



# LeadHERship in the Federal Government

Analysis of Federal Government Leadership  
Assessment Scores by Gender

OCTOBER 6, 2022

BY Emily A. Kalnicky, Ph.D.



# About the Partnership

The Partnership for Public Service is a nonpartisan, nonprofit organization that works to revitalize the federal government by inspiring a new generation to serve and by transforming the way government works. The Partnership teams up with federal agencies and other stakeholders to make our government more effective and efficient.

# Table of Contents

<b>Background</b>	4
<b>Executive Summary</b>	5
<b>Introduction</b>	7
<b>Brief background on 360 assessment tools</b>	8
<b>Who has completed the Public Service Leadership 360 assessment to date</b>	9
<b>Overview of the demographic data for individual leaders rated by the 360 tool</b>	10
<b>Findings</b>	12
<b>Conclusion</b>	19
<b>Areas for future exploration</b>	22
<b>Appendix A</b>	23

# Background

Women are underrepresented in federal leadership positions, making up just 27.3% of the U.S. Congress in 2021<sup>1</sup> and 39% of the Senior Executive Service—the highest level of our government’s career workforce—in 2022. While various sociological and psychological theories offer insight into why these persistent gender gaps exist, a potential strategy to solve this issue is for researchers and federal agencies to better understand how federal leaders view themselves, as well as how they are viewed by their colleagues, direct reports and supervisors.

Historically, the common conception of what it means to be a leader has been built upon characteristics typically associated with white men. Understanding women’s experiences in federal leadership roles, and the barriers and challenges they face, is critical to creating a federal workforce that reflects the diversity of the United States and is better equipped to serving people with different backgrounds and needs.

Collecting and analyzing this information will enable organizations that support the federal government and its institutions—and our government itself—to address the societal structures or barriers that contribute to gender disparities in federal leadership roles and build a more effective federal workplace.

This is the first brief in the Partnership for Public Service’s LeadHERship series, which explores these issues in greater depth. For more information about this series, please review our [introductory brief](#).

---

<sup>1</sup> World Economic Forum, Global Gender Gap Report 2021. Retrieved from: [www3.weforum.org/docs/WEF\\_GGGR\\_2021.pdf](http://www3.weforum.org/docs/WEF_GGGR_2021.pdf)

# Executive Summary

In 2019, the Partnership for Public Service launched the [Public Service Leadership Model](#) to set a new standard for effective federal leadership. The model is designed around two core values—commitment to public good and stewardship of public trust—and four key competencies—becoming self-aware, engaging others, leading change and achieving results—that leaders must demonstrate to lead in the federal government.

To support the application of this model for federal leaders, the Partnership developed the Public Service Leadership 360 assessment tool. This 360 assessment uses evaluations from multiple raters—from managers to peers to individual employees themselves—designed to evaluate how well a federal leader is demonstrating the Public Service Leadership Model’s values, competencies and subcompetencies. This assessment tool also provides us with unique data on the federal leadership experience for a variety of different social identities. In this brief—our first exploration of this 360 data—we examine federal employees’ leadership experiences based on gender, specifically for men and women serving as leaders at all levels across government.

As we outlined in [the introduction](#) to this series of research briefs, women remain underrepresented in certain federal leadership positions. This analysis—backed by our unique 360 data—both sheds light on and raises new questions about why this gap persists, contributing to ongoing research on the experiences of women in the federal government. Based on rating scores given to the 1,123 federal employees who completed the Public Service Leadership 360 assessment tool from December 2020 to April 15, 2022, we uncovered several trends.

First, most employees surveyed, regardless of gender, exhibited the model’s key leadership competencies most or all the time, according to both their self-assessments and their ratings by others. These findings reinforce that federal leaders possess the skills and perspectives to make an impact in government.

We also found that women scored statistically significantly higher than men on all the model’s key competencies and core values. While both men and women tended to rate themselves lower on these values and competencies than others did, women’s self-assessments and ratings by others were more aligned than they were for men. Research has shown that this alignment—between self-ratings and ratings by others—can be a predictor of leadership success and career advancement.

Nonetheless, we also discovered that women rated themselves lower than men on a few key competencies and subcompetencies: leading change—which requires leaders to help others navigate organizational transformations and support creative solutions to organization-wide challenges—as well as innovation and creativity and embracing risk and uncertainty. These findings suggest that, despite their high ratings on many key federal leadership competencies, the women in our sample may lack experience or confidence in some areas that prioritize riskier and less certain aspects of organizational management and change. It is impossible to identify the precise reason for this perceived lack of confidence in our data.

However, it is possible that long-standing structural or organizational barriers to women in leadership positions may result in these lower self-ratings.

Overall, our data demonstrates that women in federal leadership positions—at least those employees in our sample—are perceived by others as effective leaders with diverse skill sets. As such, it is likely that the broader gender gaps in certain federal leadership positions are unlikely to be based on actual competency or skill, but rather on long-standing barriers or gender stereotypes that restrict career advancement for women. These notions and obstacles may affect women’s own sense of leadership competency in some areas, perpetuating women’s underrepresentation in some government leadership roles.

# Introduction

In 2019, the Partnership for Public Service launched the [Public Service Leadership Model](#) to set a new standard for effective federal leadership, and to help leaders assess their performance and identify opportunities to enhance their skills. The model is built around two core values and four key competencies. Each competency is comprised of five subcompetencies that are specific to public service leadership.

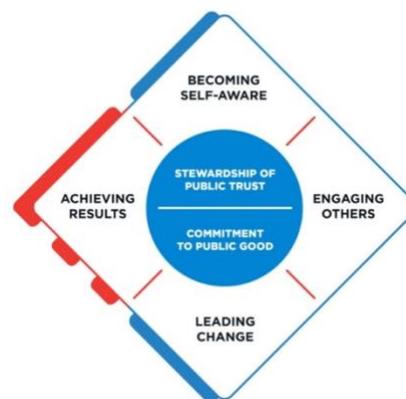
To support continued learning and growth through the application of this model, the Partnership developed the Public Service Leadership 360 assessment tool. This 360 assessment is a multi-rater tool designed to evaluate how well a federal leader is demonstrating the Public Service Leadership Model's values, competencies and subcompetencies. The tool also offers a 360-degree perspective on a leader's skills by incorporating ratings from several key perspectives: self, manager, direct reports, peers, and family or friends.

To collect meaningful data and allow for a robust assessment of their skills, individual leaders are asked to submit a minimum of three names per rater category and encouraged to provide five to six names outside their individual managers. Our [previous analyses](#) demonstrate that the tool is valid and reliable from an analytic perspective, and an effective resource that supports federal employees on their leadership journey.

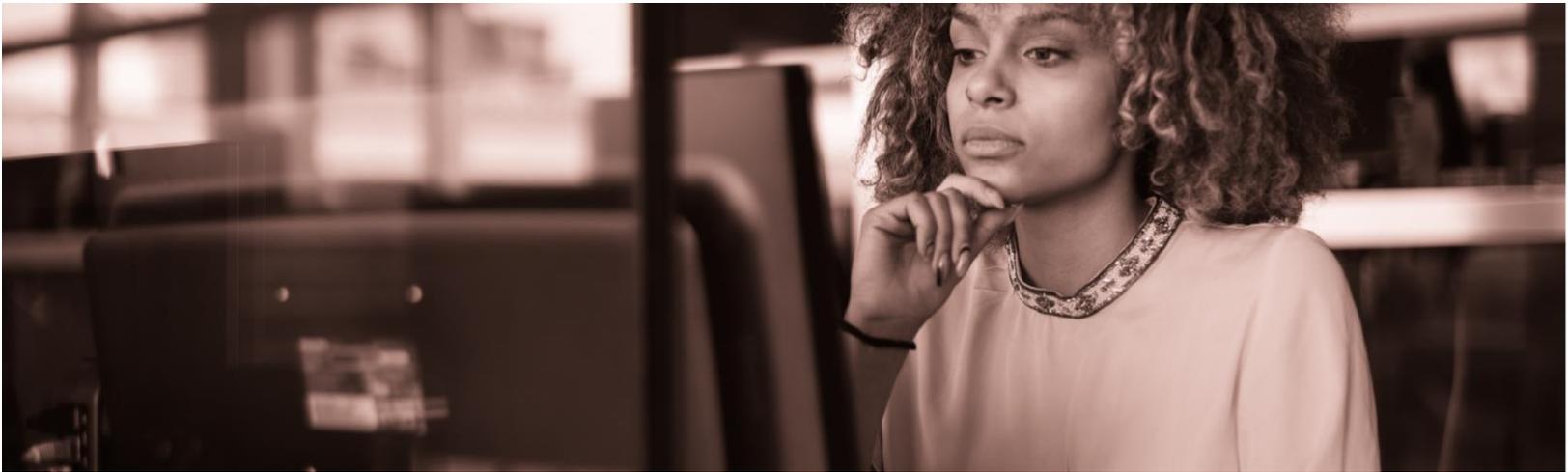
Now that the tool has been used by almost 2,000 federal employees enrolled in our leadership development programs, we identified an opportunity to explore patterns in the data to help us better understand public service leaders.

For this analysis—our first exploration of this 360 data—we compared rating scores for men and women in federal leadership positions. While examinations of gender differences in the federal government are not new, the depth and breadth of our 360 data makes this brief unique.

Our findings will contribute to ongoing research on gender gaps in federal leadership, help us refine our leadership development programming at the [Public Service Leadership Institute](#) and enable us to better support public servants in their growth. It will also enable greater awareness of the potential barriers, or societal or organizational systems that may contribute to the underrepresentation of women leaders in the federal government. Having a federal workforce that better reflects the gender diversity of the United States—while understanding and removing any barriers to increasing the representation of women leaders—is important for everyone, as it will enable our government to better serve and meet the public's needs.



**Figure 1.** Public Service Leadership Model. Circle in center demonstrates the two core values: stewardship of public trust and commitment to public good. Diamonds around the circle list the key competencies: becoming self-aware, engaging others, leading change and achieving results.



## Brief background on 360 assessment tools

Multi-rater assessments such as our 360 tool—assessments that rate employees based on feedback from different categories of individuals—have been used by organizations since at least the 1950s and 1960s. Some of the first research supporting the benefits of having multiple raters provide input on multiple leadership traits emerged in 1967.<sup>2</sup>

Multi-rater assessments have been demonstrated to be an important tool in leadership development programs designed to help individuals grow and develop and become better supervisors of people and more effective leaders of organizations.<sup>3</sup> Using multiple raters to assess employee performance and experiences provides unique and valuable perspectives,<sup>4</sup> and research has long demonstrated a connection between likelihood of promotion and consistency between how employees rate themselves and how others rate them.<sup>5</sup> However, using multiple raters is not without challenges, including possible implicit biases<sup>6</sup>, but understanding overall rating scores, as well as scores by different rater type, is key to supporting and tracking leader development.

---

<sup>2</sup> Lawler, E. E. (1967). The multitrait-multirater approach to measuring managerial job performance. *Journal of Applied Psychology*, 51(5, Pt.1), 369–381. <https://doi.org/10.1037/h0025095>

<sup>3</sup> Dean, Hazel D., et al. "Changing leadership behaviors in a public health agency through coaching and multirater feedback." *Journal of Public Health Management and Practice* 27.1 (2021): 46-54.

<sup>4</sup> Morgeson, Frederick P., Troy V. Mumford, and Michael A. Campion. "Coming Full Circle: Using Research and Practice to Address 27 Questions About 360-Degree Feedback Programs." *Consulting Psychology Journal: Practice and Research* 57.3 (2005): 196.

<sup>5</sup> Thornton, G.C. (1968). The relationship between supervisory and self-appraisals of executive performance. *Personnel Psychology*, 21, 441-456.

<sup>6</sup> Previous researchers have demonstrated that rater type can affect the scores on 360 assessment tools (for example: [Bradley, Thomas P., et al. "Leadership perception." \*Performance Improvement Quarterly\* 19.1 \(2006\): 7-23.](#)). The fundamental attribution error is one theory that explains differences in self-ratings versus ratings by others, specifically when it comes to supervisor ratings versus self-ratings (for example: [Ross, Lee. "The intuitive psychologist and his shortcomings: Distortions in the attribution process." \*Advances in experimental social psychology\*. Vol. 10. Academic Press, 1977. 173-220;](#) Mitchell, Terence R., and Laura S. Kalb. "Effects of job experience on supervisor attributions for a subordinate's poor performance." *Journal of Applied Psychology* 67.2 (1982): 181.). This theory suggests that differences in scores based on rater category may be due to differences in whether the trait being assessed is attributed to internal factors like personality and character traits, or external factors like environmental or situational effects (for example: Ross, Lee D., Teresa M. Amabile, and Julia L. Steinmetz. "Social roles, social control, and biases in social-perception processes." *Journal of personality and social psychology* 35.7 (1977): 485; [Berry, Zachariah, and Joel Frederickson. "Explanations and implications of the fundamental attribution error: A review and proposal." \*Journal of Integrated Social Sciences\* 5.1 \(2015\): 44-57.](#)).

# Who has completed the Public Service Leadership 360 assessment to date

Overall, we reviewed a total of 15,130 sets of rating scores across the model’s two values, four key competencies and 20 subcompetencies. These scores were completed by managers, direct reports, peers, and friends or family for 1,123 federal employees who completed the Public Service Leadership 360 assessment tool from December 2020 to April 15, 2022. Out of the 15,130 ratings,<sup>7</sup> 7% were self-ratings and 93% were ratings completed by others.<sup>8</sup>

We only have demographic data for the federal leaders who received ratings, not for individuals who provided the ratings. This is an area of potential future research and modification of the 360 tool.



<sup>7</sup> These analyses do not include an additional 473 individuals being rated, and 6,846 individual rating scores from 2020 when the tool did not yet include demographic questions.

<sup>8</sup> Due to the low sample sizes of “non-binary” and “not listed” categories, for the purposes of our analyses we focused on “man” and “woman.” Overall distribution can be seen in Figure 2.

# Overview of the demographic data for individual leaders rated by the 360 tool

## Gender

% of total sample for four gender categories. Additional 0.4% skipped this question.

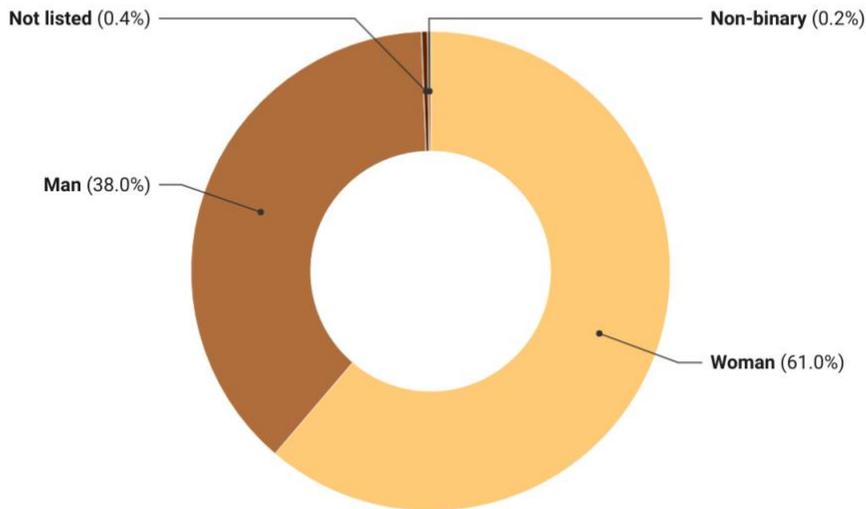


Chart: Partnership for Public Service • Created with Datawrapper

**Figure 2.** Overview of gender data.

## Leader Level

% of leaders in each category of federal leadership.

■ Emerging Leader (GS-7 to GS-11) ■ Leader of Teams or Projects (GS-12 to GS-13) ■ Leader of Leaders (GS-14 to GS-15) ■ Leader of Organizations (Senior Executive Service and political appointees)



Chart: Partnership for Public Service • Created with Datawrapper

**Figure 3.** Overview of leadership level data.

### Supervisor status

% of total sample self-reporting as a supervisor.

No Yes

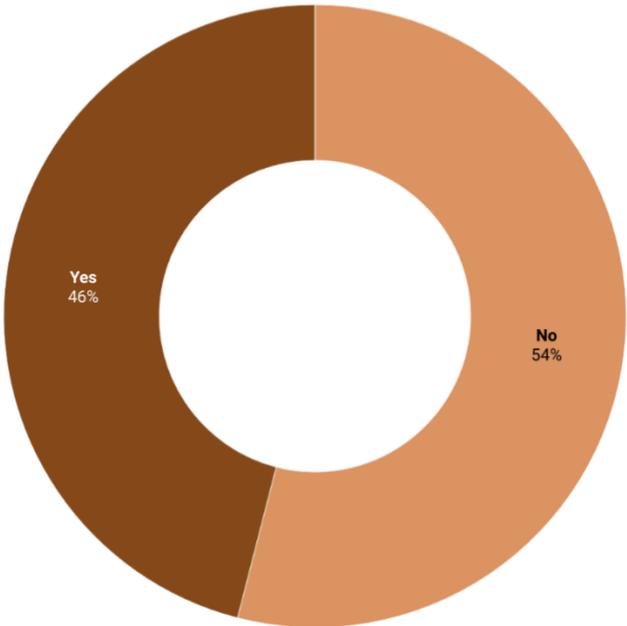
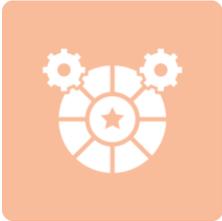


Chart: Partnership for Public Service • Created with Datawrapper

**Figure 4.** Overview of supervisor status.

# Findings



## Overall scores on the key competencies and core values of the Public Service Leadership Model

The Public Service Leadership Model identifies four key competencies that federal leaders need to exhibit to best serve our country—becoming self-aware, engaging others, leading change and achieving results. Additionally, the model outlines two core values—stewardship of public trust and commitment to public good—that federal leaders should demonstrate to live up to the highest ideals of public service.

According to our analysis of the data from the 360 assessments, federal leaders scored on average between 6.33 to 6.86 on these two core values and four competencies. The scores were issued on a scale of 1 to 7, with 1 being the lowest and 7 being the highest.<sup>9</sup> These leaders scored highest on the two core values of stewardship of public trust (6.85) and commitment to public good (6.84). The lowest average score was on the competency of leading change (6.33), which requires leaders to initiate, sponsor and implement innovative solutions in their organizations.

Key Competency or Core Value	Average Score
Stewardship of Public Trust	6.85
Commitment to Public Good	6.84
Achieving Results	6.54
Engaging Others	6.50
Becoming Self Aware	6.47
Leading Change	6.33

**Table 1.** Average scores on core competencies or core values from 1 to 7. 1-never, 2-rarely, 3-occasionally, 4-about half the time, 5-frequently, 6-usually/most of the time, 7-always. See Appendix A for the number of responses.

<sup>9</sup> The score is rating the frequency that the competency is being observed or demonstrated from 1-never, 2-rarely, 3-occasionally, 4-about half the time, 5-frequently, 6-usually or most of the time, and 7-always.



## Overall scores on the subcompetencies

Each of the four competencies in the Public Service Leadership Model includes five subcompetencies that research identifies as critical for leaders to demonstrate.<sup>10</sup> Several examples include emotional intelligence, evidence-based decision-making, equitably engaging a diverse workforce, understanding the importance of technology, and encouraging innovation and creativity.

For the 20 subcompetencies that we analyzed, average scores ranged from 6.16 for self-reflection to 6.78 for integrity. This range demonstrates that the leaders in our sample are generally perceived as exhibiting the 20 subcompetencies usually or most of the time both in their self-ratings and their ratings by others.

The high scores on integrity and diversity, equity and inclusion (DEI) are particularly important. At a time when just [four in 10 Americans](#) trust government to do what is right at least some of the time,<sup>11</sup> agencies need honest leaders who are committed to their oath of office and earn the faith of the public. In addition, the Biden administration has prioritized advancing the federal government's [support for racial equity and underserved communities](#). Leaders' understanding of DEI will be critical as agencies work toward achieving more inclusive and equitable impact of their programs and services.

The three highest and the three lowest subcompetencies are displayed in the table below:

Subcompetency	Average Score
Integrity	6.78
Diversity, Equity and Inclusion	6.66
Collaboration	6.61
Emotional Intelligence	6.21
Embracing Risk and Uncertainty	6.19
Self-reflection	6.16

**Table 2.** The three highest and three lowest subcompetency average scores. Green indicates the top three and red indicates the bottom three. See Appendix A for number of responses.

Interestingly, two of the lowest scoring subcompetencies—emotional intelligence and self-reflection—and the highest scoring subcompetency—integrity—are all within the becoming

<sup>10</sup> Giles, Sunnie. "The most important leadership competencies, according to leaders around the world." *Harvard Business Review* 15.03 (2016).

<sup>11</sup> Hitlin, Paul and Shutava, N. "A close look at public perceptions of the federal government and its employees." (2022) <https://ourpublicservice.org/publications/trust-in-government/>.

self-aware competency. The fact that two of the lowest scoring subcompetencies focus on skills needed to become self-aware may not be a surprise given the gaps in scores we identified and will be discussed below between how employees rate themselves and how others rate them.

However, a low score on self-reflection may simply indicate that leaders lack the time and space to address these gaps and build the emotional intelligence needed to become self-aware. This is something that future research and leadership development trainings might explore to better meet the needs of and strengthen public service leaders.

The other lowest scoring subcompetency—embracing risk and uncertainty—is in the leading change competency, which was also the lowest scoring of the four key competencies. As workplaces continue to move toward hybrid models because of societal shifts due in large part to the COVID-19 pandemic, and efforts to build more inclusive, equitable and accessible workplaces, it will be important to track data trends for this competency to better support leaders who need to manage evolving work patterns and expectations. Furthermore, the increasing complexity of work and ambitious missions of agencies requires leaders to embrace risk, uncertainty and lead through changes, making this an important trend to focus on.



## Values, competencies and subcompetencies: Average scores for men and women provided by all raters

We wanted to know whether the average scores on the key values, competencies and subcompetencies differed for men and women based on both their self-ratings and ratings by others—managers, direct reports, peers, and family or friends. To get at this answer, we performed a t-test—a statistical test that analyzes differences across average scores—to determine if these differences are meaningful or due to random chance.

*Overall, we found that:*

- Women scored statistically significantly higher than men on all the key competencies and core values.
- Women scored statistically significantly higher than men on all 20 subcompetencies except conflict management, embracing risk and uncertainty, adaptability and tech-savviness.
- Women still scored higher than men on three out of these four subcompetencies—all except for embracing risk and uncertainty—even though the scores for them were not statistically significantly different.

Some of these gender differences may be related to role congruity theory, which we discuss in our [introduction](#). That theory posits that women are often rated lower in these areas due to

stereotypical perceptions that they have less technical or professional expertise than men.<sup>12</sup> Another factor may be raters giving lower scores to women based on external or situational factors, such as major organizational transitions or disruptions (e.g. COVID, budget shortcomings, staff turnover, etc.), challenging teams, or other known structural barriers affecting that leader’s performance, rather than on specific attributes of a particular leader.<sup>5</sup>



## Core competencies and values: Self-ratings versus ratings by others

We also examined how individuals perceived their own leadership skills differently than others perceived them. We explored this difference across both genders and then for men and women separately.

Overall, we found that **both men and women rated themselves statistically significantly lower than others rated them for all key competencies and core values.** Differences between self-rating scores and others' rating scores is common with 360 tools and can help leadership coaches provide leaders with more accurate and targeted feedback on areas of improvement. These analyses also help leaders better understand where they may lack self-confidence and enable greater self-awareness in terms of the competencies and skills they bring to the workplace.

Key Competency or Core Value	Rater type	Average Score
<b>Commitment to Public Good</b>	Self	6.75
	Other individuals	6.85
<b>Stewardship of Public Trust</b>	Self	6.67
	Other individuals	6.87
<b>Engaging Others</b>	Self	6.13
	Other individuals	6.53
<b>Achieving Results</b>	Self	6.10
	Other individuals	6.58
<b>Becoming Self Aware</b>	Self	6.10
	Other individuals	6.51
<b>Leading Change</b>	Self	5.80
	Other individuals	6.39

**Table 3.** Average scores on key competencies and core values for self-ratings and ratings completed by others across both men and women. Red indicates the lower score for each competency or value. See Appendix A for number of responses.

<sup>12</sup> Zenger, Jack, and Joseph Folkman. "Women score higher than men in most leadership skills." Harvard Business Review 92.10 (2019): 86-93. [hbr.org/2019/06/research-women-score-higher-than-men-in-most-leadership-skills](https://hbr.org/2019/06/research-women-score-higher-than-men-in-most-leadership-skills).



## Core competencies, values and subcompetencies: Self-ratings based on gender

We found that **women consistently rated themselves slightly higher than men rated themselves** for both core values and three of the four key competencies, with differences being statistically significant for the commitment to public good value and the becoming self-aware competency.

**Women rated themselves lower on leading change than men did.** While this difference is not statistically significant, it may suggest that women have less experience than men in positions that require certain competencies related to leading organizational transformation—or that women are not as confident as men in these areas due to persistent stereotypical notions that identify men as innately possessing certain leadership qualities. Women may also be more self-critical in this category and rate themselves lower, accordingly. It may also be that the leading change competency itself is more complex and harder to grasp than the others—leading both men and women to have less confidence that they exhibit this skill.

### Average scores on core values and key competencies

Average scores for self- versus other gender (man/woman)

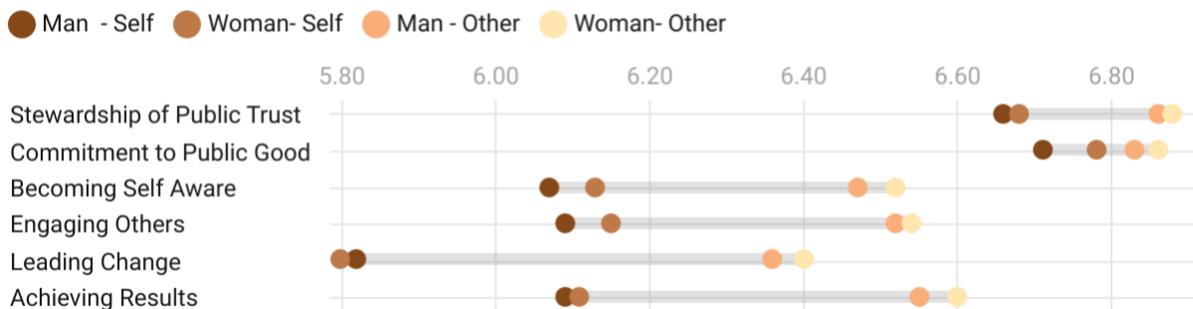


Chart: Partnership for Public Service • Created with Datawrapper

**Figure 5.** Average scores on core values and key competencies by gender (man vs woman) for self and other ratings. See Appendix A for number of responses.

Consistent with this finding, we also found that **women rated themselves slightly lower than men on innovation and creativity and embracing risk and uncertainty**—two leading change subcompetencies—although this difference was not statistically significant.

Overall, **we found that women rated themselves statistically significantly higher than men on emotional intelligence, continuous learning, empowering others, collaboration and customer experience.** Most of these subcompetencies fall under the engaging others or becoming self-

aware competencies—ones that require strong interpersonal capabilities and are often mistakenly seen as “softer” and more easy-to-master skills despite the high level of expertise and aptitude they require. Taken in context of our broader gender differences findings, it is possible—though not certain—that these ratings thus highlight that some women have internalized broader societal stereotypes that they possess these “softer” skills more than men. It is also possible that our sample of federal leaders simply has women outscoring men in these key leadership competencies. Future research might explore the causes of these self-rating differences.

Nevertheless, the finding that women rated themselves higher than men on most competencies is unique to our research. The limited research on 360 assessments that includes women and men typically suggests that women self-rate lower than men—perhaps due to gaps in self-awareness or stereotypical notions of leadership characteristics.<sup>13</sup> This brief contradicts these findings and highlights opportunities to better understand why women self-rate lower than men on certain subcompetencies and not others.

## Average scores on subcompetencies

Average scores for self- versus other and gender (man/woman)



Chart: Partnership for Public Service • Created with Datawrapper

**Figure 6.** Average scores on subcompetencies for self- and other ratings based on gender (man/woman).

<sup>13</sup> Turkel, Catherine C. Female leaders' 360-degree self-perception accuracy for leadership competencies and skills. Capella University, 2008. <https://www.proquest.com/openview/2aca7b62f67d9d95e71987c1bcadad5e/1?pq-origsite=gscholar&cbl=18750>



## Core values, competencies and subcompetencies: How others rated men and women

In most cases, women's ratings of themselves aligned with how others—their managers, direct reports, peers, and family or friends—rated them.

We found that these other raters scored **women statistically significantly higher across three of the key competencies and both core values**. While not statistically significantly different, women also scored higher than men on the fourth key competency: engaging others.

We also found that other raters scored **women higher than men on all the subcompetencies except for embracing risk and uncertainty, where men outscored women by just .01 points**. These differences were statistically significant except for conflict management, embracing risk and uncertainty, adaptability and tech savviness.

Altogether, women scored themselves higher than men—and were scored higher by others—across almost all the core values and competencies.

**The notable exception is the leading change competency.** When rating themselves on this competency, women scored themselves lower—5.80—than men—5.82 and lower than others rated them—6.40. This is a potentially significant difference. Researchers have suggested that a discrepancy between how individuals rate themselves and how they are rated by others may relate to self-confidence and be one potential reason women are underrepresented in leadership positions across sectors.<sup>14</sup> The gap may also be due to internalized stereotypes of leadership, resulting in women feeling that they do not have those skills even when they do or it could be due to self-imposed higher expectations. We will explore this topic further in a future research brief on self-efficacy.

---

<sup>14</sup> [Herbst, Tessie HH. "Gender differences in self-perception accuracy: The confidence gap and women leaders' underrepresentation in academia." SA Journal of Industrial Psychology 46.1 \(2020\): 1-8.](#)



## Conclusion

### *Overall leadership performance of federal leaders in our sample*

The top two scoring competencies or values for all rating scores regardless of gender were **stewardship of public trust and commitment to public good**. These two values are considered core to the Public Service Leadership Model and are unique to skills required for leaders in public service versus other industries. We can conclude that for the federal leaders included in our dataset they are high achieving on these two core values. This finding is important as it suggests that the federal employees in our sample are upholding the [constitutional oath](#) they take when they entered the federal service, since that is where these values derived from.<sup>15</sup>

The top two subcompetencies were **integrity, and diversity, equity and inclusion**. Having federal leaders scoring highly on these competencies is of importance for achieving the Biden administration's focus on building a federal workforce that reflects the diversity of the United States<sup>16</sup> and efforts to advance support for racial equity and underserved communities.<sup>17</sup> Scoring highly on these two competencies suggests the federal leaders in our dataset may have the knowledge, skills, or aptitude to lead with integrity and inclusively, both of which will be necessary to achieve those areas of focus for the Biden administration.

The lowest scoring key competency was **leading change**, and the three lowest rated subcompetencies were **embracing risk and uncertainty, emotional intelligence, and self-reflection**. Emotional intelligence and self-reflection are both within the becoming self-aware competency. While it is difficult to determine from our data if the lower scores in these areas are due to deficits, these lowest scoring competencies do offer opportunities to improve among the federal leaders included in our dataset. Furthermore, it may be that if federal employees are provided more resources, including time and structure, for self-reflection, they may strengthen their emotional intelligence and increase their confidence for leading change.<sup>18</sup>

---

<sup>15</sup> [McDonald, Conant & Marshall 2020](#)

<sup>16</sup> <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/06/25/executive-order-on-diversity-equity-inclusion-and-accessibility-in-the-federal-workforce/>

<sup>17</sup> <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/01/20/executive-order-advancing-racial-equity-and-support-for-underserved-communities-through-the-federal-government/>

<sup>18</sup> Drigas, Athanasios, Chara Papoutsis, and Charalabos Skianis. "Metacognitive and Metaemotional Training Strategies through the Nine-layer Pyramid Model of Emotional Intelligence." *International Journal of Recent Contributions from Engineering, Science & IT (IJES)* 9.4 (2021): 58-76.

For perspective, the range in average scores across all competencies was between 6 and 7 on the 7-point ranking scale. Having an average this high suggests that federal leaders in our sample are perceived as usually or most of the time performing the specific leadership skills in our 360 tool, even for the lowest rated competencies.

However, the range in scores is the largest when people are rating themselves compared to others' ratings. The largest range of scores for self-ratings was in the key competencies of becoming self-aware and leading change. This finding suggests opportunities for improvement for federal leaders in our sample in either their skills or their self-awareness in these areas, or both.

Furthermore, having a large range of scores and discrepancy between self and others' ratings may, in and of itself, suggest a lack of self-awareness.<sup>19</sup>

### ***Gender differences in leadership scores***

While most of the research on 360 ratings focuses on men, studies on women suggest that they tend to self-rate lower than men. Authors hypothesize that this gap exists because women leaders may be less self-aware than men or because external factors—such as systematic stereotypical bias of leadership characteristics favoring men—could result in women self-rating lower and not advancing in the workplace.<sup>12</sup>

Contrary to much of this research, our data shows that **women rated themselves—and were rated by others—higher than men** with only a few exceptions. Specifically, our analysis suggests that women federal leaders have less confidence in their ability to lead change than men. We found that women **self-rated lower** than men on two key leading change subcompetencies: **innovation and creativity** and **embracing risk and uncertainty**.

We also discovered that there was only one subcompetency where men outscored women—embracing risk and uncertainty. Overall, contrary to other historical research, we found that when rated by others, women federal leaders scored higher on all core values, key competencies, and subcompetencies except four where they scored about the same as men: embracing risk and uncertainty, adaptability, conflict management, and tech savviness. These findings suggest that women in federal leadership positions—at least those employees in our sample—are perceived as strong leaders who maintain a diverse skill set. They also may have opportunities to improve their score on the leading change competency or address structural barriers (including stereotypes) that may result in the lower scores in these categories for women.

In addition, men and women federal leaders in our sample **rated themselves significantly lower than others rated them** on all the core values and key competencies—a sign that the becoming self-aware competency is an area of growth for not just women leaders, but for men too. However, women tend to rate themselves more similarly to their raters than men do,

---

<sup>19</sup> Fleenor, John W., Cynthia D. McCauley, and Stephane Brutus. "Self-other rating agreement and leader effectiveness." *The Leadership Quarterly* 7.4 (1996): 487-506.

suggesting a more accurate assessment of their own skills in the workplace. This alignment between how employees rate themselves and how others rate them has implications on effectiveness of leadership trainings for the leader and whether they continue to develop leadership skills over time.<sup>20</sup>

Some research supports the relationship between this scoring alignment and a leader's effectiveness, but it is complex. Overall effectiveness appears to be highest for leaders who rate themselves at the same level as other raters. In contrast, leadership effectiveness tends to be lowest for those who significantly overestimate their performance.<sup>21</sup> Harsher self-ratings may be related to an individual setting high goals or standards and a commitment to continually improve, however, and the alignment of self-ratings and ratings by others may not always predict leader effectiveness.<sup>22</sup> Furthermore, it is possible these discrepancies in scores will reduce with time, driven by motivation to reduce the gap in self-perception and perceptions of others to reduce feelings of cognitive dissonance<sup>23</sup>, or by changing their self-ratings and changing their behavior over time.<sup>24</sup>

Considering that persistent gaps in representation by women exist in our federal leadership, our data is unique in providing some trends for understanding how federal leaders are perceived by themselves and by others on key leadership skills. Given that women scored higher than men across most competencies, there is clearly work to be done to remove barriers to entry for these highly qualified women into the elite leadership ranks, as well as more research to be done to better understand these trends.

---

<sup>20</sup> [Nielsen, Karina, et al. "In the eye of the beholder: How self-other agreements influence leadership training outcomes as perceived by leaders and their followers." \*Journal of business and psychology\* 37.1 \(2022\): 73-90.](#)

<sup>21</sup> Atwater, Leanne E., et al. "Self-other agreement: does it really matter?" *Personnel Psychology* 51.3 (1998): 577-598.

<sup>22</sup> Riester, Devon Nicole. *Self-other agreement and leader effectiveness: An examination of differences across rater sources and leader behaviors*. DePaul University, 2010.

<sup>23</sup> Korman, Abraham K. "Hypothesis of work behavior revisited and an extension." *Academy of Management review* 1.1 (1976): 50-63.

<sup>24</sup> [Carver, Charles S., and Michael F. Scheier. "Control theory: A useful conceptual framework for personality-social, clinical, and health psychology." \*Psychological bulletin\* 92.1 \(1982\).](#)

## Areas for future exploration

The trends we uncovered present multiple areas of future exploration, including:

- Exploring potential barriers or implicit biases that lead to lower scores in the leading change competency for women.
- Exploring if certain aspects of leading in the federal government result in women public servants rating themselves higher than men.
- Examining whether scores are affected by temporal workplace changes, or other changes due to COVID-19 or the ongoing transition to hybrid.
- Analyzing the differences or similarities in scoring based on rater type—for example, how supervisors versus direct reports rate leaders.
- Reviewing patterns in scores on the 360 assessment and any changes in scores after participation in leadership training programs based on gender or other demographic categories.
- Incorporating additional demographic information on the raters—not just the self-raters—as part of an updated 360 tool to explore whether implicit bias or other external factors contribute to differences in self-ratings versus others' ratings.

Depending upon the results of these future explorations, we plan to investigate two key interventions for supporting leadership growth and development, and improvement on 360 scores: coaching and leadership trainings. We propose ongoing evaluation and research to better understand the key drivers of leadership skill development and the alignment of self-ratings versus others' ratings over time, specifically when it comes to federal leaders.

These future analyses will continue to advance our understanding of the barriers that may contribute to the ongoing gaps in representation of women leaders in the federal government. As previously mentioned, understanding and removing these barriers is important for everyone, as it is one way to help government better reflect the diversity of the country and meet the needs of the people.

# Appendix A

Core competency or Value	Number of Responses	Average Score
Stewardship of Public Trust	11553	6.85
Commitment to Public Good	14473	6.84
Achieving Results	11452	6.54
Engaging Others	10977	6.50
Becoming Self-Aware	12280	6.47
Leading Change	10605	6.33

Subcompetency	Number of Responses	Average Score
Integrity	14293	6.78
Diversity, Equity and Inclusion	13393	6.66
Collaboration	13796	6.61
Emotional Intelligence	13980	6.21
Embracing Risk and Uncertainty	11968	6.19
Self-reflection	13605	6.16

Core Competency or Value	Rater type	Number of Responses	Average Score
Commitment to Public Good	Self	1100	6.75
	Other individuals	13372	6.85
Stewardship of Public Trust	Self	953	6.67
	Other individuals	10600	6.87
Engaging Others	Self	1012	6.13
	Other individuals	9965	6.53

<b>Achieving Results</b>	Self	1041	6.10
	Other individuals	10411	6.58
<b>Becoming Self-Aware</b>	Self	1095	6.10
	Other individuals	11185	6.51
<b>Leading Change</b>	Self	1000	5.80
	Other individuals	9605	6.39

Core Competency or Value	Gender + Rating	Number of Responses	Average Score <sup>25</sup>
<b>Stewardship of Public Trust</b>	Man- Self	370	6.66
	Woman- Self	566	6.68
<b>Commitment to Public Good</b>	Man- Self	415	6.71
	Woman- Self	664	6.78*
<b>Becoming Self-Aware</b>	Man- Self	407	6.07
	Woman- Self	667	6.13**
<b>Engaging Others</b>	Man- Self	385	6.09
	Woman- Self	611	6.15
<b>Leading Change</b>	Man- Self	385	5.82
	Woman- Self	598	5.80***
<b>Achieving Results</b>	Man- Self	397	6.09
	Woman- Self	625	6.11

Core Competency or Value	Gender + Rating	Number of Responses	Average Score <sup>26</sup>
<b>Stewardship of Public Trust</b>	Man- Other	4042	6.86
	Woman- Other	6417	6.88*
<b>Commitment to Public Good</b>	Man- Other	5056	6.83
	Woman- Other	8136	6.86**
<b>Becoming Self-Aware</b>	Man- Other	4197	6.47
	Woman- Other	6839	6.52**
<b>Engaging Others</b>	Man- Other	3790	6.52
	Woman- Other	6043	6.54***
<b>Leading Change</b>	Man- Other	3652	6.36

<sup>25</sup> \*Difference in mean score (.068) is statistically significant at p < 0.05; \*\* Difference in mean score (.059) is statistically significant at p < 0.1; \*\*\* Difference (.026) is not statistically significant.

<sup>26</sup> \*Difference in average score is statistically significant at p < 0.05; \*\* Difference in average score is statistically significant at p < 0.001; \*\*\* Difference is not statistically significant, p = .062.

	Woman- Other	5828	6.40*
<b>Achieving Results</b>	Man- Other	3933	6.55
	Woman- Other	6344	6.60**

Subcompetency	Gender + Rating	Number of Responses	Average Score <sup>27</sup>	Category
<b>Self-Reflection</b>	Man- Self	413	5.68	<b>Becoming Self Aware</b>
	Woman- Self	673	5.68	
<b>Authenticity</b>	Man- Self	412	6.18	<b>Engaging Others</b>  Range: 5.68 to 6.43
	Woman- Self	669	6.23	
<b>Emotional Intelligence</b>	Man- Self	417	5.67	
	Woman- Self	675	5.82*	
<b>Integrity</b>	Man- Self	414	6.64	
	Woman- Self	673	6.66	
<b>Continuous learning</b>	Man- Self	415	6.18	
	Woman- Self	675	6.27**	
<b>Relationship Building</b>	Man- Self	417	6.25	
	Woman- Self	674	6.31	
<b>Empowering Others</b>	Man- Self	403	5.88	
	Woman- Self	637	6.01*	
<b>Conflict Management</b>	Man- Self	408	5.68	
	Woman- Self	648	5.68	
<b>Collaboration</b>	Man- Self	412	6.17	
	Woman- Self	668	6.30*	
<b>Diversity, Equity and Inclusion</b>	Man- Self	403	6.39	
	Woman- Self	647	6.43	
<b>Vision Setting</b>	Man- Self	397	5.65	<b>Leading Change</b>  Range: 5.65 to 6.05
	Woman- Self	624	5.67	
<b>Influence</b>	Man- Self	413	5.87	
	Woman- Self	666	5.90	
<b>Innovation and Creativity</b>	Man- Self	409	5.78	
	Woman- Self	649	5.75	
<b>Embracing Risk and Uncertainty</b>	Man- Self	407	5.70	
	Woman- Self	641	5.65	
<b>Adaptability</b>	Man- Self	416	6.03	
	Woman- Self	674	6.05	
<b>Accountability</b>	Man- Self	413	6.19	<b>Achieving Results</b>
	Woman- Self	667	6.23	
<b>Evidence-based Decision Making</b>	Man- Self	417	6.17	<b>Range: 5.97 to 6.23</b>
	Woman- Self	670	6.18	
	Man- Self	515	6.05	

<sup>27</sup> \*Difference in mean score is statistically significant at p < 0.01; \*\*Difference in mean score is statistically significant at p < .05.

<b>Systems Thinking</b>	Woman- Self	667	6.06
<b>Tech Savviness</b>	Man- Self	414	5.97
	Woman- Self	664	5.99
<b>Customer Experience</b>	Man- Self	405	6.03
	Woman- Self	646	6.12**

<b>Subcompetency</b>	<b>Gender + Rating</b>	<b>Number of Responses</b>	<b>Average Score<sup>28</sup></b>	<b>Category</b>
<b>Self-Reflection</b>	Man- Other	4722	6.16	<b>Becoming Self Aware</b>  <b>Range: 6.16 to 6.81</b>
	Woman- Other	7611	6.22*	
<b>Authenticity</b>	Man- Other	4672	6.59	
	Woman- Other	7535	6.62**	
<b>Emotional Intelligence</b>	Man- Other	4826	6.21	
	Woman- Other	7862	6.28*	
<b>Integrity</b>	Man- Other	4971	6.77	
	Woman- Other	8031	6.81*	
<b>Continuous Learning</b>	Man- Other	4877	6.54	
	Woman- Other	7850	6.59*	
<b>Relationship Building</b>	Man- Other	5066	6.54	<b>Engaging Others</b>  <b>Range: 6.25 to 6.69</b>
	Woman- Other	8161	6.58*	
<b>Empowering Others</b>	Man- Other	4390	6.39	
	Woman- Other	6950	6.42**	
<b>Conflict Management</b>	Man-Other	4303	6.25	
	Woman- Other	6950	6.28***	
<b>Collaboration</b>	Man- Other	4821	6.61	
	Woman- Other	7697	6.65*	
<b>Diversity, Equity and Inclusion</b>	Man- Other	4631	6.65	
	Woman- Other	7521	6.69**	
<b>Vision Setting</b>	Man- Other	4454	6.30	<b>Leading Change</b>  <b>Range: 6.24 to 6.48</b>
	Woman- Other	7207	6.38*	
<b>Influence</b>	Man- Other	4790	6.41	
	Woman- Other	7689	6.48*	
<b>Innovation and Creativity</b>	Man- Other	4447	6.25	
	Woman- Other	7037	6.30**	
<b>Embracing Risk and Uncertainty</b>	Man- Other	4412	6.25	
	Woman- Other	6613	6.24***	
<b>Adaptability</b>	Man- Other	4817	6.42	

<sup>28</sup> \*Difference in mean score is statistically significant at p < 0.01; \*\*Difference in mean score is statistically significant at p < .05.; \*\*\*Difference is not statistically significant.

	Woman- Other	7762	6.44***	<b>Achieving Results</b>  <b>Range: 6.49 to 6.62</b>
<b>Accountability</b>	Man- Other	4581	6.55	
	Woman- Other	7400	6.62*	
<b>Evidence-based</b>	Man- Other	4908	6.52	
<b>Decision Making</b>	Woman- Other	7846	6.57*	
<b>Systems Thinking</b>	Man- Other	4814	6.54	
	Woman- Other	7733	6.58*	
<b>Tech Savviness</b>	Man- Other	4612	6.57	
	Woman- Other	7304	6.58***	
<b>Customer Experience</b>	Man- Other	4525	6.49	
	Woman- Other	7375	6.54*	

# Project Team

## PARTNERSHIP FOR PUBLIC SERVICE

**Loren DeJonge Schulman**

Vice President, Research, Evaluation and  
Modernizing Government

**Samantha Donaldson**

Vice President, Communications

**Nadzeya Shutava**

Research Manager

**Max Stier**

President and CEO

**Andrew Marshall**

Vice President, Leadership Development

**Barry Goldberg**

Senior Writer and Editor

**Christina Schiavone**

Director for Executive and Team Coaching



600 14th Street NW  
Suite 600  
Washington, DC 20005

<https://ourpublicservice.org/>  
(202) 775-9111

 [partnershipforpublicservice](https://www.facebook.com/partnershipforpublicservice)

 [@PublicService](https://twitter.com/PublicService)

 [rpublicservice](https://www.instagram.com/rpublicservice)