Defense Civilian Intelligence Personnel System (DCIPS): Evaluation of Army Employee Performance Plans

Technical Report #769

Prepared by: Kara R. Jeansonne, Kevin G. Smith, and Meredith L. Ferro

January 29, 2013

Table of Contents

List of Tables

Introduction

Given the significance of the information provided in the performance plans and appraisals for civilian employees covered under the Defense Civilian Intelligence Personnel System (DCIPS), the Undersecretary of Defense for Intelligence (USD(I)) and the Office of the Director of National Intelligence (ODNI) have jointly pursued rigorous evaluations of DCIPS implementation across the Defense Intelligence Enterprise. The purpose of this evaluation is to determine the extent to which performance plans and appraisals from the Army were prepared according to relevant performance management guidance and aligned with training.

For employees covered under DCIPS, performance plans and appraisals are used to: (a) document the key outcomes/results an employee is expected to achieve during the period of performance, (b) provide ratings on the extent to which the employee accomplished these outcomes and on the manner in which work is performed (as defined by common performance elements), and (c) provide narratives that justify these ratings.

Employees and supervisors mutually agree on an employee's performance objectives at the start of the period of performance. Then, the objectives get translated into meaningful projects or day-to-day activities. During the period of performance, at least one mid-term review of employee performance relative to the performance objectives and performance elements is conducted. At the end of the period of performance, employees complete a narrative related to the work that was conducted over the year. Then, the rater completes a final assessment during which he or she reviews that employee's self-assessment, evaluates the employee's performance relative to the objectives and elements, provides a numeric rating for the objectives and elements, and completes a narrative supporting the rating of record. Then, a higher level review is conducted. Lastly, the ratings provided in the performance appraisal are used in the pay pool panel process to determine final ratings and performance-based bonus payout.[1]

The content included in the performance plans and appraisals as well as the processes for their development and use were designed to support DCIPS' stated objectives:

1. Ensure alignment between individual performance objectives and the higher-level mission and objectives of the Intelligence Community (IC).
2. Ensure ongoing feedback between employees and supervisors regarding progress toward objectives and relative to standard behavioral elements.
3. Provide a basis for measuring and assigning accountability for individual and organizational performance for the accomplishment of these objectives.
4. Provide a fair and equitable process for appraising and evaluating employee performance within and across IC elements.
5. Maintain adherence to merit system principles[2].

---

[1] Except for NGA where the ratings still inform performance-based raises.
[2] Department of Defense. (2010). DoD Civilian Personnel Management System: Defense Civilian Intelligence Personnel System (DCIPS) Performance Management (1400.25-V2011).

The overall purpose of the evaluation study was to determine the extent to which Fiscal Year 2011 (FY11) Army performance plans and appraisals were prepared according to established performance management guidance and aligned with training. The research questions that guided each evaluation were:

- ➤ To what extent do the objectives adhere to the "SMART+" framework (i.e., specific, measurable, achievable, relevant, time-specific, can be exceeded)?

- ➤ To what degree do the objectives align with the intelligence mission, goals, and objectives of the Army?

- ➤ To what extent are the objectives consistent by occupation and level (i.e., are employees in similar jobs and at similar levels held to the same standards)?

- ➤ To what extent do objectives represent long-term outcomes versus recurring activities related to day-to-day work (i.e., recurring vs. non-recurring objectives)?

- ➤ To what extent is the information provided in the self-assessment narratives adequate to support the performance ratings?

- ➤ To what extent is the information provided in the raters' appraisal narratives adequate to support the performance ratings?

The remaining sections of this report detail the evaluation methodology, the results and conclusions of this study, and recommendations for evolving performance management at Army.

Evaluation Methodology

This section describes the process by which the samples of performance plans and accompanying appraisals were drawn and the procedures followed to evaluate them.

Sampling

PDRI, USD(I), and Army agreed upon a desired sample size of 300 plans and appraisals for the current study. These determinations were made considering a variety of factors including the size of the organization, the cost associated with reviewing the plans and appraisals, comparability to other evaluation studies, and the extent to which generalizable conclusions could be inferred based on the number of plans reviewed. For the current study, the sampling strategy was to achieve representativeness such that the conclusions drawn would reflect the agency as whole, rather than specific subpopulations (e.g., only Analysts). PDRI requested that Army draw a random stratified sample of plans and appraisals using several important background variables (e.g., occupation, gender) to meet the aforementioned goals of the study.

PDRI raters reviewed 300 plans and appraisals covering the FY11 administration cycle. The characteristics of the sample evaluated are presented in Table 1.

Table 1: Performance Plan Sample Characteristics		
	n	*Percent of Sample*
Work Category		
Technician/Administrative Support	25	8.3%
Professional	228	76.0%
Supervision/Management	47	15.7%
Work Level		
Entry	13	4.3%
Full Performance	169	56.3%
Senior	112	37.3%
Expert	6	2.0%
Pay Band		
1	13	4.3%
2	18	6.0%
3	154	51.3%
4	110	36.7%
5	5	1.7%

Table 1: Performance Plan Sample Characteristics		
	n	*Percent of Sample*
Occupational Series		
80 – Security Administration Series	57	19.0%
83 – Police Series	1	0.3%
85 – Security Guard Series	1	0.3%
86 – Security Clerical and Assistant	15	5.0%
101 – Social Sciences Series	7	2.3%
132 – Intelligence Series	110	36.7%
134 – Intelligence Aid and Clerk Series	2	0.7%
201 – Human Resources Series	2	0.7%
301 – Miscellaneous Administration and Program Series	6	2.0%
303 – Miscellaneous Clerk and Assistant Series	1	0.3%
318 – Secretary Series	2	0.7%
340 – Program Management Series	1	0.3%
341 – Administrative Officer Series	1	0.3%
343 – Management and Program Analysis Series	12	4.0%
346 – Logistics Management Series	1	0.3%
391 – Telecommunication Series	1	0.3%
501 – Financial Administration Series	1	0.3%
544 – Civilian Pay Series	1	0.3%
560 – Budget Analysis Series	3	1.0%
802 – Engineering Technical Series	1	0.3%
810 – Civil Engineering Series	1	0.3%
855 – Electronics Engineering Series	6	2.0%
905 – General Attorney Series	1	0.3%
1040 – Language Specialist Series	2	0.7%
1071 – Audiovisual Production Series	1	0.3%
1084 – Visual Information Series	1	0.3%
1102 – Contracting Series	3	1.0%
1301 – General Physical Science Series	2	0.7%
1310 – Physics Series	1	0.3%
1320 – Chemistry Series	2	0.7%
1410 – Librarian Series	1	0.3%
1411 – Librarian Technician Series	1	0.3%
1412 – Technical Information Services Series	2	0.7%
1603 – Equipment, Facilities, and Services Assistance Series	1	0.3%
1670 – Equipment Services Series	1	0.3%
1701 – General Education and Training Series	1	0.3%
1712 – Training Instruction Series	20	6.7%
1750 – Instructional Systems Series	2	0.7%
2001 – General Supply Series	2	0.7%

Table 1: Performance Plan Sample Characteristics		
	n	*Percent of Sample*
2003 – Supply Program Management Series	1	0.3%
2005 – Supply Clerical and Technician Series	1	0.3%
2010 – Inventory Management Series	1	0.3%
2210 – Information Technology Series	20	6.7%
Sex		
Male	195	65.0%
Female	105	35.0%
Race/Ethnicity		
American Indian or Alaskan Native	0	0.0%
Asian	6	2.0%
Black	47	15.7%
Hispanic/Latino	6	2.0%
Native American or Other Pacific Islander	1	0.3%
Two or More Races	11	3.7%
White	229	76.3%

Evaluation Procedure

The first step in the evaluation process was to finalize the set of evaluation criteria against which the effectiveness and adequacy of the performance plans and appraisals would be assessed. These criteria were developed to answer the pertinent research questions presented earlier in the report and were leveraged in previous evaluation studies of similar systems (i.e., National Security Personnel System – NSPS). The current criteria are very similar to the criteria used during reviews at Defense Intelligence Agency (DIA), Naval Intelligence, National Geospatial Intelligence Agency (NGA), National Security Agency (NSA), USD(I), and the National Intelligence Civilian Compensation Program (NICCP) implementation at ODNI. The final criteria are presented in Appendix A of this report.

The study team was composed of six Industrial/Organizational Psychologists who, on average, had approximately four years of experience. After the evaluation criteria were finalized, the study team members were trained on assessing the plans using the established criteria to ensure there was consistency in team members' rating approach. Training began with a review of background information on Army performance management and discussion of the rating criteria, research questions, and evaluation methodology. Then, each assessor independently evaluated the same subset of performance plans using the evaluation criteria. The assessors then met as a group, compared ratings, discussed any discrepancies, and came to agreement on a final rating for each performance plan.

The sample of performance plans was then divided by occupational series (e.g., 0132 – Intelligence Series) and each assessor evaluated between 20 and 80 plans. Most assessors were responsible for evaluating multiple occupations; two assessors evaluated only one occupation each because of the large amount of plans associated with those occupations.

One member of the assessor team served as the lead assessor. The lead assessor monitored project completion, and ensured that the assessors remained calibrated throughout the task and did not diverge in their interpretations of the evaluation criteria. If any rating discrepancies arose, the lead assessor brought these discrepancies to the attention of the assessor team. Throughout the rating period, assessors discussed issues that arose and came to agreement on how to apply the evaluation criteria in subsequent cases when the same issue was encountered. The lead assessor communicated these developments to any team member who was not present at the time to ensure common understanding. If necessary, the assessors re-evaluated plans that involved the same issues. These steps were necessary to ensure that the assessor team applied the criteria consistently throughout the evaluation process.

Results

The sample of 300 FY11 performance plans yielded a total of 1083 objectives, with a mean of 3.6 objectives per performance plan. This section presents the results of the plan evaluations in accordance with the primary research questions.

SMART Objectives

Each objective was evaluated on the extent to which it met each element of the SMART framework *(i.e., specific, measurable, achievable, relevant, time-bound)*. These elements are further explained below.

- ➢ **Specific:** The job objective describes, clearly and concisely, what the employee is expected to accomplish in the coming year.

- ➢ **Measurable:** The result of the employee's job objective is observable or verifiable through appropriate quantity, quality, resources, or time measurements.

- ➢ **Achievable:** The job objective can be realistically accomplished over the rating period considering the employee's work category and band level.

- ➢ **Relevant:** The job objective represents meaningful outcomes for the employee's a) job and b) organization/mission. Job and organization/mission relevance were included as separate variables.

- ➢ **Time-bound:** The job objective identifies time-frames for completion.

The results of the evaluation of the objectives against the SMART framework are presented in Tables 2 through 7.

Table 2: Specific Element of SMART Framework[3]		
Frequency of Rating n (%)		
Not Specific	*Moderately Specific*	*Highly Specific*
32 (3.0%)	410 (37.9%)	641 (59.2%)

The results in Table 2 indicate that:

➢ In general, the objectives reviewed were moderately or highly specific. Highly specific objectives included very specific observable actions, behaviors, or achievements that clearly identified key results and outcomes while moderately specific objectives were somewhat vague or broad and written in generic terms. When objectives were rated as moderately specific, they included enough information to understand approximately what needed to be accomplished but did not fully address issues such as how many, how often, and other detailed standards to which subsequent actions, behaviors, and outcomes would be compared. For example, "Create an electronic Standard Operating Procedure (SOP) for all major tasks in the Customer Service Branch." Performance objectives that were not specific were generally very short in length and lacked descriptiveness or referenced generic job duties. For example, "Throughout the rating period, perform identified extra duties as assigned to facilitate adherence to requirements..."

Table 3: Measurable Element of SMART Framework		
Frequency of Rating n (%)		
Not Measurable	*Moderately Measurable*	*Highly Measurable*
23 (2.1%)	567 (52.4%)	493 (45.5%)

The results in Table 3 indicate that:

➢ In general, the objectives reviewed tended to be moderately measurable or highly measurable. Highly measurable objectives included an exact and precise method for assessing or verifying the quality, quantity, or timeliness of the objective. For example, "Manage course documentation and perform test control functions in accordance with TRADOC regulations and 111th MI Brigade Organization Inspection Program (OIP) standards; receive a "GO" on at least 80% of the 111th MI Brigade OIP Checklist."

➢ A very small proportion of objectives were not measurable in the performance plans. "Not measurable" objectives did not include a method or procedure for assessing or verifying the quality of the objective. These objectives typically included generic tasks or duties found in position descriptions or job announcements.

---

[3] For the full description of each scale point on the criteria, please refer to Appendix A.

➢ The extent to which performance objectives included percentages as a measurement component was examined. Early versions of DCIPS training emphasized quantitative measures which were sometimes adopted despite there being no reasonable method for tracking them. Forty-one percent (41.0%) of objectives included percentages. Of those, nearly a third (29.7%) appeared to be arbitrarily defined and unrealistic to track (e.g., meet proposed study goals to 90-95% accuracy) which could be improved upon in the future.

Table 4: Achievable Element of SMART Framework			
Frequency of Rating n (%)			
Not Available	*Too Low*	*Appropriate*	*Too High*
7 (0.6%)	24 (2.2%)	1047 (96.7%)	5 (0.5%)

The results in Table 4 indicate that:

➢ The objectives appeared to be achievable considering the employee's work category and band level.

Table 5: Job Relevance Element of SMART Framework			
Frequency of Rating n (%)			
Not Available	*Not Relevant Outcome to Job*	*Moderately Relevant Outcome to Job*	*Very Meaningful Outcome to Job*
0 (0.0%)	1 (0.1%)	96 (8.9%)	986 (91.0%)

The results in Table 5 indicate that:

➢ Objectives typically represented very meaningful outcomes given the individual's occupation. When objectives represented moderately meaningful outcomes it was usually because the objective related to training or performing very generic duties without much explanation as to how the training actually applied to an individual's position. The moderately relevant outcomes generally covered other duties as assigned and training requirements that were not clearly described as job-relevant. With respect to training and development, 23.3% of the plans contained objectives with training or developmental components. Training or professional development was usually described in a generic way or was sometimes limited to mandatory requirements. There were few, if any, instances of other experiential learning activities that are effective for development such as self-development or mentoring cited in these objectives.

Table 6: Organizational Relevance Element of SMART Framework			
Frequency of Rating n (%)			
Not Available	*Not Aligned*	*Aligned-Goal Referenced*	*Specific Goal Stated*
67 (6.2%)	0 (0.0%)	745 (68.8%)	271 (25.0%)

The results in Table 6 indicate that:

➢ Most performance plans listed relevant organizational/mission goals before presenting individual performance objectives. When those specific organizational/mission goals were referenced or described in the performance objectives, the team evaluated whether a clear linkage could be made from the objective to the organizational goal. If so, this was rated as "Specific Goal Stated." If the objective was not aligned to the specific goal it was rated as "Not Aligned." One-quarter of the objectives directly linked to specific organizational/mission goals. None of the objectives reviewed were not aligned with identified higher level objectives. When there was no mention of a specific organizational/mission goal within a performance objective, the study team judged whether the content of the objective related to any of the organizational/mission goals stated at the front of the plan (if that information was available); if a relationship could be inferred, the objective was rated as "Aligned-Goal Referenced." This was the most common rating (68.8%) made for this criterion. When no higher level goals were stated anywhere in the plan, and no reference to any higher level goals was made in the objectives, a rating of "Not Available" was assigned; about 6.2% of the objectives fell into this category.

Table 7: Time-bound Element of SMART Framework	
Frequency of Rating n (%)	
Did Not Include Time Element	*Included Time Element*
377 (34.8%)	706 (65.2%)

The results in Table 7 indicate that:

➢ A majority of the objectives (65.2%) included time elements. The time-bound criterion can be met with a specific statement (e.g., June 1, 2011) or a general statement (e.g., "on a monthly basis"). Objectives rated as not having a timeframe did not include any mention of a time period whether general or specific. Some objectives did state that outcomes would be achieved in a "timely manner," but that was deemed too subjective to meet the general time-bound standard.

Adhering to the SMART framework is necessary, but not sufficient, for effective objectives. Objectives need to fulfill some additional requirements (i.e., SMART+) to provide maximum

utility. For example, objectives need to be written at the "successful" level of performance to enable employees to exceed the defined measurement properties within the objective.

> ➢ A relatively small number (12.8%) of objectives across the sample were written such that the measurement properties identified within the objective precluded the individual from exceeding the objective. In other words, these objectives were written such that they resulted in a pass/fail scenario. Examples of objectives written as pass/fail include statements that products must be accurate 100% of the time as a measure of success or those that were written in such general terms that it would be nearly impossible to justify how they might be exceeded (e.g., "Submit all TDY travel vouchers in accordance with applicable policies and regulations").

Similarity of Objectives

Plans were organized by occupation or function and assigned to raters such that a rater reviewed all or most of the plans from a given occupation or function. Due to the small sample sizes across many of the occupational groups a quantitative comparison of within group similarity was not feasible. However, the study team did a qualitative review of the degree to which objectives within each occupation or function were similar. The results of this review indicated that objectives tended to be highly similar within occupations, with some objectives for several individuals within a given job being very similar or identical and related to their specialty area.

Progression of Objective Difficulty

To evaluate whether objectives progressed in difficulty from lower bands to higher bands, the data collected regarding the achievable element of the SMART framework was analyzed across work levels. The achievable portion of the criteria captures the extent to which the objective aligned with the work category and work level of the employee for which the objective was written. The work level descriptions present work in a hierarchy of increasing levels of difficulty and complexity as one moves into higher levels; thus if objectives were rated "Appropriate" across the different levels they progressed in difficulty across band level. The results showed that an overwhelming number of performance objectives (96.7%) were written at the appropriate level. There was no observable pattern related to the work level.

Leadership Objectives

Supervisors/managers are expected to focus their efforts on the people they manage and supervise rather than producing/completing technical work on their own. At the request of a key stakeholder, the study team examined the extent to which the objectives prepared for managers actually addressed leading/supervising. The team examined the number of supervisory plans in which all of the objectives addressed supervisor/leadership outcomes. Slightly more than a third (34.0%) of the supervisory plans met this criterion. Thus, approximately two-thirds of the supervisory/managerial performance plans still contained some objectives that addressed individual contributions.

Recurring versus Non-Recurring Objectives

As described above, performance objectives are an expectation-setting tool focused on planning an employee's work over the course of a yearly performance cycle. IC guidance requires that objectives be based on critical outcomes and not day-to-day tasks. However, some individuals find this guidance challenging because their work consists primarily of performing the same set of tasks on a recurring basis. Examples include individuals who work IT Help Desks or other largely reactive positions. For other employees, work is more project-based in nature (e.g., finding and reporting on requested information). Therefore, the study team examined the proportion of objectives that involved recurring versus non-recurring activities. Recurring objectives are based on activities that occur on an ongoing basis and are often repeated multiple times. Non-recurring objectives are focused on activities that generally are not repeated and which have a defined beginning and end.

The proportion of recurring versus non-recurring objectives was of interest to the study team because it can be difficult to write recurring objectives that adhere to the SMART+ framework. In particular, specificity and measurement are often challenging because it may be difficult to predict anticipated quantity (though quality measures are still relevant). Moreover, as the notion of recurring objectives is a bit at odds with guidance that states that objectives should not include "everyday" activities, the study team examined the number of objectives that represented recurring tasks. Table 8 presents the results of this analysis.

Table 8: Recurring versus Non-Recurring Objectives	
Frequency of Rating n (%)	
Non-Recurring	***Recurring***
98 (9.0%)	985 (91.0%)

The results in Table 8 indicate that:

➢ Objectives were largely written to capture recurring activities. An example of a recurring objective is, "COR responsibilities for the mail and courier service contract. Ensures security, financial, administrative reporting requirements are completed within 5 days of tasking. Ensures official file documentation is maintained. Is a Trusted Agent (TA) to ensure timely receipt of Command Access Cards (CAC) for Contractors within the Mail and Distribution Center." An example of a non-recurring objective is, "Complete three Overflight mission evaluations, evaluate three Open Skies Missions, and determine if they impact APG." It should be noted that the rating team had some difficulty making these ratings within the current sample. Many objectives included multiple sub-objectives, behaviors, or strategies which had both non-recurring and recurring features. When this was the case, the raters considered the entire objective and made a rating that best represented the objective as a whole. This finding is noteworthy because it suggests that the majority of work is not amenable to setting specific, long-term objectives that span the entire period of performance. For example, it may be far more practical to require an intelligence analyst to produce reports that adhere to professional

standards and are completed in a standard timeframe in response to customer demands rather than to specify a required number or type of reports at the beginning of the performance period. The results show that many DCIPS employees within Army are already writing objectives in this manner, which points to a potential disconnect between the current guidance on SMART objectives and the reality of the work.

Narrative Justifications

At the end of the period of performance, employees are required to provide a detailed narrative that highlights the key accomplishments and resulting organizational impact relative to each performance objective. Additionally, employees describe their performance relative to the standardized performance elements. Then, the rating official completes an assessment during which he or she reviews the employee's self-assessment and evaluates the employee's performance relative to the objectives and elements, and provides a numeric rating for each objective and element. Those ratings become the rating of record.

Since this information serves as the foundation for important decisions related to compensation, it was important to assess the adequacy of the information provided. Employees and rating officials had training resources available to them in advance of the rating cycle to ensure that there was a common awareness regarding the types of information that should be included when documenting performance. Thus, rating criteria were drafted to measure the extent to which the narratives addressed the training content and policy guidance. In this study, the team examined both the self-assessment narratives and the raters' narratives.

Performance Objective Narrative Justifications

The results of the evaluation as they pertain to the performance objective self-assessment narratives are presented in Table 9. The results of the evaluation relative to the performance objective rater assessment narratives are presented in Table 10.

Table 9: Adequacy of Information Provided in Objective Self-Assessments			
Self-Assessment Specificity for Objectives Frequency of Rating n (%)			
----	*Not Available*	*Vague*	*Specific*
----	40 (3.7%)	89 (8.2%)	954 (88.1%)

Table 9: Adequacy of Information Provided in Objective Self-Assessments			
Self-Assessment and Measurement Relationship Frequency of Rating n (%)			
Not Available	*Relationship is Not Apparent*	*Apparent but Not Fully Substantiated*	*Apparent and Fully Substantiated*
39 (3.6%)	8 (0.7%)	326 (30.1%)	710 (65.6%)

The results in Table 9 indicate that:

➢ In most cases, the self-assessments supporting performance objective ratings were specific in describing accomplishments and outcomes. It was rare for self-assessment narratives to address development areas or deficiencies in performance. Only 8.2% of the performance objective narratives included vague or general support for the rating that was made. When narratives were vague, they were generally very short in length, used jargon, or contained language like, "successfully met this objective," with little elaboration on the accomplishments.

➢ The relationship between the self-assessment narratives and the associated performance objectives was obvious for nearly all narratives that were reviewed. For the most part, the measurement properties for the objectives were fully represented in the self-assessment (65.6%). Thirty percent of the self-assessments did not fully substantiate the relationship; in these cases, the narratives supported the rating by highlighting accomplishments and outcomes, but did not address every aspect of the measurement component described in the original objectives.

Table 10: Adequacy of Information Provided in Objective Rating Official Assessments			
Rater Specificity for Objectives Frequency of Rating n (%)			
----	*Not Available*	*Vague*	*Specific*
----	54 (5.0%)	231 (21.3%)	798 (73.7%)

Table 10: Adequacy of Information Provided in Objective Rating Official Assessments			
Rater Narrative and Measurement Relationship Frequency of Rating n (%)			
Not Available	*Relationship is Not Apparent*	*Apparent but Not Fully Substantiated*	*Apparent and Fully Substantiated*
54 (5.0%)	14 (1.3%)	518 (47.8%)	497 (45.9%)
Rater Narrative Support for the Performance Objective Rating Frequency of Rating n (%)			
Not Available	*Not Supported*	*Partially Supported*	*Fully Supported*
51 (4.7%)	25 (2.3%)	451 (41.6%)	556 (51.3%)

The results in Table 10 indicate that:

➢ The majority of rater narratives were specific in describing accomplishments, outcomes or deficiencies; when narratives were vague, they were generally very short in length, used jargon, or contained language like, "successfully met this objective," with little elaboration on the accomplishments. At times, raters would simply write, "concur with employee assessment," which did not provide any unique information to consider and was thus was rated "Not Available."

➢ The relationship between the rater narrative and the associated performance objectives was obvious; however, compared to the self-assessments, fewer supervisory narratives fully addressed the measurement properties present within the performance objective, while about another half did address every important measurement property noted in the performance objective.

➢ Half of the narratives contained ratings that were fully supported. When ratings were partially supported or unsupported, there was often one main cause: the justification indicated a rating of "3" or "Successful" was appropriate, but the rating associated with that objective documented a higher level of performance (i.e., Excellent or Outstanding). However, there were still some instances when a rating of "Successful" was provided but higher levels of performance (e.g., Excellent) were described within the narrative statements.

Performance Element Narrative Justifications

The results of the evaluation as they pertain to the performance element self-assessment narratives are presented in Table 11. The results of the evaluation as they pertain to the performance element rater narratives are presented in Table 12.

17

Table 11: Adequacy of Information Provided in Performance Element Self-Assessments			
Frequency of Rating n (%)			
Not Available	*Vague*	*Specific but Unrelated to Elements*	*Specific and Related to Elements*
6 (2.0%)	74 (24.7%)	28 (9.3%)	192 (64.0%)

The results in Table 11 indicate that:

➢ For the most part, self-assessment narratives were specific and related to the performance elements. About one-quarter of the narratives were vague, and a small percent, although specific, were unrelated to the performance elements.

Table 12: Adequacy of Information Provided in Performance Element Rating Official Assessments			
Rater Specificity Frequency of Rating n (%)			
Not Available	*Vague*	*Specific but Unrelated to Elements*	*Specific and Related to Elements*
5 (1.7%)	116 (38.7%)	28 (9.3%)	151 (50.3%)
Rating Support Frequency of Rating n (%)			
Not Available	*Not Supported*	*Partially Supported*	*Fully Supported*
5 (1.7%)	8 (2.7%)	148 (49.3%)	139 (46.3%)

The results in Table 12 indicate that:

➢ Half of the performance element narratives made by rating officials were both specific and related to the performance elements. However, nearly 40% were vague. A small percent, although specific, were unrelated to the performance elements.

➢ Almost half of the narratives contained ratings that were fully supported. When ratings were partially supported or unsupported, there was often one main cause: the justification indicated a rating of "3" or "Successful" was appropriate, but the rating associated with that objective documented a higher level of performance (i.e., Excellent or Outstanding).This finding is similar to that of the rating officials' narrative justifications for the performance objectives. However, there were still some instances when a rating

of "Successful" was provided but higher levels of performance (e.g., Excellent) were described within the narrative statements.

Qualitative Themes

The study team noticed that there were many objectives that covered emerging requirements or other duties as assigned. While there were attempts to adhere to the SMART criteria, these objectives were still very vague and measures were sometimes arbitrary (e.g., focusing only on suspense dates) because nothing more specific could be developed at the time. Supervisors/Managers should either use the flexibility of DCIPS timelines to develop new performance objectives later in the rating period and make those meaningful or actively manage performance on these ancillary requirements as they arise. It was never the intent of DCIPS for performance objectives to cover an entire job. Rather, objectives should reflect the position's most critical outcomes or functions. In most cases, other duties as assigned should not meet that threshold. In addition to this theme, there were several objectives written with standards that included the phrase "at least" before defining a specific measure. This is problematic because this type of standard creates pass / fail scenarios. Standards should be set at the "Successful" level of performance and language such as "at least" or "greater than" should be considered carefully before inclusion.

Conclusions

This section presents conclusions based on reviewing 300 performance plans and accompanying appraisals for the FY11 performance cycle at Army. Short-term recommendations for improving adherence to the current set of DCIPS policies and guidance are also provided.

Research Questions

➢ *To what extent do the objectives adhere to the "SMART+" framework (i.e., specific, measurable, achievable, relevant, time-bound, can be exceeded)?*

Army's objectives largely adhere to the important elements of the SMART+ criteria. There is still room for improvement with respect to setting highly specific and highly measurable objectives.

➢ *To what extent are the objectives consistent by occupation and level (i.e., are employees in similar jobs and at similar levels held to the same standards)?*

Objectives tended to be highly similar within occupations, with some objectives for several individuals within a given job being very similar or identical and related to their specialty area.

➢ *To what extent do objectives represent long-term outcomes versus recurring activities related to day-to-day work (i.e., recurring vs. non-recurring objectives)?*

The vast majority of Army's performance objectives reflect repeatable activities where it is impossible to predict in advance a specific volume or quantity of production across the entire annual period of performance. Thus, the guidance that states that performance objectives should reflect highly specific but long-term outcomes may not accurately reflect how work is accomplished in the IC.

➢ *To what extent is the information provided in the self-assessment narratives adequate to support the performance rating?*

The information provided in the self-assessment narratives tended to be specific and related to measurement properties in the original performance objectives. Although there was often more than one measure included within a performance objective, for the most part, the associated self-assessment narratives addressed every one of those measures. In addition, the information provided in the performance elements justifications was specific.

➢ *To what extent is the information provided in the rating official narratives adequate to support the performance rating?*

The information provided in the rating official narratives tended to be specific and related to measurement properties in the original performance objectives. However, almost half

20

of the associated rating official narratives did not address every measure. Supervisors and managers justified their rating on about half of the performance objectives in a clear manner such that text descriptions aligned with performance rating expectations. There was, however, a sizable portion of narratives that did not fully align with performance expectations.

Related to the performance elements, about half of the supervisors'/managers' narratives were specific in describing performance. Similar to the findings for the performance objectives, almost half of the performance element narratives aligned with performance rating expectations, while the other half were provided partial or insufficient support.

Recommendations

In light of this study's findings, it may be beneficial for Army to evaluate their performance objective and accomplishment writing training to ensure that several topics are adequately covered. Alternatively, a tip sheet could be developed for managers/employees to use that highlights some areas to focus on improving during the development of objectives in the next evaluation period. Specific topics to address would be: 1) the appropriate use of percentages as a measurement component, 2) how a time element can be easily added to an objective and why it is important, 3) avoiding pass/fail language (e.g., "at all times," "at least 80%," "100%"), 4) focusing the objectives for supervisors/managers on supervisory and leadership activities, 5) avoiding the use of "other duties as assigned" language and instead using the flexibility of the DCIPS system to make adjustments to objectives as needed, 6) ensuring that both employee and supervisor accomplishment narratives address each of the measurement properties in an objective. Addressing these topics will further enhance the quality of the objectives written and assessed by Army employees.

Limitations

It is important to acknowledge some of the limitations of the evaluation study summarized in this report. Although the study team followed a rigorous analytical methodology and took care to rate plans consistently, the methodology itself does limit the ability to assess how well important aspects of the performance management system are implemented. For example, this methodology cannot directly assess how much ongoing coaching and feedback occurred over the course of the performance period or whether developmental goals are being assigned. While the plans do contain information that is critically important for a number of reasons, they are still only artifacts of the supervisory relationship and not a direct assessment of the relationship itself. It is impossible to know whether the information in the plans is accurate or whether the documentation matched the employee's observable performance. We are only able to examine the extent to which the information in the plans met the requirements set forth in performance management guidelines and elaborated on in formal training, frequently asked questions, and tips available through the DCIPS website.

Despite these limitations, it is also important to note that performance management system evaluations are best conducted using a variety of data sources that include independent measurements to address the methodological limitations associated with single method studies. This report should serve as an excellent source of information to be considered in conjunction with other sources such as the NAPA study (Dorn, Blair, Disney, Faga, Hawley, Hazlewood, Lachance, & Massiah, 2010) and any other more recent evaluations Army may have completed internally.

References

Dorn E., Blair, D.G., Disney D.M., Faga, M.C., Hawley K., Hazlewood L., & Lachance J., Massiah M.G. (2010). The Defense Civilian Intelligence Personnel System: An Independent Assessment of Design, Implementation, and Impact. *A report by a Panel of the National Academy of Public Administration,* Washington, DC.

Appendix A: Evaluation Criteria

1. **Performance plan ID**

SECTION 1 – PERFORMANCE OBJECTIVES

The following questions address the objectives. After completing the Rating Official Evaluation for an objective, return to this section until all objectives have been rated.

2. **Objective Rating**

3. **Is this objective a recurring or non-recurring objective?**

A. **Recurring.** Objectives that are written for activities that are part of an employee's routine tasking which are unforeseen in terms of volume and timing.
B. **Non-recurring.** Objectives that are written for activities and projects where it is possible to specify in advance the timing, volume, type, and/or quantity, associated with the completion of planned work over the rating period.

SMART OBJECTIVES

4. **To what extent does the job objective conform to the specific aspect of the "SMART" framework?**

A. **Not Specific.** Did not include specific and observable actions, behaviors, or achievements; objectives were written in very general terms (e.g., generic job duties or work activities).
B. **Moderately Specific.** Included moderately specific and observable actions, behaviors, or achievements; objectives were somewhat vague or broad.
C. **Very Specific.** Included very specific and observable actions, behaviors, or achievements that clearly identified key results and outcomes.

5. **To what extent does the job objective conform to the measurable aspect of the "SMART" framework?**

A. **Not Measurable.** Did not include a method or procedure for assessing or verifying the quality, quantity, or timeliness of the objective.
B. **Moderately Measurable.** Included an approximate, vague, or imprecise method or procedure for assessing or verifying the quality, quantity, or timeliness of the objective.
C. **Highly Measurable.** Included an exact and precise method or procedure for assessing or verifying the quality, quantity, or timeliness of the objective.

6. **Does the objective use percentages to assess the measurable aspect of the "SMART" framework?**

A. **Yes.** Percentages were used in the objective.
B. **No.** Percentages were not used in the objective.

7. **(If <u>Yes.</u> *above)* Is it plausible that there would be a method to track the percentage?**

A. <u>Yes.</u> It is reasonable to believe that there is a method (e.g., automated system, thorough review by a reviewers) to track or assess the aforementioned percentage.
B. <u>No.</u> It would be difficult or unreasonable (e.g., 96% accuracy across reports etc.) to have a developed a method to track or assess the aforementioned percentage.
C. <u>Not Applicable.</u> No percentage was used in the objective.

8. **Is this objective written such that ratings:**

A. <u>Not Available.</u> Not enough information provided to make a rating.
B. <u>Objective is written as pass/fail.</u> Objective is written such that the measurement component makes it impossible to exceed expectations (e.g., 100% accuracy).
C. <u>Can be met and exceeded.</u> Objective is written such that the measurement component does not preclude the ability to exceed expectations.

9. **To what extent does the job objective conform to the <u>achievable</u> aspect of the "SMART" framework?**

A. <u>Not Available.</u> Not enough information provided to make a rating.
B. <u>Too Low.</u> Objective was written at a level that is not challenging enough for the work level considering the work category; objective would be more appropriate for a lower work level within the same category.
C. <u>Appropriate.</u> Objective was written at an appropriate difficulty/challenge level for the work level considering the work category given the amount responsibility expected for that work level within the same category.
D. <u>Too High.</u> Objective was written at a level that is too high or challenging for the work level considering the work category; objective would be more appropriate for a higher level within the same category (may require more responsibility than expected for that level).

To what extent does the job objective conform to the <u>relevant</u> aspect of the "SMART" framework?

10. **Job/Occupational Relevance**

A. <u>Not Available.</u> Not enough information provided to make a rating.
B. <u>Not Relevant Outcome.</u> Objective did not reflect a meaningful outcome for the job series; objective was either irrelevant or insignificant to the tasks of the job.
C. <u>Moderately Relevant Outcome.</u> Objective reflected a moderately meaningful outcome for the job series; objective was relevant and reflected a core work area for the job series.
D. <u>Very Meaningful Outcome.</u> Objective reflected a very meaningful outcome for the job series; objective was relevant, key, and significant to the job series.

11. Organizational/Mission Relevance

A. **Not Available.** Not enough information provided to make a rating.
B. **Not Aligned.** A linkage was made but the performance objective and higher level goal were not aligned with one another.
C. **Aligned – Goal Referenced.** Did not specifically identify a linkage in the objective, but alignment between performance objective and higher level goal can be inferred.
D. **Aligned – Specific.** Identified the specific higher level goal that a performance objective supports. Alignment between performance objective and higher level goal is apparent.

12. To what extent does the job objective conform to the time specific aspect of the "SMART" framework?

A. **No time element.** Did not include timeline for accomplishing the objective.
B. **Included time element.** Included timeline for accomplishing the objective. This timeline may or may not have been specific (e.g., "within project timelines" vs. "on a monthly basis" or "by July 2008").

Continue on to the next section.

UNCLASSIFIED

EMPLOYEE SELF-REPORT OF ACCOMPLISHMENTS: Objectives

13. The self-assessment narrative for the objectives was:

A. **Not Available.** Not enough information provided to make a rating.
B. **Vague.**   Included vague examples of accomplishments, outcomes, and/or deficiencies.
C. **Specific.** Included specific examples of accomplishments, outcomes, and/or deficiencies.

14. To what extent does the self-assessment relate to the original performance objective?

A. **Not Available.** Not enough information provided to make a rating.
B. **Relationship is not apparent.** Objective and self-assessment appear to be unrelated.
C. **Relationship is apparent but not fully substantiated** Objective and self-assessment are related but the justification does not address some important aspects of objective attainment (e.g., types of measurement, specifics, etc.)
D. **Relationship is apparent and fully substantiated** Objective and self-assessment are related and the justification addresses all aspects of objective attainment (e.g., types of measurement, specifics, etc.)

Continue to the next section.

RATING OFFICIAL EVALUATION: Objectives

15. The Rating Official Narrative was:

A. **Not Available.** Not enough information provided to make a rating.
B. **Vague.** Included vague examples of accomplishments, outcomes, and/or deficiencies.
C. **Specific.** Included specific examples of accomplishments, outcomes, and/or deficiencies.

16. To what extent does the Rating Official Narrative relate to the original performance objective?

A. **Not Available.** Not enough information provided to make a rating.
B. **Relationship is not apparent.** Objectives and narrative appear to be unrelated.
C. **Relationship is apparent but not fully substantiated.** Objectives and narrative are related but the justification does not address some important aspects of objective attainment (e.g., types of measurement, specifics).
D. **Relationship is apparent and fully substantiated.** Objectives and narrative are related and the justification addresses all aspects of objective attainment (e.g., types of measurement, specifics).

17. To what extent does the Rating Official Assessment support the rating made on the performance objective:

A. **Not Available.** Not enough information provided to make a rating.
B. **Not supported.** The ratings received and associated justifications do not align with the performance expectations guidance.
C. **Partially supported.** The ratings received and associated justifications partially aligned with the performance expectations guidance across some objectives.
D. **Fully Supported.** The ratings received and associated justifications aligned with the performance expectations guidance across the majority of the objectives.

Go back to Criteria Point #2 and repeat for the next set of performance objectives, self-assessments, and supervisor assessments. If you have completed making these ratings across the set of performance objectives and justifications, continue to the next section.

<u>*OVERALL OBJECTIVE RATINGS*</u>

The following questions relate to summary ratings that consider the entire set of performance objectives. Consider all objectives to make your rating.

18. Number of objectives in performance plan

19. Are any of the objectives developmental in nature?

A. <u>Yes - Enter in Objective #(s).</u>
B. <u>No.</u> None of the objectives were developmental in nature.

20. Did all of the objectives relate to leadership or supervisory objectives (only if the performance plan under review was prepared for a supervisory employee)?

A. <u>Yes.</u> All of the objectives related to a leadership or supervisory objectives.
B. <u>No.</u> Not all of the objectives related to a leadership or supervisory objective.
C. <u>Not Applicable.</u> <u>(Plan was not for a Supervisor).</u>

21. Did any of the objectives specifically address customer or employee (for supervisors) perspectives?

A. <u>Yes - Enter in Objective #(s).</u>
B. <u>No.</u> None of the objectives addressed customer or employee perspectives.

Examples of measures of employee perspective include determining the level of:
- *Performance management and recognition employees receive for doing quality work;*
- *Information employees receive from management on what's going on in the organization; and*
- *Training and career development opportunities that employees have.*

Customers may include a wide variety of public groups, other agencies and elements, other governments, and Congress. Examples of measures of customer perspective for line organizations are:
- *A defense unit reports that the intelligence provided to them met or exceeded their needs for successful mission accomplishment.*
- *Most customers report that critical mission support activities, like computer network management, are exceptionally competent and helpful.*
- *Agencies or elements respond that the regulations proposed by a policy work unit are flexible, easier to read and follow, and take into account the agency's or element's previous comments.*

Support organizations also will have goals and measures, but most of the customers of support organizations will likely be employees who work for the same agency or element. Examples of measures of customer perspective for various support organizations are:
- *Level of assistance received from their human resources office.*
- *Satisfaction of employees regarding the heating and cooling service provided in their building.*
- *Adequate instruction provided by the procurement office on new contracting procedures.*

SECTION 2 – PERFORMANCE ELEMENTS

PERFORMANCE ELEMENTS SUMMARY

22. The self-report of accomplishments for the performance elements (overall) were:

A. **Not Available.** Not enough information provided to make a rating.
B. **Vague.** Included vague examples of accomplishments, outcomes, and/or deficiencies.
C. **Specific but unrelated to elements.** Included specific examples of accomplishments, outcomes, and/or deficiencies that did not relate to the key behaviors in the performance elements.
D. **Specific and related to elements.** Included specific examples of accomplishments, outcomes, and/or deficiencies that relate to the key behaviors in the performance elements.

23. The Rating Official Evaluation supporting the performance element ratings overall was:

A. **Not Available.** Not enough information provided to make a rating.
B. **Vague.** Included vague examples of accomplishments, outcomes, and/or deficiencies.
C. **Specific but unrelated to elements.** Included specific examples of accomplishments, outcomes, and/or deficiencies that did not relate to the key behaviors in the performance elements.
D. **Specific and related to elements.** Included specific examples of accomplishments, outcomes, and/or deficiencies that relate to the key behaviors in the performance elements.

24. To what extent do the Rating Official's overall comments support the overall performance element rating?

A. **Not Available.** Not enough information provided to make a rating.
B. **Not supported.** The ratings received and associated justifications do not align with the performance expectations guidance.
C. **Partially supported.** The ratings received and associated justifications aligned with the performance expectations guidance.
D. **Fully Supported.** The ratings received and associated justifications aligned with the performance expectations guidance.

Appendix B: Performance Rating Descriptors

Performance Objectives and Element Rating Descriptors		
Performance Rating	**Objectives Descriptors**	**Element Descriptors**
5 – OUTSTANDING	The employee far exceeded expected results on the objective such that organizational goals were achieved that otherwise would not have been.	The employee consistently performed all key behaviors at an exemplary level on the element.
4 – EXCELLENT	The employee surpassed expected results in a substantial manner on the objective.	The employee demonstrated mastery-level performance of the key behaviors on the element.
3 – SUCCESSFUL	The employee achieved expected results on the assigned objective.	The employee fully demonstrated effective, capable performance of key behaviors for the performance element.
2 – MINIMALLY SUCCESSFUL	The employee only partially achieved expected results on the performance objective.	The employee's performance requires improvement on one or more of the key behaviors for the objective.
1 – UNACCEPTABLE	The employee failed to achieve expected results in one or more assigned performance objectives.	The employee' failed to adequately demonstrate key behaviors for the performance element.
NOT RATED (NR)	The employee did not have the opportunity to complete the objective because it became obsolete due to changing mission requirements or because of extenuating circumstances beyond the control of the employee and supervisor (e.g., resources diverted to higher priority programs, employee in long-term training, deployed, on leave without pay).	Not used for performance elements.