Words

	Likability			Respectability		
Zero-hurdle Model	b (SE)	z	p	b (SE)	z	p
Obituary Length	0.01 (0.00)	7.57	**<.001**	0.02 (0.00)	8.94	**<.001**
Newspaper (conservative)	-0.22 (0.10)	-2.23	**.026**	-0.06 (0.13)	0.45	.651
Academic leader	0.85 (0.60)	1.42	.156	0.73 (0.55)	1.31	.190
Business leader	0.87 (0.59)	1.48	.139	1.32 (0.53)	2.48	**.013**
Charity leader	1.11 (0.61)	1.83	.067	0.58 (0.56)	1.05	.296
Political leader	0.70 (0.61)	1.14	.255	1.37 (0.61)	2.26	**.024**
Leadership level (high)	-0.51 (0.10)	-4.89	**<.001**	-0.28 (0.15)	-1.86	.063
Founder	-0.11 (0.16)	-0.72	.473	-0.04 (0.19)	-0.23	.817
Decade	-0.05 (0.03)	-1.72	.086	0.01 (0.04)	0.19	.847
Gender (women leader)	-0.58 (0.25)	-2.34	**.020**	-0.27 (0.13)	-2.04	**.042**
Gender × Decade	0.12 (0.05)	2.32	**.020**	-	-	ns[1]
Count Model						
Obituary Length	0.01 (0.00)	5.64	**<.001**	0.01 (0.00)	17.71	**<.001**
Newspaper (conservative)	-0.11 (0.13)	-0.80	.421	0.04 (0.04)	1.07	.286
Academic leader	0.22 (1.06)	0.21	.836	-0.03 (0.21)	-0.13	.894
Business leader	0.12 (1.04)	0.12	.905	0.19 (0.21)	0.94	.349
Charity leader	0.12 (1.06)	0.12	.908	-0.07 (0.22)	-0.30	.765
Political leader	0.15 (1.07)	0.14	.887	0.11 (0.21)	0.51	.612
Leadership level (high)	-0.52 (0.14)	-3.80	**<.001**	-0.09 (0.04)	-2.12	**.034**
Founder	0.01 (0.25)	0.04	.971	-0.06 (0.07)	0.85	.394
Decade	0.08 (0.06)	2.02	**.043**	0.02 (0.01)	1.49	.137
Gender (women leader)	-0.48 (0.42)	-1.15	.249	-0.08 (0.04)	-1.84	.067
Gender × Decade	0.07 (0.08)	0.94	.349	-	-	ns
Pseudo R^2	.082			.250		
Unique R^2 Gender	.000			.003		
Unique R^2 Gender × Decade	.003			-		
Hypothesis testing	**H2a: Contradicted**			**H2b: Supported**		

[1]*Note.* In the initial zero-hurdle model on **respectability**, gender (b = -0.51, z =- 1.61, p = .107) and the interaction Gender × Decade (b = 0.06, z = 0.84, p = .404) were not significant. Similarly, in the initial count model on agency, gender (b = -0.18, z = -1.57, p = .116) and the interaction Gender × Decade (b = 0.02, z = 0.94, p = .347) were not significant.

Supplement 5.

Robustness Check2: Analyses Using an Agency Dictionary That Includes Leadership Titles

The leader obituaries included 2,015 (1.17%) leadership titles, such as supervisor, manager, chairperson, and executive board member. Following the example of English agency dictionaries (CITE), in our main analyses, we did not include specific leadership titles in our German agency dictionary (we did however maintain words referring to leading more generally, such as lead, direct, guide). In the context of obituaries, however, the frequencies with which specific leadership titles are mentioned can have a somewhat ambiguous meaning. On one hand, the inclusion of many leadership titles may merely indicate that a leader had many (different) leadership roles within an organization. On the other hand, emphasizing a leaders' (many different) leadership roles may convey perceptions of agency, as leadership continues to be associated with agency. To account for the latter, we examined changes in agency (Hypothesis 1a) with an agency dictionary that includes leadership titles.

The results from these analyses, shown in the Supplement Table 7, were consistent with the results from our main analyses. In linear negative binomial regression, women were described with fewer agency words (including leadership titles) than men leaders, but this gender gap decreased over time. Non-linear negative binomial regression with two natural cubic splines revealed that the gender gap in agency (with leadership titles) only decreased in the second spline, thus from the 1980s.

Like the results reported in the main manuscript, these results support Hypothesis 1a. Although gender differences persist, from the 1980s, women leaders have been described as increasingly agentic, also upon the inclusion of leadership titles.

Supplement 5: Table 1.

Linear and Non-linear Negative Binomial Regression Predicting the Frequencies of Agency Words Including Leadership Titles

Linear	b	SE	z	p
Obituary Length	0.01	0.00	36.15	**<.001**
Newspaper (conservative)	0.11	0.03	4.02	**<.001**
Academic leader	0.17	0.15	1.14	.254
Business leader	0.39	0.15	2.63	**.009**
Charity leader	0.24	0.15	1.60	.111
Political leader	0.12	0.16	0.75	.453
Leadership level (high)	0.10	0.03	3.33	**.001**
Founder	0.17	0.04	4.03	**<.001**
Decade	0.02	0.01	3.00	**.003**
Gender (woman leader)	-0.39	0.08	-5.07	**<.001**
Gender × Decade	0.05	0.02	3.01	**.003**
Pseudo R^2		.416		
Hypothesis testing		**H1a: Supported**		
Non-linear model				
Obituary Length	0.01	0.00	36.65	**<.001**
Newspaper (conservative)	0.10	0.03	3.83	**<.001**
Academic leader	0.15	0.15	1.01	.312
Business leader	0.36	0.15	2.47	**.014**
Charity leader	0.23	0.15	1.53	.126
Political leader	0.09	0.15	0.58	.563
Leadership level (high)	0.11	0.03	3.46	**.001**
Founder	0.17	0.04	4.23	**<.001**
Decade Spline 1	0.60	0.12	4.90	**<.001**
Decade Spline 2	0.25	0.08	3.02	**.003**
Gender (woman leader)	-0.40	0.09	-4.27	**<.001**
Gender × Decade Spline 1	0.34	0.25	1.36	.173
Gender × Decade Spline 2	0.60	0.17	3.66	**<.001**
Pseudo R^2		.422		
Hypothesis testing		**H1a: Supported**		