

Kathryn Townsend, Commissioner  
Arizona Commission on Judicial Performance Review  
631 N Manzanita Pl  
Saint David, AZ 85630  
520-861-4533  
[KathrynDTownsend@Protonmail.com](mailto:KathrynDTownsend@Protonmail.com)

In the Matter of:	)	
	)	SUPREME COURT NO. R- __ - __
PETITION TO AMEND RULE 5,	)	
RULE 6, RULE 7, AND RULE 12	)	<b>Petition to Amend Rule 5, Rule</b>
OF RULES	)	<b>6, Rule 7, and Rule 12 of Rules</b>
OF PROCEDURE FOR	)	<b>of Procedure for Judicial</b>
JUDICIAL PERFORMANCE	)	<b>Performance Review in the</b>
REVIEW IN THE STATE OF	)	<b>State of Arizona</b>
ARIZONA	)	

---

Pursuant to Rule 28, Rules of the Supreme Court, Petitioner respectfully petitions this Court to adopt amendments to the Rules of Procedure for Judicial Performance Review in the State of Arizona.

### **I. Background and Purpose of the Proposed Rule Amendments**

The Judicial Performance Review (JPR) process in Arizona evaluates judges by conducting a census survey of all eligible individuals. A census survey, in theory, is designed to collect feedback from every person within a defined population, ensuring comprehensive representation and inclusivity. For such a method to be effective, the survey must achieve high participation rates in order to accurately reflect the experiences and opinions of the population. Unless non-

responses are minimized, the results of a census survey will be skewed and non-representative of the people surveyed.

The most glaring issue of the current census method is the low response rates some judges receive. For example, during the 2024 cycle, response rates below 15% were not uncommon and in one case, there was a judge with a response rate of zero from jurors who received surveys. Another judge had a response rate of 3% from litigant/witnesses. The lack of participation makes the census incomplete, introducing a systemic flaw called non-response bias.

Non-response bias occurs when the individuals who choose not to participate in a survey differ systematically from those who do respond. This bias is particularly problematic in the JPR process where it is clear there is some kind of non-response bias, but no effort is made to identify its nature. Common sense says that individuals with strong opinions, whether positive or negative, are more likely to respond, but that supposition has not been verified with data because the commission has made no effort to determine how those who respond differ from those who do not. What is clear, however, is that in far too many cases, the paucity of responses renders the data in some judges' surveys unreliable. It's also clear that the way the census survey respondents self-select whether or not to participate has the result of disproportionate representation, with some groups' voices amplified and others entirely erased from the evaluation process.

In spite of efforts to increase response rates, the census method employed by the current JPR process consistently fails to achieve high participation rates and non-responses are all too common. Worse, because the commission doesn't know the opinions of those who fail to respond, there's no legitimate way within the context of a commissioner casting a vote or inviting a judge to appear in front of the commission for anyone serving on the commission to determine if those who are pleased, disappointed, or neutral are more likely to return surveys. This lack of ability to determine the nature of the non-response bias is more than troubling. It transforms the JPR process from one that is mathematically, statistically, or morally robust into something that is more akin to Yelp for judges.

The effects of non-response bias extend beyond disproportionate representation. When participation rates are low, the data collected no longer reflects the true distribution of opinions in the broader population. Instead, the survey results are skewed toward the perspectives of the most vocal respondents. For instance, a judge might receive overwhelmingly negative feedback because dissatisfied individuals are more likely to respond, even if the majority of non-respondents hold a neutral or positive view. Conversely, a judge might appear to perform exceptionally well if satisfied respondents dominate the survey pool while critics abstain. In both scenarios, the results provide a distorted view of the judge's performance and, because there has been no effort to determine how those who

don't respond differ from those who do, there is no way for anyone to correct the data to account for non-responses. Further, there's no empirical evidence that has been provided to the commission that definitively demonstrates that there is a way for the commission to realistically achieve the kind of universal response rate that legitimate census surveys strive for.

To put the problem of non-responses in perspective, think of the United States' decennial census survey and the allotment of Congressional seats. If only 3% of Arizonans responded, instead of the 7,151,502 people that were counted in 2020, Arizona would be credited as having a population of 214,545 and instead of having nine Congressional seats, the state would have the minimum number. One. The validity of the US Census depends entirely on the cooperation of the American people and their compliance with the requirement that they return the census form. The United States Government's commitment to getting census forms back means that it tracks those surveys and knows who has and who hasn't returned the form and will send a census taker to the front door of people who don't comply in order to get the data it needs. Unlike the United States Government, the Arizona Commission on Judicial Performance review does not aggressively follow up on survey forms. As a result, the return rates are often far too low to draw any reasonable conclusion from the data collected.

The consequences of publishing such skewed data are profound. Judicial performance reviews play a crucial role in shaping public opinion and informing decisions about judicial retention. When non-response bias distorts the data, the public is misled, and the integrity of the review process is compromised. A judge with biased survey results may face undue criticism or unearned praise, leading to decisions that do not reflect the reality of their performance. A method that consistently produces such unreliable outcomes undermines trust in the judiciary and fails to fulfill its purpose of promoting accountability and public confidence.

In the most recent cycle, to mitigate the impact of low response rates on reported numbers, the Commission, on advice of the AOC, voted to aggregate all of the scores across respondents and report a composite percentage for each evaluation area. However, that was a single vote in a specific circumstance, not a rules change and the census data with the low response rates were also reported, giving the public a false picture of at least some judges' performances.

Reforming the JPR survey method is not merely an administrative necessity but a moral imperative. The flaws inherent in the current census approach are evident and pervasive. A reliable system must address these issues by ensuring participation, reducing bias, and producing representative data. Fortunately, there are multiple viable alternatives that can achieve these goals, each offering a clear and practical solution to the problems posed by the census method.

## **II. Summary of Proposed Rule Amendments.**

Section II briefly outlines five ways the challenges of the census survey can be addressed and summarizes the proposed rules changes. The language of the proposed rule amendments can be found in Appendices A through E.

### **Alternative Proposal Appendix A – Adopt a Sampling Survey Approach**

One option to address the failings of the census surveys currently in use is to implement a sampling survey method. By using statistically valid sampling techniques the process can achieve representativeness without the inefficiencies of a full census. Sampling should reduce the number of surveys that must be sent, allow for follow up to achieve compliance, and minimize non-response bias while preserving the ability to evaluate judges fairly across all relevant categories. This method balances the need to hear from multiple constituencies with practicality, ensuring that the evaluation process remains efficient and effective.

#### *Overview of Changes Proposed in Appendix A*

##### *Rule 5(b)(1)(A) and 5(b)(1)(B): Surveys, Purpose, and Sampling Methodology*

- Introduces a requirement for surveys to use a random sampling methodology with a 95% confidence level and a  $\pm 5\%$  margin of error.
- Adds language to flag and exclude survey data that fails to meet these statistical thresholds.

- Specifies that invalid survey data must not influence evaluations, with alternative evaluation tools used when necessary.

*Rule 5(b)(1)(C) and 5(b)(1)(D): Confidentiality of Surveys and Data Handling*

- Clarifies that respondent identities may be collected for validation and tracking purposes but must remain confidential and anonymized before analysis.
- Expands on the rule to align with Rule 12(b), ensuring full compliance with confidentiality standards.

*Rule 5(c)(1) through 5(c)(3): Data Report Preparation, Dissemination, and Publication of results*

- Specifies that survey data failing to meet statistical standards must be flagged as invalid and excluded from the data report.
- Adds a requirement to disclose the limitations of invalid data in the report.

*Rule 6(a): Relevant Information*

- Requires the survey data used for decision making meet the standards set forth in 5(b)(1)(A) and 5(b)(1)(B).
- Includes a provision allowing the Commission to rely on other tools, such as public comments and disciplinary records, when survey data is deemed statistically invalid.

*Rule 7(a)(1)(D): Commission Report Content*

- In the event insufficient data is collected, adds language requiring the report to state, "The Commission for Judicial Performance Review failed to collect sufficient data to fairly evaluate this judge."

*Rule 12(b): Anonymity and Confidentiality*

- Revises the rule to explicitly allow for the collection of survey respondent identities for validation purposes, provided they remain anonymized during analysis and reporting stages.

**Alternative Proposal Appendix B – Limit Surveys to Attorneys**

Another alternative is to limit surveys to attorneys while mandating their participation as a condition of bar membership. Attorneys are uniquely positioned to provide informed feedback on judicial performance based on their professional interactions with judges. Requiring attorneys to respond would ensure robust and representative input from a knowledgeable group. This approach not only addresses the issue of low response rates but also focuses the evaluation process on those most qualified to assess judicial competency. The JPR rules petition only specifies limiting surveys to attorneys, because modifying bar requirements is outside the scope of the JPR rules.

*Overview of changes in Appendix B*

*Rule 5(b)(1): Surveys*

- The purpose and scope of surveys focus specifically on licensed attorneys who meet the eligibility criteria outlined in Rule 5(b)(1)(D)(i).

*Rule 5(b)(1)(D): Respondents*

- Survey participation is limited to licensed attorneys who directly interact with the judge during the evaluation period.

**Alternative Proposal Appendix C – Adopt Measurable Data Standards**

A third approach is to rely on measurable data, as Utah does, while still retaining the use of some non-measurable data. The measurable data consists of metrics judges must meet in order to demonstrate they have achieved expected performance standards. These objective performance metrics eliminate the variability and unreliability inherent in subjective surveys. Measurable indicators provide consistent benchmarks for evaluating judicial performance, offering a data-driven foundation that sidesteps the issue of non-response entirely. This method ensures that evaluations are grounded in tangible outcomes rather than subjective perceptions.

*Overview of Changes in Appendix C*

*Rule 5(a)(1): Measurable Standards*

- Introduces measurable standards for judicial performance, including:
  - Timeliness: Resolving cases within specific timeframes (60 days for trial judges, six months for appellate judges).

- Judicial Education: Completion of 16 hours of continuing education annually.
- Case Management: Maintaining a 90% case clearance rate.
- Physical and Mental Competence: Judges self-certify annually and may be required to provide additional verification if concerns arise.

*Rule 5(a)(2): Non-Measurable Standards*

- Defines non-measurable standards, including:
  - Disciplinary Records: Evaluation of records since the last survey period in accordance with the Arizona Commission on Judicial Conduct rules.
  - Public Comments: Consideration of oral and written comments received during the relevant survey period.

*Rule 5(a)(3): Exceptions to Measurable Standards*

- Allows exceptions to measurable standards in extraordinary circumstances, such as personal illness, family emergencies, case complexity, or unanticipated workload increases.
- Judges may submit supporting statements to explain deviations but are not required to.

*Rule 5(a)(4): Support for Exceptions*

- Allows judges requesting an exception to submit supporting statements from individuals with knowledge of the circumstances but prohibits using the absence of such statements as a basis to deny an exception.

*Rule 5(b): Tools for Evaluating Performance*

- Clarifies when surveys will occur.
- Removes conflicting language regarding subjective surveys.
- Outlines who will be surveyed to obtain measurable data and information regarding non-measurable standards and what additional tools may be used and when.

*Rule 5(c): Data Report Preparation and Dissemination*

- Requires the Commission to prepare data reports that include measurable metrics.
- Judges can review their data, address deviations, and request exceptions before publication.
- Requires the Commission to refer breaches of professional standards to the Commission on Judicial Conduct.
- Addresses how data regarding measurable metrics and non-measurable metrics will be reported.

- Addresses how reports on judges granted an exception to the measurable standards will be written.

### **Alternative Proposal Appendix D – Treat Non-Responses as Neutral**

For those concerned about the implications of non-responses, two additional methods provide alternative solutions. One is to treat non-responses as neutral, assigning a midpoint score that neither penalizes nor rewards judges for low participation. This mitigates the impact of non-responses and ensures that the data remains balanced and fair.

#### *Overview of Changes in Appendix D*

##### *Rule 5(c): Treatment of Non-Responses*

- Defines a "non-response" as the failure to return a completed survey within the prescribed timeframe.
- Specifies that non-responses are treated as neutral:
  - For numerical scoring (e.g., 1–5 scale), a neutral score is assigned as the midpoint of the scale (e.g., 3 on a 1–5 scale).
  - For binary questions (e.g., yes/no), non-responses are excluded from calculations.
- Requires public disclosure of the total number of respondents, non-responses, and the impact of non-responses on final scores. A statement accompanies the results to explain the neutral treatment.

### *Rule 5(c)(d): Data Report*

- Updates references to integrate the new Rule 5(c) on non-responses.
- Ensures that data reports account for non-responses and reflect their treatment in the calculation of results.

### **Alternative Proposal Appendix E – Treat Non-Responses as Positive.**

Alternatively, non-responses could be treated as positive. This approach assumes that individuals who do not respond are generally satisfied, addressing the tendency for dissatisfied individuals to dominate the respondent pool. This approach also mitigates the impact of non-responses and ensures that the data remains balanced and fair.

### *Overview of Changes in Appendix E*

#### *Rule 5(c): Treatment of Non-Responses*

- Defines "non-response" as the failure to return a completed survey within the prescribed timeframe.
- Treats non-responses as positive evaluations:
  - For numerical scoring (e.g., 1–5 scale), assigns the highest score (e.g., 5 on a 1–5 scale).
  - For binary questions (e.g., yes/no), includes non-responses in a way that gives the judge a positive response.

- Requires public disclosure of the total number of respondents, non-responses, and the impact of non-responses on final scores. A statement accompanies the results to explain the positive treatment.

*Rule 5(c)(d): Data Report*

- Updates the data reporting process to incorporate the treatment of non-responses.
- Ensures reports include the impact of non-responses on aggregate scores and discloses the rationale for their positive treatment.

### **III. Similar Petitions Filed Within The Last Five Years.**

Over the past five years, two petitions have addressed changes to the JPR Rules. The most recent, Petition R-23-0036, replaced the previous JPR Rules and introduced a new set of rules effective October 1, 2023, which are now under consideration in this petition. Earlier, Petition R-19-0039 proposed amendments developed by a Task Force to formalize the existing policies and procedures of the Commission on Judicial Performance Review. Those amendments were adopted on an emergency basis on September 1, 2019, and later permanently approved by the Court on December 12, 2019.

#### **IV. Conclusion**

Each of these alternatives represents a distinct solution to the weaknesses of the current census approach. While they differ in methodology, they share a commitment to improving transparency in judicial performance evaluations. Moreover, the existence of multiple viable options underscores the feasibility of reform. The problems posed by the census method are not insurmountable; they can, and should, be addressed through thoughtful and deliberate action.

Beyond being Constitutionally mandated, the JPR process is critical to maintaining public trust in Arizona's judiciary. It ensures accountability, transparency, and confidence in the courts. However, these goals cannot be achieved without a reliable and equitable evaluation method. By adopting one of the proposed alternatives, this Court can strengthen the integrity of the JPR process and restore confidence in the system. Reform is not just a possibility—it is an imperative.

RESPECTFULLY SUBMITTED this 9<sup>th</sup> day of January, 2025

By /s/ Kathryn Townsend

Kathryn Townsend, Member  
Commission on Judicial Performance Review  
631 N Manzanita Pl  
Saint David, AZ 85630  
520-861-4533  
KathrynDTownsend@Protonmail.com