The Leadership Quarterly xxx (xxxx) xxx



Contents lists available at ScienceDirect

The Leadership Quarterly



journal homepage: www.elsevier.com/locate/leaqua

Full length article

Text-mining obituaries between 1953 and 2019 revealed that women leaders are described increasingly like men leaders, but yet evaluated differently

Miriam K. Zehnter^{a,*}, Christoph U. Wolfmayr^b, Leona A. Andriopoulos^b, Erich Kirchler^{c,d}, Martin Voracek^b, Michelle K. Ryan^{a,e}

^a Department of Psychology, University of Exeter, UK

^b Department of Cognition, Emotion, and Methods in Psychology, Faculty of Psychology, University of Vienna, Austria

^c Department of Occupational, Economic and Social Psychology, Faculty of Psychology, University of Vienna, Austria

^d IHS-Institute for Advanced Studies, Vienna, Austria

^e Global Institute for Women's Leadership, The Australian National University, Australia

ARTICLE INFO

Keywords: Gender stereotypes Descriptions Evaluations Women leaders Obituaries Change Constructive Replication Text-mining

ABSTRACT

Replicating and extending previous research on changes in gender stereotypes in the context of leadership (Zehnter et al, 2018), we text-mined 2,283 obituaries of leaders published between 1953 and 2019 in (Western) Germany. Using a rigorously developed dictionary with substantial internal reliability, coverage, convergent, and predictive validity, we counted descriptive words signifying agency, competence, and communality alongside evaluative words signifying likability and respectability. Over time, women leaders were described more like men leaders in terms of agency and competence, but continued to be described as more communal. Moreover, women leaders were evaluated as increasingly likable, but continued to be evaluated as less respectable than men leaders. Penalizations of agency with reduced likability initially disappeared, but re-emerged after the millennial shift. Ultimately, these results highlight that despite some changes towards greater equality, disparaging views of women and men leaders persist.

Introduction

Gender stereotypes are one of the greatest obstacles to gender parity in leadership (Eagly & Heilman, 2016). Descriptive gender stereotypes generalize women as less agentic and competent than men, stipulating that women lack the traits associated with successful leadership. Concurrently, prescriptive gender stereotypes dictate that women should be communal, and that agency and competence are unlikable traits in women (Eagly & Karau, 2002; Heilman, 2001). Yet, psychological theory is not only divided over the question of whether gender stereotypes change over time, but there is also a paucity of research on how change – or a lack thereof – may affect women leaders. After all, women with leadership positions occupy a counter-stereotypical role that remains highly dominated by men (Catalyst, 2022; Edwards et al., 2020) and associated with masculine attributes (Koenig et al., 2011).

Examining gender stereotypes in the context of leadership, Zehnter and colleagues (2018) proposed that descriptions of women leaders may change over time, while their evaluations may be more stable. While this work offers an intriguing reconciliation to divided theory, significant shortcomings limit its empirical contribution. In this article, we present a constructive replication of Zehnter and colleagues' work tracking the descriptions and evaluations of women and men leaders between 1953 and 2019 by text-mining their obituaries.

Social role theory proposes that observing increasing numbers of individuals in counter-stereotypical roles, such as women leaders, changes gender stereotypes (Eagly and Steffens, 1984; Wood & Eagly, 2012). Indeed, research revealed that stereotypes about women tend to be dynamic as women's social roles change to a greater degree than those of men (Diekman & Eagly, 2000; Gustafsson Sendén et al., 2019). A meta-analysis of public opinion polls between 1946 and 2018 revealed at least some change in gender stereotypes (Eagly et al., 2020): While men's relative advantage in agency remained stable, an increasing number of people believed women to be equally or more competent than men. The stereotype of women being more communal than men became

https://doi.org/10.1016/j.leaqua.2025.101884

Received 14 June 2023; Received in revised form 17 April 2025; Accepted 2 May 2025

1048-9843/© 2025 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

^{*} Corresponding author at: Washington Singer Laboratories, University of Exeter, Perry Road, Prince of Wales Road, Exeter EX4 4QG, UK. *E-mail address*: M.K.Zehnter@exeter.ac.uk (M.K. Zehnter).

M.K. Zehnter et al.

stronger, presumably because women tend to be overrepresented in communal roles (e.g., as healthcare professionals) (Eagly et al., 2020).

Contrastingly, backlash hypothesis argues that gender stereotypes do not change, as any behavior violating stereotypes tends to be penalized (Rudman et al., 2012; Williams & Tiedens, 2016). Concretely, women who display agency tend to be disliked (Brescoll et al., 2010; Heilman et al., 2004), whereas men who display communality were disrespected (Heilman & Wallen, 2010; Moss-Racusin et al., 2010). Indeed, comparing gender stereotypes of U.S. college students in 1983 with those of U.S. adults in 2014 found very little change (Haines et al., 2016). In both years, men were described as more agentic and women as more communal. Likewise, there was stability of perceptions of what constitutes stereotypical female versus male behaviors, occupations, and physical characteristics.

Reconciling divided theory and examining gender stereotypes in the context of leadership, Zehnter and colleagues (2018) proposed that descriptions of women leaders may change over time, while their evaluations may be more stable. Re-analyzing words extracted from leader obituaries between 1974 and 2016, they found first evidence for this claim. Women, but not men leaders were described as increasingly agentic. However, agency was negatively associated with likability across the obituaries of women leaders, indicating that counterstereotypical descriptions were related to less favorable evaluations. Similarly, a recent study that examined word embeddings from Google Books across 200 years found change in the descriptions associated with women and men, but persistence in the evaluative tone of these descriptions. That is, women continued to be associated with more negative words than men (Charlesworth et al., 2022).

Zehnter and colleagues (2018) did not find significant decreases in the communality descriptions of women leaders. This was consistent with the Stereotype Content Model, according to which individuals with high status (e.g., leaders) tend to be ascribed agency and communality (Fiske, 2018; Fiske et al., 2007). Other recent research also showed that in organizations with women leaders, women were more strongly associated with agency without reducing associations with communality (Lawson et al., 2022).

While Zehnter and colleagues (2018) offered an intriguing solution to one of psychology's open questions, significant shortcomings limit their empirical contribution. Specifically, they based their conclusions on the re-analysis of previously collected and aggregated data, limiting their ability to perform robust statistical tests. Most importantly, this approach did not allow them to examine temporal change of the hypothesized associations of leader descriptions (e.g., as agentic) with leader evaluations (e.g., as likable). Moreover, the authors did not include any control variables in their models, limiting their ability to control for potential confounds and uncover alternative mechanisms.

Finally, like most previous research on backlash against counterstereotypical women (Rudman et al., 2012; Williams & Tiedens, 2016), Zehnter et al (2018) only examined the likability of women leaders. In the context of leadership however, assessing the respectability of women would be of particular interest. Walking on a tightrope, women leaders must not only adhere to the standards of ideal womanhood set by gender stereotypes, but also to the standards of ideal – masculine – leadership. Failing in the former may cost women leaders likability, while failing in the latter may cost them respectability. That is, women leaders who fail to fulfil the masculine role of leadership may be penalized with a loss of respect in the same way that men who fail to perform masculinity are penalized by losing respect (Bosak et al., 2018; Heilman & Wallen, 2010; Moss-Racusin et al., 2010).

Present research

The aim of this research was to conduct a constructive replication of Zehnter et al (2018) and undertake a thorough investigation of changes in the descriptions and evaluations of women and men in the context of leadership. Building on the strengths of the replicated work, we used

obituaries dedicated to leaders as data source. Unlike opinion polls and questionnaires, obituaries allow for an unobtrusive analysis avoiding issues related to response biases among research participants (Podsakoff et al., 2003; Webb et al., 2000). Moreover, obituaries tend to reflect gender inequalities. In past work, fewer and shorter obituaries were dedicated to women (Moremen & Cradduck, 1999), fewer lines were dedicated to women's careers (Ogletree et al., 2005), and women were described as less able leaders than men (Curşeu & Boroş, 2011). Alongside the descriptions of the deceased leaders' character and demeanor, obituaries also contain evaluative information on whether a deceased leader was liked and appreciated. Drawing on the Stereotype Content Model (Cuddy et al., 2007), the lack of positive evaluations in obituaries may resemble passive harm (e.g., neglecting to acknowledge a leader's achievement) and can thus be used to study backlash against women leaders (Zehnter et al., 2018).

Following the principles of substantial constructive replication (Köhler & Cortina, 2021), the present work was designed to overcome the shortcomings of the replicated research. Specifically, we (1) introduced a distinction between likability and respectability to examine leader evaluations with more nuance, (2) we collected a new and larger sample of obituaries, (3) included obituaries from a longer time span, 1953 – 2019, and (4) collected obituaries in more narrow time intervals (every three instead of every six years). Thus, unlike Zehnter and colleagues' aggregated dataset, our data was based on individual cases, which allowed us to conduct more robust and fine-grained analyses. To control for potential confounds and alternative mechanisms, we also introduced additional variables, that is, (a) the political leaning of the newspapers in which the obituaries were published (conservative versus liberal), (b) the context of leadership (i.e., academia, business, charity, politics), (c) the level of leadership position (i.e., leaders of organizations versus department leaders), (d) whether a leader had founded the organization they had led, and (e) the masculinity of the leaders' industries.

Most importantly, our approach to extract gender stereotypes from leader obituaries was radically different. Using an inductive approach, Zehnter and colleagues (2018) re-categorized 58 previously derived word groups into categories representing gender stereotypes (see also, Kirchler, 1992; Rodler et al., 2001). In contrast, in this research, we chose a theory-driven, deductive approach. Hereto, we rigorously developed a text-mining dictionary and conducted three studies to assess its internal reliability, coverage, and convergent and predictive validity. On the basis of this dictionary, we utilized cutting-edge methodology to text-mine leader obituaries across seven decades following best-practice recommendations for text-mining (Banks et al., 2018). With this approach, we join a growing body of research that utilizes modern technology to study changes in stereotypes (Charlesworth et al., 2022; Garg et al., 2018).

Despite replicating past work, the present research is, to the best of our knowledge, first to (1) conduct a thorough analysis of changes in the descriptions and evaluations of women and men in the context of leadership and (2) assess evaluations of women leaders not only in terms of likability, but also respectability. This work is also at the forefront of using modern text-mining technology to study gender stereotypes and covers an unusually long time-span of seven decades.

Hypotheses: Broadly, we expected that, over time, women leaders would be described as increasingly agentic and competent, while they would continue to be evaluated as less likable and respectable than men leaders. Moreover, we expected that adhering to the masculine ideals of leadership (i.e., agency, competence) would cost women leaders likability, but gain them respect. Vice versa, showing feminine qualities (i.e. communality) should gain women leaders likability, but cost them respect. Concretely, this should translate into (1) negative associations of agency (respectively competence) with likability, (2) positive associations of agency (respectively competence) with respectability, (3) positive associations of communality with likability, and (4) negative associations of communality with respectability among women leaders.

M.K. Zehnter et al.

Table 1

Overview Over the Tested Hypotheses and Conclusions Based on the Research Presented Below

Hypotheses	Conclusions
Descriptions of Leaders <i>Hypothesis 1a</i> Women leaders are described with lower agency than men leaders (main effect of gender), but the magnitude of this difference decreases over time (significant interaction Gender × Decade).	 Supported by linear and non-linear negative binomial regression Contradicted upon controlling for masculinity of industry (Table 12) Contradicted by hurdle regression (OS 4) Supported by negative binomial regression with an agency dictionary with leadership titles (OS 5) Significant main effect of gender (Table 8) Significant interaction Gender × Decade (Table 8)
Hypothesis 1b Women leaders are described with lower competence than men leaders (main effect of gender), but the magnitude of this difference decreases over time (significant interaction Decade \times Gender).	 Supported by linear and non-linear negative binomial regression Supported upon controlling for masculinity of industry (Table 12) Supported by hurdle regression (OS 4) Significant main effect of gender (Table 8) Significant interaction Gender × Decade (Table 8)
Hypothesis 1c Women leaders are described with greater communality than men leaders (main effect of gender) and the magnitude of this difference does not change over time (non- significant interaction Decade \times Gender).	 Supported by linear and non-linear negative binomial regression Contradicted upon controlling for masculinity of industry (Table 12) Supported by hurdle regression (OS 4) Significant main effect of gender (Table 8) Non-significant interaction Gender × Decade (Table 8)
Evaluations of Leaders	
Hypothesis 2a Women leaders are described with lower likability than men leaders (main effect of gender) and the magnitude of this difference does not change over time (non- significant interaction Decade \times Gender).	 Contradicted by linear and non-linear negative binomial regression Supported upon controlling for masculinity of industry (Table 12) Contradicted by hurdle regression (OS 4) Significant main effect of gender (Table 10) Significant interaction Gender × Decade (Table 10)
Hypothesis 2b Women leaders are described with lower respectability than men leaders (main effect of gender) and the magnitude of this difference does not change over time (non- significant interaction Decade \times Gender).	 Supported by linear and non-linear negative binomial regression Supported upon controlling for masculinity of industry (Table 12) Supported by hurdle regression (OS 4) Significant main effect of gender (Table 10) Non-significant interaction Gender × Decade (Table 10)
Associations Between Descriptions and Evaluations	
In women leaders' obituaries, likability is associated <i>Hypothesis 3a</i> : negatively with agency.	<i>Supported</i> by partial correlation, controlling for obituaries length By decade: Significant negative correlation in the 1050 s, 2000er, and 2010er years (Table 11)
Hypothesis 3b: negatively with competence.	Contradicted by partial correlation By decade: Significant negative correlation in 1950 s (Table 11)
Hypothesis 3c: positively with communality.	Supported by partial correlation
In women leaders' obituaries, respectability is associated	By decade: Significant positive correlation in 2010er years (Table 11)
Hypothesis 4a: positively with agency.	<i>Contradicted</i> by partial correlation By decade: Significant positive correlation in 1980 s (Table 11)
Hypothesis 4b: positively with competence.	<i>Contradicted</i> by partial correlation By decade: Significant positive correlation in 1980 s (Table 11)
Hypothesis 4c: negatively with communality.	Contradicted by partial correlation.

While we did not formulate hypotheses for these associations in the obituaries of men leaders, we explored them for comparison. Table 1 specifies the hypotheses and provides a directory for the results presented below.

Method

Identification of obituaries

Between 1953 and 2019, we collected leader obituaries in three-year intervals (1953, 1959, and so on). The obituaries were published in two widely circulated, daily (Western) German newspapers, the Frankfurter Allgemeine Zeitung (conservative) and the Sueddeutsche Zeitung (liberal). Both newspapers have a long history of publishing obituaries in specific sections, in which organizations can pay to publish obituaries.

Hence, the obituaries analyzed here were announced by the organizations in which the deceased had held leadership positions. They stemmed from different organizations, including business firms, academic and political institutions, and charity organizations, and covered a range of different industries (e.g., construction, finance and insurance, health and social services). Thus, rather than covering publicly highly visible leaders (e.g., state leaders) with obituaries that are curated by the editorial staff of a newspaper, the obituaries analyzed in this research were dedicated to women and men with a wide range of different leadership roles.

These were the inclusion criteria for leadership obituaries: First, the obituary had to be published by the organization in which the deceased leader had held a leadership position. Second, the obituary had to explicitly state that the deceased had held a formal leadership position. This includes CEO's, managers, chairmen and –women, directors, heads

The Leadership Quarterly xxx (xxxx) xxx

Table 2

Exemplary Obituaries of Academic Leaders, Both Published in the Liberal Newspaper in 2004 by Major Medical Associations.

Woman Leader	Man Leader
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 loveable 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 with the *Name of Award* from the *Name of Awarding Institution*.	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, level-headed 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.

¹Note: The obituaries were translated from German by the first author.

of departments, supervisors, etc., but excludes informal leadership roles, such as team leaders. To select obituaries, we used a systematic random sampling strategy (Neuendorf, 2002), which was developed in previous obituaries research (Kirchler, 1992). For each year and newspaper, we screened 26 issues for obituaries dedicated to leaders. Specifically, we screened the Monday issue of the second calendar week, the Tuesday issue of the fourth calendar week, and so forth. As the number of obituaries written for women leaders were very small (N = 69), we conducted a second search screening all issues of the relevant years for additional obituaries of women leaders. In total, we identified 2,941 obituaries, dedicated to 2,283 leaders. Seeking to include only one obituary per leader, we selected one randomly for leaders with two or more obituaries. Table 2 shows two exemplary obituaries (one for a woman leader, one for a man leader).

Sample

In total, we analyzed 2,283 leader obituaries between 1953 and 2019. Eight hundred eight (35.39%) obituaries were dedicated to

Table 3

Number of Leader Obituaries.

women leaders and 1,475 (64.61%) to men leaders. Notably, only 69 obituaries of women leaders were found when screening randomly selected newspaper issues. Seven hundred thirty-nine additional obituaries were found in an extended search including all newspaper issues in the years of analysis. Thus, our analyses are based on the full population of women leaders' obituaries and a random sample of men leaders' obituaries for the years of analysis. Nevertheless, the numbers of obituaries for women leaders remained low, especially in the early years of analysis. To increase test power, we thus chose to analyze the obituaries' content by decade, rather than year. A Poisson regression indicated a significant increase in the obituaries of women leaders over the seven decades of analysis, b = 0.16, z = 8.93, p < .001, $R^2 = .60$.

Obituaries were predominantly dedicated to business leaders heading business or law firms (75.47%) and leaders with higher-level leadership positions, that is, leaders of organizations, corporate board members, and senior leaders (73.41%). Few obituaries (10.34%) named the deceased as founder of the organization they had led. Many obituaries (40.78%) were written for leaders in men-dominated industries (more than 70% men in 2019); fewer obituaries (18.40%) were written

Decade	Wom N	en Leaders N = 808	Men leaders $N = 1,475$
1950 s (1953, 1956, 1959)	9 ^a	65	179
1960 s (1962, 1965, 1968)	8	81	231
1970 s (1971, 1974, 1977)	2	56	187
1980 s (1980, 1983, 1986, 1989)	16	165	335
1990 s (1992, 1995, 1998)	16	142	213
2000er (2001, 2004, 2007)	4	113	165
2010er (2010, 2013, 2016, 2019)	14	186	165
Newspaper		N (%)	N (%)
Frankfurter Allgemeine Zeitung (conservative)	37-	4 (46.29)	912 (61.83)
Sueddeutsche Zeitung (liberal)	43	4 (53.71)	563 (38.17)
Type of leader			
Academic (heads of schools, universities, and research institutes)	95	5 (11.76)	130 (8.81)
Business (leaders of business firms, partners in law firms)	53	0 (65.59)	1,193 (80.88)
Charity (leaders of charity foundations, religious leaders)	12	5 (15.47)	63 (4.27)
Political (leaders of governmental institutions, parliament, majors)	5	0 (6.19)	78 (5.29)
Level of leadership			
Higher-level (leaders of organizations, board members, senior leaders)	56	1 (69.43)	1,115 (75.59)
Lower-level (department leaders)	24	7 (30.57)	360 (24.41)
Founder	10	2 (12.62)	134 (9.08)
Masculinity of industry (by % men in 2019)			
Men dominated (> 70 % men: manufacturing, transport and storage, mining, energy, water, and waste disposal, construction)	24	8 (30.69)	683 (46.31)
Gender balanced (30-70 % men: e.g., finance and insurance, trade and maintenance of motor vehicles, information and communication)	28	8 (35.64)	532 (36.07)
Women dominated (<30 % men: health and social services, school and education)	22	4 (27.27)	196 (13.29
Information not available	4	8 (5.94)	64 (4.34)
Mean Length of Obituaries (SD)	66.4	40 (37.93)	80.68 (40.36)

^a Note: Number of obituaries before screening all newspaper issues in the years of analysis.

M.K. Zehnter et al.

ARTICLE IN PRESS



Fig. 1. Flowchart of the Development and Quality Assessment of the Text-mining Dictionary.

for leaders in women-dominated industries (more than 70% women in 2019). Only 67 percent of obituaries included the leaders' age, which mean was 73.34 years (SD = 15.54) for women and 71.02 years (SD = 12.89) for men. Notably, 2,240 (98.12%) of obituaries were written by different organizations. Only nine organizations wrote several obituaries, with the highest numbers of obituaries written by the Max-Planck Research Institute (N = 14) and Pharmaceutical Producer Bayer AG (N = 14).

Overall, the obituaries contained a total of 172,659 words. Between the 1950s and the 2010er years, the mean length of obituaries of women leaders increased from 56.51 (SD = 28.52) to 83.26 (SD = 45.31) words. For men leaders, it increased from 74.71 (SD = 29.93) to 97.64 (SD = 57.78) words. Ultimately, in the 2010er years, the obituaries of women leaders were about 15 percent shorter than those of their men counterparts. Poisson regression revealed an increasing length of obituaries over time and longer obituaries in the conservative (versus liberal

M.K. Zehnter et al.

newspaper), for political (versus all other) leaders, for leaders with higher level (versus lower level) leadership positions, for non-founders (versus founders), and for more men-dominated industries. Women's obituaries were shorter than men's, b = -0.42, z = -28.51, p < .001, but this gender gap decreased over time, as indicated by the interaction Gender × Decade, b = 0.05, z = 16.09, p < .001. Table 3 shows the number of leader obituaries by gender, decade, newspaper, type of leader, level of leadership, being a founder, and the masculinity of industry (in 2019) in which the leaders were active.

Identification of relevant words through text-mining

We identified relevant words using text-mining following bestpractice recommendations (Banks et al., 2018) and using the R-package UDPipe (Wijffels, 2021). As a basis for the text-mining, we created a dictionary including the five categories agency, competence, communality, likability, and respectable. Following recommendations for the development of gender stereotype dictionaries (Nicolas et al., 2021), we assessed the quality of this dictionary through indicators, such as its internal reliability and coverage, and conducted three studies to assess its convergent and predictive validity. Notably, the creation and quality assessment of the text-mining dictionary was an iterative process including several rounds of assessment and improvement. Fig. 1 visualizes this process and below, we summarize key results. The online supplement provides a detailed description of the creation and validation of our text-mining dictionary alongside a qualitative demonstration of the text-mining procedure.

The final dictionary included 2,622 nouns, adjectives, and verbs with 1,468 unique roots – 35.56 percent signified agency, 19.35 percent competence, 32.70 percent communality, 4.77 percent likability, and 7.23 percent respectability. Here, we present key results concerning the quality of the *final* text-mining dictionary, which had substantial internal reliability, Cohen's Kappa = .61 [.58,.64].

Coverage. Coverage is the overall proportion of words that textmining identifies in comparison to human coders. To satisfy this criterion, our text-mining dictionary should identify at least 90 percent of context-relevant words identified by a human coder, but count less than 10 percent context-irrelevant words. 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. To assess its coverage, we examined whether text-mining a random sub-sample of 50 obituaries (Study 2) using the final dictionary would yield similar results than using traditional human coding. In comparison to a human coder (baseline), text-mining identified 94 percent agency, 95 percent competence, 90 percent communality, 100 percent likability, and 103 percent respectability words. 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. To meet the criterion of convergent validity, there should be substantial agreement, intra-class-correlations (ICC) > .75, between the specific words identified through text-mining versus a

human coder. Comparing text-mining and human coding in a random sub-sample of 50 obituaries (Study 2) revealed very strong agreements between the two methods for agency (ICC = .96), competence (ICC = .91), communality' (ICC = .90), likability (ICC = .93), and respectability (ICC = .96).

Predictive validity. As a last step, we tested the predictive validity of the text-mining dictionary (Study 3), examining whether the word counts obtained through text-mining predicted how research participants assessed leaders based on their obituaries. Using stratified sampling (Neuendorf, 2002), we randomly selected four obituaries for women leaders and four obituaries for men leaders, which varied in the frequencies of words signifying agency, competence, communality, likability, and respectability. We revised these obituaries to conceal the leaders' gender. Then, 200 participants (36% women, 67% men) each read one obituary and assessed the deceased leader based on the five dimensions of the 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 also asked participants to identify the leader's gender based on their obituary (financially rewarding correct responses).

To test whether the word counts obtained from text-mining would predict research participants' assessment of the deceased leaders, we used 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 textmining 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 indicats 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 indicats that participants assessed leaders based on their own gender stereotypes.

At large, the text-mining counts were significant predictors of participants' assessment of the deceased leaders as agentic, competent, communal, and likable, but not as respectable (See Table 4 for the results). Notably, the leaders' true gender also predicted participants competence assessments, such that men leaders were rated as more competent than women leaders. This indicates that obituaries include gendered competence signals (e.g., job titles) that are not captured by our text-mining approach. Participants' perceptions of leaders' gender 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 that participants' assessments are influenced by their own gender stereotypes and potentially a 'women are wonderful' bias (Eagly et al., 1991; Rosette & Tost, 2010). Neither textmining counts, nor leaders' true gender, nor perceived leaders' gender predicted participants' assessment of leaders as respectable. In general, respectability ratings were high across all obituaries. Just receiving an obituary may have signaled respectability to research participants.

Table 4

General Linear Regression Analyses Predicting Research Participants' Assessment of Leaders.

	Age	ency	Comp	etence	Comm	unality	Likab	oility	Respect	ability
	z	р	z	р	z	р	z	р	Z	р
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).

The Leadership Quarterly xxx (xxxx) xxx

M.K. Zehnter et al.

Data analysis

Upon the completion of text-mining, we tested our hypotheses. Following the replicated article (Zehnter et al., 2018), we tested changes in the descriptions (Hypotheses 1a-c) and the evaluations of leaders (Hypotheses 2a-b) with linear negative binomial regressions using the R package 'MASS' (Venables & Ripley, 2002). Negative binomial regressions can handle count data where the observed variance exceeds the mean counts (over-dispersion) and are thus a suitable analysis strategy for zero-inflated data in which the zeros are part of a continuous count variable (Lee et al., 2023). In addition, we conducted non-linear negative binominal regression including natural cubic splines, which allowed us to identify inflection points where temporal trends accelerate or decelerate (Shumway & Stoffer, 2017). In each non-linear model, we tested the inclusion of two, three, four, and five natural cubic splines that were equally distributed across our observed time-span of seven decades, and chose the model with the best fit, or the simpler model in case that model fit was similar.

In each linear and non-linear model, we regressed the outcome variable (agency, competence, communality, likability, respectability) on gender, decade, and the interaction Gender \times Decade, and included several control variables. When Gender \times Decade was not statistically significant, we removed the interaction from the regression equation to avoid that non-significant regression terms suppress meaningful main effects (Beck & Bliwise, 2014; Engqvist, 2005).

Support for our hypotheses was indicated by statistically significant predictors (gender, Gender × Decade), whereby statistical significance was defined as probability of less than 5 percent that the observed effects were due to chance (p < .05) (Dahiru, 2011). Pseudo R-squares indicated the total amount of variance explained by our linear and non-linear models. To calculate the unique effect sizes of our predictors, we calculated differences in R-squares between the models including and excluding them (Aguinis & Gottfredson, 2010).

To test the hypothesized associations between descriptive and evaluative words (Hypotheses 3a-c, 4a-c) we calculated Partial Pearson's product-moment correlations, controlling for the length of obituaries. For each predicted association, we first examined its overall correlation (across the obituaries of all women / men leaders), followed by analyses by decade.

For a subset of obituaries, we could identify to which industry the organization that published the obituary belonged. To explore whether gender gaps in word frequencies and changes thereof depended on the masculinity (i.e., men's dominance) of the industries in which the deceased leaders were active, we repeated linear negative binomial

Table 5

(

	Correlation Matrix	of All	Variables	Included	in Hy	potheses	Testing	and	Robustness	Checks.
--	--------------------	--------	-----------	----------	-------	----------	---------	-----	------------	---------

			51		0										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Agency	1														
Competence	.39	1													
Communality	.34	.28	1												
Likability	.06	.09	.20	1											
Respectability	.35	.33	.26	.07	1										
Gender (women leader)	12	23	02	02	14	1									
Decade	.18	.00	.00	.07	.06	.18	1								
Length	.64	.56	.51	.20	.51	17	.13	1							
Newspaper (conservative)	.20	.18	.14	02	.12	15	07	.23	1						
Academic leader	03	.02	09	.02	05	.05	.21	.02	08	1					
Business leaders	.00	.04	.02	03	.06	17	30	08	.11	58	1				
Charity leader	.01	06	.07	.03	07	.19	.17	.03	01	10	53	1			
Political leader	.03	02	01	.01	.05	.02	.08	.07	08	08	43	07	1		
Leadership level	.07	.02	.04	12	.00	07	01	.09	.21	07	.05	.06	06	1	
Founder	.06	03	02	04	04	.06	.10	06	.05	.06	09	.13	05	.19	1
Masculinity of industry	.02	.04	.03	05	.06	20	31	02	.16	48	.74	38	27	.08	03

p < .05 are displayed in bold

regression analyses on the outcome variables (agency, competence, communality, likability, respectability). In these models. We included the Masculinity of Industry as a main factor, the two-way interactions Gender \times Masculinity of Industry and Decade \times Masculinity of Industry, and the three-way interaction Gender \times Decade \times Masculinity of Industry.

To test the robustness of our results, we re-tested parts of our hypotheses (Hypotheses 1a-c, and 2a-b) using negative binomial hurdle regressions to account for the zero-inflation present in our data. Hurdle regressions combine a binomial zero-inflated model predicting the odds of zero versus at least one count with a model predicting counts above zero (Blevins et al., 2015; Feng, 2021). While negative binomial regression (which treats zero counts as part of a natural count variable) maintains greater variability in the count data, the logistic zero-hurdle model is less susceptible to the influence of extreme cases (as all counts are coded as one). Moreover, Hurdle regression allowed us to include the same predictors and control variables in both the zero-hurdle and the count model.

In our main analyses, we did not include leadership titles (e.g., supervisor, manager, chairperson, executive board member) in our agency dictionary (following other agency dictionaries). As robustness check, we also examined changes in agency (Hypothesis 1a) with an agency dictionary that includes leadership titles.

Our data included few outliers (see online supplement), which were maintained in the analyzed data. Our online repository (https://osf. io/7p29f/?view_only https://osf.io/7p29f/?view_only=07061ecd9 ddf4f6e814b55963dccde36) includes the anonymized data (excluding the original German obituaries, which are available upon request from the corresponding author) and the R code underlying the text-mining and data analyses.

Results

A first look at the correlations between all variables (Table 5), revealed that, overall, being a woman leader was negatively associated with agency, competence, and respectability, but not with communality and likability. Women leaders received increasing numbers of obituaries over time, but shorter obituaries, and fewer obituaries in the conservative newspaper. Women were less often business leaders and more often academic and charity leaders; they held more often lower-level leadership positions and less often positions in men-dominated industries. Table 6 provides an overview of the absolute and relative frequencies of descriptive and evaluative words in obituaries by gender and decade.

and Darada

hir Condor

and Evoluative Words

of Morde Signifizing Deceriptive

nod Dalative Erect

Table

		Number ¿	and Lengt	h		Age	ncy			Compet	tence			Commui	ıality			Likabil	ity			Respectal	bility	
		ML		ML	M	L	MI		ML		ML	_	ML		ML		ML		ML		ML		ML	
Decade	z	M (SD)	z	M (SD)	M (SD)	%	M (SD)	%	M (SD)	%	M (SD)	%	M (SD)	%	M (SD)	%	M (SD)	%	M (SD)	%	M (SD)	%	M (SD)	%
1950 s	65	56.51 (28.52)	179	74.71 (29.93)	1.11 (1.28)	1.79	1.88 (1.79)	2.36	0.77 (1.13)	1.27	1.66 (1.54)	2.15	2.00 (1.92)	3.32	2.32 (1.98)	3.12	0.34	0.78	0.51	0.70	1.54 (1.28)	2.83	1.93 (1.40)	2.67
1960 s	81	59.04 (28.11)	231	83.78 (36.45)	1.41	2.17	2.31	2.64	0.74	1.13	1.88	2.13	2.06	3.53	2.22 (1.95)	2.66	0.43	0.74	0.52	0.62	1.70	3.08	2.23 (1.55)	2.85
1970 s	56	51.14 (24.62)	187	74.04	1.25	2.06	2.29	2.88	0.89	1.45	(1.42) (1.42)	1.86	(1.62)	2.80	1.77 (1.78)	2.27	0.25	0.75	0.39	0.56	1.41	2.78	1.95 (1.25)	2.85
1980 s	165	54.16 (28.08)	335	76.13	1.54 (1.64)	2.46	2.55	3.19	0.55	0.82	1.70	2.11	1.41	2.39	1.70	2.21	0.28	0.65	0.44	0.61	1.44	2.85	1.99 (1.46)	2.74
1990 s	142	(35.22)	213	79.88 (38.40)	2.08 (1.95)	2.92	2.85 (2.12)	3.44	0.97	1.23	1.53 (1.54)	1.90	1.48	2.11	1.58 (1.54)	2.03	0.44	0.76	0.76)	0.65	1.45 (1.31)	2.39	2.23 (1.48)	2.93
2000er	113	77.74 (43.47)	165	83.70 (43.86)	2.73 (2.46)	3.35	2.87 (2.20)	3.54	1.15 (1.40)	1.28	1.67	1.92	2.21	2.68	1.95 (2.05)	2.25	0.63	1.00	0.51	0.66	1.92	2.63	2.15 (1.79)	2.64
2010er	186	83.26 (45.31)	165	97.64 (57.78)	2.73 (2.43)	3.00	3.27 (2.65)	3.40	1.25 (1.44)	1.39	1.83 (1.77)	1.92	2.28 (2.28)	2.57	2.28 (1.95)	2.48	0.64 (1.01)	0.85	0.62 (1.07)	0.69	2.06 (1.66)	2.69	2.47 (1.79)	2.80
Note. WL	= Wome	en leaders, N	ML = Me	n leaders.																				

The Leadership Quarterly xxx (xxxx) xxx

Table 7
Most Frequently Used Description

Agency (N)	Competence (N)	Communality (N)
leader / lead/ guide (657)	knowledge / know (282)	humane (367)
action / active (452)	(life's) achievement (278)	connect / connectedness (337)
commitment / committed (392)	professional (255)	service /serve (222)
dedication / dedicate (355)	outstanding (244)	family / familial (151)
influence / exert influence (307)	experience / experienced (218)	loyal / loyalty (137)
relentless (255)	merit (210)	candor / candid (133)
to shape / shaping (229)	being capable (166)	social (125)
creator / create / creative (214)	skill (137)	contribute (93)
entrepreneurial (209)	foresight / foresighted (120)	understanding (88)
decisive (194)	academic (109)	compassion / compassionate (82)

Changes in the descriptions of leaders

First, we present changes in the descriptions of leaders as agentic, competent, and communal. Table 7 shows the most frequently used descriptions. Fig. 2 visualizes these changes, showing linear trends over time.

Hypothesis 1a: Agency. Overall, 5,393 (3.12%) words signified agency. Across seven decades, the mean count of agency words in the obituaries of women leaders increased from 1.11 (1.79%) to 2.73 (3.00%). In the obituaries of men leaders, the mean count of agency words also increased from 1.88 (2.36%) to 3.27 (3.40%). In linear negative binomial regression, women leaders were described with fewer agency words, but this gender gap decreased over time as indicated by the significant interaction Gender × Decade. Non-linear negative binomial regression with two natural cubic splines further revealed that the gender gap in agency decreased only in the second spline, thus from the 1980s. Table 8 shows the results from linear and non-linear regression analyses.

These results support Hypothesis 1a. While gender differences persist, from the 1980s, women leaders have been described as increasingly agentic.

Hypothesis 1b: Competence. A total of 3,235 (1.87%) words signified competence. Across seven decades, the mean count of competence words in the obituaries of women leaders increased from 0.77 (1.27%) to 1.25 (1.39%). In the obituaries of men leaders, the use of competence words changed little from 1.66 (2.15%) to 1.83 (1.92%). In linear negative binomial regression, women leaders were described with fewer competence words, but this gender gap decreased over time as indicated by Gender × Decade. Non-linear negative binomial regression with two natural cubic splines showed that, again, decreases in the gender gap of competence only occurred in the second spline, from the 1980s. See Table 8 for the results.

These results support Hypothesis 1b. Although gender differences persist, from the 1980s, women leaders have been described as increasingly competent.

Hypothesis H1c: Communality. Overall, 4,366 (2.53%) words signified communality. Across seven decades, the mean count of communality words in the obituaries of women leaders changed from 2.00 (3.32%) to 2.28 (2.57%). In the obituaries of men leaders, the mean count of communality words changed from 2.32 (3.12%) to 2.28 (2.48%). In the initial linear regression model, there were neither gender differences in the frequencies of communality words, nor changes thereof over time. Upon removing the non-significant interaction Gender \times Decade, the main effect of gender emerged as statistically significant, with women being described with more communality words than men. Non-linear regression analyses mirrored these results (See

The Leadership Quarterly xxx (xxxx) xxx

Women Leaders Men Leaders



Fig. 2. Absolute and Relative Mean Frequencies of Agency, Competence, and Communality Words by Gender and Decade with Linear Regression Lines.

Table 8).

These results support Hypothesis 1c. Gender differences in communality descriptions persist, such that women leaders are described as more communal than their men counterparts.

Changes in the evaluations of women and men leaders

Next, we present changes in the evaluations of women and men leaders as likable and respectable. Table 9 shows the most frequent evaluations. Fig. 3 visualizes these changes.

Hypothesis 2a: Likability. Overall, 1,085 (0.63%) words signified likability. Across seven decades, the mean counts of likability words in the obituaries of women leaders increased from 0.34 (0.78%) to 0.64 (0.85%). In the obituaries of men leaders, the mean counts of likability words changed little from 0.51 (0.70%) to 0.62 (0.69%). In linear negative binomial regression, women leaders were described with fewer likability words than men leaders, but this gap decreased over time as indicated by Decade × Gender. Non-linear negative binomial regression with two natural cubic splines further revealed that, again, the gender gap in likability decreased only in the second spline (from the 1980s). Table 10 shows the results from linear and non-linear regression analyses.

These results contradict Hypothesis 2a. Although gender differences persist, from the 1980s, women leaders have been described as increasingly likable.

Hypothesis 2b: Respectability. A total of 4,490 (2.60%) words signified respectability. Across seven decades, the mean counts of respectability words in the obituaries of women leaders changed from

1.54 (2.83%) to 2.06 (2.69%). In the obituaries of men leaders, respectability changed from 1.93 (2.67%) to 2.47 (2.80%) words. In the initial linear regression, there were neither gender differences in the frequencies of respectability words, nor changes thereof over time. Upon removing the non-significant interaction Gender \times Decade, the main effect of gender emerged as statistically significant, with women leaders being described with less respectability words than men leaders. Non-linear negative binomial regression with two natural cubic splines mirrored these results. See Table 10.

These results support Hypothesis 2b. Significant gender differences persist, such that women leaders are described with fewer respectability words and this gender gap did not decrease over time.

Associations between descriptive and evaluative words

To test our hypotheses about the associations of descriptive and evaluative words, we conducted partial Pearson's product-moment correlation, controlling for the length of obituaries. For each association, we first examined the overall correlation (across the obituaries of all women leaders), followed by analyses by decade. Table 11 shows all correlations by gender and decade and Fig. 4 visualizes them.

Likability (Hypotheses 3a-c). Supporting Hypothesis 3a, in the obituaries of women leaders, agency and likability correlated negatively, r(806) = -.16, p < .001. Analyses by decade specified that the agency-likability association was significantly negative in the 1950s and again in the 2000er and 2010er years. Contradicting Hypothesis 3b, competence and likability were uncorrelated, r(806) = -.04, p = .272. Analyses by decade found the predicted negative association only in the

Table 8

Linear and Non-linear Negative Binomial Regression Predicting the Frequencies of Descriptive Words.

		Agency		C	ompetence		Co	mmunality	
Linear model	b (SE)	Z	р	b (SE)	Z	р	b (SE)	Z	р
Obituary Length	0.01 (0.00)	32.30	<.001	0.01 (0.00)	27.40	<.001	0.01 (0.00)	26.84	<.001
Newspaper (conservative)	0.16 (0.03)	5.58	<.001	0.17 (0.04)	3.72	<.001	0.06 (0.04)	1.66	.097
Academic leader	0.12 (0.19)	0.62	.534	0.71 (0.27)	2.65	.008	-0.17 (0.21)	-0.83	.407
Business leader	0.36 (0.18)	2.01	.044	0.67 (0.26)	2.55	.011	0.17 (0.20)	0.87	.382
Charity leader	0.21 (0.19)	1.11	.268	0.44 (0.27)	1.61	.108	0.24 (0.21)	1.18	.239
Political leader	0.19 (0.19)	0.99	.324	0.35 (0.28)	1.28	.201	-0.08(0.21)	-0.37	.715
Leadership level (high)	-0.05 (0.04)	-1.23	.218	-0.15 (0.05)	-3.12	.002	-0.06 (0.04)	-1.34	.182
Founder	0.28 (0.05)	5.49	<.001	0.08 (0.07)	1.18	.239	0.06 (0.06)	0.97	.332
Decade	0.09 (0.02)	5.89	<.001	0.04 (0.02)	-3.25	.001	-0.04 (0.01)	-3.80	<.001
Gender (women leader)	-0.32 (0.10)	-3.31	<.001	-0.77 (0.13)	-6.03	<.001	0.14 (0.04)	3.36	<.001
$\textbf{Gender} \times \textbf{Decade}$	0.04 (0.02)	2.30	.021	0.09 (0.03)	3.41	.001	-	-	ns ^a
Pseudo R ^b		.341			.280			.223	
Unique R ^b Gender		.003			.017			.005	
Unique \textbf{R}^{b} Gender \times Decade		.002			.002			.000	
··· ··		1			a . 1			1	

Hypotheses testing Non-linear model	H1a	: Supported		H1b	: Supported		H1c	: Supported		
	b (SE)	z	р	b (SE)	z	р	b (SE)	z	р	
Obituary Length	0.01 (0.00)	32.47	<.001	0.01 (0.00)	27.11	<.001	0.01 (0.00)	26.15	<.001	
Newspaper (conservative)	0.15 (0.03)	4.44	<.001	0.17 (0.04)	3.74	<.001	0.08 (0.04)	1.95	.052	
Academic leader	0.10 (0.19)	0.55	.586	0.71 (0.27)	2.65	.008	-0.15 (0.21)	-0.75	.455	
Business leader	0.34 (0.18)	1.91	.056	0.67 (0.26)	2.54	.011	0.19 (0.20)	0.98	.211	
Charity leader	0.20 (0.19)	1.07	.287	0.43 (0.27)	1.58	.190	0.24 (0.20)	1.15	.250	
Political leader	0.17 (0.19)	0.89	.373	0.35 (0.28)	1.26	.210	-0.06 (0.20)	-0.31	.758	
Leadership level (high)	-0.04 (0.04)	-1.14	.253	-0.15 (0.05)	-3.11	.002	-0.06 (0.04)	-1.46	.145	
Founder	0.28 (0.05)	5.63	<.001	0.08 (0.07)	1.16	.248	-0.05 (0.04)	-1.28	.201	
Decade spline 1	0.73 (0.16)	4.68	<.001	-0.10 (0.18)	-0.56	.578	-0.80 (0.14)	-5.75	<.001	
Decade spline 2	0.53 (0.11)	5.04	<.001	-0.43 (0.13)	-3.26	.001	-0.37 (0.10)	-3.80	<.001	
Gender (women leader)	-0.30 (0.12)	-2.57	.010	-0.55 (0.14)	-3.84	<.001	0.12 (0.04)	3.05	.002	
Gender \times Decade spline 1	0.23 (0.31)	0.72	.474	-0.11 (0.40)	-0.28	.777	-	-	ns	
Gender \times Decade spline 2	0.55 (0.21)	2.67	.008	0.77 (0.26)	2.95	.003	_	-	ns	
Pseudo R ^b		.344			.281			.227		
Unique R ^b Gender		.002			.018			.002		
Unique R^b Gender $ imes$ Decade		.002			.004			-		
Hypotheses testing	H1a	: Supported		H1b	: Supported		H1c	H1c: Supported		

^a *Note*. In the initial linear regression on communality, neither gender (b = 0.06, z = 0.59, p = .554), nor the interaction Gender × Decade (b = 0.02, z = 0.91, p = .363) were significant.

^b *Note.* In the initial non-linear regression on communality, neither gender (b = 0.08, z = 0.74, p = .459), nor the interaction Gender × Decade (Spline 1: b = 0.14, z = 0.45, p = .654; Spline 2: b = 0.08, z = 0.41, p = .685) were significant.

Table 9

Words signifying likability and respectability by frequency.

Likability (N)	Respectability (N)
friend / friendship (357) colleague (233) love/ lovable/ loved (182) dear (47) popular (45) likable (28) father (24) charisma (22)	appreciation/ appreciate (821) honor / (highly) honored (808) role model (388) success/ successful (304) extraordinary (225) value / valuable (204) esteem / esteemed (199) reverence / revere / revered (189) recognition / recognize / recognized (146)

1950s. Supporting Hypothesis 3c, communality and likability correlated positively, r(806) = .12, p < .001. Analyses by decade, however, revealed that the communality-likability association was only significantly positive in the 2010er years.

Respectability (Hypothesis 4a-c). Contradicting Hypothesis 4a, in the obituaries of women leaders, agency and respectability were uncorrelated, r(806) = .01, p = .740. Analyses by decade only found the predicted positive association in the 1980s. Contradicting Hypothesis 4b, competence and respectability were uncorrelated, r(806) = .06, p = .069. Analysis by decade only found the predicted positive association in the 1960s and 1980s. Finally, contradicting Hypothesis 4c, communality and respectability were uncorrelated, r(806) = -.03, p = .335, and within each decade.

What about men leaders? Exploring these associations in the obituaries of men leaders revealed some similarities and some differences. Agency and likability correlated negatively, r(1,473) = .06, p = .020, but analyses by decade found significant negative agency-likability associations only in the 1960s and 1980s. Competence and likability were uncorrelated, overall, r(1,473) = .02, p = .446, and within each decade. Communality and likability correlated positively, r(1,473) = .11, p < .001. Again, this association was only significantly positive in the 2010er years. Agency and respectability were uncorrelated, r(1,473) = .04, p = .095, with the exception of a significantly positive correlation in the 1990s. Competence and respectability were also uncorrelated, r(1,473) = .05, p = .058, but for a significant positive correlation in the 2010er years. Communality and respectability were uncorrelated, r(1,473) = .02, p = .362, but for a significant positive correlation in the 1950s.

Exploring Context: The Masculinity of industries

For 2,171 (95.09%) of obituaries, we could identify to which of 14 industries the organization that published the obituary belonged. Based on the distribution of women and men in these industries from 2019 (Bundesagentur für Arbeit, 2022), we ranked them according to the extent to which they were dominated by men:

Health and social services (least dominated by men)
 School and education

M.K. Zehnter et al.

Table 10

Linear and Non-linear Negative Binomial Regression Predicting the Frequencies of Likability and Respectability Words.

		Likability		Respectability					
Linear model	b (SE)	Z	р	b (SE)	Z	р			
Obituary Length	0.01 (0.00)	10.16	<.001	0.01 (0.00)	24.07	<.001			
Newspaper (conservative)	-0.17 (0.07)	-2.31	.021	0.03 (0.03)	0.99	.322			
Academic leader	0.73 (0.48)	1.52	.128	0.08 (0.18)	0.45	.653			
Business leader	0.74 (0.47)	1.58	.114	0.30 (0.17)	1.75	.081			
Charity leader	0.87 (0.48)	1.82	.068	0.03 (0.18)	0.17	.865			
Political leader	0.64 (0.49)	1.32	.187	0.25 (0.18)	1.39	.166			
Leadership level (high)	-0.46 (0.08)	-6.15	<.001	-0.09 (0.04)	-2.66	.008			
Founder	-0.08 (0.13)	-0.65	.514	0.03 (0.05)	0.50	.620			
Decade	0.09 (0.03)	2.66	.008	0.01 (0.01)	1.38	.169			
Gender (women leader)	-0.49 (0.20)	-2.51	.012	-0.09 (0.04)	-2.54	.011			
Gender \times Decade	0.10 (0.04)	2.44	.015	-	-	ns ^a			
Pseudo R ^b		.072		.221					
Unique R ^b Gender		.000		.002					
Unique R^b Gender $ imes$ Decade		.003			-				
TT		10 C			Hole Commented				

Hypotheses testing	g
--------------------	---

Hypotheses testing	F	12a: Contradicted		H2D: Supported				
Non-linear model	b (SE)		р	b (SE)	Z	р		
Obituary Length	0.01 (0.00)	9.76	<.001	0.01 (0.00)	23.58	<.001		
Newspaper (conservative)	-0.16 (0.07)	-2.19	.028	0.03 (0.03)	1.06	.288		
Academic leader	0.75 (0.48)	1.57	.116	0.08 (0.18)	0.47	.636		
Business leader	0.77 (0.47)	1.65	.099	0.31 (0.17)	1.79	.074		
Charity leader	0.88 (0.48)	1.82	.068	0.03 (0.18)	0.17	.867		
Political leader	0.67 (0.49)	1.37	.171	0.26 (0.18)	1.43	.154		
Leadership level (high)	-0.46 (0.08)	-6.19	<.001	-0.10 (0.04)	-2.69	.007		
Founder	-0.09 (0.13)	-0.72	.473	0.02 (0.05)	0.45	.656		
Masculinity of industry	-0.02 (0.01)	-1.85	.065	0.00 (0.01)	0.06	.955		
Decade spline 1	-0.60 (0.32)	-1.89	.059	-0.06 (0.12)	-0.49	.627		
Decade spline 2	-0.05 (0.22)	-0.23	.816	0.12 (0.09)	1.41	.158		
Gender (women leader)	-0.38 (0.22)	-1.74	.083	-0.09 (0.04)	-2.63	.009		
Gender \times Decade spline 1	0.47 (0.61)	0.76	.446	_	-	ns		
Gender \times Decade spline 2	0.82 (0.40)	2.06	.039	-	-	ns		
Pseudo R ^b		.074			.222			
Unique R ^b Gender		.000			.002			
Unique $R^{\rm b}$ Gender \times Decade		.002			-			
Hypotheses testing	ŀ	I2a: Contradicted		H2b: Supported				

^a Note. In the initial linear regression on respectability, gender (b = -0.17, z = -1.98, p = -0.48), but not the interactions Gender × Decade (b = 0.02, z = 1.06, p = -0.17, z = -1.98, p = -0.048), but not the interactions Gender × Decade (b = 0.02, z = 1.06, p = -0.048), but not the interactions Gender × Decade (b = 0.02, z = 1.06, p = -0.048), but not the interactions Gender × Decade (b = 0.02, z = 1.06, p = -0.048), but not the interactions Gender × Decade (b = 0.02, z = 1.06, p = -0.048). .290) was significant.

^b Note. In the non-linear initial regression on respectability, neither gender (*b* = -0.04, *z* = -0.38, *p* = .707), nor the interactions Gender × Decade (Spline 1: *b* = -0.37, z = -1.33, p = .184; Spline 2: b = 0.08, z = 0.47, p = .639) were significant.

- 3. Public administration, defense, and national insurance
- 4. Other and private services
- 5. Finance and insurance
- 6. Restaurant and hotel
- 7. Trade and maintenance of motor vehicles
- 8. Business-related services
- 9. Information and communication
- 10. Agriculture, forest, and fishing
- 11. Manufacturing
- 12. Transport and storage
- 13. Mining, energy, water, waste disposal
- 14. Construction (most dominated by men)

Then, we re-examined changes in agency (H1a), competence (H1b), communality (H1c), likability (H2a), and respectability (H2b), controlling for the masculinity of industry and its interactions with gender and decade. In particular, we sought to explore whether gender gaps in word frequencies and changes thereof depended on the masculinity (i.e., men's dominance) of the industries in which the deceased leaders were active. Including these interactions had significant effects on the frequencies of agency, competence, communality, and likability words, but not on the frequencies of respectability words (See Table 12 for all results). Fig. 5 shows the frequencies of agency, competence, communality, likability, and respectability words by gender, decade, and masculinity of industry.

Agency and likability showed similar patterns (Table 12). Decreases in the gender gap in agency and likability words over time did not depend on the masculinity of industries, as indicated by the statistically non-significant three-way interactions Gender \times Decade \times Masculinity. Upon removing this interaction from the regression equations, the interaction Gender \times Masculinity of Industry emerged as statistically significant, signaling that gender differences (to the disadvantage of women leaders) in agency and likability descriptions were greater the more men-dominated an industry was. Notably, Gender × Masculinity of Industry suppressed the effect of Gender \times Decade in agency and likability, indicating that gender differences in men-dominated industries are a stronger predictor for agency and likability descriptions than any changes thereof over time. These results contradict Hypotheses 1a (which predicted change in agency), but supported Hypothesis 2a (which predicted persistence in likability). In men-dominated professions, women leaders were described with fewer agency and likability words, and these gender gaps persisted over time.

Decreases in the gender gap in competence words over time depended on the masculinity of industry, as indicated by the statistically significant three-way interaction Gender \times Decade \times Masculinity (See Table 12). A visual analysis of Fig. 5 shows that, over time, the gender gap in competence decreased in gender-balanced industries, but persisted in particular in men-dominated industries. While these results support Hypothesis 1b, they add important boundary conditions for its

The Leadership Quarterly xxx (xxxx) xxx

- Women Leaders - Men Leaders



Fig. 3. Absolute and Relative Mean Frequencies of Words Signifying Likability and Respectability by Gender and Decade with Linear Regression Lines.

able 11
artial Pearson Correlations of Agency, Competence, and Communality with Likability and Respectability Controlled for the Length of Obituaries.

	Likability						_		tability				
	Age	ency	Comp	etence	Comm	unality	_	Age	ncy	Comp	etence	Comm	unality
	WL	ML	WL	ML	WL	ML	_	WL	ML	WL	ML	WL	ML
1950s	26	.01	25	03	19	.04		13	01	.21	12	06	.15
1960s	.02	14	19	04	.21	.01		.18	01	.26	02	07	01
1970s	.07	04	03	.01	08	.07		03	01	06	.06	10	.09
1980s	10	13	.03	05	.05	.11		.18	.06	.18	.11	12	05
1990s	13	.07	13	.01	.13	.11		.05	.16	03	06	.07	.08
2000er	28	10	14	.01	.04	.09		.12	06	.05	.05	10	01
2010er	21	05	.11	06	.25	.23		14	06	.01	.28	01	.04

Note. WL = Women leaders, ML = Men leaders.

Statistically significant correlations (p < .05) are displayed in bold black.

validity. Over time, women leaders are described as increasingly competent in more gender-balanced and women-dominated industries, but they continue to be described as less competent than their men counterparts in men-dominated industries.

While neither the main effect of masculinity of industry nor its interactions with gender and decade were statistically significant, their inclusion in the regression model suppressed the main effect of gender. Thus, these results contradict Hypothesis 1c. Upon controlling for the masculinity of industry, there were neither gender differences, nor changes thereof in the descriptions of leaders as communal. Instead, decreases in communality descriptions over time were a general trend.

Robustness checks

Detailed results of all robustness checks are provided in the online supplement. Here, we highlight key findings. The results from negative binomial hurdle regression analyses were consistent with the abovedescribed results from negative binomial regressions, except for agency. Upon separating the zero counts from the counts above one, women leaders were more likely described with zero agency words than men leaders and this gender gap did not decrease over time. Simultaneously, those women leaders who were described as agentic (whose obituaries included at least one agency word) were described as equally agentic than their men counterparts across time (See Supplement 4).

Overall, the leader obituaries included 2,015 (1.17%) leadership titles, such as supervisor, manager, chairperson, and executive board member. Following several English agency dictionaries (Abele et al., 2016; Hanges & Dickson, 2004; Hentschel et al., 2019; Pietraszkiewicz et al., 2019), in our main analyses, we did not include specific leadership titles as agency words in our German text-mining dictionary (we did however maintain words referring to leading more generally, such as lead, direct, guide). In the context of obituaries, 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'

The Leadership Quarterly xxx (xxxx) xxx

- Women Leaders - Men Leaders



Fig. 4. Partial Pearson Correlations of Agency, Competence, and Communality with Likability and Respectability Controlled for the Length of Obituaries.

Table 12 Negative Binomial Regression Predicting the Frequencies of Descriptive and Evaluative Words Controlled for the Masculinity of Industry and its Interactions with Gender and Decade.

	Agency		Competence		Communality		Likability		Respectability	
	z	р	z	р	z	р	z	р	z	р
Obituary Length	31.64	<.001	27.04	<.001	26.16	<.001	9.57	<.001	23.45	<.001
Newspaper (conservative)	4.42	<.001	3.78	<.001	1.60	.109	-2.11	.035	1.31	.258
Academic leader	0.97	.334	2.63	.009	-0.89	.376	1.43	.152	0.35	.730
Business leader	2.26	.024	3.03	.002	0.99	.325	1.86	.063	1.56	.119
Charity leader	1.36	.174	1.27	.204	1.27	.205	1.44	.150	0.17	.868
Political leader	1.37	.171	1.44	.149	-0.38	.706	1.33	.184	1.27	.205
Leadership level (high)	-1.21	.225	-3.10	.002	-1.48	.140	-5.74	<.001	-3.14	.002
Founder	5.37	<.001	1.22	.222	1.33	.183	-0.61	.543	0.36	.718
Decade	2.19	.029	-1.32	.186	-2.62	.009	0.30	.765	0.28	.780
Gender (women leader)	-1.02	.310	-3.58	<.001	-0.95	.342	-0.32	.749	-2.18	.029
$\textbf{Gender} \times \textbf{Decade}$	1.69	.090	2.80	.005	1.61	.108	1.56	.118	1.31	.190
Masculinity (of industry)	-0.03	.975	-0.87	.383	-1.29	.220	-0.14	.889	-0.18	.855
Gender \times Masculinity	-2.04	.041	1.70	.089	1.75	.080	-2.20	.028	1.13	.258
Decade \times Masculinity	0.67	.503	-0.22	.830	0.39	.700	-0.39	.700	-0.08	.938
$G \times D \times M$	_	ns ^a	-1.97	.048	_	ns ^b	_	ns ^c	_	ns ^d
Pseudo R	.347		.289		.224		.73		.226	
Hypotheses testing	H1a:		H1b:		H1c:		H2a:		H2b:	

^a Note. In the initial regression on **agency**, neither Gender × Masculinity (z = -1.63, p = .104), nor Decade × Masculinity (z = -0.01, p = .993), nor G × D × M (z = 0.98, p = .330) were significant.

^b Note. In the initial regression on **communality**, neither Gender × Masculinity (z = -1.77, p = .078), nor Decade × Masculinity (z = 1.04, p = .298), nor G × D × M (z = -1.18, p = .238) were significant.

^c Note. In the initial regression on **likability**, Gender × Masculinity (z = -2.59, p = .010), but neither Decade × Masculinity (z = 1.44, p = .149), nor G × D × M (z = -1.91, p = .056) were significant.

^d Note. In the initial regression on **respectability**, neither Gender × Masculinity (z = 0.76, p = .449), nor Decade × Masculinity (z = 0.14, p = .892), nor G × D × M (z = -0.36, p = .716) were significant.

(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 Supplement 5, were consistent with the results presented above. Linear negative binomial regression revealed that women were described with fewer agency words (including leadership titles) than men leaders, but this gender gap

Men Leaders

Women Leaders

The Leadership Quarterly xxx (xxxx) xxx



Decade

Fig. 5. Absolute Mean Frequencies of Words Signifying Agency, Competence, Communality, Likability and Respectability by Gender, Decade, and Masculinity of Industry with Linear Regression Lines.

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.

Discussion

Replicating and extending previous work on change in gender stereotypes in the context of leadership (Zehnter et al., 2018), we used textmining to examine how women and men leaders were described and evaluated in obituaries between 1953 and 2019. As constructive replication (Köhler & Cortina, 2021), this research adds a novel contribution to the literature by (1) conducting an analysis of changes in the descriptions and evaluations of women and men leaders over time, (2) assessing the evaluations of leaders not only in terms of likability, but also respectability, and (3) utilizing text-mining to identify descriptive and evaluative words in obituaries. Three studies assessed the quality of our text-mining dictionary, attesting its internal reliability, coverage, convergent, and predictive validity. Hence, this research joins a growing body of literature employing cutting-edge methodology to investigate social phenomena (Charlesworth et al., 2022; Garg et al., 2018).

Partly replicating the findings from Zehnter and colleagues' (2018), our results show a more complex and nuanced picture of change alongside persistence in the descriptions and evaluations of women and men leaders. Overall, significant gender differences persist in all leader descriptions and evaluations. Even after controlling for the length of obituaries, the political leaning of the newspaper in which the obituaries were published, types of leadership (academic, business, charity, political), the level of leadership, and foundership, women leaders were described in line with gender stereotypes (less agentic, less competent, and more communal than men leaders) and evaluated differently (less likable, less respectable). However, some of these differences significantly decreased over time.

Changes in the descriptions of women leaders

Consistent with the replicated work and our hypotheses, we found that, over time, women leaders were described as increasingly agentic and competent (attenuating gender differences in ascribed agency and competence), but continued to be described as more communal compared to their men counterparts. Non-linear analyses further revealed that decreases in the gender gaps in agency and competence accelerated from the 1980s, a time characterized by the influx of women in the working sphere (Allmendinger et al., 2008; Toossi & Morisi, 2017). This finding supports social role theory, according to which gender stereotypes change as a result of a dwindling segregation of women as housemakers versus men as paid employees (Eagly & Steffens, 1984; Wood & Eagly, 2012).

In regard to competence and communality, our findings were consistent across robustness checks with zero-hurdle regressions (which we conducted because of the zero-inflation of our data). Regarding changes in agency, our analyses produced mixed results. Following the analysis strategy of the replicated work (i.e., negative binomial regression) showed the predicted changes in agency: Women leaders (but less so men leaders) were described as increasingly agentic. However, results from zero-hurdle regression suggest that increases in the frequencies of

M.K. Zehnter et al.

agency words were a general trend in the obituaries of women and men leaders. But significant gender differences persisted with women leaders being more likely described with zero agency words. Distinguishing only between zero counts versus counts above zero (which are all coded as one), the logistic zero-hurdle model is more robust against extreme cases than other count models. Thus, decreases in the agency gender gap over time may be driven by extraordinary women leaders who were described as particularly agentic.

For about 95 percent of obituaries, we could identify to which industries the organization that published the obituary belonged. Analyses controlling for the masculinity of industries revealed important boundary conditions in regard of the observed changes. Concerning agency and competence, significant gender gaps (to the disadvantage of women leaders) persisted in more men-dominated industries (e.g., construction, transport and storage). In more women-dominated industries, gender gaps in agency were small across all decades, and gender gaps in competence decreased over time. Although the masculinity of industry did not affect descriptions of leaders as communal, it suppressed the effect of gender on communality. That is, upon controlling for the masculinity of industry, gender differences in communality (to the advantage of women) disappeared.

At large, our results are remarkably consistent with previous research that examined gender stereotypes across a similar time span, but used a different research design. Eagly and colleagues (2020) analyzed explicit responses from opinion polls, in which U.S. Americans described average women and men, whereas we used an unobtrusive text analysis in the context of leadership in Germany. Nevertheless, both Eagly 's and our work uncovered (1) a persisting agency advantage of men over women (in our data this advantage decreased over time, although not in men-dominated industries), (2) most pronounced changes towards the advantage of women in competence (in our data these changes were more pronounced in women-dominated and genderbalanced industries), and (3) a persistent communality advantage of women over men (although in our data this advantage disappeared upon controlling for the masculinity of industry).

Our finding that women leaders continue to face greater hurdles to be seen as agentic, especially in men-dominated industries, is also consistent with the continued strong association of leadership with masculinity (Koenig et al., 2011). Women leaders' persistent communality advantage is consistent with the Stereotype Content Model, according to which women with high status (e.g., leaders) are ascribed agency and communality (Fiske, 2018; Fiske et al., 2007), and other recent research showing that in organizations with women leaders, women are more strongly associated with agency without reducing associations with communality (Lawson et al., 2022).

Changes in the evaluation of women leaders as likable

Unlike Zehnter et al. (2018), we found change in the evaluations of women leaders as likable. Over time, women leaders were described as increasingly likable. Again, non-linear analyses revealed that decreases in the gender gap in likability accelerated from the 1980s, coinciding with women's influx in the working sphere. This finding was also confirmed in a robustness check with zero-hurdle regressions (the category likeability was characterized by pronounced zero-inflation).

Additional analyses on a subset of obituaries revealed, however, that in men-dominated industries, women leaders were persistently described with fewer likability words than their men counterparts. It is noteworthy though that only about 34 percent of obituaries included likability words at all, indicating that, overall, the attribute 'likable' was rarely bestowed to deceased leaders.

Replicating Zehnter and colleagues (2018), agency and likability were associated negatively in the obituaries of women leaders. Analyses by decade specified that the agency-likability association was significantly negative in the 1950s, and surprisingly also in the 2000er and the 2010er year. In the obituaries of men leaders, agency and likability were also associated negatively, but this association was not significant in the past two decades. These results support Zehnter and colleagues' (2018) claim that negative evaluations of agentic women are more stable over time. More than that, our results suggest that the dislike of agentic women made a comeback in the past two decades. This result is consistent with backlash hypothesis according to which counter-stereotypical women are evaluated negatively (Rudman et al., 2012; Rudman & Fairchild, 2004; Rudman & Glick, 2001; Rudman & Phelan, 2008; Williams & Tiedens, 2016). However, our analyses suggest, that rather than being a persistent trend, the negative agency-likability association of women leaders is a returning trend, supporting the notion that backlash against counter-stereotypical women and greater gender equality more generally have increased since the millennial shift (Flood et al., 2021).

Unlike predicted, the competence-likability association was only significantly negative in the obituaries of women leaders in the 1950s. According to backlash hypothesis (Rudman et al., 2012; Williams & Tiedens, 2016), the dislike of counter-stereotypical women reinforces gender stereotypes and thus hinders any change in this regard. Contrary to this claim, the disappearance of the once negative competence-likability association suggests that change is possible. Notably, as above described the negative agency-likability association from the 1950s also disappeared for decades, but made a comeback after the millennial shift, indicating that change is not necessarily linear, but has the potential to reverse.

As predicted, communality and likability were associated positively in the obituaries of women leaders, but only in the 2010er years. The same pattern was found in the obituaries of men leaders. These results contradict previous research according to which communality is not a desirable characteristic in leaders (Bongiorno et al., 2014) and support the increasing advantage of modern leadership styles – for women and men leaders – that emphasize communality, such as transformational, authentic, and relationship-oriented leadership (Antonakis & Day, 2012).

Paying respect: Respectability words in leader obituaries

Unlike previous research, including the replicated work, we examined not only the likability, but also the respectability of leaders. Consistent with our hypotheses, the obituaries of women leaders contained fewer respectability words than those of men leaders and this gender gap persisted over time. Men's persistent respectability advantage was supported by our main analyses, analyses controlling for the masculinity of industry, and a robustness check with hurdle regression making this our most robust result. Other than predicted, in the obituaries of women leaders, agency and competence correlated only positively with respectability in the 1980s. Perhaps, in a decade, in which women started to enter the workplace in unprecedented numbers, agency and competence was particularly valued in women leaders. In the obituaries of men leaders, agency and respectability were positively associated in the 1990s, and competence and respectability in the 2010er years. Especially the latter findings indicate that disparities in the evaluations of women and men leaders may not only stem from disadvantages of women leaders but may also be based on advantages of men leaders.

Overall, the gender gap in respectability is astonishing given that obituaries are essentially written to pay respect (Fowler & Bielsa, 2007). Previous research has mostly framed respectability as reward for men in stereotype-conforming roles and the loss of respect as penalization of men in counter-stereotypical roles (Heilman & Wallen, 2010; Moss-Racusin et al., 2010). The present research highlights that respectability is also relevant in the evaluations of women leaders, and potentially of women in counter-stereotypical roles more generally. Moreover, our results suggest that backlash against women leaders may not only be expressed by denying them the psychological reward (likability) commonly allocated to ideal women, but also through

denying them the psychological reward (respectability) allocated to ideal, masculine leaders.

Limitations

Naturally, this research has limitations. As we were interested in increases of descriptive and evaluative words in the obituaries of women compared to men leaders over time, we chose relatively simple linear and non-linear regression models over more complex time series analyses and higher-order autoregressive models. As a consequence, despite our longitudinal data and the inclusion of control variables, our models support predictive relationships, but not causal effects. There may be alternative and additional explanations than time and gender for the observed variations in the descriptions and evaluations of leaders, such as women's overall representation in the workforce and in leadership roles. The introduction of legislation to increase gender parity in leadership, such as gender quotas, may also affect views of women leaders as suggested by past experiments (Heilman et al., 1992; Heilman & Welle, 2006; Nater et al., 2023). Additionally, increasing preferences for leadership styles that emphasize communality may also change views of women leaders (Antonakis & Day, 2012). More complex statistical analyses (e.g., time series analyses, higher-order autoregressive models) may have allowed us to better capture delayed or cumulative influences of the above-mentioned societal transformations.

Despite including all obituaries of women leaders in the years of analysis, this sample remained relatively small, particularly in the early years. Despite some amendments (i.e., aggregating the data across decade), the scarcity of obituaries for women may have decreased test power and the robustness of results. Due to this scarcity, outstanding women leaders, that is, those who were described and evaluated in extraordinary terms (e.g., with particularly great frequencies of agency) may have affected our results to a greater extent than outstanding men leaders.

Our results may also be limited by the selection of newspapers in which the obituaries were published. We chose two widely distributed national German newspapers with a long history of publishing obituaries written and paid for by organizations (one liberal, one conservative). This led to an obituaries sample that was heterogenous in terms of the organizations (i.e., academic, business, charity, and political) and industries (e.g., women-dominated, gender-balanced, men-dominated), in which the deceased leaders were active. However, three quarter of obituaries were written for higher-level leaders who led organizations, were corporate board members, and senior leaders. The inclusion of obituaries from smaller newspapers (e.g., local and industry-specific newspaper) might have increased the number of obituaries for lowerlevel leaders (who headed departments), which might have attenuated some of the persistent gender differences that we observed. In our data, lower-level leaders were described with greater frequencies of competence, likability, and respectability words than higher-level leaders, suggesting a more appreciative tone towards lower-level leaders. Notably though, the lower-level leaders who were honored in obituaries in national newspapers may have been exceptionally liked and respected individuals.

Another limitation is the positive bias of obituaries. We sought to infer the penalization of counter-stereotypical attributes (i.e., agency, competence) from the absence of positive evaluations (i.e., likability, respectability). However, displaying counter-stereotypical attributes may be penalized through blatant hostility, rarely expressed in obituaries. Moreover, women leaders were extraordinary in the early years of analysis and to this day remain the exception, especially in high-level leadership positions. This may have introduced an additional positive bias into our analysis, as outstanding women leaders tend to be evaluated more favorably than their outstanding men counterparts (Rosette & Tost, 2010). Finally, our data did not allow the analysis of effects based on intersecting leader identities, such as ethnicity, age, socio-economic status, and sexual orientation.

Future Directions

Our work paves promising avenues for future work. To increase the robustness and generalizability of our findings, future research could investigate whether similar patterns of change and persistence in the descriptions and evaluations of women and men leaders can be observed across other text resources than obituaries, such as newspaper articles and social media comments about public leadership figures. Especially, the longitudinal analysis of text resources that include negative language (e.g., abusive, derogatory words) could illuminate whether there are gender differences in blatant hostility towards leaders, and in which direction such gender differences change over time.

Future research could also augment our obituaries data, draw on other longitudinal datasets or curate new data to examine the effects of additional exogenous variables, such as national indices for gender equality, characteristics of the workforce (e.g., percentage of working women and women leaders), and changes in gender equality legislations (e.g., the introduction of gender quotas) on views about women and men leaders. Including such variables in future analyses would help to determine whether additional or alterative mechanisms than time and gender underlie views about leaders.

Additional research to establish causal effects is also warranted. Randomized field experiments would allow researchers to examine causal effects of leader characteristics, such as gender, on views about leaders (Antonakis et al., 2010). For incidence, to test whether women, compared to men leaders face greater hurdles to be seen as agentic and respectable (as suggested by our analyses), one could match a woman and man leader on important characteristics (e.g., education, profession, career trajectory) and investigate whether they are described and evaluated differently. Similar designs could be used to investigate the effect of context factors (e.g., the gender composition of an industry or organization) and organizational culture (flat versus steep hierarchies, leadership styles) on views about leaders.

Laboratory experiments would provide additional insights into the causal relationships between the descriptions and evaluations of leaders. Past research manipulated the extent to which women and men were described as agentic leaders and examined downstream effects on likability (Heilman & Okimoto, 2007; Rudman et al., 2012; Rudman & Fairchild, 2004). A similar design would be useful to examine the effects of agency, competence, and communality on evaluations of women and men leaders as respectable.

Practical implications

For organizations striving towards a better representation of women leaders, our results may be a valuable reminder that change towards greater equality does not necessarily mean that the journey is complete. Despite some changes over time, significant gender differences persisted in all leader descriptions and evaluations. Even in the most recent decade, the obituaries of men leaders were longer (and thus more expensive) than those of women leaders and included greater frequencies of agency and competence words, which tend to be associated with successful leadership (Koenig et al., 2011), as well as greater frequencies of respectability words. These differences were particularly pronounced in men-dominated industries.

Notably, these effects were relatively small (Bosco et al., 2015), explaining small percentages of variance in the descriptions and evaluations of leaders. However, these effects emerged despite (1) the strong positivity bias of obituaries, which are designated to commemorate, praise and honor a person and (2) a selection bias by which the mere fact of receiving an obituary made our sample of women leaders somewhat exceptional. Ultimately, finding diverging descriptions and evaluations of women and men leaders in the positively biased context of obituaries may be a powerful demonstration of how pervasive this phenomenon is (Cortina & Landis, 2009).

The Leadership Quarterly xxx (xxxx) xxx

M.K. Zehnter et al.

Conclusion

Using a rigorously developed dictionary (which quality was assessed in three studies), we text-mined 2,283 leader obituaries published between 1953 and 2019 to examine change in the descriptions and evaluations of women and men leaders. While significant gender differences persist in all leader descriptions and evaluations, in particular in mendominated industries, some of these differences decreased over time. Women leaders were described more similarly to men leaders in terms of agency, competence and likability, but they continued to be described as more communal, and less respectable. Ultimately, these results highlight that despite some changes towards greater equality, disparaging views of women and men leaders persist.

CRediT authorship contribution statement

Miriam K. Zehnter: Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. Christoph U. Wolfmayr: Writing – review & editing, Software, Methodology, Formal analysis, Data curation. Leona A. Andriopoulos: Writing – review & editing, Visualization, Data curation, Conceptualization. Erich Kirchler: Writing – review & editing, Supervision, Resources, Project administration, Data curation, Conceptualization. Martin Voracek: Writing – review & editing, Supervision, Resources, Methodology. Michelle K. Ryan: Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Patrizia Hubner, Stephanie Lüer, Catherine Mayrhofer, and David Wehner contributed to data collection, and Kirsten Handschuch to the supplementary material. This work was supported, in part, through the European Commission's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101024885 (awarded to Zehnter), and by the European Research Council Consolidator Grant No 725128 (awarded to Ryan).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.leaqua.2025.101884.

Data availability

An OSF link to the data and R-script used in this this reserach is shared in the manuscript.

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/fpsys.2016.01810
- Aguinis, H., & Gottfredson, R. K. (2010). Best-practice recommendations for estimating interaction effects using moderated multiple regression. *Journal of Organizational Behavior*, 31(6), 776–786. https://doi.org/10.1002/job.686

Allmendinger, J., Leuze, K., & Blanck, J. M. (2008). Jahre Geschlechtergerechtigkeit und Arbeitsmarkt [50 years of gender equality and labour market]., 50.

- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21(6), 1086–1120. https:// doi.org/10.1016/j.leaqua.2010.10.010
- Antonakis, J., & Day, D. (2012). Leadership: Past, present, and future. In J. Antonakis, & D. Day (Eds.), *The Nature of Leadership* ((2nd ed.,, pp. 3–26). SAGE. https://doi.org/ 10.4135/9781506395029.n1.
- Banks, G. C., Woznyj, H. M., Wesslen, R. S., & Ross, R. L. (2018). A review of best practice recommendations for text analysis in R (and a user-friendly app). *Journal of Business* and Psychology, 33(4), 445–459. https://doi.org/10.1007/s10869-017-9528-3

- Beck, C. W., & Bliwise, N. G. (2014). Interactions Are Critical. CBE—Life Sciences. Education, 13(3), 371–372. https://doi.org/10.1187/cbe.14-05-0086
- Blevins, D. P., Tsang, E. W. K., & Spain, S. M. (2015). Count-based research in management. Organizational Research Methods, 18(1), 47–69. https://doi.org/ 10.1177/1094428114549601
- Bongiorno, R., Bain, P. G., & David, B. (2014). If you're going to be a leader, at least act like it! Prejudice towards women who are tentative in leader roles. *British Journal of Social Psychology*, 53(2), 217–234. https://doi.org/10.1111/bjso.12032
- Bosak, J., Kulich, C., Rudman, L., & Kinahan, M. (2018). Be an advocate for others, unless you are a man: Backlash against gender-atypical male job candidates. *Psychology of Men and Masculinity*, 19(1). https://doi.org/10.1037/men0000085
- Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. Journal of Applied Psychology, 100(2), 431–449. https://doi. org/10.1037/a0038047
- Brescoll, V. L., Dawson, E., & Uhlmann, E. L. (2010). Hard won and easily lost: The fragile status of leaders in gender-stereotype-incongruent occupations. *Psychological Science*, 21(11), 1640–1642. https://doi.org/10.1177/0956797610384744
- Bundesagentur für Arbeit. (2022). Die Arbeitsmarktsituation von Frauen und Männern 2021 [Labourmarket conditions for women and men in 2021].
- Catalyst. (2022). Women in management: Quick take. https://www.catalyst.org/research/ women-in-management/.
- Charlesworth, T. E. S., Caliskan, A., & Banaji, M. R. (2022). Historical representations of social groups across 200 years of word embeddings from Google Books. *Proceedings* of the National Academy of Sciences, 119(28). https://doi.org/10.1073/ pnas.2121798119
- Cortina, J. M., & Landis, R. S. (2009). When small effect sizes tell a big story, and when large effect sizes don't. In C. E. Lance, & R. J. Vandenberg (Eds.), *Statistical and Methodological Myths and Urban Legends* (pp. 287–308). Routledge.
- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, 92(4), 631–648. https://doi.org/10.1037/0022-3514.92.4.631
- Curşeu, P. L., & Boroş, S. (2011). Gender stereotypes in management: A comparative study of communist and postcommunist Romania. *International Journal of Psychology*, 46(4), 299–309. https://doi.org/10.1080/00207594.2011.554554
- Dahiru, T. (2011). P-Value, a true test of statistical significance? a cautionary note. Annals of Ibadan. Postgraduate Medicine, 6(1). https://doi.org/10.4314/aipm. v6i1.64038
- Diekman, A. B., & Eagly, A. H. (2000). Stereotypes as Dynamic Constructs: Women and Men of the Past, Present, and Future. *Personality and Social Psychology Bulletin, 26* (10), 1171–1188. https://doi.org/10.1177/0146167200262001
- Eagly, A. H., & Heilman, M. E. (2016). Gender and leadership: Introduction to the special issue. *The Leadership Quarterly*, 27(3), 349–353. https://doi.org/10.1016/j. leaqua.2016.04.002
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573–598. https://doi.org/10.1037/0033-295X.109.3.573
- Eagly, A. H., Mladinic, A., & Otto, S. (1991). Are women evaluated more favorably than men? An analysis of attitudes, beliefs, and emotions. *Psychology of Women Quarterly*, 15(2), 203–216. https://doi.org/10.1111/j.1471-6402.1991.tb00792.x
- Eagly, A. H., Nater, C., Miller, D. L., 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. doi: 10.1037/ amp0000494.
- Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, 46(4), 735–754. https://doi.org/10.1037/0022-3514.46.4.735
- Edwards, R., Guzzo, R., Jackson, C., Knoepflmacher, A., & Nalbantian, H. (2020). Let's get real about equality: When women thrive 2020 global report.
- Engqvist, L. (2005). The mistreatment of covariate interaction terms in linear model analyses of behavioural and evolutionary ecology studies. *Animal Behaviour*, 70(4), 967–971. https://doi.org/10.1016/j.anbehav.2005.01.016
- Feng, C. X. (2021). A comparison of zero-inflated and hurdle models for modeling zeroinflated count data. *Journal of Statistical Distributions and Applications*, 8(1), 8. https://doi.org/10.1186/s40488-021-00121-4
- Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. Current Directions in Psychological Science, 27(2), 67–73. https://doi.org/10.1177/ 0963721417738825
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, 11(2), 77–83. https://doi.org/ 10.1016/j.tics.2006.11.005
- Flood, M., Dragiewicz, M., & Pease, B. (2021). Resistance and backlash to gender equality. Australian Journal of Social Issues, 56(3), 393–408. https://doi.org/ 10.1002/ajs4.137
- Fowler, B., & Bielsa, E. (2007). The Lives We Choose to Remember: A Quantitative Analysis of Newspaper Obituaries. *The Sociological Review*, 55(2), 203–226. https:// doi.org/10.1111/j.1467-954X.2007.00702.x
- Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, 115(16). https://doi.org/10.1073/pnas.1720347115
- Gustafsson Sendén, M., Klysing, A., Lindqvist, A., & Renström, E. A. (2019). The (not so) changing man: Dynamic gender stereotypes in Sweden. *Frontiers in Psychology*, 10. https://doi.org/10.3389/fpsyg.2019.00037
- Haines, E. L., Deaux, K., & Lofaro, N. (2016). The times they are a-changing or are they not? A comparison of gender stereotypes, 1983–2014. Psychology of Women Quarterly, 40(3), 353–363. https://doi.org/10.1177/0361684316634081

M.K. Zehnter et al.

The Leadership Quarterly xxx (xxxx) xxx

- Hanges, P. J., & Dickson, M. W. (2004). The development and validation of the GLOBE culture and leadership scales. In R. J. House, M. W. Hanges, P. W. Javidan, P. W. Dorfman, & V. Gupta (Eds.), *Culture, leadership, and organization: The GLOBE study of 62 societies* (pp. 122–151). Sage.
- 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
- Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, 57(4), 657–674. https://doi.org/10.1111/0022-4537.00234
- Heilman, M. E., Block, C. J., & Lucas, J. A. (1992). Presumed incompetent? Stigmatization and affirmative action efforts. *Journal of Applied Psychology*, 77(4), 536–544.
- Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks? The implied communality deficit. *Journal of Applied Psychology*, 92(1), 81–92. https://doi.org/10.1037/0021-9010.92.1.81
- Heilman, M. E., & Wallen, A. S. (2010). Wimpy and undeserving of respect: Penalties for men's gender-inconsistent success. Journal of Experimental Social Psychology, 46(4), 664–667. https://doi.org/10.1016/j.jesp.2010.01.008
- Heilman, M. E., Wallen, A. S., Fuchs, D., & Tamkins, M. M. (2004). Penalties for success: Reactions to women who succeed at male gender-typed tasks. *Journal of Applied Psychology*, 89(3), 416–427. https://doi.org/10.1037/0021-9010.89.3.416
- Heilman, M. E., & Welle, B. (2006). Disadvantaged by diversity? The effects of diversity goals on competence perceptions. *Journal of Applied Social Psychology*, 36(5), 1291–1319. https://doi.org/10.1111/j.0021-9029.2006.00043.x
- Kirchler, E. (1992). Adorable woman, expert man: Changing gender images of women and men in management. European Journal of Social Psychology, 22(4), 363–373. https://doi.org/10.1002/ejsp.2420220405
- Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*, 137 (4), 616–642. https://doi.org/10.1037/a0023557
- Köhler, T., & Cortina, J. M. (2021). Play it again, Sam! An analysis of constructive replication in the organizational sciences. *Journal of Management*, 47(2), 488–518. https://doi.org/10.1177/0149206319843985
- Lawson, M. A., Martin, A. E., Huda, I., & Matz, S. C. (2022). Hiring women into senior leadership positions is associated with a reduction in gender stereotypes in organizational language. In *Proceedings of the National Academy of Sciences*. https:// doi.org/10.1073/pnas.2026443119
- Lee, K. H., Pedroza, C., Avritscher, E. B. C., Mosquera, R. A., & Tyson, J. E. (2023). Evaluation of negative binomial and zero-inflated negative binomial models for the analysis of zero-inflated count data: Application to the telemedicine for children with medical complexity trial. *Trials*, 24(1), 613. https://doi.org/10.1186/s13063-023-07648-8
- Moremen, R. D., & Cradduck, C. (1999). "How will you be remembered after you die?" Gender discrimination after death twenty years later. OMEGA - Journal of Death and Dying, 38(4), 241–254. https://doi.org/10.2190/0KY0-P54V-KFDB-9VCR
- Moss-Racusin, C. A., Phelan, J. E., & Rudman, L. A. (2010). When men break the gender rules: Status incongruity and backlash against modest men. *Psychology of Men & Masculinity*, 11(2), 140–151. https://doi.org/10.1037/a0018093

Nater, C., Heilman, M. E., & Sczesny, S. (2023). Footsteps I would like to follow? How gender quotas affect the acceptance of women leaders as role models and inspirations for leadership. *European Journal of Social Psychology*, 53(1), 129–146. https://doi.org/10.1002/ejsp.2892

Neuendorf, K. A. (2002). The content analysis guidebook. Sage.

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

Ogletree, S. M., Pena, D., & Figueroa, P. (2005). A double standard in death? Gender differences in obituaries. OMEGA - Journal of Death and Dying, 51(4), 337–342. https://doi.org/10.2190/V443-74PW-7066-0173

- 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
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. https://doi.org/10.1037/ 0021-9010.88.5.879
- Rodler, C., Kirchler, E., & Hölzl, E. (2001). Gender stereotypes of leaders: An analysis of the contents of obituaries from 1974 to 1998. Sex Roles, 45(11–12), 827–843. https://doi.org/10.1023/A:1015644520770
- Rosette, A. S., & Tost, L. P. (2010). Agentic women and communal leadership: How role prescriptions confer advantage to top women leaders. *Journal of Applied Psychology*, 95(2), 221–235. https://doi.org/10.1037/a0018204
- Rudman, L. A., & Fairchild, K. (2004). Reactions to counterstereotypic hehavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology*, 87(2), 157–176. https://doi.org/10.1037/0022-3514.87.2.157
- Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, 48(1), 165–179. https:// doi.org/10.1016/j.jesp.2011.10.008
- Shumway, R. H., & Stoffer, D. S. (2017). Time series analysis and its applications: With R examples. Springer.
- Toossi, M., & Morisi, T. L. (2017). Women In the workforce before, during, and after the great recession. Bureau of Labor Statistics: U.S.
- Venables, W. N., & Ripley, B. D. (2002). Modern Applied Statistics with S. Springer. Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (2000). Unobtrusive Measures (revised edition). Sage Publications.
- Wijffels, J. (2021). udpipe: Tokenization, parts of speech tagging, lemmatization and dependency parsing with the "UDPipe" "NLP" Toolkit [R package].
 Williams, M. J., & Tiedens, L. Z. (2016). The subtle suspension of backlash: A meta-
- Williams, M. J., & Tiedens, L. Z. (2016). The subtle suspension of backlash: A metaanalysis of penalties for women's implicit and explicit dominance behavior. *Psychological Bulletin*, 142(2), 165–197. https://doi.org/10.1037/bul0000039
- Wood, W., & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. In Advances in Experimental Social Psychology (pp. 55–123). doi: 10.1016/B978-0-12-394281-4.00002-7.
- 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