

American Federation of Government Employees

Affiliated with the AFL-CIO

Council 260

June 2017

Evaluation of Recently Issued Performance Plans in Research Services (R)

Introduction

Over the past several years, Union representatives have processed numerous individual grievances related to the performance management system. In the course of this work, we have identified systemic problems in the development, issuance, and management of performance plans NARA-wide. This is an analysis of performance plans issued to NARA's bargaining unit employees in various components of Research Services. It is meant to begin a discussion between the AFGE NARA Council and NARA's leadership about improving performance plans not just in Research Services, but NARA-wide.

Background

The Union has typically challenged performance plans individually through the negotiated grievance procedure when employees were harmed by them, i.e. rated below fully successful, when the plan failed the standards set forth in the collective bargaining agreement, NARA 430, or applicable administrative law. Several years of this approach have failed to bring about a large-scale review of NARA's performance plans to remedy the common errors and deficiencies. In some cases, performance plans were not changed even after final disposition of a grievance successfully challenged a rating based on them. There has, however, been an improvement in some aspects of the performance management culture at NARA. For example, the Agency has improved significantly in issuing timely mid-year reviews. In addition, Human Capital officials have held a series of informative presentations for employees on the performance management system.

Improvement in wording and structure of performance plans and performance standards is harder to find. The Union suggests that the time has come for an Agency-wide review of performance plans. We are issuing this document rather than filing a grievance because we recognize that this is a massive project, and one that we would like to be collaborative rather than adversarial. Since inadequate performance plans are a condition of employment that renews every day the plan is in effect, the Union may at some point file group or individual grievances should this initial approach fail.

Our anecdotal experience suggested that many poor performance plans are simply a result of neglect. For instance, a particular job might have a plan associated with it that was drafted 15 years ago and never updated because nobody took the initiative to do so. However, a new plan is not necessarily a better plan. The recently issued performance plans in Research Services suggest that plans just now being created suffer from some of the same problems we see in old plans. Using R as an example, we have determined that too many revisions to plans at the office level

appear to have been done without sufficient rigor, and without adequate knowledge of the intricacies of performance management in the federal sector. Aside from being riddled with typographical, agreement, and punctuation errors; poor word choice; and improper capitalization--an ironic feature in light of the functions designated in the new R system as Communication 1 and Reporting 3¹--the plans fail the basic tests established through collective bargaining, NARA 430, in Merit System Protection Board case law, and as expressed in OPM's most recent Handbook for Measuring Employee Performance.²

R Performance Plans

For FY 2017, bargaining unit employees in R were issued new plans using a new model. The plans have multiple elements, with one or more "functions" falling underneath those elements. These plans appear to have been pieced together using a master list or matrix of all possible functions under the various elements. It is not clear how the various functions falling under each element are weighed in determining a rating on that element.

We have noticed several problems that are common to the new R plans. First we will discuss the basis of our analysis before moving on to specific problematic elements from the plans.

1) No differentiation between types of elements

The plans fail to convey to the employee which elements are critical elements and which are non-critical elements. A critical element is "a component of a position consisting of one or more duties and responsibilities that contribute toward accomplishing organizational goals and objectives and which is of such importance that unacceptable performance on the element would result in unacceptable performance in the position."³ This means that the employer bears a special burden to clearly articulate standards for critical elements. Other elements could be classified as non-critical, such as if the elements measure group work or set forth general principles but are not established as critical for success in the position. The plans do not convey to employees which elements are of which kind.

2) Insufficient regard for objectivity of performance criteria

Under 5 USC §4302(b)(1), each performance appraisal system shall provide for establishing performance standards which will "to the maximum extent feasible, permit the accurate evaluation of job performance on the basis of objective criteria."

¹ Fully Successful Standard for Communications 1 is: "Written and oral work, regardless of intended recipients, rarely needs revision for clarity, accuracy, responsiveness, conciseness, courtesy, conformity with relevant guidelines, standards, policies, procedures, and contains no typos." The Fully Successful standard for Reporting 3 is "Written assignments are well researched and written in clear grammatically correct language."

² See United States Office of Personnel Management, *A Handbook for Measuring Employee Performance*, March 2017, available at https://www.opm.gov/policy-data-oversight/performance-management/measuring/employee_performance_handbook.pdf

³ NARA Personnel 300, Chapter 430

NARA 300 Chapter 430 states that “To the extent possible, performance standards include objective measures of performance such as quality, quantity, cost efficiency, timeliness, and manner of performance, and may include other measures of performance as appropriate.”

We believe the examples below suggest an insufficient regard for developing objective criteria when possible.

3) Plans must be SMART when possible

The SMART principle is that the Agency must endeavor to create performance standards that are Specific, Measurable, Achievable, Relevant, and Time-bound is part of our collective bargaining agreement.

NARA and AFGE have established the procedures for applying the performance management system in Article 16 of the National Agreement. According to Article 16, Sec 2: “To the extent possible, standards will be specific, measurable, achievable, relevant, and time-bound.”

Article 16, Section 3 further states that “Each employee will be issued a performance plan in accordance with Chapter 430 on NA Form 3020 with critical elements and performance standards specifically defined for her or his position.” The method used to develop these plans calls into question whether they really are standards that are “specifically defined” for an employee’s position.

4) The MSPB Test

The Merit Systems Protection Board has issued a significant body of case law involving performance standards for agencies to consider when developing standards.

In sum, critical element standards should:

- Be "reasonable, realistic, attainable, and clearly stated in writing. . . .specific enough to provide an employee with a firm benchmark toward which to aim her performance, . . . and must be sufficiently precise so as to invoke general consensus as to their meaning and content.” Gladys Towne v. Dept. of Air Force, 2013 MSPB 81 (2013)
- Be reasonable and valid (Jackson v. Dept of VA, 97 MSPR 13 (2004)) which can mean that a standard may not demand too high a degree of error-free work. Blain v. VA, 36 MSPR 322 (1988) and Walker v. Dept.of Treasury, 28 MSPR 227 (1985) or too much work be done in the time available.
- Be capable of being “applied in a verifiable fashion.” Stone v. HHS, 35 MSPR 603 (1987)
- Create a “firm benchmark.” Melnick v. HUD, 42 MSPR 93 (1989)
- Be realistic and attainable. Greer v. Dept of the Army, 79 MSPR 477 (1998)
- Describe with clarity how an employee can attain a particular rating in any given element other than fully successful. Van Pichard v. Dept. of Defense, 2011 MSPB 94 (2011)

Critical element standards should not:

- Be stated backwards in terms of what the employee should not do. See Jackson-Francis v. OGE, 2006 MSPB 255.
- Use vague measurements such as “somewhat short.” 38 MSPR. 634 (1988)
- Be subject to various interpretations and subjective evaluation. Eibel v. Dept. of Navy, 857 F.2d1439 (1988)
- Be written as an elusive goal which the agency may find the employee met or failed to meet at its pleasure. Donaldson v. Department of Labor, 27 MSPR. 293, 298 (1985)

While MSBP is interested chiefly in the adequacy of retention standards, we believe it is in both management’s and employees’ interest that performance plans be sufficiently clear to inform employees of exactly what is expected of them.

5) Examples from the R Plans

The following critical element components from the R plans either fail to conform to the positive criteria, or suffer from some of the deficiencies, that we’ve described above. This is not an exhaustive list of problems; rather, they are examples that are emblematic of widespread problems.

Collaboration 2

Under the Communication/Collaboration element, the Collaboration 2 function includes the following benchmarks:

Minimally Successful

Contributes to the development of formal and informal training for archival employees to promote consistent, effective archival practices.

Fully Successful

Participates in the development and delivery of formal and informal training for archival employees to promote consistent, effective archival practices.

Highly Successful

Leads in the development and delivery of formal and informal training for archival employees to promote consistent, effective archival practices.

It is unclear what differentiates “contribut[ing] to the development of formal and informal training for archival employees” from “participat[ing] in the development and delivery of formal and informal training for archival employees. These standards do not convey a meaningful measurement that communicates to employees what is expected of them. The criterion for highly successful, meanwhile, specifies that an employee performing in this category “leads in the

development and delivery of formal and informal training for archival employees...” The additional problem with these standards is that they are too dependant on the work that was assigned to the employee. For instance, an employee who is assigned to lead a training activity inevitably has an advantage over one who is not. Conversely, an employee who improperly asserts a leadership over training development when someone else is assigned that task risks acting contrary to the instructions of the supervisor.

Collaboration 11

Under the Communication/Collaboration element, the Collaboration 11 function includes the following benchmarks:

Minimally Successful

Usually shows respect for others [*sic*] backgrounds, experiences, talents, abilities, thoughts, and opinions. Exceptions are rare and do not adversely affect staff morale.

Fully Successful

Consistently shows respect for others [*sic*] backgrounds, experiences, talents, abilities, thoughts and opinions.

Highly Successful

Always shows respect for others [*sic*] backgrounds, experiences, talents, abilities, thoughts, and opinions. Actively seeks feedback from all staff to ensure all perspectives are considered.

Collaboration 11 is among the most bizarre standards we have seen included in a performance plan. It fails the test of specificity, measurability, and achievability. It is unclear how to differentiate the fully successful benchmark from the highly successful one. First, the difference between the highly successful and fully successful standards depends in part on the difference between the words “consistently” and “always” This is not clear, specific, or measurable, since consistently means “unchanging in nature, standard, or effect over time” or, in other words, always. Second, it is unclear how an employee not gifted with extrasensory perception shows respect for the thoughts of another. Third, it should go without saying that some thoughts, if expressed, do not deserve and should not receive respect.

Collaboration 12

Under the Communication/Collaboration element, the Collaboration 12 function includes the following benchmarks:

Minimally Successful

Demonstrates awareness of need to cooperate with personnel in other agencies.

Fully Successful

Maintains positive working relationships with personnel in other agencies.

Highly Successful

Creates positive working relationships with personnel in other agencies.

Similar to a problem noted in Collaboration 11, the distinction between creating and maintaining a positive relationship is too meager to be meaningful to employees, or to be meaningfully measured by the supervisor. It is also likely to be driven by events outside the control of the employee. For example, an employee engaged in contacts with four agencies that have a turnover in personnel of their points of contact will have a far greater opportunity to “create positive working relationships” than a similarly situated employee whose POCs remain unchanged over the same period.

Communication 8

Under the Communication/Collaboration element, the Communication 8 function includes the following benchmarks:

Minimally Successful

Asks questions of supervisor when problems are encountered and sometimes follows prescribed instructions correctly.

Fully Successful

Asks appropriate questions of supervisor when problems are encountered and consistently follows prescribed instructions correctly.

Highly Successful

Asks meaningful questions of supervisor when problems are encountered, suggests possible solutions, and always follows prescribed instructions correctly.

Again, the distinction between fully successful and highly successful is not specific and measurable: It depends on distinguishing appropriateness from meaningfulness. Can a question be appropriate but not meaningful, and vice versa? Are supervisors really going to ponder such questions when evaluating employees, and explain the difference come appraisal time? The benchmark for this function includes the incomprehensible distinction between “consistently” and “always” described above.

Knowledge 1

Under the Technical Capability element, the Knowledge 1 function includes the following benchmarks:

Minimally Successful

Demonstrates general working knowledge of NARA and unit policies and procedures relating to archival work.

Fully Successful

Demonstrates in-depth knowledge of NARA and unit policies and procedures relating to archival work.

Highly Successful

Shows a thorough understanding of NARA and unit objectives and requirements and policies and procedures regarding archival work.

Here again, the fully successful and highly successful standards are not adequately differentiated. How do rating officials explain what differentiates “demonstrat[ion] of in-depth knowledge of NARA and unit policies and procedures” from “show[ing] a thorough understanding of NARA and unit objectives and requirements and policies and procedures.” If this difference can be explained, such detail should be conveyed in the plan itself.

Knowledge 1

Under the Technical Capability element, the Communications 1 function includes the following benchmarks:

Minimally Satisfactory

Written and oral work, regardless of intended recipients, regularly needs revisions for clarity, accuracy, responsiveness, conciseness, courteousness, conformity with relevant guidelines, standards, policies, procedures, or typos. The pattern of revisions may persist, or may change, showing improvement in one aspect but deficiencies in new aspects.

Fully Successful

Written and oral work, regardless of intended recipients, rarely needs revisions for clarity, accuracy, responsiveness, conciseness, courtesy,⁴ conformity with relevant guidelines, standards, policies, procedures, and contains no typos.

Highly Successful

Written and oral work exceeds Fully Successful standards. Drafts and submits updates or revisions to established local or divisional guidelines, standards, and/or procedures as circumstances dictate.

The standard establishes the need for regular revision as minimally successful, describing the negative rather than the positive characteristics that must be reached for the retention standard.

Knowledge 3

Under the Technical Capability element, the Knowledge 3 function includes the following benchmarks:

⁴ It is unclear why the minimally satisfactory standard evaluates employees on courteousness--a mindset--while the fully successful standard evaluates them on courtesy--the behavior resulting from that mindset. If this discrepancy is unintentional, this standard itself expresses the below fully-successful standard of work it attempts to describe.

Minimally Satisfactory

Does not apply established procedures or policy to situations, or frequently requires assistance to resolve problems.

Fully Successful

Deals with problems in local procedures or policy as they arise, and/or occasionally requires assistance to resolve them.

Highly Successful

Anticipates problems in local procedures or policy, regardless of functional area, and shows initiatives to resolve them.

It is unclear why the minimally satisfactory standard involves “established procedures or policy” while the fully successful one is limited to “local procedures or policy.” Moreover, the minimally successful standard fails the backwards test by describing what the employee should not do rather than what he must do to meet the retention standard.

Knowledge 4

Under the Technical Capability element, the Knowledge 4 function includes the following benchmarks:

Minimally Satisfactory

Recurring reports contain significant mistakes, or incomplete analysis.

Fully Successful

Consistently prepares accurate, informative recurring reports.

Highly Successful

Prepares special studies or reports that clearly explain and analyze current situation, anticipates [*sic*] future needs and developments, and makes [*sic*] appropriate recommendations.

Highly Successful and Fully Successful standards do not measure the same things. The Fully Successful standard measures recurring over time while the Highly Successful one measures the quality of individual special studies and reports. The Highly successful standard is dependant on factors beyond the control of the employee if special studies or reports are not assigned. Finally, the minimally successful standard solely describes the negative qualities of the work rather than the positive attributes that make it satisfactory for job retention.

General 1

Under the Execution and Results element, the General 1 function includes the following benchmarks:

Minimally Satisfactory

Completes accessioning of fewer than 10 transfers per year.

Fully Successful

Completes accessioning of between 10 and 15 transfers per year.

Highly Successful

Completes accessioning of more than 15 transfers per year.

Because these standards take into account only quantity, the fully successful or higher standard may not be achievable for reasons beyond the control of the employee. For example, an employee could spend the same amount of time on three large and complex projects that it would take to accomplish 16 routine, straightforward transfers. The minimally successful standard also leaves no room for unsuccessful performance. An employee who accomplishes 0 accessions would meet the retention standard because 0 is fewer than 10.

Accessioning 1

Under the Execution and Results element, the Accessioning 1 function includes the following benchmarks:

Minimally Satisfactory

Displays a general understanding of the requirements for accessioning records and maintaining archival control over them, but minor misunderstandings occasionally occur.

Fully Successful

Demonstrates good understanding and application of the principles of archival accessioning.

Highly Successful

Demonstrates an excellent understanding and application of the principles of archival accessioning, and shows good initiative in independently suggesting processes for improvements.

This standards use vague measurements, are subject to various interpretations and subjective evaluation, and are elusive goals which the agency may find the employee met or failed to meet at its pleasure. To be meaningful to employees, the standards should describe the characteristics of general vs. good vs. excellent understanding.

Professional Relationship 4

Under the Customer Service element, the Relationship 4 function includes the following benchmarks:

Minimally Satisfactory

Usually meets the Fully Successful standards with some exceptions. The exceptions are infrequent and have only short term or minimal effect on the overall performance of the individual or unit.

Fully Successful

Encourages a positive attitude among co-workers and assists in creating a positive work atmosphere.

Highly Successful

Actively encourages a positive attitude among co-workers and shows initiative in creating a positive work atmosphere.

It is not a measurable standard to differentiate fully from highly successful performance on whether an employee “encourages a positive attitude among co-workers” or whether an employee “*actively* encourages a positive attitude among co-workers.” Encouragement by its nature can scarcely be a passive activity. The standards use vague measurements, are subject to various interpretations and subjective evaluation, and are elusive goals which the agency may find the employee met or failed to meet at its pleasure.

Professional Relationship 7

Under the Customer Service element, the Professional Relationship 7 function includes the following benchmarks:

Minimally Satisfactory

May not use all available resources to find solutions in published material or fellow staff.

Fully Successful

Seeks assistance from supervisors and specialists when necessary.

Highly Successful

Requires little direct supervision in accomplishing tasks.

The minimally successful standard is written backwards. It expresses the standard in terms of what the employee does not do

Knowledge 18

Under the Technical Capability element, the Knowledge 18 function includes the following benchmarks:

Minimally Satisfactory

Accessioning paperwork is completed accurately with some exceptions.

Fully Successful

Accessioning paperwork is usually completed accurately.

Highly Successful

accurately with additional documentation provided as appropriate [*sic*].

The Highly Successful standard for this function is an unintelligible sentence fragment.

5) Conclusion

The standards in the R plans examined here read as if they were created by officials who knew they needed to word levels differently, but did not approach the task with the requisite care or knowledge of the underpinnings of the federal performance management system. In the area of best practices, OPM has developed a significant body of guidance on plan development of which NARA might avail itself. While NARA's plans are too often grounded in measurement of behaviors or competencies, OPM encourages the development of elements and standards that instead measure employee and work unit achievements.⁵ In contrast to the recommended approach, NARA standards too often describe activities rather than accomplishments.⁶ Per the OPM report, if an element does not lend itself to being measured with numbers and can only be described, the official should ask: Who could judge that the element was done well? What factors would that person look for? The report makes clear that "distinguishing between Fully Successful and levels above or below Fully Successful requires careful planning and forethought."⁷ Evidence of such planning and forethought are not to be found in NARA's new R plans.

AFGE recommends that NARA officials review NARA's performance plans Agency-wide with the following questions in mind:

- Are the standards consistent with NARA 430 and the requirements of the collective bargaining agreement?
- Are the critical elements truly critical? Does failure on the critical element mean that the employee's overall performance is unacceptable? If not, what non-critical elements can be included on the plans
- Is the range of acceptable performance clear? Are the performance expectations quantifiable, observable, and/or verifiable?
- Are the standards applicable? Can the rating official actually use the standards to appraise performance?
- Will work units and employees understand what is required of them at each level? Can rating officials explain the standards to employees?
- Are retention standards correctly formulated?

⁵ See *Handbook for Measuring Employee Performance*, p. 2.

⁶ See *Ibid*, pp. 15-16.

⁷ See *Ibid*. pp. 49, 57.

For questions on this report, please contact:

Sharmila Bhatia
Executive Vice President
AFGE Council 260

Ashby Crowder
Executive Vice President
AFGE Local 2578